Cargèse Autumn School “Liquids at interfaces”

October 13th to October 19th 2019

LECTURERS :

Lydéric Bocquet *(LPS, CNRS/ENS Paris/PSL, Paris)* : "Nanofluidics and osmotic processes"

Ramin Golestanian *(Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany)* **:** "Phoretic active matter"

Elise Lorenceau *(LIPhy, CNRS/Grenoble Alpes University, France)* : "Flows in foams"

Frieder Mugele (University of Twente, The Netherlands) : "Experimental probes of surface interactions"

Roland Netz *(Free University of Berlin, Germany)* : "Liquid at interfaces from a numerician viewpoint"

Jacco Snoeijer *(University of Twente, The Netherlands)* : "Soft wetting"

ORGANIZERS:

Lionel Bureau, Sophie Grandet, Abdelhamid Maali, Josiane Parzych, Thomas Salez.

INSTITUT D'ÉTUDES SCIENTIFIQUES DE CARGÈSE

LAI

cnrs
dépasser les frontières

Liquids at interfaces

October 13th to October 19th 2019

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
08:00 - 08:30		Breakfast Institute	Breakfast Institute	Breakfast Institute	Breakfast Institute	Breakfast Institute	DEPARTURE Village IESC to Airport
08:30 - 09:00		Welcome	R.Golestanian				
09:00 - 09:30		R. Netz		E. Lorenceau	F. Mugele	F. Mugele	
09:30 - 10:00							
10:00 - 10:30			Coffee break				
10:30 - 11:00		Coffee break	R. Golestanian	Coffee break	Coffee break	Coffee break	
11:00 - 11:30		R. Golestanian		L. Bocquet	J. Snoeijer	J. Snoeijer	
11:30 - 12h00							
12:00 - 12h30			Lunch - Institute				
12:30 - 14:00		Lunch - Institute		Lunch - Institute	Lunch - Institute	Lunch - Institute	
14:00 - 14:30		R. Netz	R. Netz	F. Mugele	E. Lorenceau	E. Lorenceau	
14:00 - 15:00							
15:00 - 15:30							
15:30 - 16:00		Coffee break	Coffee break	Coffee break	Coffee break	Coffee break	
16:00 - 16:30		L. Bocquet	L. Bocquet	Discussions	Discussions	J. Snoeijer	
16:30 - 17:00							
17:00 - 17:30	Airport to Institute						
17:30 - 18h00		Discussions	Discussions			Closing	
18:00 - 21:00		Welcome Reception	Conference Dinner				

Lectures Outlines

Roland Netz *:* “Liquid at interfaces from a numerician viewpoint”

1. Interfacial water (theory & simulations)
2. Charged surfaces (theory & simulations)
3. Liquid Flow at surfaces (theory & simulations)

Lydéric Bocquet: "Nanofluidics and osmotic processes"

1. Ionic transport at the nanoscales: linear and non-linear transport
2. Flows at surfaces and in confinement (experiments)
3. Osmotic processes: fundamentals and applications

Ramin Golestanian : “Phoretic active matter”

1. What is Diffusiophoresis
2. Microscopic Theory of Diffusiophoresis
3. Self-diffusiophoresis
4. Stochastic Dynamics of Phoretically Active Particles
5. Experiments on Self-phoresis
6. Mixtures of Apolar Active Colloids: Active Molecules
7. Mixtures of Apolar Active Colloids: Stability of Suspensions
8. Polar Active Colloids: Moment Expansion
9. Non-equilibrium Dynamics of Active Enzymes
10. Phoresis on the Slow Lane: Trail-following Bacteria
11 Chemotaxis and Cell Division

Elise Lorenceau: “Flows in foams”

1. Foam structure.
2. Dissipation in foams
3. Non aqueous flow in foams

Frieder Mugele : “Experimental probes of surface interactions”

1. Physical principles of ion adsorption and charge regulation in electric double layers (focus on modeling starting with double layer theory, derivation of PB equation from variational energy minimization)
2. Probing surfaceforces and surface charges using Atomic Force Microscopy (focus on experimental techniques and specific results)
3. Physical principles of electrowetting (1h basic principles; 1h application examples)

Jacco Snoeijer *:* “Soft wetting”

1. Elastocapillary interfaces
2. Soft wetting statics
3. Soft wetting dynamics

Liquid at interfaces from a numerician viewpoint

Roland Netz

- Interfacial water (theory & simulations)
- Charged surfaces (theory & simulations)
- Liquid Flow at surfaces (theory & simulations)

○ Monday: 09h00-10h30
○ Monday : 14h00-15h30
○ Tuesday : 14h00-15h30

NANO LETTERS

Letter

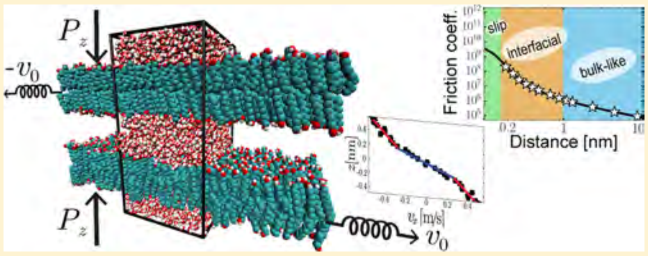
pubs.acs.org/NanoLett

Hydration Friction in Nanoconfinement: From Bulk via Interfacial to Dry Friction

Alexander Schlaich, Julian Kappler, and Roland R. Netz*

Fachbereich Physik, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany

Supporting Information



ABSTRACT: The viscous properties of nanoscopically confined water are important when hydrated surfaces in close contact are sheared against each other. Numerous experiments have probed the friction between atomically flat hydrated surfaces in the subnanometer separation regime and suggested an increased water viscosity, but the value of the effective viscosity of ultraconfined water, the mechanism of hydration layer friction, and the crossover to the dry friction limit are unclear. We study the shear friction between polar surfaces by extensive nonequilibrium molecular dynamics simulations in the linear-response regime at low shearing velocity, which is the relevant regime for typical biological applications. With decreasing water film thickness we find three consecutive friction regimes: For thick films friction is governed by bulk water viscosity. At separations of about a nanometer the highly viscous interfacial water layers dominate and increase the surface friction, while at the transition to the dry friction limit interfacial slip sets in. Based on our simulation results, we construct a confinement-dependent friction model which accounts for the additive friction contributions from bulklike water, interfacial water layers, and interfacial slip and which is valid for arbitrary water film thickness.

KEYWORDS: *Friction, hydration, nanoconfinement, viscosity, water, molecular dynamics simulations*

The viscous properties of confined water are relevant for colloidal jamming, surface drainage, lubrication of joints, and nanofluidics.[1−4] Consensus has been reached that the standard hydrodynamic description using bulk water viscosity works for surface separations above a few nanometers,[5−8] but the precise value of the effective viscosity of ultraconfined water in the subnanometer separation is not clear.[7,9−13] Apart from the value of the viscosity of ultraconfined water, a fundamental question that is difficult to address experimentally is how the friction between wet surfaces crosses over to the dry-friction limit, which is obtained for high loads when all water is squeezed out from between the surfaces. In fact, standard hydrodynamic theory predicts the shear friction between surfaces that are separated by a fluid layer to diverge as the fluid film thickness goes to zero. On the other hand, it is expected that friction forces are finite in the dry limit. It is difficult to probe this transition from wet to dry friction experimentally, because the high pressures needed to squeeze all water out cannot be easily generated with common experimental techniques. However, in biolubrication applications, the dry limit is important. For example, in joint lubrication, forces of the order of a few grams are sufficient to squeeze the water completely out if the surface area is small enough, as we show at the end of this paper. Thus, a microscopic picture of water-mediated nanoscale friction at the border between hydrodynamic friction and dry friction is clearly needed but at present missing.

Classical lubrication theory[14] describes the friction between wet solids as the crossover between the nonlinear hydrodynamic regime, where a finite surface separation under load is maintained by hydrodynamic lift and friction is governed by bulklike fluid viscosity, and the boundary regime, where surfaces touch at isolated contact zones and the friction properties are similar to dry friction as described by Amontons' heuristic law.[15] With the advent of experimental nanoscale techniques the so-called hydration friction scenario has moved into the focus: here, the water film is of nanoscopic thickness and stabilized not by hydrodynamic lift forces but rather by the hydration repulsion between hydrophilic surfaces.[16,17] This regime persists up to extremely high loads since the hydration pressure between sufficiently polar surfaces reaches several kbars at surface contact. Many experiments considered the idealized geometry of weakly curved flat surfaces and determined the force response to lateral or normal surface motion. Early measurements showed that the fluid viscosity stays bulklike down to surface separations of a few nanometers.[5−7] Later experiments also probed the subnanometer separation range and demonstrated the friction to change dramatically as the last water layers are squeezed out.[10−13,18−20] However, the experimental results are diverse, some showing a crossover between elastic and viscous response,[21] some

Received: May 12, 2017
Revised: September 11, 2017
Published: September 14, 2017

DOI: 10.1021/acs.nanolett.7b02000

suggesting confinement-induced freezing,[22] while others demonstrating pronounced oscillatory viscosity profiles at small surface separations.[23] The plethora of conflicting results and interpretations is not surprising: To measure friction between surfaces at nanometer separations, the surfaces must be atomically smooth, clean, and free of defects, since surface inhomogeneities give rise to strong friction contributions that typically are nonlinear in the shearing velocity.[12,24,25] In fact, only some of the experiments demonstrated to be in the linear friction regime,[10,26] which is a prerequisite to infer the Newtonian fluid viscosity from the shear-force response. As an additional complication, it was recently demonstrated that surface elasticity leads to a non-negligible and even dominating contribution to the measured strain−stress relation at nanometer separations, in particular if fluid drainage is imposed by normal relative surface motion.[27]

Simulations however can be designed to sidestep these pitfalls, as surface irregularities and contaminations can be avoided, and the perfect shear geometry of planar parallel surfaces is straightforwardly imposed. Indeed, early atomistic simulations demonstrated that water diffusivities are reduced at polar surfaces[19,28−32] and that the viscosity of water increases when under confinement.[29] The advantage of simulations in the idealized planar shear geometry is that not only the mean viscosity, averaged over the entire fluid film, can be deduced but that mechanistic insight into hydration friction can be gained from the locally resolved shear velocity profile across the confined water film.[31,33,34] In fact, effective viscosity profiles can be derived from time-averaged shear velocity profiles on the linear-response level (see Methods) and have broad applicability to other interface-dominated flow scenarios such as Poiseuille, electrophoretic/osmotic, and shear or drainage flow in ultra confinement. This forms the motivation and conceptual pivot point of the present work.

In our nonequilibrium molecular dynamics simulations we shear two polar decanol $CH_3(CH_2)_9OH$ bilayers against each other at imposed velocities $\pm v_0$, see Figure 1 (a) for the setup that closely mimics experiments with self-assembled monolayers.[13,24,25] The number N_w of water molecules is adjusted for each box height L_z such that the chemical potential equals bulk water[35] (see Methods). We are particularly interested in the transition from the hydrodynamic friction regime, obtained at large surface separation and where the friction force is dominated by the bulk water viscosity, to the dry boundary friction regime, where the normal pressure is of the order of 10 kbar and thus sufficient to squeeze all water out. In between we obtain the interfacial friction regime, where the friction is dominated by interfacial water layers with a substantially increased viscosity. Interestingly, and this is one of the main results of our paper, the effective viscosity within the interfacial layers, which have a thickness of about 0.4 nm for large surface separation, increases significantly as the water layer thickness goes down. Our simulations thus indicate a dynamic shear freezing transition of interfacial water upon ultra confinement. The effect of this dramatic viscosity increase on the shear friction is however rather mild, since with decreasing water layer thickness interfacial slip sets in, thus producing a smooth transition to the dry-friction regime that is characterized by a finite surface friction coefficient.

All our results are obtained in the linear friction regime, which we make painstakingly sure by monitoring the shear-velocity dependence of our results. In this limit, the standard friction coefficient,[15] defined as the ratio between friction and normal force, $\mu = F_f/F_n$, is linear in the shear velocity and has only anecdotal relevance in the present context: Trivially, it can be made arbitrarily low by decreasing the shear velocity. Rather, we here concentrate on the linear shear friction coefficient, $\gamma = F_f/(v_0 A)$, which results from the friction force by dividing by the shear velocity and the area A. The linear regime is for smooth surfaces obtained below shear velocities v_0 in the cm/s to m/s range and thus relevant for most biological shear scenarios.[3]

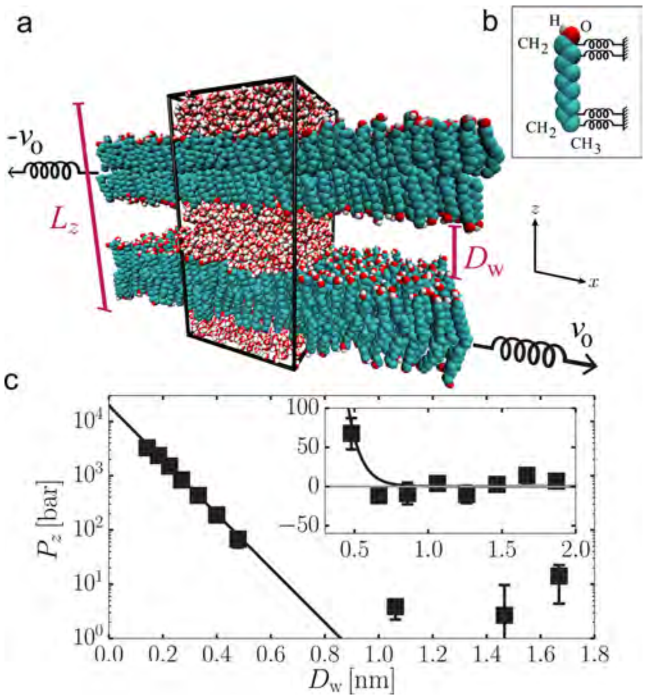


Figure 1. Simulation setup and confinement-dependent normal pressure. (a) Snapshot of the periodically replicated simulation system for total box height $L_z = 9.9$ nm and surface separation $D_w = 2.27$ nm. Shear is imposed by moving the two decanol bilayers at prescribed velocities v_0 against each other. (b) Chemical structure of decanol, positional restraints are indicated by springs. (c) Pressure-separation curve at fixed water chemical potential. The solid line is an exponential fit, and the inset shows the normal pressure P_z at large D_w in lin-lin presentation.

Hydration Pressure. Figure 1 (c) shows the normal pressure P_z versus the surface separation D_w, defined as $D_w = N_w v_w / A$, where $v_w = 0.0304$ nm^3 is the molecular bulk water volume, N_w is the number of water molecules per water slab, and $A = 5.198 \times 4.502$ nm^2 is the lateral area of the simulation box. This definition of D_w is equivalent to the distance between the water Gibbs-dividing surfaces (GDS) and allows for direct comparison with experiments, where the same definition is used.[17] In contrast to rigid surfaces, where oscillatory pressure profiles are seen in experiments[16] and in simulations,[36,37] our soft polar surfaces give rise to an exponential monotonic repulsion

$$P_z(D_w) = P_0 e^{-D_w/\lambda} \tag{1}$$

indicated by a solid line, with a decay length $\lambda = 0.87 \pm 0.02$ Å and an extrapolated dry pressure of $P_0 = 19 \pm 1$ kbar, quite similar to pressures in dehydrated phospholipid bilayer stacks.[17] For larger separation $D_w > 0.5$ nm, the interlayer pressure is essentially zero within our accuracy of about ±10 bar, as seen in the inset of Figure 1 (c).

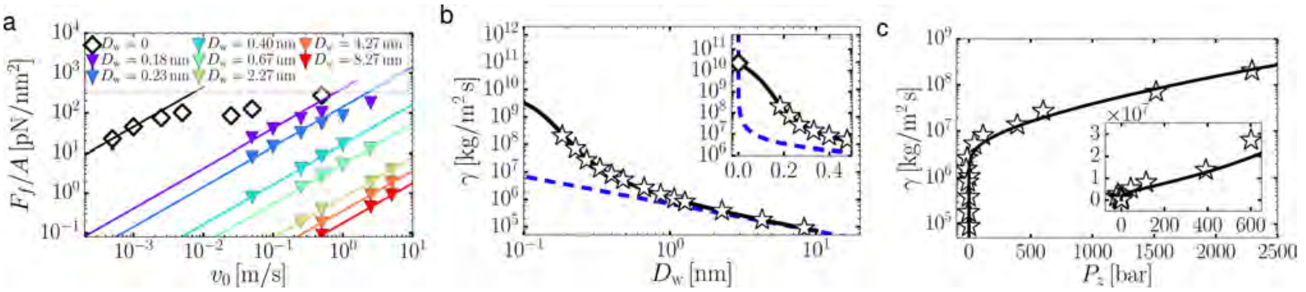


Figure 2. Confinement-dependent shear friction. (a) Friction force F_f versus shear velocity v_0 for different surface separations D_w. Solid lines are linear fits. (b) Simulated linear friction coefficient $\gamma = F_f/(v_0 A)$ (open symbols) as a function of D_w, compared with the bulk prediction $\gamma = 2\eta_b/D_w$ (dashed blue line) and the friction model eq 6 (solid black line). The inset shows the transition to dry friction at $D_w = 0$. (c) Friction coefficient γ versus normal pressure P_z; the line denotes the friction model eq 6 combined with eq 1.

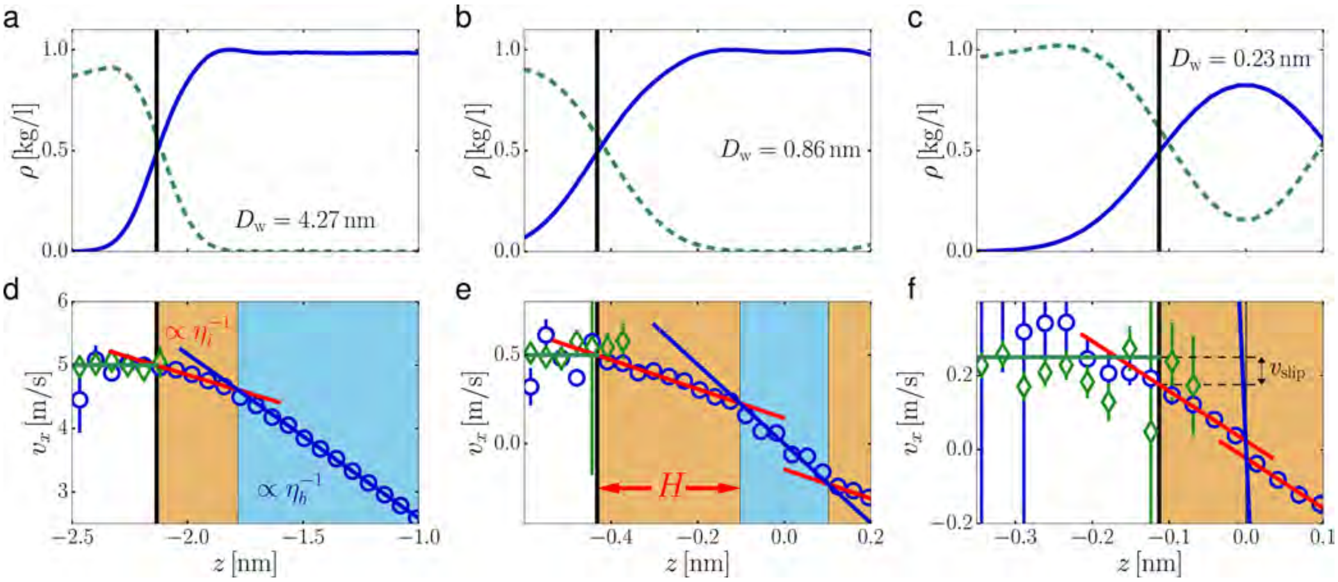


Figure 3. Simulated density and shear velocity profiles. Water (blue solid lines) and decanol (green broken lines) mass density profiles are shown in (a)−(c) for different surface separations D_w. Corresponding velocity profiles for water (blue spheres) and decanol (green diamonds) are shown in (d)−(f); the horizontal green lines denote the imposed shear velocity v_0. The blue straight lines denote the expected Couette bulklike velocity profile $v_x(z) = -zF_f/(A\eta_b)$, and the red straight lines denote fits to the interfacial velocity profile from which the interfacial viscosity η_i is deduced. The crossing between the blue and red lines defines the width H of the interfacial layer (shown in orange). For $D_w = 0.23$ nm in (f) a finite slip velocity $v_{\rm slip}$ at the decanol−water interface at $z = -D_w/2$ is observed.

Linear Friction Regime. The linear regime, where the friction force is proportional to shear velocity, is relevant for most practical situations, but in simulations (as well as in experiments) it is often not easily achieved. In Figure 2 (a) we plot the mean friction force F_f acting on the surfaces for a few values of D_w as a function of the imposed shear velocity v_0. For all D_w, we find for small enough v_0 a linear regime where F_f is proportional to v_0, as is expected from linear response theory. Decreasing D_w (corresponding to an increased vertical pressure load, as follows from Figure 1 (c)) at fixed v_0 leads to an enormous increase in the friction force by several orders of magnitude. Dry friction for $D_w = 0$ is studied in the absence of water at a pressure of $P_z \approx 10$ kbar (see the Supporting Information), and the linear response limit is reached for velocities smaller than about 1 mm/s. A simple scaling theory based on the thermally activated crossing over energetic barriers that have a typical length scale b predicts the transition from linear to nonlinear friction to occur at a characteristic friction force per area of $F_f^*/A \sim k_BT/b^3$.[38] For our system the length scale b corresponds to the lateral separation between decanol head groups, which is $b \simeq 0.5$ nm. With the thermal energy $k_BT \simeq 4 \times 10^{-21}$ J we obtain a crossover force of $F_f^*/A \sim 3 \times 10^7$ N/m^2 = 30 pN/(nm)^2. Looking at the data in Figure 2 (a) we see that this simple estimate correctly predicts the order of magnitude of the transition between the linear and nonlinear regimes. In particular, since it is the friction force that governs the onset of nonlinear effects, the actual shear velocities at the crossover between the linear and nonlinear regimes sensitively depend on the surface separation D_w. Using the simple scaling estimate for F_f^*/A it also transpires why the linear friction regime can be difficult to obtain experimentally: Assuming rough surfaces with a roughness wavelength of about $b \simeq 1$ μm, which is a typical value, the crossover force decreases down to $F_f^*/A \sim 4 \times 10^{-3}$ N/m^2, and the corresponding crossover shear velocity would be (assuming the same friction coefficient γ we find for dry decanol layers) $v_0^* \sim 10^{-13}$ m/s. Experiments at such low shear velocities are practically impossible to perform, thus our study is most relevant for atomically flat surfaces.

DOI: 10.1021/acs.nanolett.7b02000
Nano Lett. 2017, 17, 5969−5976

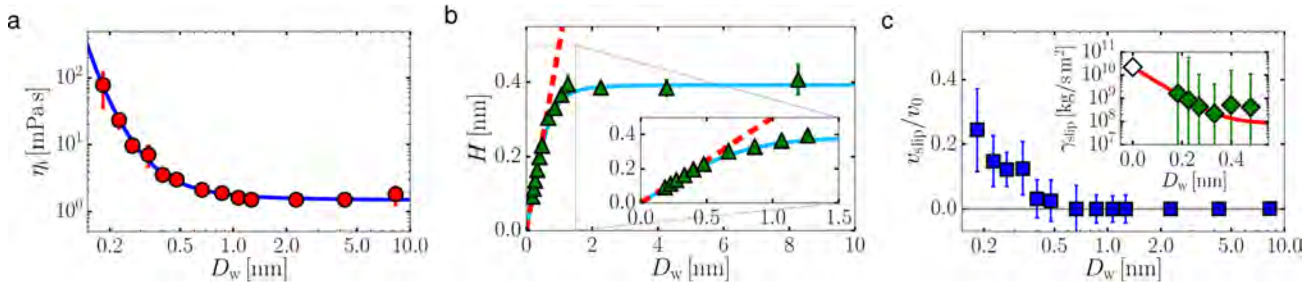


Figure 4. Parameters of the confinement-dependent friction model. (a) Interfacial viscosity η_i versus surface separation D_w. The solid line shows the function eq 3. (b) Interfacial water layer thickness H. The dashed red line shows the maximal value $H = D_w/2$; the solid line shows the function eq 4. (c) Interfacial slip v_{slip}/v_0. The inset shows the slip coefficient $\gamma_{slip} = (F_f/A)/v_{slip}$ together with the function eq 5.

Hydration-Dependent Linear Friction Coefficient. In Figure 2 (b) we show the linear friction coefficient $\gamma = F_f/(v_0 A)$ (data points) as a function of the surface separation D_w (the curve follows from our confinement-dependent friction model to be explained below). The results for γ, which are averaged over shear velocities in the linear regime where $F_f \sim v_0$, increase by 5 orders of magnitude as D_w decreases from $D_w = 8.1$ nm down to $D_w = 0$. The blue dashed line shows the hydrodynamic prediction based on water bulk viscosity, $\gamma = 2\eta_b/D_w$, where $\eta_b = 0.695 \pm 0.001$ mPa s is the SPC/E bulk water viscosity (independently determined from bulk simulations, see the Supporting Information, and close to the values reported by other authors[39−42]). Whereas the D_w^{-1} law is excellent for $D_w \gg 1$ nm, for smaller separations it fails in two aspects: First, γ increases stronger than the bulk prediction for $D_w < 1$ nm, which suggests an interfacial water layer with increased viscosity.[29,33,43] Second, the D_w^{-1} law diverges as $D_w \to 0$, in contrast to the finite dry friction coefficient $\gamma_{dry} = 2.2 \times 10^{10}$ kg m^{-2} s^{-1} obtained for $D_w = 0$ at the extrapolated contact pressure (see inset of Figure 2 (b)), which reflects a finite surface slip. Clearly, to understand these two opposing tendencies, we need an in-depth local understanding of interfacial friction.

From Shear-Velocity to Effective Viscosity Profiles. Figure 3 shows mass density (top) and velocity profiles (bottom) for a few different separations D_w, where $z = 0$ denotes the water slab center. The water (in blue) and decanol (in green) densities exhibit pronounced interfacial broadening and overlap over an interfacial thickness of about 0.5 nm around the GDS position $z = \pm D_w/2$ (indicated by vertical black lines). The water density shows no oscillations, as expected for our relatively soft surfaces,[36,37] and for $D_w = 4.27$ nm and $D_w = 0.86$ nm reaches bulk water density in the water slab center. For $D_w = 0.23$ nm in Figure 3 (c), the maximal water density is significantly lower than in bulk, and the opposing decanol headgroups overlap.

The decanol velocity profiles in Figure 3 (d)−(f) (green diamonds) scatter around the imposed shear velocity v_0, as expected for an elastic solid in the stationary shear limit. The water velocity profiles $v_x(z)$ (blue spheres) are more complex and allow for defining an effective water viscosity profile $\eta_{eff}(z)$ in the following manner: In the stationary limit and sufficiently far from the surface, the shear stress is constant in space and time, equals the externally applied friction force, and is given by

$$F_f = -A\eta_{eff}(z)\partial_z v_x(z) \tag{2}$$

In the Methods, eq 2 is rigorously derived from the nonlocal linear stress−strain relation using the conditions of stationarity and lateral homogeneity, also accounting for the finite width of the decanol−water interface. Contrary to an idealized hydrodynamic boundary, the surface water forces in our system act over a finite range, depending on the extent of decanol−water intermixing, as illustrated in Figure 3 (a)−(c), and on the range of intermolecular interactions. The resulting shear force profiles that act on the water molecules are discussed in the Supporting Information. In particular, $\eta_{eff}(z)$ and $v_x(z)$ are properly defined as averages over fluctuating microscopic quantities, and the derivation needs no locality or continuity assumption.

For large separation $D_w = 4.27$ nm in Figure 3 (d), the velocity profile in the water slab center corresponds to Couette flow, $v_x(z) = -zF_f/(A\eta_b)$, as indicated by the blue solid line. Close to the surface, the water velocity profile exhibits a smaller slope, which according to eq 2 indicates an increased interfacial effective water viscosity η_i. A linear fit denoted by the red line yields η_i as well as the interfacial layer width H, indicated by an orange slab that extends from the GDS position $z = -D_w/2$ to the crossing with the linear bulklike velocity profile at $z = -D_w/2+H$. Alternative fit models for the viscosity profile are possible but do not change our conclusions (see the Supporting Information). Note that for $D_w = 0.23$ nm in Figure 3 (f) the thickness of the bulklike water region is basically zero, while the water velocity at the GDS position $z = -D_w/2$ does not equal the imposed surface velocity v_0 but exhibits significant slip $v_{slip} = v_0 - v_x(-D_w/2)$. The emergence of a finite slip is not really surprising since in the dry limit $D_w = 0$ nm, i.e. in the absence of water, the opposing surfaces necessarily show complete slip, i.e. $v_{slip} = v_0$. Our simulations reveal that the dry friction limit is approached smoothly by a gradually increasing slip as $D_w \to 0$.

The derived interfacial viscosity η_i is shown in Figure 4 (a) as a function of D_w together with the heuristic expression

$$\eta_i = \eta_{i,0}(e^{\lambda_i/D_w} - 1) + \eta_{i,\infty} \tag{3}$$

which for small D_w increases significantly. This confirms previous reports of frozen water in strong confinement.[22,44] The fit to the data yields $\eta_{i,0} = 0.1$ mPa s and $\lambda_i = 1.2$ nm. For large D_w the interfacial viscosity saturates at $\eta_{i,\infty} = 1.5 \pm 0.1$ mPa s, significantly higher than the bulk value, $\eta_b = 0.7$ mPa s, in qualitative agreement with previous simulation results for water at a single hydrophilic diamond surface.[33] An increase of water viscosity in the hydration layer around polar objects has also been inferred from electrokinetic[45] and NMR experiments[46] and thus seems to be a general feature.

The D_w-dependent interfacial water layer thickness H is shown in Figure 4 (b) and fitted with the heuristic expression

$$H = ((D_w/2)^{-3} + H_\infty^{-3})^{-1/3} \tag{4}$$

For small D_w, we obtain $H \simeq D_w/2$, as indicated by the dashed red line in Figure 4 (b). At large D_w, H saturates at $H_\infty = 0.39 \pm 0.01$ nm, similar to previous simulations.[33]

In Figure 4 (c) we show the decanol–water slip velocity v_{slip}, rescaled by the shear velocity v_0. For the smallest separation $D_w = 0.18$ nm significant slip of $v_{slip}/v_0 \simeq 0.24$ is observed. Clearly, full slip $v_{slip}/v_0 = 1$ is obtained in the dry limit $D_w \to 0$. In the inset we show the slip coefficient $\gamma_{slip} = (F_f/A)/v_{slip}$ together with the fit function

$$\gamma_{slip} = (\gamma_{dry} - \gamma_{slip,\infty})e^{-D_w/\lambda_{slip}} + \gamma_{slip,\infty} \tag{5}$$

At $D_w = 0$ the dry friction coefficient $\gamma_{dry} = F_f/(Av_0) = 2.2 \times 10^{10}$ kg m^{-2} s^{-1} is obtained from simulations of decanol bilayers in the absence of water at a pressure of $P_z \approx 10$ kbar (open diamond in Figure 4 (c)). The slip velocities v_{slip} are zero within our accuracy for $D_w > 0.5$ nm, we thus only fit γ_{slip} for $D_w < 0.5$ nm and obtain $\gamma_{slip,\infty} = 8 \times 10^7$ kg m^{-2} s^{-1} and $\lambda_{slip} = 0.07$ nm.

For our layered planar shear geometry, the inverse friction coefficient by integration of eq 2 separates into slip, interfacial, and bulk contributions according to

$$\frac{1}{\gamma} \equiv \frac{v_0}{F_f/A} = \frac{1}{\gamma_{slip}} + \frac{H}{\eta_i} + \frac{D_w/2 - H}{\eta_b} \tag{6}$$

This prediction is, using the heuristic expressions for η_i, H, and γ_{slip} in eqs 3–5, compared with the simulation data for γ in Figure 2 (b). By construction, the agreement is excellent. In Figure 5 we plot the simulated inverse friction coefficient $1/\gamma$ (white stars) and the prediction according to eq 6 (black line) as a function of D_w on logarithmic scales. Now the slip, interfacial, and bulk friction contributions in eq 6 are additive and denoted by green, red, and blue lines, respectively. The dominant friction mechanism depends on the hydration level: At large separations $D_w > 1$ nm, corresponding to more than $N_w/N_d \approx 3$ water molecules per decanol (see upper scale), bulk friction dominates, and the bulk hydrodynamic friction law $\gamma = 2\eta_b/D_w$ (blue broken line) is valid. In the ultraconfined regime, for $D_w < 0.15$ nm or below half a hydration water per surface group, slip becomes dominant and furnishes the smooth crossover to the dry friction limit, denoted by a horizontal green broken line to the left. For intermediate separations 0.15 nm < D_w < 1 nm the inverse friction is smaller than predicted by bulk hydrodynamics (broken blue line) and dominated by the interfacial water layers of thickness H with an increased viscosity η_i. The confinement-dependent friction between planar smooth surfaces thus features three distinct friction modes.

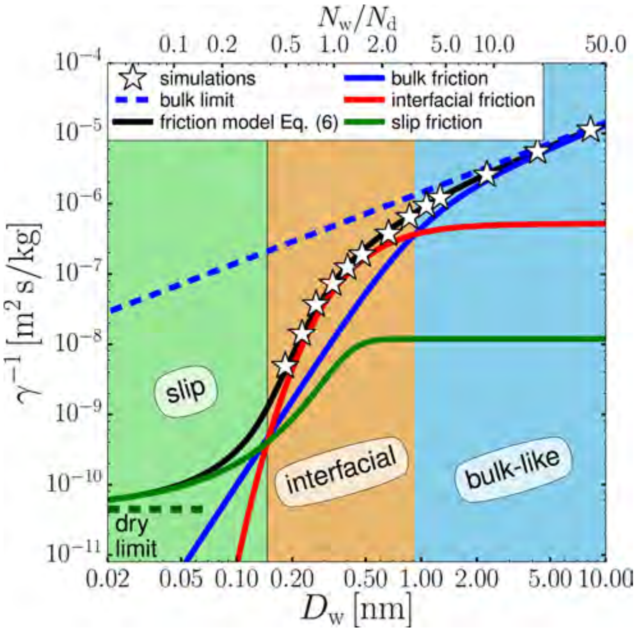


Figure 5. Confinement-dependent friction regimes. Comparison of the simulated inverse friction coefficient $1/\gamma$ (white stars) with the confinement-dependent friction model eq 6 (black line). Crossings between the slip, interface, and bulk friction contributions (green, red, and blue lines) define three distinct friction regimes. The bulk friction prediction is denoted by a blue broken line; the dry friction limit $D_w = 0$ is denoted by a horizontal green broken line.

In experiments the friction force is typically reported for fixed shearing or sliding velocity as a function of the normal load.[24] To connect to our results, in Figure 2 (c) we plot the velocity-independent linear friction coefficient γ versus the normal pressure P_z. In the logarithmic main plot γ is seen to vary steeply for small pressures $P_z < 50$ bar. The inset reveals quasi-Amontons-like scaling $\gamma \sim P_z$ for intermediate pressures 50 bar < P_z < 500 bar, similar to experimental reports.[24] The reason for this is that both P_z and γ decrease for small D_w in an approximately exponential manner (see Figure 1 (c) and inset in Figure 2 (b)). We note that the mechanism by which we observe the quasi-Amontons scaling $\gamma \sim P_z$ is very different from the mechanism that governs the friction between macroscopic rough dry surfaces, which is described by the heuristic Amontons law.[15] In the latter case it is the combination of roughness and plasticity that makes the friction force proportional to the normal force.

While Amontons law holds for dry as well as weakly wetted rough surfaces at elevated shear velocities, where the friction force does not depend on shearing velocity, the friction between rough surfaces in the presence of a lubricating fluid typically shows a pronounced shear velocity dependence. In the so-called boundary lubrication regime of the Stribeck curve, the friction force decreases drastically with the sliding velocity within a characteristic finite velocity window.[14] This regime is caused by hydrodynamic lift forces which push the surfaces away from each other and thereby eliminate contact zones between the rough surfaces. At even larger shearing velocities the friction force typically increases with the shearing velocity due to the balance of normal load and hydrodynamic lift force.[14] We note that these effects are caused by the combination of roughness and hydrodynamic lift effects at large shear velocities, whereas our results are relevant in the linear response regime of small shear velocities where the surface separation results from the balance of normal load and hydration repulsion.

Although derived for planar surfaces, our confinement-dependent viscosity model eq 6 is also relevant for rough and weakly curved surfaces, which we illustrate next. The crossover between the slip and interfacial friction regimes occurs close to $D_w \approx 0.2$ nm (see Figure 5), corresponding to an enormous hydration pressure of about $P_z \approx 1$ kbar (see Figure 1 (c)). In fact, such a pressure is produced for a contact area of $A = 100$ μm^2 by the modest weight of 1 g. Thus, our crossover scenario from bulk over interfacial to slip friction is very relevant for rough surfaces, since local pressures in small contact zones can be enormous. With a friction coefficient value $\gamma \approx 10^8$ kg/(m^{2}s) for $D_w \approx 0.2$ nm (from Figure 5) we obtain for a shear velocity

$v_0 = 1$ m/s (which for $D_w \approx 0.2$ nm is at the border between the linear and nonlinear friction regimes, as seen in Figure 2 (a)) a friction force per area of $F_f/A = \gamma v_0 = 10^8$ Pa. The ratio of friction and normal force is therefore equal to $\mu = F_f/(AP_z) = 1$ for $P_z \approx 1$ kbar and $v_0 = 1$ m/s and is linear in v_0. Thus, for a reduced shear velocity of $v_0 = 1$ mm/s and the same pressure we obtain the drastically reduced ratio $\mu = F_f/(AP_z) = 0.001$, a typical value measured in hydration friction experiments.[47] Similar to these estimates, friction forces for arbitrary loads and shear velocities can be derived directly from our confinement-dependent viscosity model eq 6.

Methods. *MD Simulations.* We simulate two bilayers consisting of 4 × 100 polar decanol molecules, separated by two water slabs, see Figure 1 (a) for a snapshot. Force-field parameters are based on GROMOS53A6,[48] decanol hydroxyl groups are represented in atomistic detail, and CH_2 and CH_3 groups are represented as united atoms. The repulsion between headgroup oxygens is increased to reduce intrasurface hydrogen bonding.[37] The SPC/E water model[49] is employed. Simulations are performed using version 5.0 of the GROMACS simulation package[50] at $T = 300$ K with periodic boundary conditions and are analyzed using the MDAnalysis package.[51]

The decanol molecules are fixed on a hexagonal lattice by harmonic position restraints. To allow for shear in the x direction, the first two and last two carbon-groups are fixed only in y and in z directions with harmonic constants $k_y = 500$ kJ/mol and $k_z = 10$ kJ/mol, as indicated in Figure 1 (b). The lateral area of $A = 5.198$ nm × 4.502 nm of the simulation box corresponds to the tensionless state in vacuum with an area per chain $a = A/N_d = 0.234$ nm^2, where $N_d = 100$ is the number of decanol molecules per monolayer. Lennard-Jones interactions are truncated at $r_c = 0.9$ nm; for the electrostatic interactions the Particle Mesh Ewald method[52] is employed with a real-space cutoff $r_c = 0.9$ nm.

Constant Velocity Pulling. Shear is achieved by pulling the two bilayers against each other at constant velocity v_0. To this end the time-dependent potential

$$U(t) = \frac{k}{2}(l_x(t) - (l_x(0) - 2v_0 t))^2 \tag{7}$$

acts on the center of mass (COM) of each bilayer, where k is the force constant, l_x is the distance between the COMs of both bilayers in the x direction, and t is the simulation time. The resulting shear force $F_f(t) = -\mathrm{d}U(t)/\mathrm{d}l_x(t)$, weighted by the atomic mass, acts on all bilayer atoms. The force constant k has to be carefully adjusted, see the Supporting Information for technical details. All simulations are performed for at least 1 μs, for $D_w < 0.25$ nm typically for 5 μs. Error estimates for the shear force F_f and the velocity profile $v_x(z)$ are obtained via block-averaging.

Thermodynamic Extrapolation. The total number of water molecules in each water slab is varied between $N_w = 0$ and $N_w = 6368$. First, simulations without shear are performed in the $N_w A L_z T$ ensemble at constant volume $V = AL_z$, and the number of water molecules N_w is adjusted via thermodynamic extrapolation to yield a constant water chemical potential (see the Supporting Information and refs 35 and 53 for details). In order to obtain a pressure resolution of about 10 bar, the accumulated simulation time per data point in Figure 1 (c) is about 6 μs. Simulations at finite shear are then performed at the same box size and water number.

Definition of a Local Effective Viscosity Profile. In the literature one can often read the statement that hydrodynamic approaches become invalid at scales corresponding to the molecular size, which for the case of water would be the nanometer scale. Although the statement is correct in many situations, it is typically based on the spurious notions that hydrodynamic formulations are necessarily (i) continuous and (ii) described by homogeneous, bulklike, and local shear-strain response functions. We here show how hydrodynamic equations, which reflect the conservation of momentum, a condition that arguably holds at any length scale, in conjunction with nonlocal linear-response theory, allow to define a spatially dependent effective viscosity profile without invoking any locality or continuum assumption. As we show, the main reason that the standard hydrodynamic description fails at the nanoscale is not related to the discreteness of water molecules or to fluctuating densities but rather to the finite range of surface-fluid interactions, for which we propose a simple fix in terms of a suitable definition of an effective viscosity profile.

We start with the momentum balance equation of the i-th Cartesian component for a viscous fluid,

$$\rho(\mathbf{r}, t)\frac{\mathrm{d}v_i(\mathbf{r}, t)}{\mathrm{d}t} = f_i(\mathbf{r}, t) + \nabla_j \sigma_{ij}(\mathbf{r}, t) \tag{8}$$

where $\mathbf{f}$ is the local external force density acting on the fluid, σ is the stress tensor, and $\mathrm{d}/\mathrm{d}t = \partial/\partial t + v_j(\mathbf{r},t)\nabla_j$ is the substantial time derivative. Note that indices that doubly appear are summed over. We consider a stationary state but at the same time allow the instantaneous velocity, stress tensor, density, and force to fluctuate around their stationary values. Averaging the entire equation over the lateral coordinates x and y as well as over time, we are left with the time independent equation for the averaged x-component

$$0 = \bar{f}_x(z) + \nabla_z \bar{\sigma}_{xz}(z) \tag{9}$$

where the averaged force and stress profiles are denoted by $\bar{f}_x(z)$ and $\bar{\sigma}_{xz}(z)$. All other components are zero because of the surface boundary conditions and lateral translational invariance. Integrating this equation once we obtain

$$0 = \bar{G}_x(z) + \bar{\sigma}_{xz}(z) \tag{10}$$

where the external shear stress $\bar{G}_x(z) = \int_{-\infty}^{z} \mathrm{d}z' \bar{f}_x(z')$ results from the integral over the force density profile. We consider here the simplified geometry of a single interface where the moving solid is located at $z < 0$ and the fluid at $z > 0$. For the shear scenario, the total friction force F_f acting on the liquid is given by $\bar{G}_x(\infty) = F_f/A$ (note that for pressure or electric-field driven flow scenarios, which we do not explicitly consider here, the external force acting on the fluid is counteracted by the stationary surface and in that case $\bar{G}_x(\infty) = 0$). From eq 10 we see that in the region where the force density $\bar{f}_x(z)$ is zero, the external stress $\bar{G}_x(z)$ and thus also the fluid stress $\bar{\sigma}_{xz}(z)$ are constant.

The linear but nonlocal response relation between velocity gradient and stress is (for the relevant xz-component) given by

$$\partial_z v_x(\mathbf{r}, t) = \int \mathrm{d}\mathbf{r}'\mathrm{d}t' \eta_{\mathrm{nl}}^{-1}(\mathbf{r}, \mathbf{r}', t, t')\sigma_{xz}(\mathbf{r}', t') \tag{11}$$

We assume that the nonlocal response function is homogeneous in time and the lateral spatial coordinates and thus can be written as $\eta_{\mathrm{nl}}^{-1}(x - x', y - y', z, z', t - t')$, which is a nontrivial assumption. Averaging the entire equation over x, y, t we obtain

$$\partial_z \bar{v}_x(z) = \int \mathrm{d}\mathbf{r}'\mathrm{d}t' \eta_{\mathrm{nl}}^{-1}(z, z')\sigma_{xz}(\mathbf{r}', t') \tag{12}$$

We can now perform the remaining integrals over x', y', and t' and obtain

$$\partial_z \bar{v}_x(z) = \int \mathrm{d}z' \eta_{\mathrm{nl}}^{-1}(z, z')\bar{\sigma}_{xz}(z') \tag{13}$$

If we assume the stress $\bar{\sigma}_{xz}(z')$ to be constant in space, we can perform the remaining integral over z' and obtain

$$\bar{\sigma}_{xz} = \eta(z)\partial_z \bar{v}_x(z) \tag{14}$$

where

$$1/\eta(z) \equiv \int \mathrm{d}z' \eta_{\mathrm{nl}}^{-1}(z, z') \tag{15}$$

defines the local viscosity profile. We see from eq 10 that in this scenario the external stress profile $\bar{G}_x(z)$ is constant and given by $\bar{G}_x(z) = F_\mathrm{f}/A$, in other words, the surface-fluid interactions are assumed to have a vanishing range. By combining this with eq 10 and eq 14 we arrive at eq 2 in the main text, where in this simplified case there is no difference between the effective viscosity profile $\eta_{\mathrm{eff}}(z)$ that appears in eq 2 and the viscosity profile $\eta(z)$ defined in eq 15. Note that in order to arrive at this equation, we did not assume the viscosity response function to be local, rather, the viscosity profile $\eta(z)$ involves an integral of the nonlocal viscosity response function over space and thus fully includes interfacial effects and the nonlocality of the viscosity response. If on the other hand $\bar{G}_x(z)$ and thus $\bar{\sigma}_{xz}(z)$ vary on length scales comparable to the length scales over which $\partial_z \bar{v}_x(z)$ varies, we define

$$1/\eta(z) \equiv \int \mathrm{d}z' \eta_{\mathrm{nl}}^{-1}(z, z')\bar{\sigma}_{xz}(z')/\bar{\sigma}_{xz}(z) \tag{16}$$

and from eq 13 obtain the modified formula

$$\bar{\sigma}_{xz}(z) = \eta(z)\partial_z \bar{v}_x(z) \tag{17}$$

Note that in this case the viscosity profile $\eta(z)$ depends on details of the external shear stress profile $\bar{G}_x(z) = -\bar{\sigma}_{xz}(z)$, which reflects the finite range of the surface-fluid interactions. Combining eq 17 with eq 10 we see that the inverse slope of the shear velocity profile corresponds in this more realistic scenario to an effective viscosity profile $\eta_{\mathrm{eff}}(z)$ which incorporates the effects of the external stress profile $\bar{G}_x(z)$ according to

$$\eta_{\mathrm{eff}}(z) \equiv \frac{\eta(z)F_\mathrm{f}/A}{\bar{G}_x(z)} = -\frac{F_\mathrm{f}/A}{\partial_z \bar{v}_x(z)} \tag{18}$$

This is the general derivation of eq 2 in the main text. In the limit $G_x(z) = F_\mathrm{f}/A$ we see that the effective viscosity profile $\eta_{\mathrm{eff}}(z)$, which includes the effects of a spatially varying external stress profile $\bar{G}_x(z)$, and the bare viscosity profile $\eta(z)$ defined in eq 15 are the same. Note that the definition of the effective viscosity profile $\eta_{\mathrm{eff}}(z)$ is valid at arbitrary scales since all quantities are averages over microscopic, fluctuating quantities. Also, the effective viscosity profile does not need to be continuous but instead can consist of multiple steps and discontinuities.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.nanolett.7b02000.

> Further data on dry friction, calculation of the bulk viscosity and results from alternative fitting models, details on constant velocity pulling and the thermodynamic extrapolation technique, influence of the chemical potential on friction, details of shear force and viscosity profiles (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: rnetz@physik.fu-berlin.de.

ORCID

Alexander Schlaich: 0000-0002-4250-363X

Author Contributions

A.S. and R.R.N. designed the research project; A.S. performed the simulations; A.S., J.K., and R.R.N. developed analytical tools and analyzed the data; A.S. and R.R.N. wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the Deutsche Forschungsgemeinschaft (DFG) for support via grants NE810/10 and SFB 1114.

■ REFERENCES

(1) Comtet, J.; Chatté, G.; Niguès, A.; Bocquet, L.; Siria, A.; Colin, A. *Nat. Commun.* **2017**, *8*, 15633.

(2) Lin, N. Y. C.; Guy, B. M.; Hermes, M.; Ness, C.; Sun, J.; Poon, W. C. K.; Cohen, I. *Phys. Rev. Lett.* **2015**, *115*, 228304.

(3) Wang, H.; Ateshian, G. A. *J. Biomech.* **1997**, *30*, 771−776.

(4) Urbakh, M.; Meyer, E. *Nat. Mater.* **2010**, *9*, 8−10.

(5) Chan, D. Y. C.; Horn, R. G. *J. Chem. Phys.* **1985**, *83*, 5311−5324.

(6) Israelachvili, J. N. *J. Colloid Interface Sci.* **1986**, *110*, 263−271.

(7) Raviv, U.; Laurat, P.; Klein, J. *Nature* **2001**, *413*, 51−54.

(8) Ma, L.; Gaisinskaya-Kipnis, A.; Kampf, N.; Klein, J. *Nat. Commun.* **2015**, *6*, 6060.

(9) Ortiz-Young, D.; Chiu, H.-C.; Kim, S.; Voïtchovsky, K.; Riedo, E. *Nat. Commun.* **2013**, *4*, 2482.

(10) Zhu, Y.; Granick, S. *Phys. Rev. Lett.* **2001**, *87*, 096104.

(11) Antognozzi, M.; Humphris, A. D. L.; Miles, M. J. *Appl. Phys. Lett.* **2001**, *78*, 300−302.

(12) Pittenger, B.; Fain, S. C.; Cochran, M. J.; Donev, J. M. K.; Robertson, B. E.; Szuchmacher, A.; Overney, R. M. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2001**, *63*, 134102.

(13) Kim, H. I.; Kushmerick, J. G.; Houston, J. E.; Bunker, B. C. *Langmuir* **2003**, *19*, 9271−9275.

(14) Bowden, F.; Tabor, D. *The Friction and Lubrication of Solids*; Clarendon Press: Oxford, 1950.

(15) Gao, J.; Luedtke, W. D.; Gourdon, D.; Ruths, M.; Israelachvili, J. N.; Landman, U. *J. Phys. Chem. B* **2004**, *108*, 3410−3425.

(16) Israelachvili, J. N.; Pashley, R. M. *Nature* **1983**, *306*, 249−250.

(17) Parsegian, V.; Fuller, N.; Rand, R. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76*, 2750−2754.

(18) Sakuma, H.; Otsuki, K.; Kurihara, K. *Phys. Rev. Lett.* **2006**, *96*, 046104.

(19) Li, T.-D.; Gao, J.; Szoszkiewicz, R.; Landman, U.; Riedo, E. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2007**, *75*, 115415.

(20) Labuda, A.; Kobayashi, K.; Suzuki, K.; Yamada, H.; Grütter, P. *Phys. Rev. Lett.* **2013**, *110*, 066102.

(21) Khan, S. H.; Matei, G.; Patil, S.; Hoffmann, P. M. *Phys. Rev. Lett.* **2010**, *105*, 106101.
(22) Dhopatkar, N.; Defante, A. P.; Dhinojwala, A. *Science Advances* **2016**, *2*, e1600763.
(23) Maali, A.; Cohen-Bouhacina, T.; Couturier, G.; Aimé, J.-P. *Phys. Rev. Lett.* **2006**, *96*, 086105.
(24) Clear, S. C.; Nealey, P. F. *J. Colloid Interface Sci.* **1999**, *213*, 238−250.
(25) Chen, J.; Ratera, I.; Park, J. Y.; Salmeron, M. *Phys. Rev. Lett.* **2006**, *96*, 236102.
(26) Major, R. C.; Houston, J. E.; McGrath, M. J.; Siepmann, J. I.; Zhu, X.-Y. *Phys. Rev. Lett.* **2006**, *96*, 177803.
(27) Villey, R.; Martinot, E.; Cottin-Bizonne, C.; Phaner-Goutorbe, M.; Léger, L.; Restagno, F.; Charlaix, E. *Phys. Rev. Lett.* **2013**, *111*, 215701.
(28) Lee, S. H.; Rossky, P. J. *J. Chem. Phys.* **1994**, *100*, 3334−3345.
(29) Leng, Y.; Cummings, P. T. *Phys. Rev. Lett.* **2005**, *94*, 026101.
(30) Lane, J. M. D.; Chandross, M.; Stevens, M. J.; Grest, G. S. *Langmuir* **2008**, *24*, 5209−5212.
(31) Lorenz, C. D.; Chandross, M.; Lane, J. M. D.; Grest, G. S. *Modell. Simul. Mater. Sci. Eng.* **2010**, *18*, 034005.
(32) Fedyanin, I.; Pertsin, A.; Grunze, M. *J. Chem. Phys.* **2011**, *135*, 174704.
(33) Sendner, C.; Horinek, D.; Bocquet, L.; Netz, R. R. *Langmuir* **2009**, *25*, 10768−10781.
(34) Boţan, A.; Joly, L.; Fillot, N.; Loison, C. *Langmuir* **2015**, *31*, 12197−12202.
(35) Schneck, E.; Sedlmeier, F.; Netz, R. R. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14405−14409.
(36) Lorenz, C. D.; Lane, J. M. D.; Chandross, M.; Stevens, M. J.; Grest, G. S. *Langmuir* **2009**, *25*, 4535−4542.
(37) Kanduč, M.; Schneck, E.; Netz, R. R. *Chem. Phys. Lett.* **2014**, *610−611*, 375−380.
(38) Erbaş, A.; Horinek, D.; Netz, R. R. *J. Am. Chem. Soc.* **2012**, *134*, 623−630.
(39) Song, Y.; Dai, L. L. *Mol. Simul.* **2010**, *36*, 560−567.
(40) González, M. A.; Abascal, J. L. F. *J. Chem. Phys.* **2010**, *132*, 096101.
(41) Fanourgakis, G. S.; Medina, J. S.; Prosmiti, R. *J. Phys. Chem. A* **2012**, *116*, 2564−2570.
(42) Tazi, S.; Boţan, A.; Salanne, M.; Marry, V.; Turq, P.; Rotenberg, B. *J. Phys.: Condens. Matter* **2012**, *24*, 284117.
(43) Thompson, P. A.; Robbins, M. O. *Phys. Rev. A: At., Mol., Opt. Phys.* **1990**, *41*, 6830−6837.
(44) Zangi, R.; Mark, A. E. *Phys. Rev. Lett.* **2003**, *91*, 025502.
(45) Bonthuis, D. J.; Netz, R. R. *Langmuir* **2012**, *28*, 16049−16059.
(46) Halle, B.; Davidovic, M. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 12135−12140.
(47) Briscoe, W. H. *Curr. Opin. Colloid Interface Sci.* **2017**, *27*, 1−8.
(48) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656−1676.
(49) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(50) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. *SoftwareX* **2015**, *1−2*, 19−25.
(51) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. *J. Comput. Chem.* **2011**, *32*, 2319−2327.
(52) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(53) Kanduč, M.; Schlaich, A.; Schneck, E.; Netz, R. R. *Langmuir* **2016**, *32*, 8767−8782.

DOI: 10.1021/acs.nanolett.7b02000
Nano Lett. 2017, 17, 5969−5976

LANGMUIR
Article
Cite This: *Langmuir* 2019, 35, 551−560
pubs.acs.org/Langmuir

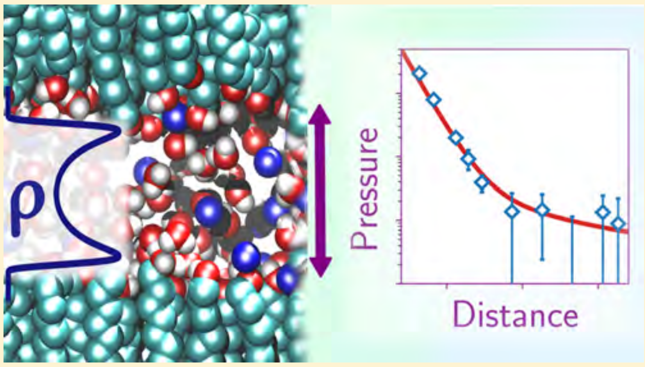
Simulations of Nanoseparated Charged Surfaces Reveal Charge-Induced Water Reorientation and Nonadditivity of Hydration and Mean-Field Electrostatic Repulsion

Alexander Schlaich,*[†,‡] Alexandre P. dos Santos,[‡,§] and Roland R. Netz*[‡]

[†]Laboratoire Interdisciplinaire de Physique, CNRS and Université Grenoble Alpes, UMR CNRS 5588, 38000 Grenoble, France
[‡]Fachbereich Physik, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany
[§]Instituto de Física, Universidade Federal do Rio Grande do Sul, CEP 91501-970, 15051 Porto Alegre, Brazil

Supporting Information

ABSTRACT: We perform atomistic simulations of nanometer-separated charged surfaces in the presence of monovalent counterions at fixed water chemical potential. The counterion density profiles are well described by a modified Poisson−Boltzmann (MPB) approach that accounts for nonelectrostatic ion−surface interactions, while the effects of smeared-out surface-charge distributions and dielectric profiles are found to be relatively unimportant. The simulated surface interactions are for weakly charged surfaces well described by the additive contributions of hydration and MPB repulsions, but already for a moderate surface charge density of $\sigma = -0.77$ e/nm^2 this additivity breaks down. This we rationalize by a combination of different effects, namely, counterion correlations as well as the surface charge-induced reorientation of hydration water, which modifies the effective water dielectric constant as well as the hydration repulsion.



INTRODUCTION

Many biologically and industrially relevant surfaces are charged in water; classical examples are lipid membranes,[1−3] ionic surfactant layers,[4] and solid surfaces such as glass, silica, or mica.[5−8] The experimental and theoretical descriptions of the interaction between charged surfaces across aqueous electrolytes forms the foundation of colloidal science. The celebrated Poisson−Boltzmann (PB) theory[9] relies on a mean-field approximation for the ion distribution and becomes valid when surface charge density and ion valencies are low and thus ion correlations are negligible. The usual formulation of PB theory treats water as a dielectric continuum described by a local and spatially homogeneous dielectric constant. According to PB theory, the interaction pressure between similarly charged surfaces is always repulsive. Numerous experiments confirmed this prediction and found an exponentially decaying surface pressure for large surface separation with a decay length that depends on the added salt concentration.[5−8]

Surface separations in the nanometer range have been investigated in experiments and simulations for systems such as silica,[10−14] clay,[15,16] or membrane stacks.[17−21] At such low surface separations, an additional, exponentially decaying repulsive pressure contribution is present, which is similar to the hydration pressure found for soft polar surfaces with a zero net charge.[17−20] Experimental pressures between charged surfaces have been successfully fitted by assuming additivity of hydration and PB contributions.[22−25] However, such fits are of only limited persuasive power since the surface charge density and its location are in most cases mere fitting parameters.

Indeed, additional effects, that are particularly important for nanometer surface separations, suggest essential modifications of the traditional PB theory: (i) Water confined in nanometer slabs exhibits dielectric properties distinctly different from bulk.[26−29] This is also suggested by a modified interfacial water structure inferred from nonlinear spectroscopy.[30−32] (ii) Surface charge distributions are neither laterally homogeneous nor sharply peaked normal to the surface, as typically assumed in PB modeling, but rather are discrete[33] and broadly distributed.[26] (iii) Ions interact with charged as well as uncharged surface groups via ion-surface interactions which involve surface-induced partial ion dehydration[34,35] and lead to ion-specific Hofmeister effects[36] and modifications of surface interactions.[14,37] (iv) Finally, ion correlation effects, not included in PB theory, become relevant for highly charged surfaces and high ion valency, and make similarly charged walls attract each other,[38] in strong contrast to the mean-field PB predictions.

Received: October 16, 2018
Revised: December 14, 2018
Published: December 20, 2018

DOI: 10.1021/acs.langmuir.8b03474

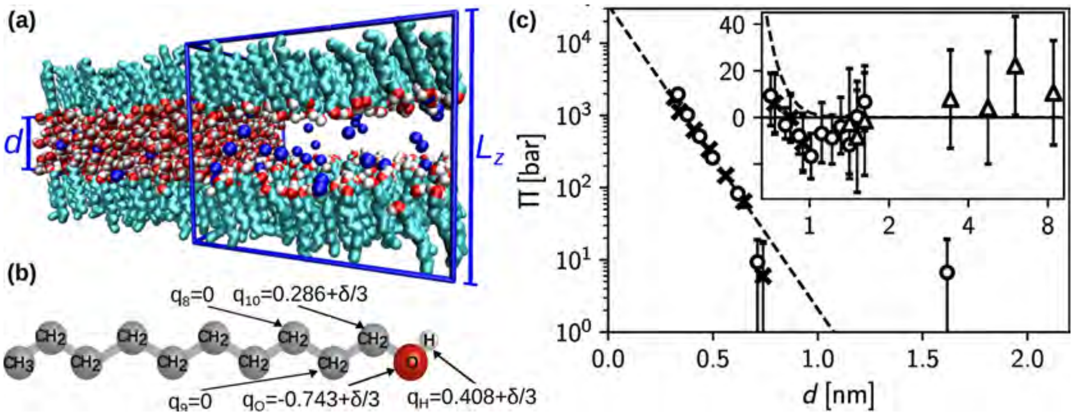


Figure 1. (a) Simulation snapshot for surface charge density $\sigma = -0.77$ e/nm^2 with $N_{ion} = 18$ Na$^+$ ions and box height $L_z = 4.05$ nm, leading to a surface separation $d = 1.5$ nm. Water molecules are not shown in the central box. (b) Partial charge distribution on a decanol headgroup with a net charge δ. (c) Interaction pressure between neutral decanol surfaces for $N_d = 100$ (circles) and $N_d = 196$ decanols per surface (crosses). The dashed line is a fit of eq 6 to all positive pressure data yielding a decay length $\lambda = 0.10$ nm. In the large distance regime, in the inset, the pressure is zero within numerical accuracy. Triangles denote results for the equivalent osmotic pressure $\Pi = \Pi_0 + (\mu_b - \mu_w)/v_w^b$ in the alternative $N_w A \Pi_0 T$ ensemble for fixed $\Pi_0 = 1$ bar; the see Supporting Information for details.

Water-explicit simulations of charged surfaces in principle include all these effects and should thus allow for a crucial test of the assumption of additive hydration and PB pressures, in particular since the surface charge and its location are precisely known. Previous simulations reported ion density profiles between nanometer-separated charged surfaces.[10−14,26] However, technical difficulties to impose a constant water chemical potential, which is the experimentally realized ensemble, precluded so far the quantitative comparison of surface pressures from water-explicit simulations with theoretical predictions.

In our simulations, we fix the water chemical potential by the thermodynamic extrapolation method,[39,40] which we previously introduced for the study of the hydration repulsion between neutral lipid membranes,[39] and compare simulated pressures between charged surfaces in the presence of monovalent counterions with theoretical predictions. For this we introduce a modified PB (MPB) theory that includes a general ion-surface interaction potential, an inhomogeneous dielectric profile and a surface charge distribution that is smeared-out in the direction normal to the surface. For low surface charge density we demonstrate that the simulated total pressure is very accurately described by the sum of the MPB pressure and the hydration pressure, the latter being extracted from corresponding simulations between uncharged surfaces. But already for a moderate surface charge density of $\sigma = -0.77$ e/nm^2, this additive description breaks down. Monte Carlo simulations of corresponding water-implicit model systems suggest the deviations between water-explicit simulations and MPB theory to be only partly due to ion correlation effects. By analysis of the water orientation profile, which is significantly perturbed by the presence of surface charges, the breakdown of the additivity assumption is suggested to be also due to a change of the effective water dielectric constant in confinement and presumably also due to a modification of the effective hydration pressure. In contrast to the interaction pressures, the counterion density profiles are for all studied surface charge densities well described by the MPB theory. Our study suggests that water structural effects are crucial for the quantitative modeling of the interaction between charged surfaces at the nanoscale, while MPB theory with suitably defined ion−surface interactions quantitatively describes ion-density profiles. This is important for the correct interpretation of experimental results for ion density profiles and interaction pressures between charged surfaces.

■ METHODS

Simulation Model. Our model surfaces consist of decanol bilayers with variable charges added to the head groups. Similar to experimental self-assembled monolayers (SAMs) on gold substrates, we fix the molecules on a centered rectangular lattice at a tilt angle of 30°; see Figure 1a for a snapshot. Two different lateral areas $A = L_x \times L_y$ with $N_d = 100$ or 196 decanols per monolayer are studied; see Table 1 for parameters. The water number in the $N_w A L_z T$ ensemble at constant temperature $T = 300$ K is adjusted for each box height L_z such that the water chemical potential μ_w equals the bulk value $\mu_b = (-48.25 \pm 0.02)$ kJ/mol, determined such that at large separations the pressure is zero within the error bars; see the Supporting Information for details. For an accurate interaction pressure Π, the error in μ_w has to be below $0.01 k_B T$, requiring at least 6 μs simulation time per data point. Force-field parameters are based on GROMOS53A6[41] where the decanol hydroxyl groups are represented in atomistic detail, and CH_2 and CH_3 groups as united atoms. The SPC/E water model is employed,[42] and Na$^+$ parameters are taken from ref 43. The repulsion between headgroup oxygens is increased to reduce intrasurface hydrogen bonding.[44] We consider counterions only so that the ion number N_{ion} is determined by total charge neutrality. All simulations are performed using version 5.0 of the GROMACS simulation package[45] with periodic boundary conditions. Lennard−Jones interactions are truncated and shifted at $r_c = 0.9$ nm; for the electrostatic interactions, the particle mesh Ewald method[46] is employed. As shown in the united atom representation of decanol in Figure 1b, a negative charge of up to $\delta = -0.18$ is evenly distributed

Table 1. Simulation Parameters for Neutral and Charged Surfaces[a]

δ	σ [e/nm^2]	Ξ	N_{ion}	$L_x \times L_y$ [nm^2]	N_d
0	0			4.83 × 4.83	100
0	0			6.77 × 6.77	196
−0.0255	−0.109	0.43	10	6.77 × 6.77	196
−0.09	−0.385	1.53	18	4.83 × 4.83	100
−0.18	−0.770	3.06	36	4.83 × 4.83	100

[a] The coupling parameter Ξ is calculated using the SPC/E bulk dielectric constant $\varepsilon_b = 70$.[28]

over the three COH headgroup atoms, producing surface charge densities up to $\sigma = -0.77$ e/nm^2. Such a homogeneous surface charge distribution prevents specific ion binding and can experimentally be realized by potentiometric SAM setups.[47−52] The electrostatic coupling parameter in Table 1, defined by $\Xi = 2\pi q^3 l_B^2 |\sigma|/e$ with the Bjerrum length $l_B = e^2/(4\pi\varepsilon_0\varepsilon_b kBT)$, is rather small and in a range where deviations from PB theory are moderate;[53] see Results and Discussion. Our choice of a polar surface ensures that even for vanishing net charge the water slab is stable. For further simulation details, see the Supporting Information.

Definition of Surface Charge Position. When comparing simulations or experiments with PB models the surface charge distribution is important. Figure 2 (a) shows water and decanol oxygen density profiles for neutral and charged surfaces. Water density oscillations are absent since the surfaces are relatively soft.[44,54] The surface separation d is defined as the mean distance between decanol oxygen atoms in opposing layers. In Figure 2b, we show the decanol charge density profile $\rho_{\mathrm{surf}}(z)$, which for the neutral surface (solid black line) reveals a pronounced orientation of the OH headgroups. With decreasing net surface charge density $\sigma = \int_0^{L_z/2} \mathrm{d}z\, \rho_{\mathrm{surf}}(z)$, the surface charge profile shifts downward. The plot of the excess surface charge profile $\Delta\rho_{\mathrm{surf}}(z) = \rho_{\mathrm{surf}}(z)|_\sigma - \rho_{\mathrm{surf}}(z)|_{\sigma=0}$ in Figure 2c demonstrates that d equals the distance between the excess charge extrema quite accurately, we therefore use d to characterize the surface charge position whenever needed.

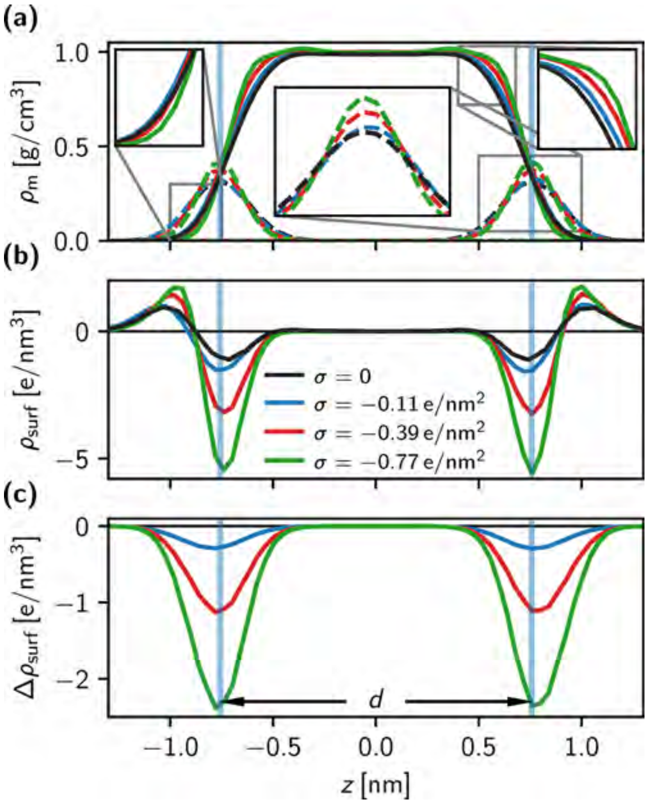


Figure 2. (a) Mass density distribution ρ_m of water (solid lines) and decanol oxygens (dashed lines) for different surface charges σ at fixed box height $L_z = 4.05$ nm. Vertical lines denote the surface position given by the separation $d = 1.51$ nm which follows from the mean decanol oxygen separation and differs by less than 0.02 nm for different σ. (b) Decanol charge density distribution $\rho_{\mathrm{surf}}(z)$. (c) Excess surface charge density distribution $\Delta\rho_{\mathrm{surf}}(z)$.

Modified PB Theory. The normal electric $E_\perp(z)$ and displacement fields $D_\perp(z)$ are related by the nonlocal inverse perpendicular dielectric function $\varepsilon_{\mathrm{nl},\perp}^{-1}$ according to

$$E_\perp(z) = \varepsilon_0^{-1} \int \mathrm{d}z' \varepsilon_{\mathrm{nl},\perp}^{-1}(z, z')\, D_\perp(z') \tag{1}$$

where we used lateral homogeneity and averaged over the xy coordinates. For slowly varying $D_\perp(z)$ (i.e., for low σ), a systematic gradient expansion yields to leading order

$$E_\perp(z) = \varepsilon_0^{-1} \varepsilon_\perp^{-1}(z) D_\perp(z) \tag{2}$$

with the dielectric profile defined as $\varepsilon_\perp^{-1}(z) = \int \mathrm{d}z' \varepsilon_{\mathrm{nl},\perp}^{-1}(z, z')$. Defining the electrostatic potential $\Psi(z)$ by $\mathrm{d}\Psi(z)/\mathrm{d}z = -E_\perp(z)$, the PB mean-field counterion density distribution reads

$$\rho_{\mathrm{ion}}(z) = qec_0\, \mathrm{e}^{-[qe\Psi(z)+U(z)]/k_BT} \tag{3}$$

where $q = 1$ is the counterion valence and the potential $U(z)$ accounts for interfacial effects not included in the electrostatic potential $\Psi(z)$. The factor c_0 ensures that the total charge is zero, $\int_{-L_z/2}^{L_z/2} \mathrm{d}z \rho_{\mathrm{ion}}(z) \overset{!}{=} -2\sigma$. The displacement field is due to the sum of ion and surface charges according to

$$\frac{\mathrm{d}D_\perp(z)}{\mathrm{d}z} = \rho(z) = \rho_{\mathrm{ion}}(z) + \rho_{\mathrm{surf}}(z) \tag{4}$$

Combining the derivative of eq 2 with eqs 3 and 4 results in the integro-differential MPB equation[55−57] which is numerically solved (see Supporting Information).

Monte Carlo Simulations. In order to assess the validity of the mean field PB model to predict the interaction pressures, we perform canonical Monte Carlo (MC) simulations of $N_{\mathrm{ion}} = 50$ point-like monovalent counterions in a box of lateral size L, which is determined by electroneutrality, $L = \sqrt{N_{\mathrm{ion}}/(2|\sigma|)}$. The counterions are confined between two surfaces located at $z = 0$ and $z = d$. The electrostatic energy is obtained using a generalized efficient 3D Ewald summation method with correction for slab geometry.[58] In this method, an empty layer which is 3 times thicker than the wall separation[59] is present in the region $d < z < 4d$. After 1×10^6 MC steps for equilibration, we save 1×10^6 uncorrelated counterion configurations every 1000 MC steps for further analysis.

■ RESULTS AND DISCUSSION

Ion Density Profiles. The simulated ion density profiles for different surface charge densities σ and different separations d are shown as blue lines in Figure 3. The predictions from standard PB theory, using a constant bulk dielectric profile $\varepsilon_\perp^{-1}(z) = \varepsilon_b^{-1}$, vanishing ion-surface interactions $U(z) = 0$ and sharply localized surface charges $\rho_{\mathrm{surf}}(z) = \delta(|z| - d/2)\sigma$, are shown as dashed black lines (see the Supporting Information for details). For the lowest surface charge $\sigma = -0.11$ e/nm^2 and large separation $d = 6$ nm in Figure 3 (c) the two profiles agree well apart from the interfacial region. However, for the smaller separations in Figure 3 (a) and (b) the standard PB model obviously fails.

We now turn to the MPB that features an exponential repulsive ion-surface interaction of the form

$$U(z) = a\exp\left(\frac{z - d/2}{\xi}\right) + a\exp\left(\frac{-z - d/2}{\xi}\right) \tag{5}$$

The decay length $\xi = (0.16 \pm 0.01)$ nm and the potential strength $a/(k_BT) = 2.6 \pm 0.1$ are obtained from a fit of the MPB predictions to the simulation data for $d > 1$ nm at $\sigma = -0.11$ e/nm^2; see the Supporting Information. Using the same $U(z)$ for all surface separations with $\varepsilon_\perp^{-1}(z) = \varepsilon_b^{-1}$ and sharply localized surface charges, the MPB results shown by dashed red lines in Figure 3 describe the simulation data well for all separations and surface charges.

DOI: 10.1021/acs.langmuir.8b03474
Langmuir 2019, 35, 551−560

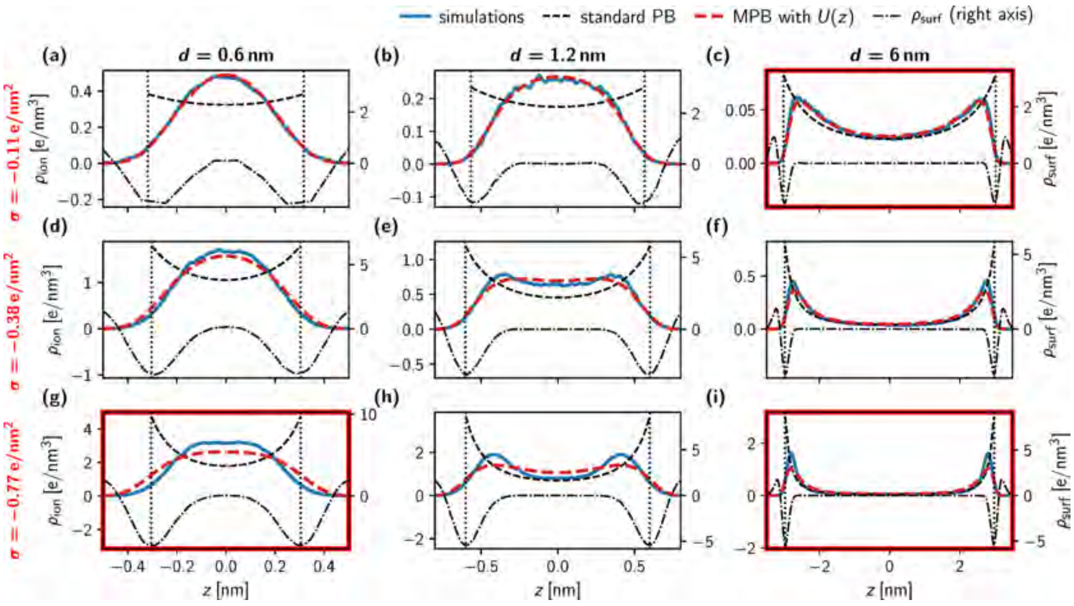


Figure 3. Simulation results for the ion density profiles $\rho_{ion}(z)$ (solid blue lines) for different surface separations d (columns) and surface charge densities σ (rows). Dashed black lines show the standard PB prediction with the surface charge positions indicated by vertical dashed lines at $\pm d/2$, and red dashed lines the MPB predictions including an ion-surface repulsion $U(z)$ according to eq 5. Dashed-dotted black lines show the surface-charge profiles $\rho_{surf}(z)$ (right axis). Results highlighted by red boxes are discussed in more detail in Figure 4.

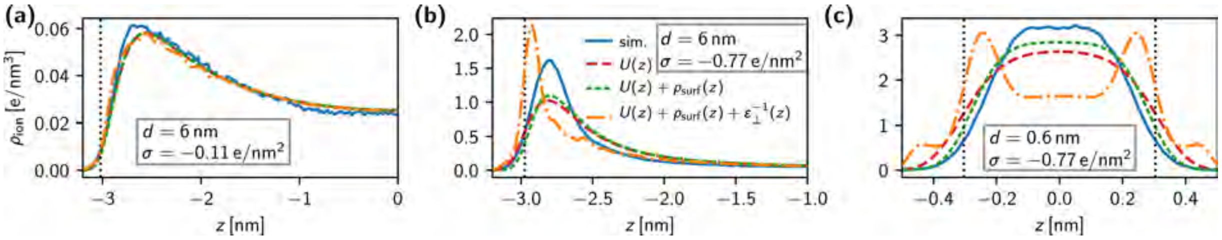


Figure 4. Simulation results for the ion density profiles $\rho_{ion}(z)$ (solid blue lines) for a few selected values of d and σ compared with MPB predictions including only the ion−surface interaction $U(z)$ (red broken lines), $U(z)$ and the surface charge distribution $\rho_{surf}(z)$ (green dotted lines), $U(z)$, $\rho_{surf}(z)$, and the dielectric profile $\varepsilon_\perp^{-1}(z)$ (orange dashed-dotted lines). Vertical dashed lines indicate the mean surface-charge positions $\pm d/2$.

The MPB that additionally includes the surface charge distribution $\rho_{surf}(z)$ extracted from simulations, shown as dashed-dotted black lines in Figure 3, does not significantly change the results, as we demonstrate for selected distances and surface charges by dotted green lines in Figure 4. This is not surprising since $\rho_{surf}(z)$ is peaked in the region where the surface repulsion $U(z)$ is large.

We now check the influence of a dielectric profile $\varepsilon_\perp^{-1}(z)$, which we extract from polarization fluctuations for uncharged surfaces;[28] see the Supporting Information for details. The MPB solution including the previously determined ion-surface potential $U(z)$, the surface charge distribution $\rho_{surf}(z)$, and the dielectric profile $\varepsilon_\perp^{-1}(z)$ is shown as orange dashed-dotted lines in Figure 4. For small surface charge and large separation in Figure 4a, where linear dielectric theory is expected to be reliable, the influence of the dielectric profile on the predicted ion density is negligible; for the high surface charge densities in Figure 4b and c, for which we show below that linear dielectric theory is not expected to work, including $\varepsilon_\perp^{-1}(z)$ does not improve the agreement with the simulation data. We checked the robustness of our fitting procedure for $U(z)$ by an alternative procedure where we determine the parameters of eq 5 in the presence of the $\rho_{surf}(z)$ and $\varepsilon_\perp^{-1}(z)$ profiles; see the Supporting Information. From this alternative fit, we obtain modified parameters $a^*/(k_BT) = 3.8 \pm 0.1$ and $\xi^* = (0.13 \pm 0.01)$ nm. However, the agreement between the MPB and simulation counterion density profiles shown in Figures S4 and S5 does not improve and at small separation and surface charges the deviations are even stronger; see Figure S4a. In particular, as we show in Figure S4g−i, taking into account the $\varepsilon_\perp^{-1}(z)$ profiles drastically overestimates the ion density close to the surface for $\sigma = -0.77$ e/nm^2, in contrast to the MPB curves shown in Figure 3g−i. This is not surprising, since linear dielectric theory is not expected to work for these high surface charges as we argue further below. In conclusion, the MPB theory which only includes the ion-surface potential $U(z)$ reproduces the simulated ion density profiles very well, while additional effects due to dielectric effects and surface charge distributions are negligible.

Interaction pressure. The simulated interaction pressure between charge-neutral polar surfaces in Figure 1c decays exponentially according to

$$\Pi_{hyd} = \Pi^* e^{-d/\lambda} \tag{6}$$

with a decay length $\lambda = 0.10$ nm, and corresponds to the hydration repulsion.[28,40,44,60] The simulated pressures for finite surface charges are shown in Figure 5 (blue diamonds)

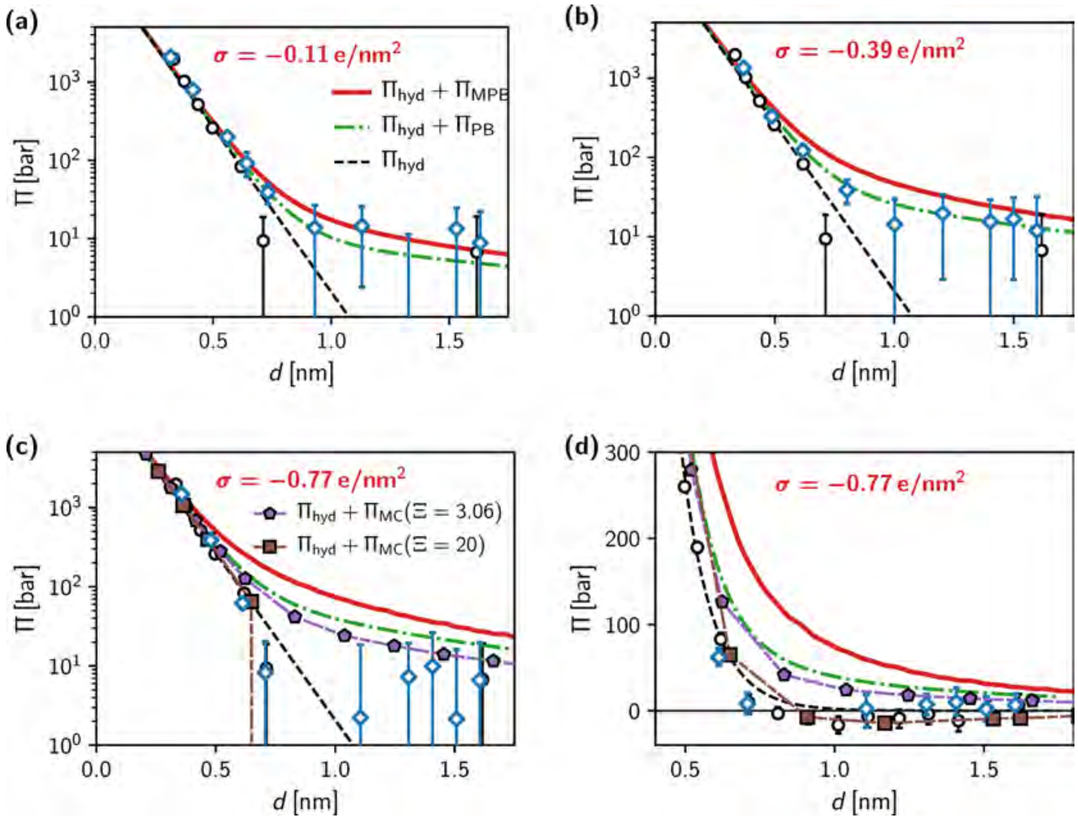


Figure 5. Simulated interaction pressures (blue diamonds) for three different surface charge densities σ are compared with results for uncharged surfaces (black circles, including the exponential fit $\Pi_{\rm hyd}(z)$, eq 6, from Figure 1c as a broken black line) and the additive predictions $\Pi_{\rm hyd} + \Pi_{\rm MPB}$ (red lines) and $\Pi_{\rm hyd} + \Pi_{\rm PB}$ (green lines). The plot in (d) shows the same data as in (c) on a linear scale. In (c) and (d), we also include the additive prediction $\Pi_{\rm hyd} + \Pi_{\rm MC}$ from Monte Carlo simulations for coupling parameters $\Xi = 3.06$ (purple pentagons) and $\Xi = 20$ (brown squares).

together with the neutral-surface result (black spheres). For small separations, the simulated pressures for finite surface charge agree perfectly with the hydration pressure between neutral surfaces, confirming one essential assumption made in the description of experimental data.

The MPB pressure follows from the free energy

$$\frac{\mathcal{F}_{\rm MPB}(d)}{Ak_{\rm B}T} = -\frac{1}{2k_{\rm B}T}\int_{-L_z/2}^{L_z/2} \rho(z)\ \Psi(z){\rm d}z + \int_{-L_z/2}^{L_z/2} c(z)\ (\ln(c(z)) - 1){\rm d}z + \int_{-L_z/2}^{L_z/2} U(z)\ c(z){\rm d}z \qquad (7)$$

by differentiation, $\Pi_{\rm MPB}(d) = -\partial\mathcal{F}_{\rm MPB}(d)/(A\partial d)$. The first term in eq 7 is the electrostatic energy, the second accounts for the counterion entropy, and the third results from ion−surface interactions, where $c(z) = \rho_{\rm ion}(z)/(qe)$ is the ion number density, see Supporting Information.

In Figure 5, we compare the simulated pressure to the sum of hydration and MPB pressures (red lines) and the sum of hydration and standard PB pressures (green broken lines). At low surface charge $\sigma = -0.11\ {\rm e/nm^2}$ in Figure 5a, the simulation data is excellently described by $\Pi_{\rm hyd} + \Pi_{\rm MPB}$. The ion−surface repulsion $U(z)$ makes the MPB pressure slightly more repulsive compared to standard PB theory, but this difference is rather small, which is noteworthy in light of the pronounced differences between the MPB and PB ion density profiles in Figure 3a−c. This reflects that the PB pressure for these small surface separations is essentially entropic and due to ion confinement; see the Supporting Information for details. The simulated pressures for $\sigma = -0.39\ {\rm e/nm^2}$ in Figure 5b are slightly smaller than $\Pi_{\rm hyd} + \Pi_{\rm MPB}$; for $\sigma = -0.77\ {\rm e/nm^2}$ in Figure 5c and d, they are significantly smaller. These simulation results can in fact be explained by ion-correlation effects, which invalidate the mean-field approximation inherent to the MPB model, as we demonstrate below by Monte Carlo (MC) simulations of counterions in a homogeneous dielectric medium between planar charged walls with the same surface charge densities as employed in the MD simulations, if we take the modification of the water dielectric constant due to the surface-charge induced water orientation into account. This suggests that the additivity assumption of hydration and mean-field MPB repulsive pressures, with the hydration pressure corresponding to the pressure acting between charge net-neutral surfaces and which is assumed to solely account for the surface-induced water structure modification, breaks down already for moderate surface charge densities.

Water Uptake. To emphasize the importance of a precise control of the water chemical potential in atomistic simulations of charged surfaces, we report in Figure 6a the surface-charge induced rescaled water number change in the slab at given separation d, $\Delta N_{\rm w}/N_{\rm w}^0 = (N_{\rm w} - N_{\rm w}^0)/N_{\rm w}^0$, where $N_{\rm w}^0$ denotes the water number present between uncharged polar surfaces. As seen in Figure 6a, $\Delta N_{\rm w}/N_{\rm w}^0$ for constant water chemical potential is for large surface separation d and large surface charge density significantly positive. This water uptake has drastic consequences for the interaction pressure, as shown in Figure 6b. If the water number is fixed to $N_{\rm w}^0$, corresponding to

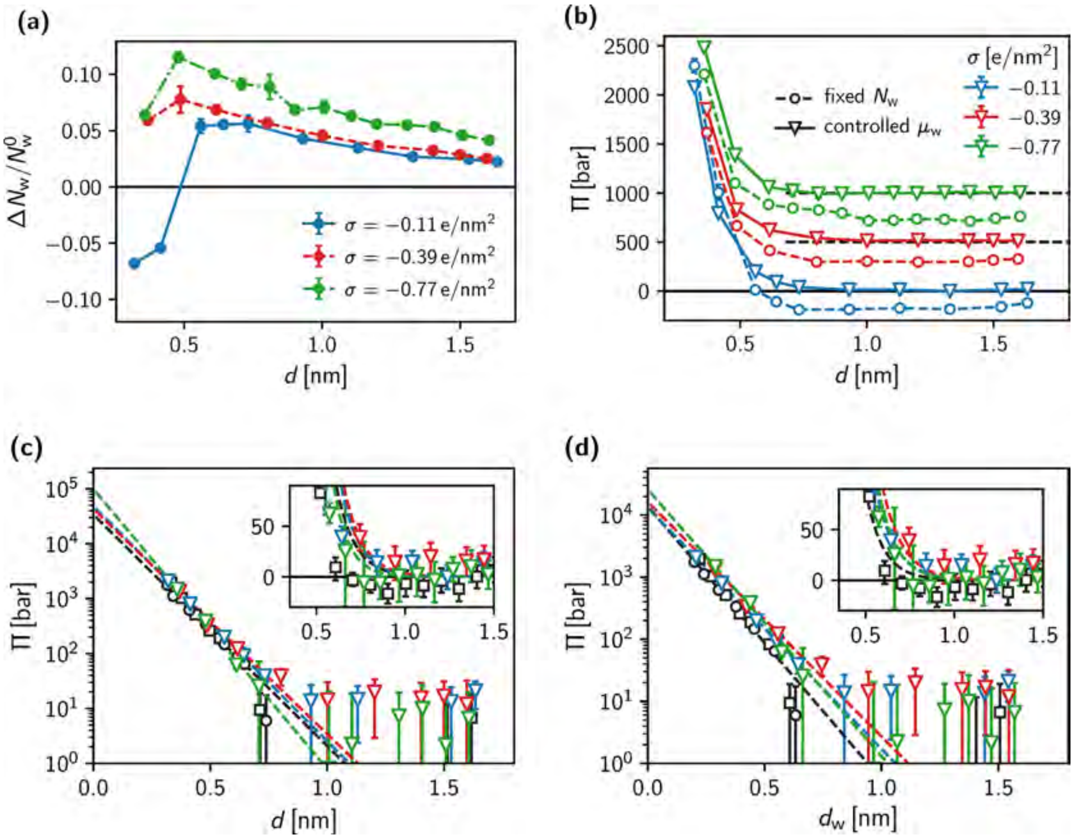


Figure 6. (a) Relative water uptake $\Delta N_w/N_w^0$ compared to the uncharged system for different surface charge densities $\sigma = -0.11, -0.39$, and $-0.77\ e/nm^2$ (blue, red and green data, respectively). (b) Interaction pressure at fixed water number N_w^0 corresponding to the uncharged system (circles) and at fixed bulk water chemical potential (triangles). Lines serve as guide to the eye. For clarity, the data is shifted by 500 bar for $\sigma = -0.39\ e/nm^2$ and 1000 bar for $\sigma = -0.77\ e/nm^2$, respectively, indicated by the dashed black lines to the right. (c) Interaction pressure at fixed bulk water chemical potential for the neutral and charged surfaces as a function of the geometric water slab thickness d. Dashed lines denote exponential fits to the simulation data. (d) Same data as in (c) but now plotted as a function of the equivalent water slab thickness $d_w = N_w v_w/A$.

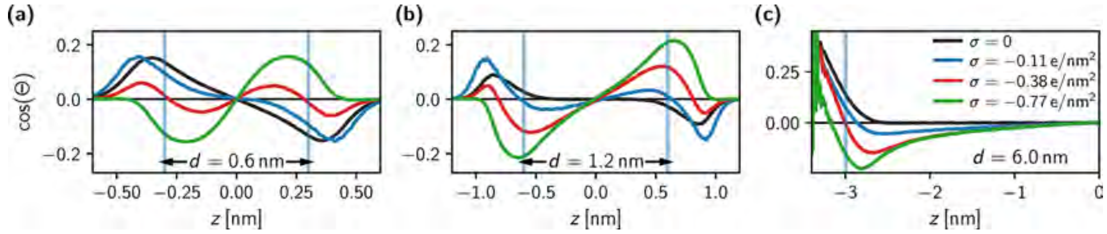


Figure 7. Water orientation profiles cos(Θ) for different surface charge densities σ at surface separation (a) d = 0.6 nm, (b) 1.2 nm, and (c) 6 nm. For clarity in (c), only the left half-space is shown.

the water number between uncharged surfaces, the interaction pressures are strongly attractive at large separations (circles). In contrast, if the water chemical potential is kept constant at the bulk value, thus allowing for water uptake (triangles), the pressures decay to zero at large separations and are generally larger by about 200 bar. Thus, the quantitative analysis of simulation pressures in Figure 5 only makes sense if the chemical potential is precisely controlled.

We note that the finite water uptake ΔN_w also influences the analysis of experimental pressure−distance curves, where often the equivalent water slab thickness $d_w = N_w v_w/A$, which is proportional to the number of water molecules, is reported and used as a proxy for the surface separation.[17] A comparison of the simulated pressures for different surface charge densities when plotted as a function of the geometric surface separation d in Figure 6c reveals that the short-range pressure between charged surfaces, fitted to exponential functions denoted by colored dashed lines, does not change significantly compared to the uncharged surface pressure (black dashed line). However, when using the equivalent water slab thickness d_w as the surface separation in Figure 6d, the amplitude of the repulsion at small distances acquires a spurious dependence on the surface charge. This adds an additional complication to the analysis of experimental pressure−distance data of charged systems, since as we show here, the definition of the surface separation will not only influence the fit value of the effective surface charge but also the fit values of the hydration repulsion amplitude.

Water Orientation at Charged Interfaces. The hydration repulsion is in literature typically attributed to surface-induced water orientation effects.[61−63] In Figure 7 we show profiles of the cosine of the orientation angle Θ between

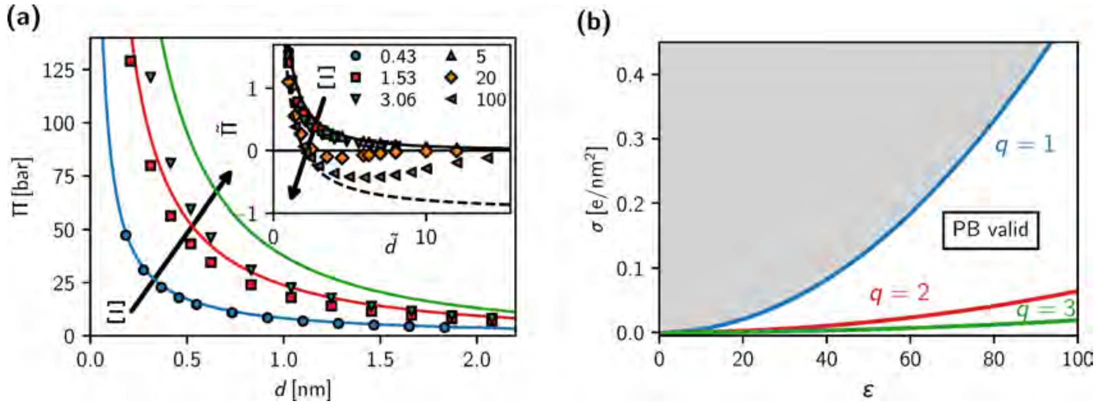


Figure 8. (a) Comparison of interaction pressures from Monte Carlo simulations of point-like counterions between homogeneously charged surfaces (data points) and PB theory (solid lines) for the parameters considered in the MD simulations. The inset shows MC simulation results, including additional results for $\Xi = 5$, $\Xi = 20$, and $\Xi = 100$, in rescaled units $\tilde{\Pi} = \Pi/(k_BT2\pi l_B\sigma^2)$ and $\tilde{d} = d/b$ (data for $\Xi = 5$ and $\Xi = 20$ are taken from ref 67). Solid line shows the PB prediction, and dashed line denotes the strong coupling prediction[66] valid in the limit $\Xi \to \infty$. (b) Scaling diagram of the surface charge density range where PB theory is expected to be valid as a function of the dielectric constant for monovalent, divalent and trivalent counterions (at room temperature).

water dipoles and the surface normal for different surface separations. For neutral surfaces, $\sigma = 0$, and at large separation, $d = 6$ nm, shown in Figure 7c, $\cos(\Theta) > 0$, meaning that interfacial water molecules point with their oxygens toward the decanols due to favorable hydrogen-bonding interactions with the surface hydrogens (solid black line), further away from the interface the average orientation quickly decays to zero. In strong confinement, Figure 7a, the water orientation profile is strongly perturbed compared to the free interface; this is one contribution to the hydration repulsion. For weak surface charge $\sigma = -0.11$ e/nm^2, the orientation profile is not drastically changed compared to $\sigma = 0$, which explains the observed additivity of hydration and MPB pressures in Figure 5a. For elevated surface charges $\sigma = -0.38$ and -0.77 e/nm^2, the water orientation changes profoundly and water hydrogens point to the surface due to the negative surface charge, in agreement with previous experiments and simulations.[30−32,64,65] This suggests that the failure of the pressure additivity assumption for $\sigma = -0.77$ e/nm^2 in Figure 5c could partly result from a modification of the hydration force due to surface-charge induced water reorientation. We discuss the influence of ion correlation effects, which are not captured by PB theory, below and show that they give rise to substantial modifications of the interaction pressure if the confinement-induced decrease of the water dielectric constant is accounted for.

Corrections to PB Theory. The PB theory describes the pressure between charged surfaces correctly for small coupling parameter $\Xi = 2\pi l_B^2 q^3|\sigma|/e < 1$, that means for low surface charge density σ and ions with low valency q.[66] Note that for very small surface separation d the mean-field approximation used in deriving PB theory breaks down even for small values of Ξ, this however does not invalidate the pressure predicted by PB since it is predominantly due to the counterion confinement pressure.[66] For the three surface charge densities considered in this work, $\Xi < 1$ is strictly fulfilled only for $\sigma = -0.11$ e/nm^2; see Table 1.

To check whether the breakdown of PB theory could be the cause of the deviations between the pressures from MD simulations and the MPB predictions in Figure 5c and d, we compare in Figure 8a pressures from PB theory (solid lines) with water-implicit Monte Carlo simulations of point-like monovalent counterions between homogeneously charged plates (data points). This comparison allows one to exclusively judge the effect of counterion correlations on the pressure without complications due to nonelectrostatic ion−surface interactions, smeared-out surface charge distributions, and dielectric and hydration effects, which are all present in the MD simulations. The surface charge densities used in Figure 8a correspond exactly to the ones in the MD simulations. As expected, the agreement between PB theory and simulations is nearly perfect for the lowest surface charge, i.e., coupling parameter $\Xi = 0.43$, whereas for the system with the highest surface charge, $\Xi = 3.06$, deviations are significant. However, the reduction of the surface pressure by about 30% for $\Xi = 3.06$ observed in Figure 8a is not large enough to fully explain the pressure reduction seen in the MD simulations compared to the MPB + hydration pressure prediction; this is demonstrated in Figure 5c and d by the purple pentagons which represent the sum of the hydration pressure and the MC simulation result for $\Xi = 3.06$.

In fact, even more enhanced deviations from PB are expected since the dielectric constant of water in planar confinement decreases,[28] which would further increase the coupling parameter since it scales with the inverse square of the dielectric constant, i.e., $\Xi \sim 1/\varepsilon^2$. In the inset of Figure 8a, we show the rescaled pressure $\tilde{\Pi} = \Pi/(2\pi k_BTl_B\sigma^2/e^2)$ versus the rescaled distance $\tilde{d} = d/b$, where $b = 1/(2\pi q l_B|\sigma|/e)$ is the Gouy−Chapman length. It is seen that for large coupling parameters the pressure decreases and around $\Xi = 20$ becomes attractive in an intermediate distance range. The perpendicular dielectric constant of water confined between neutral polar surfaces has been demonstrated to decrease for surface separations below $d \approx 0.7$ nm from the SPC/E bulk value of $\varepsilon_\perp \approx 70$ down to values of $\varepsilon_\perp \approx 10$ at $d \approx 0.3$ nm, which indeed would suggest a massively increased effective coupling parameter for small surface separations. To make a conclusive prediction for the effective coupling parameter Ξ that describes our MD simulations of charged surfaces, one would need to derive the effective water dielectric constant in the presence of surface charges and counterions, which is nontrivial even in bulk electrolytes.[68] As a test, we present in Figure 5d by brown squares the sum of the hydration pressure and the MC simulation pressure for $\Xi = 20$, which matches the MD

DOI: 10.1021/acs.langmuir.8b03474
Langmuir 2019, 35, 551−560

pressure quite nicely. This comparison suggests that the change of the dielectric constant, caused by the surface-induced water structure perturbation, is in strong confinement sufficient to significantly increase the effective electrostatic coupling parameter and thereby to invalidate the mean-field PB prediction for the pressure. The perturbation on the counterion density profiles is expected to be less severe since it is dominated by surface−ion interactions, which enter the PB and the strong-coupling description (valid for high coupling parameters) in a similar fashion.

We present the criterion for the applicability of PB theory, $\Xi < 1$, in Figure 8b as a function of the dielectric constant for monovalent, divalent, and trivalent counterions. For the SPC/E bulk dielectric constant $\varepsilon = 70$ and $q = 1$, we find a threshold surface charge density of $|\sigma| = 0.25\ \mathrm{e/nm^2}$, for the experimental value $\varepsilon = 80$, one obtains the slightly increased value $|\sigma| = 0.33\ \mathrm{e/nm^2}$. For a reduced dielectric constant, which should apply to water in strong confinement, the threshold surface charge density decreases and deviations from PB theory are expected for a wider range of surface charge densities. Clearly, for divalent and trivalent counterions the threshold surface charge density is further decreased.

■ SUMMARY AND CONCLUSIONS

Our water-explicit simulations at fixed water chemical potential show that monovalent counterion density profiles at soft charged surfaces are well described by a modified Poisson−Boltzmann (MPB) approach that includes nonelectrostatic ion−surface interactions. The effects of dielectric profiles and smeared-out surface charge distributions on the counterion density profiles are less important. At low surface charge densities, our simulations confirm the additivity of the hydration repulsion (extracted from simulations of uncharged surfaces) and the surface-charge induced MPB pressure contributions down to sub-nanometer surface separations. However, already for a moderate surface charge density $\sigma = -0.77\mathrm{e/nm^2}$, this additivity breaks down, which we rationalize by the combined effects of a counterion-correlation induced modification of the MPB predictions (presumably enhanced by a decrease of the water dielectric constant in confinement) and a modification of the hydration repulsion due to the surface-charge induced reorientation of interfacial water. Corrections to PB theory due to ion correlation effects, which have been extensively discussed,[38] can thus even for moderate surface charge densities and monovalent counterions become relevant. That interfacial water reacts sensitively to the presence of surface charges is known from simulations and experiments,[30−32,64,65] and we suggest that this restructuring modifies the surface interaction pressure between charged surfaces significantly.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.langmuir.8b03474.

> Further simulation details including the control of the water chemical potential, dielectric profiles, details on the MPB including a free energy decomposition and the analytical solution of the PB equation between planar walls (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: alexander.schlaich@univ-grenoble-alpes.fr.
*E-mail: rnetz@physik.fu-berlin.de.

ORCID

Alexander Schlaich: 0000-0002-4250-363X

Author Contributions

A.S. and R.R.N. designed the research project; A.S. and A.P.d.S. performed simulations; A.S., A.P.d.S., and R.R.N. analyzed the data; A.S. and R.R.N. wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the European Innovative Training Network "Transport of Soft Matter at the Nanoscale" (NANOTRANS). A.P.d.S acknowledges funding from the Alexander von Humboldt Foundation through the CAPES program.

■ REFERENCES

(1) Suetsugu, S.; Kurisu, S.; Takenawa, T. Dynamic Shaping of Cellular Membranes by Phospholipids and Membrane-Deforming Proteins. *Physiol. Rev.* **2014**, *94*, 1219−1248.

(2) Leventis, P. A.; Grinstein, S. The Distribution and Function of Phosphatidylserine in Cellular Membranes. *Annu. Rev. Biophys.* **2010**, *39*, 407−427.

(3) Di Paolo, G.; De Camilli, P. Phosphoinositides in cell regulation and membrane dynamics. *Nature* **2006**, *443*, 651−657.

(4) Smulders, E.; von Rybinski, W.; Sung, E.; Rähse, W.; Steber, J.; Wiebel, F.; Nordskog, A. *Ullmann's Encyclopedia of Industrial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2000.

(5) Israelachvili, J. N.; Adams, G. Measurement of forces between two mica surfaces in aqueous electrolyte solutions in the range 0−100 nm. *J. Chem. Soc., Faraday Trans. 1* **1978**, *74*, 975−1001.

(6) Pashley, R. M. DLVO and hydration forces between mica surfaces in Li+, Na+, K+, and Cs+ electrolyte solutions: A correlation of double-layer and hydration forces with surface cation exchange properties. *J. Colloid Interface Sci.* **1981**, *83*, 531−546.

(7) Pashley, R. Hydration forces between mica surfaces in aqueous electrolyte solutions. *J. Colloid Interface Sci.* **1981**, *80*, 153−162.

(8) Horn, R. G.; Smith, D. T.; Haller, W. Surface forces and viscosity of water measured between silica sheets. *Chem. Phys. Lett.* **1989**, *162*, 404−408.

(9) Andelman, D. In *Handbook of Biological Physics: Structure and Dynamics of Membranes*; Lipowsky, R., Sackmann, E., Eds.; North-Holland, 1995; Vol. *1*, pp 603−642.

(10) Argyris, D.; Cole, D. R.; Striolo, A. Ion-Specific Effects under Confinement: The Role of Interfacial Water. *ACS Nano* **2010**, *4*, 2035−2042.

(11) Videla, P. E.; Sala, J.; Martí, J.; Guàrdia, E.; Laria, D. Aqueous electrolytes confined within functionalized silica nanopores. *J. Chem. Phys.* **2011**, *135*, 104503.

(12) Bonnaud, P. A.; Coasne, B.; Pellenq, R. J.-M. Solvated calcium ions in charged silica nanopores. *J. Chem. Phys.* **2012**, *137*, No. 064706.

(13) Ho, T. A.; Argyris, D.; Cole, D. R.; Striolo, A. Aqueous NaCl and CsCl Solutions Confined in Crystalline Slit-Shaped Silica Nanopores of Varying Degree of Protonation. *Langmuir* **2012**, *28*, 1256−1266.

(14) Hartkamp, R.; Siboulet, B.; Dufrêche, J.-F.; Coasne, B. Ion-specific adsorption and electroosmosis in charged amorphous porous silica. *Phys. Chem. Chem. Phys.* **2015**, *17*, 24683−24695.

(15) Bordallo, H. N.; Aldridge, L. P.; Churchman, G. J.; Gates, W. P.; Telling, M. T. F.; Kiefer, K.; Fouquet, P.; Seydel, T.; Kimber, S. A. J. Quasi-Elastic Neutron Scattering Studies on Clay Interlayer-Space

Highlighting the Effect of the Cation in Confined Water Dynamics. *J. Phys. Chem. C* **2008**, *112*, 13982−13991.

(16) Malikova, N.; Longeville, S.; Zanotti, J.-M.; Dubois, E.; Marry, V.; Turq, P.; Ollivier, J. Signature of Low-Dimensional Diffusion in Complex Systems. *Phys. Rev. Lett.* **2008**, *101*, 265901.

(17) Parsegian, V.; Fuller, N.; Rand, R. Measured work of deformation and repulsion of lecithin bilayers. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76*, 2750−2754.

(18) Rand, R.; Parsegian, V. Hydration forces between phospholipid bilayers. *Biochim. Biophys. Acta, Rev. Biomembr.* **1989**, *988*, 351−376.

(19) Israelachvili, J. N.; Pashley, R. M. Molecular layering of water at surfaces and origin of repulsive hydration forces. *Nature* **1983**, *306*, 249−250.

(20) Marra, J.; Israelachvili, J. N. Direct measurements of forces between phosphatidyl- choline and phosphatidylethanolamine bilayers in aqueous electrolyte solutions. *Biochemistry* **1985**, *24*, 4608−4618.

(21) Herrmann, L.; Johner, A.; Kékicheff, P. Interactions between Charged Lamellae in Aqueous Solution. *Phys. Rev. Lett.* **2014**, *113*, 268302.

(22) McIntosh, T. J.; Magid, A. D.; Simon, S. A. Interactions between charged, uncharged, and zwitterionic bilayers containing phosphatidylglycerol. *Biophys. J.* **1990**, *57*, 1187−1197.

(23) Parsegian, V. A.; Rand, R. P.; Fuller, N. L. Direct osmotic stress measurements of hydration and electrostatic double-layer forces between bilayers of double-chained ammonium acetate surfactants. *J. Phys. Chem.* **1991**, *95*, 4777−4782.

(24) Popa, I.; Sinha, P.; Finessi, M.; Maroni, P.; Papastavrou, G.; Borkovec, M. Importance of Charge Regulation in Attractive Double-Layer Forces between Dissimilar Surfaces. *Phys. Rev. Lett.* **2010**, *104*, 228301.

(25) Trefalt, G.; Palberg, T.; Borkovec, M. Forces between colloidal particles in aqueous solutions containing monovalent and multivalent ions. *Curr. Opin. Colloid Interface Sci.* **2017**, *27*, 9−17.

(26) Faraudo, J.; Bresme, F. Anomalous Dielectric Behavior of Water in Ionic Newton Black Films. *Phys. Rev. Lett.* **2004**, *92*, 236102.

(27) Bonthuis, D. J.; Gekle, S.; Netz, R. R. Dielectric Profile of Interfacial Water and its Effect on Double-Layer Capacitance. *Phys. Rev. Lett.* **2011**, *107*, 166102.

(28) Schlaich, A.; Knapp, E. W.; Netz, R. R. Water Dielectric Effects in Planar Confinement. *Phys. Rev. Lett.* **2016**, *117*, No. 048001.

(29) Fumagalli, L.; Esfandiar, A.; Fabregas, R.; Hu, S.; Ares, P.; Janardanan, A.; Yang, Q.; Radha, B.; Taniguchi, T.; Watanabe, K.; Gomila, G.; Novoselov, K. S.; Geim, A. K. Anomalously low dielectric constant of confined water. *Science* **2018**, *360*, 1339−1342.

(30) Dreier, L. B.; Nagata, Y.; Lutz, H.; Gonella, G.; Hunger, J.; Backus, E. H. G.; Bonn, M. Saturation of charge-induced water alignment at model membrane surfaces. *Science Advances* **2018**, *4*, No. eaap7415.

(31) Cyran, J. D.; Backus, E. H. G.; Nagata, Y.; Bonn, M. Structure from Dynamics: Vibrational Dynamics of Interfacial Water as a Probe of Aqueous Heterogeneity. *J. Phys. Chem. B* **2018**, *122*, 3667−3679.

(32) Hosseinpour, S.; Tang, F.; Wang, F.; Livingstone, R. A.; Schlegel, S. J.; Ohto, T.; Bonn, M.; Nagata, Y.; Backus, E. H. G. Chemisorbed and Physisorbed Water at the TiO2/Water Interface. *J. Phys. Chem. Lett.* **2017**, *8*, 2195−2199.

(33) Moreira, A. G.; Netz, R. R. Counterions at charge-modulated substrates. *Europhysics Letters (EPL)* **2002**, *57*, 911−917.

(34) Schwierz, N.; Netz, R. R. Effective Interaction between Two Ion-Adsorbing Plates: Hofmeister Series and Salting-In/Salting-Out Phase Diagrams from a Global Mean- Field Analysis. *Langmuir* **2012**, *28*, 3881−3886.

(35) Schwierz, N.; Horinek, D.; Netz, R. R. Specific Ion Binding to Carboxylic Surface Groups and the pH Dependence of the Hofmeister Series. *Langmuir* **2015**, *31*, 215−225.

(36) Jungwirth, P.; Tobias, D. J. Specific Ion Effects at the Air/Water Interface. *Chem. Rev.* **2006**, *106*, 1259−1281.

(37) Schwierz, N.; Horinek, D.; Netz, R. R. Anionic and Cationic Hofmeister Effects on Hydrophobic and Hydrophilic Surfaces. *Langmuir* **2013**, *29*, 2602−2614.

(38) Moreira, A. G.; Netz, R. R. Binding of Similarly Charged Plates with Counterions Only. *Phys. Rev. Lett.* **2001**, *87*, No. 078301.

(39) Schneck, E.; Sedlmeier, F.; Netz, R. R. Hydration repulsion between biomembranes results from an interplay of dehydration and depolarization. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14405−14409.

(40) Kanduč, M.; Netz, R. R. From hydration repulsion to dry adhesion between asymmetric hydrophilic and hydrophobic surfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 12338−12343.

(41) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656−1676.

(42) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.

(43) Smith, D. E.; Dang, L. X. Computer simulations of NaCl association in polarizable water. *J. Chem. Phys.* **1994**, *100*, 3757−3766.

(44) Kanduč, M.; Schneck, E.; Netz, R. R. Attraction between hydrated hydrophilic surfaces. *Chem. Phys. Lett.* **2014**, *610−611*, 375−380.

(45) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.

(46) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(47) Hillier, A. C.; Kim, S.; Bard, A. J. Measurement of Double-Layer Forces at the Electrode/Electrolyte Interface Using the Atomic Force Mircoscope: Potential and Anion Dependent Interactions. *J. Phys. Chem.* **1996**, *100*, 18808−18817.

(48) Rentsch, S.; Siegenthaler, H.; Papastavrou, G. Diffuse Layer Properties of Thiol- Modified Gold Electrodes Probed by Direct Force Measurements. *Langmuir* **2007**, *23*, 9083−9091.

(49) Papastavrou, G. Combining electrochemistry and direct force measurements: from the control of surface properties towards applications. *Colloid Polym. Sci.* **2010**, *288*, 1201−1214.

(50) Kuznetsov, V.; Papastavrou, G. Ion Adsorption on Modified Electrodes as Determined by Direct Force Measurements under Potentiostatic Control. *J. Phys. Chem. C* **2014**, *118*, 2673−2685.

(51) Tivony, R.; Klein, J. Modifying surface forces through control of surface potentials. *Faraday Discuss.* **2017**, *199*, 261−277.

(52) Pashazanusi, L.; Oguntoye, M.; Oak, S.; Albert, J. N. L.; Pratt, L. R.; Pesika, N. S. Anomalous Potential-Dependent Friction on Au(111) Measured by AFM. *Langmuir* **2018**, *34*, 801.

(53) Moreira, A. G.; Netz, R. R. Strong-coupling theory for counter-ion distributions. *Europhysics Letters (EPL)* **2000**, *52*, 705−711.

(54) Lorenz, C. D.; Lane, J. M. D.; Chandross, M.; Stevens, M. J.; Grest, G. S. Molecular Dynamics Simulations of Water Confined between Matched Pairs of Hydrophobic and Hydrophilic Self-Assembled Monolayers. *Langmuir* **2009**, *25*, 4535−4542.

(55) Boström, M.; Deniz, V.; Franks, G. V.; Ninham, B. W. Extended DLVO theory: Electrostatic and non-electrostatic forces in oxide suspensions. *Adv. Colloid Interface Sci.* **2006**, *123−126*, 5−15.

(56) Paillusson, F.; Blossey, R. Slits, plates, and Poisson-Boltzmann theory in a local formulation of nonlocal electrostatics. *Phys. Rev. E* **2010**, *82*, No. 052501.

(57) Ben-Yaakov, D.; Andelman, D.; Podgornik, R.; Harries, D. Ion-specific hydration effects: Extending the Poisson-Boltzmann theory. *Curr. Opin. Colloid Interface Sci.* **2011**, *16*, 542−550.

(58) dos Santos, A. P.; Girotto, M.; Levin, Y. Simulations of Coulomb systems with slab geometry using an efficient 3D Ewald summation method. *J. Chem. Phys.* **2016**, *144*, 144103.

(59) Yeh, I.-C.; Berkowitz, M. L. Ewald summation for systems with slab geometry. *J. Chem. Phys.* **1999**, *111*, 3155−3162.

(60) Kanduč, M.; Schlaich, A.; Schneck, E.; Netz, R. R. Water-Mediated Interactions between Hydrophilic and Hydrophobic Surfaces. *Langmuir* **2016**, *32*, 8767−8782.

(61) Marčelja, S.; Radić, N. Repulsion of interfaces due to boundary water. *Chem. Phys. Lett.* **1976**, *42*, 129−130.

DOI: 10.1021/acs.langmuir.8b03474
Langmuir 2019, 35, 551−560

(62) Cevc, G.; Podgornik, R.; Žekš, B. The free energy,enthalpy and entropy of hydration of phospholipid bilayer membranes and their difference on the interfacial separation. *Chem. Phys. Lett.* **1982**, *91*, 193−196.

(63) Kanduč, M.; Schlaich, A.; Schneck, E.; Netz, R. R. Hydration repulsion between membranes and polar surfaces: Simulation approaches versus continuum theories. *Adv. Colloid Interface Sci.* **2014**, *208*, 142−152.

(64) Du, Q.; Freysz, E.; Shen, Y. R. Vibrational spectra of water molecules at quartz/water interfaces. *Phys. Rev. Lett.* **1994**, *72*, 238−241.

(65) Dewan, S.; Carnevale, V.; Bankura, A.; Eftekhari-Bafrooei, A.; Fiorin, G.; Klein, M. L.; Borguet, E. Structure of Water at Charged Interfaces: A Molecular Dynamics Study. *Langmuir* **2014**, *30*, 8056−8065.

(66) Netz, R. R. Electrostatistics of counter-ions at and between planar charged walls: From Poisson-Boltzmann to the strong-coupling theory. *Eur. Phys. J. E: Soft Matter Biol. Phys.* **2001**, *5*, 557−574.

(67) Moreira, A. G.; Netz, R. R. Simulations of counterions at charged plates. *Eur. Phys. J. E: Soft Matter Biol. Phys.* **2002**, *8*, 33−58.

(68) Rinne, K. F.; Gekle, S.; Netz, R. R. Dissecting ion-specific dielectric spectra of sodium-halide solutions into solvation water and ionic contributions. *J. Chem. Phys.* **2014**, *141*, 214502.

Nanofluidics and osmotic processes

Lydéric Bocquet

- Ionic transport at the nanoscales: linear and non-linear transport

- Flows at surfaces and in confinement (experiments)

- Osmotic processes: fundamentals and applications

  - Monday : 16h00-17h30
  - Tuesday : 16h00-17h30
  - Wednesday : 11h00-12h30

Nanofluidics, from bulk to interfaces†

Lydéric Bocquet* and Elisabeth Charlaix

Received 30th June 2009
First published as an Advance Article on the web 1st December 2009
DOI: 10.1039/b909366b

Nanofluidics has emerged recently in the footsteps of microfluidics, following the quest for scale reduction inherent to nanotechnologies. By definition, nanofluidics explores transport phenomena of fluids at nanometer scales. Why is the nanometer scale specific? What fluid properties are probed at nanometric scales? In other words, why does 'nanofluidics' deserve its own brand name? In this *critical review*, we will explore the vast manifold of length scales emerging for fluid behavior at the nanoscale, as well as the associated mechanisms and corresponding applications. We will in particular explore the interplay between bulk and interface phenomena. The limit of validity of the continuum approaches will be discussed, as well as the numerous surface induced effects occurring at these scales, from hydrodynamic slippage to the various electro-kinetic phenomena originating from the couplings between hydrodynamics and electrostatics. An enlightening analogy between ion transport in nanochannels and transport in doped semi-conductors will be discussed (156 references).

I. Nanofluidics, surrounding the frame

Nanofluidics, the study of fluidic transport at nanometer scales, has emerged quite recently in the footsteps of microfluidics. Pushing the limits of fluidic downsizing further is an attractive goal, in the spirit of scale reduction inherent to all micro- and nano-technologies. Various reasons may be seen to motivate these novel developments. First, from the point of view of biotechnological ("lab on a chip") applications, decreasing the scales considerably increases the sensitivity of analytic techniques, with the ultimate goal of isolating and studying individual macromolecules.[1,2] But also, from the point of view of fluidic operations, nanometric scales allow new fluidic functionalities to be developed, using the explicit benefit of the predominance of surfaces. Typical examples involve preconcentration phenomena,[3] the development of nanofluidic transistors[4,5] or the recently proposed nanofluidic diodes.[6,7] But the analogy to micro-electronics is somewhat limited: fluid molecules are not electrons and the notion of large scale integration for fluidic devices, *i.e.* nanofluidics as a way of increasing the density of fluidic operations on a chip even further, is probably not a pertinent goal to reach for fluidic operations.

But from a different perspective, nanofluidics also carries the hope that new properties will emerge by benefiting from the specific phenomena occurring at the smallest scales: new solutions may be obtained from the scales where the behavior of matter departs from common expectations. The great efficiency of biological nanopores (in terms of permeability or selectivity) is definitely a great motivation to foster research in this direction.[8,9] There is indeed a lot of room for improvement at these smallest scales: the example of aquaporins (AQP) is interesting in this context. Aquaporin channels are a key component of many biological processes[10] and play the role

Lydéric Bocquet

Prof. Lydéric Bocquet is head of the "Liquids at Interfaces" research group in the condensed matter laboratory of the University of Lyon. His interests—mainly curiosity driven—extend to domains where fluid dynamics connect to surface science. He combines theory, experiments and simulations to explore the intimate mechanisms of surface phenomena at the molecular level.

Elisabeth Charlaix

Elisabeth Charlaix is professor at the University of Lyon 1. She runs a research group on surface forces, and the behavior of liquids at the smallest scales. Her field of interest ranges from transport in porous materials, granular materials, to nanofluidics and bio-lubrication.

Laboratoire de Physique de la Matière Condensée et des Nanostructures, Université Lyon 1 and CNRS, UMR 5586, 43 Bvd. du 11 Nov. 1918, 69622 Villeurbanne Cedex, France. E-mail: lyderic.bocquet@univ-lyon1.fr

† Part of the themed issue: From microfluidic application to nanofluidic phenomena.

of water filters across biological membranes. These channels fulfill the conflicting tasks of being both extremely permeable to water, while extremely selective for other species.[8] To give an order of magnitude, the permeability of the water channel is typically *3 orders of magnitude larger* than what would be expected on the basis of the classical fluid framework for the same pore size.[11,12] A potential answer is that AQP, although a filter for water, is mainly *hydrophobic*, *i.e.* water repellent! Of course some hydrophilic polar nodes are distributed along the pore, so as to keep water in the mainly hydrophobic environment. Quoting the terms of Sui *et al.* in ref. 8, "*the availability of water-binding sites at these nodes reduces the energy barrier to water transport across this predominantly hydrophobic pathway, while the relatively low number of such sites keeps the degree of solute–pore interaction to a minimum. In balancing these opposing factors the aquaporins are able to transport water selectively while optimizing permeability*." Beyond this elementary picture, understanding how AQP fulfills its challenging properties would definitely be a source of inspiration and open new perspectives for technological breakthrough in filtration, desalination, power conversion,... However reproducing such a delicate composite (patched) architecture in bio-mimetic membranes is a great challenge, which requires breakthrough in the conception of its elementary constituents. But it points out that surfaces and their chemical engineering are key to optimizing fluid properties at the nano-scale.

However, in its roots, nanofluidics is not a new field: as judiciously recalled by Eijkel and van den Berg in their pioneering review on the subject,[1] many 'old' fields of physics, chemistry and biology already involve the behavior of fluids at the nanoscale. Like Monsieur Jourdain in "Le bourgeois gentilhomme" by Molière, one has done 'nanofluidics' for more than forty years without knowing it.[154] One may cite for example the domains of electro-kinetics (electro-osmosis or -phoresis,...) with applications in chemistry and soil science, membrane science (ultra-filtration, reverse osmosis, fuel cells,...), colloid chemistry, and of course physiology and the study of biological channels.[1] An interesting question is accordingly whether—on the basis of the novel 'nanofluidic' point of view—one may go beyond the traditional knowledge in those 'old' fields and obtain unforeseen results, *e.g.* allowing for better optimization of existing technologies? Our belief is that the answer to this question is already positive and this is one of the key aspects that we shall discuss in this review.

Finally it is also important to note that nanofluidics has emerged recently as a scientific field (*i.e.* naming a field as "nanofluidics") also because of the considerable progress made over the last two decades in developing nano-fabrication technologies, now allowing specifically designed nanofluidic devices to be fabricated, as well as the great development of new instruments and tools which give the possibility of investigating fluid behavior at the nanometer scale. One may cite for example: new electrical detection techniques, Surface Force Apparatus (SFA), Atomic Force Microscopy (AFM), nano-Particle Image Velocimetry (nano-PIV) coupling PIV to TIRF set-up (Total Internal Reflection Fluorescence), as well as the considerable progress made in computational techniques, like Molecular Dynamics simulations. It is now possible to *control/design* what is occuring at these scales, and *observe/measure* its effects. This is the novelty of the field, and the reason why nanofluidics now deserves its own terminology.

The paper is organized as follows: In the second section we will replace nanofluidics in the perspective of the various length scales at play in fluid dynamics. We shall in particular discuss the limits of validity of continuum (*e.g.* hydrodynamic) descriptions. In the third section we discuss the dynamics of fluids at interfaces and the nanofluidic tools which have been developed recently to investigate it. In the fourth section we will explore various transport phenomena occuring in diffuse layers. In the fifth section we raise the question of thermal noise in nanofluidic transport. Finally we will conclude by exploring some general considerations and expectations about nanofluidics, especially in terms of energy conversion and desalination.

As a final remark, this review, like any review, is our subjective and personal view of the field of nanofluidics and the perspectives one may foresee. We organized our exploration around the length scales underlying fluid dynamics at the nanometric scales and how nanofluidics allows us to probe the corresponding mechanisms. Accordingly, our aim is not to explore exhaustively the—already large—literature of the domain, but merely to disentangle the various effects and length scales underlying the behavior of fluids at nanometer scales. In doing so, we certainly hope that this review will raise new questions, open new directions and attract people to this fascinating domain.

II. Limits of validity of continuum descriptions and nanofluidic length scales

The introduction of the terminology "nanofluidics" (furthermore to define a specific scientific field) suggests that something special should occur for the transport of a fluid when it is confined in a channel of nanometric size. This leads to an immediate question: why should the nanometer length scale have anything specific for fluidic transport? Nanofluidics "probe" the properties of fluids at the nanoscale: so, what does one probe specifically in the nanometer range?

Actually one may separate two different origins for finite-size effects associated with nanometer scales: *bulk* and *surface* finite-size effects. The former, *bulk* effects are intimately associated with the question of validity of the classical continuum framework, in particular the Navier–Stokes equations of hydrodynamics: when do such descriptions break down? Can one then expect 'exotic' fluid effects associated with the molecular nature of the fluid? On the other hand, *surface*

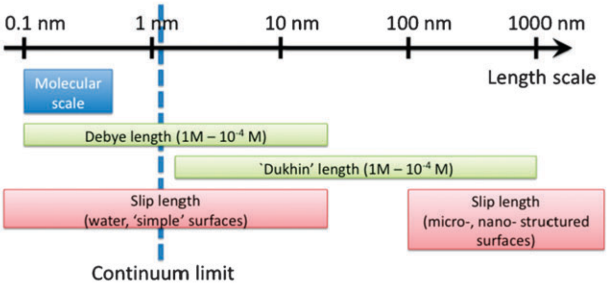


Fig. 1 Various length scales at play in nanofluidics.

effects play an increasingly important role as the "surface to volume ratio" increases (*i.e.* as the confinement increases). We have already pointed out the importance of such effects on the example of AQP water channels. As we show below, the surface effects occur at much larger scales than the 'bulk' deviations from continuum expectations.

The discussion on the length scales at play, to be explored in this section, is summarized in Fig. 1.

A Validity of bulk hydrodynamics

We first start with a short discussion on the validity of bulk hydrodynamics. In practice, this raises the question: when do the Navier–Stokes (NS) equations break down? These equations were developed in the 19th century to describe fluid flows at, say, the human scale. What comes as a surprise is their incredible robustness when applied to ever smaller scales.

As a fact, for simple liquids, the continuum framework of hydrodynamics is apparently valid down to the nanometer scale. In other words, there is no expected deviation to the bulk NS equations for confinement larger than ~ 1 nm. This very surprising result is actually suggested by a number of experimental and molecular simulations studies. On the experimental side, one may cite the early work by Chan and Horn[13] and later by Georges *et al.*[14] using Surface Force Apparatus, where the prediction of hydrodynamics (the Reynolds formula in their case) was verified to be valid for confinement larger than typically ten molecular diameters. More recently, and specifically for water, the works by Klein *et al.*[15] and E. Riedo *et al.*[16] showed that water keeps its bulk viscosity down to ~ 1–2 nm, with a drastic change of behavior for stronger confinements, where the wettability of the confining surface plays a role.[16] A similar behavior was found for other liquids like octamethylcyclotetrasiloxane (OMCTS).[17,18]

This threshold for the applicability of continuum hydrodynamics was also investigated using Molecular Dynamics simulations of confined water, and the same value of about 1 nm came out of the simulation results.[19–21] Furthermore, it is interesting to note that beyond the validity of continuum equations, the value of the viscosity also remains quantitatively equal to its bulk value. We show in Fig. 2 the results of MD simulations for the viscosity η of water (SPC-E model) measured in various confinements. The value of the viscosity is obtained by measuring the shear-stress on the confining plates for a given shear-rate (corrected for slippage effects). As shown in this figure, the shear viscosity of water keeps its bulk value down to confinements of ≈ 1 nm (typically 3 water layers).

Now, in contrast to the viscosity, other transport coefficients may be more strongly affected by confinement. This is the case for example of the (self-)diffusion coefficient which was shown to depend algebraically on the confinement width.[22,23] As a consequence, the diffusion coefficient in confinement strongly departs from its Stokes–Einstein prediction, $D = k_BT/(3\pi\eta\sigma)$, with σ the molecule diameter. The latter is accordingly not a correct measure of the viscosity, as sometimes assumed.[19,24]

Fundamentally the validity of NS equations down to typically 1 nm (for water) is *a priori* unexpected. Navier–Stokes (NS) equations have been developed to account for the fluid dynamics at 'large' scales and they rely on an assumption of a local and linear relationship between stress and velocity gradients.

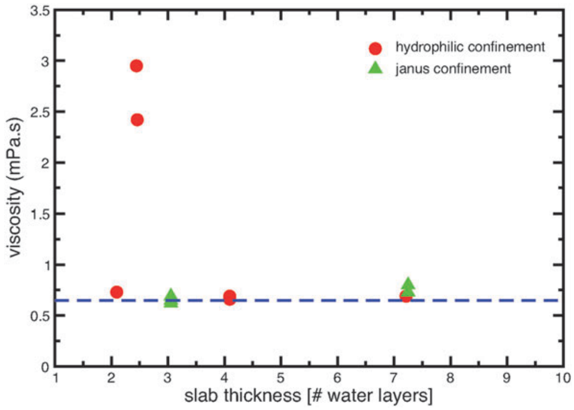


Fig. 2 Viscosity of water *versus* confinement, as measured from MD simulations of water confined in a nano-slit made of either two hydrophilic walls, or one hydrophilic and one hydrophobic wall (janus confinement) [courtesy of D. Huang].

To our knowledge, there is no firmly grounded argument for this validity and we propose here a few tentative leads. NS equations, like any continuum framework, rely actually on the key assumption of a separation of length- and time-scales between the investigated length scale and the 'molecular' dynamics. This is the "hydrodynamic limit", in which the continuum framework should hold. Under this time- and length- scale separation, the microscopic dynamics associated with a huge number of degrees of freedom $\mathcal{N} \approx 10^{23}$ reduces to equations with just a few degrees of freedom (velocity field, pressure, density, temperature, *etc.*), while all the complexity is hidden in just a few phenomenological coefficients. Achieving this averaging out of the "fast variables" is a huge challenge which can be achieved systematically only in a few limiting cases, see *e.g.*ref. 25 for an explicit example. In general, the elimination of fast variables is summarized in the Green–Kubo relationship for the transport coefficients: for example the shear viscosity η can be expressed as the integral of the stress–stress correlation function according to

$$\eta = \frac{1}{Vk_BT}\int_0^\infty \langle \sigma_{xy}(t)\sigma_{xy}(0)\rangle_{\text{equ}}\, dt \qquad (1)$$

with σ_{xy} a non-diagonal component of the stress tensor[26] and $\langle\cdot\rangle_{\text{equ}}$ denotes an equilibrium average.

The time-scale separation underlying the validity of eqn (1) requires the microscopic time scales, τ_σ, characterizing the stress–stress correlation function, to be much smaller than a hydrodynamic time scale. This time scale is *e.g.* the relaxation time of momentum, which for a given wave vector q has the expression $\tau_q = (\nu q^2)^{-1}$, with $\nu = \eta/\rho$ the kinematic viscosity and ρ the mass density. An implicit condition of validity is accordingly $\nu q^2\tau_\sigma < 1$. This fixes the limit for time-scale separation at confinements w larger than a viscous length scale:

$$w > \ell_c = \sqrt{\nu\cdot\tau_\sigma} \qquad (2)$$

Putting numbers for water, a typical correlation time for the stress–stress correlation function is in the picosecond range $\tau_\sigma \sim 10^{-12}$ s, while $\nu = 10^{-6}$ m^2 s^{-1}. Altogether this gives for water

$$\ell_c \approx 1 \text{ nm} \quad (3)$$

A nanometric characteristic length scale thus emerges naturally as a lower boundary for the validity of the notion of viscosity, and thus for the application of standard NS hydrodynamics. This is indeed the experimental value below which strong departure from the continuum framework is observed.

Of course, it would be interesting to explore further the physical contents of this condition: what occurs for smaller confinements? Also, can one tune ℓ_c in specific systems? *etc.*

In the extreme limit of single file transport, the transport behavior is indeed expected to strongly deviate from the bulk expectations, with the predicted occurrence of anomalous diffusion, and of fast stochastic transport, to cite some examples.[28–30] Apart from biological systems—where molecular pores are the rule—fabricating channels with molecular size is however still out of reach with state-of-the-art nanofabrication technologies.

As a side remark, it would be interesting to extend this discussion to other transport phenomena, like for example heat transfer at nanoscales. Heat transfer from nanoparticles and nanostructures to a fluid is an actively growing field, *e.g.* in the context of cooling enhancement by structured surfaces or local heating of fluids *via* nanoparticles.[27] Similar arguments as above may be proposed, and lead to an equivalent sub-nanometer heat length scale (for water at room temperature). This suggests that the continuum phenomenological picture for heat transport in fluids, involving Fourier law, is expected to hold down to nanoscales, in agreement with observations.[27] However a full investigation of the limit of applicability of the corresponding continuum phenomenological laws remain to be developed in general.

A conclusion of the above discussion is that for most nanofluidic applications involving water in supra-nanometric confinements, bulk NS equations can be safely used to account for the fluid transport.

B A broad spectrum of length scales

Aside from ℓ_c introduced above, there are a number of length scales which enter nanofluidic transport. These length scales are all related to surface effects in one way or another. We review in this section a panel of these length scales, climbing up from the smallest—molecular—scales to the largest, but still nanometric, scales. This panel is not exhaustive, and we merely insist on the main length scales which appear generically in nanofluidic problems. Our aim is to show that there are a number of pertinent scales which lie indeed in the nanometric range. This implies that specific nanofluidic phenomena will show up when the confinement compares with these values.

1. Molecular length scales. At the smallest scales, the granularity of the fluid and its components (solvent, ions, dissolved species,…) should play a role. This is defined by the molecular scale, associated with the diameter of the molecules σ, typically in the Angström scale, 3 Å for water. As we discussed above, this is the scale where the validity of hydrodynamics breaks down, typically for confinements of 3–5 molecular diameter. Various phenomena then come into play. First, the structuring and ordering of the fluid at the confining walls plays a role. Such ordering was indeed shown experimentally to induce oscillatory dissipation in liquid films with a width of several molecules diameter.[17,18]

Similarly non-local rheological properties have been predicted in strongly confined situations,[31] associated with a length characterizing non-locality typically in a few diameters range (indeed in line with the above limit of NS hydrodynamics). This result echoes recent findings of non-local flow curves in confined soft glasses.[32] An interesting consequence of non-locality, as pointed out in ref. 32, is that the specific nature of surfaces does influence the global flow behavior in confinement, in line with recent observations in AFM measurements by Riedo *et al.* for strongly confined water (less than 1 nm) films at hydrophilic and hydrophobic surfaces.[16]

But even more drastic are the deviations occurring in the extreme limit where only one molecule can enter the confining pore. Very strong correlations and collective motion build up in the liquid dynamics, leading to the so-called "single-file" transport.[33,34] The associated transport differs strongly from the bulk hydrodynamic predictions and various specific phenomena show up, like non-fickian transport,[34] fast water transport and ion transfer, stochastic flow with the occurrence of "bursts", *etc.*, see *e.g.*ref. 33. Simulations of water transport through single-file carbon nanotubes (where only one water molecular can pass through the channel) showed indeed flow rates much larger than those predicted from the Poiseuille law of hydrodynamics.[28]

Note however that this behavior is limited to nanochannels where molecules cannot cross each other, *i.e.* with diameters not exceeding the water molecular length scale, like in biological channels. Up to now, artificially produced nanochannels do not reach this limit. This point is interesting to discuss in the context of recent experimental results obtained for flow through carbon nanotube membrane.[9,35,36] A recent experimental first was achieved in Bakajin's group, who investigated the flow through a membrane constituted of carbon nanotube with a size in the range 1.3–2 nm,[9] following the work on larger nanotubes (∼7 nm) by Majumder *et al.*[35] These works demonstrated massively enhanced flow permeability, as compared to bulk predictions (in terms of slippage, to be discussed below, this would correspond to slip lengths up to a micron, far above any expected result for these surfaces). More recent experiments performed with membranes made of wider carbon nanotubes suggest a large, but still smaller, enhancement of the permeability.[36] Therefore, while it would be tempting to discuss these results in the context of single-file transport and the predictions by Hummer *et al.* on fast single-file transport in sub-nanometer carbon nanotube,[33] this point of view is difficult to justify: as discussed above, fluid flow in nanotube pipes with a supra-nanometric diameter can be safely described by continuum approaches (Stokes equation) and the dynamics is definitely not single file. Molecular Dynamics simulations of flow through similar carbon nanopipes indeed suggest large flow enhancement,[24,37] but still far below the experimental results. There is therefore a strong need to perform flow experiments through a single object, *i.e.* through a single carbon nanotube. This is still an experimental challenge.

Altogether transport at the molecular scale offers a rich panel of transport behavior. This is definitely the scale where, as we quoted in the introduction, new solutions and technological breakthrough could emerge as the behavior of matter departs from common expectations. This suggests that the molecular scale would be the ultimate scale to reach for nanofluidics. But at present, producing channels of subnanometric size remains a technological challenge.

2. Electrostatic length scales and their dynamic influence. The scales associated with the interaction of the flowing matter with its environment also enter the game. And of course, electrostatics plays a key role in this context. We leave aside interactions deriving indirectly from electrostatics to focus on the interactions between charged species. These obviously concern ions dissolved in the fluid: ion–ion interactions, and interactions of ions with charged surfaces. This also involves the dipolar interaction among water molecules, as well as H-bonding.

Electrostatics involves a rich panel of length scales. And more importantly, due to the long range nature of electrostatics, these length scales are intimately coupled: effects occurring at the smallest electrostatic lengths do climb the scales to span the whole range.

a. Bjerrum length (and some derivatives). We start with the Bjerrum length, ℓ_B, which is defined as the distance at which the electrostatic interaction $\mathcal{V}_{el}$ between two charged species becomes of the order of the thermal energy, $k_{\mathrm{B}}T$: $\mathcal{V}_{el}(\ell_B) \approx k_{\mathrm{B}}T$. For two ions, with valence Z, embedded in a dielectric medium with dielectric constant ε, this takes the form

$$\ell_B = \frac{Z^2 e^2}{4\pi\varepsilon k_{\mathrm{B}} T} \tag{4}$$

(e the elementary charge). For bulk water at ambient temperature and a valency $Z = 1$, this gives $\ell_B = 0.7$ nm. The Bjerrum length therefore compares with the molecular range in this case. Note however that for multi-valent ions or for organic solvent with lower dielectric constants, ℓ_B may be much larger and thus separate from the molecular range.

Other length scales could be introduced along the same lines by comparing thermal energy to *e.g.* charge–dipole, dipole–dipole, *etc.* interactions, taking into account the multi-pole interaction under consideration. This leads *a priori* to length scales smaller than the Bjerrum length introduced above. For the charge–dipole interaction, this gives for example $\ell_d = (\ell_B \frac{p}{e})^{1/2}$, with p the dipole strength: for water parameters, ℓ_d is in the Angström range. It does correspond to the typical thickness of a hydration layer around an ion.[38]

By definition, the Bjerrum length (and other Bjerrum-like lengths such as ℓ_d above) is the scale below which direct electrostatic interactions dominate over thermal effects. Its consequences on nanofluidic transport is thus expected at the molecular scale. For example, for confinement below ℓ_B one expects a large free-energy cost to undress an ion from its hydration layer and make it enter a molecular pore. This has therefore immediate consequences on the filtering process of charged species, as in biological ion channels.

But as pointed out above, a strong interplay with upper scales is expected due to the long range nature of electrostatic interactions. For example, as we will discuss in section IV, ion-specific effects at interfaces, the origin of which occurs at a scale ℓ_B, have strong effects on the electro-kinetics at nanometric, micrometric and even larger scales.

b. Debye length scale. A central concept of electrostatics is the notion of the electrostatic diffuse layer (EDL). At a charged interface, the EDL is the region where the surface charge is balanced by the cloud of counterions. In this region the ion concentration profile depart from their bulk values due to the interaction of ions with the surface charge. This is therefore the region where local electro-neutrality is not obeyed.

The Debye length emerges naturally from the Poisson–Boltzmann theory and characterizes the electrostatic screening in the bulk electrolyte.[39] It is defined in terms of the salt concentration ρ_s according to the expression:

$$\lambda_D = (8\pi\ell_B\rho_s)^{-1/2} \tag{5}$$

Note that in this expression, as in the following, ρ_s denotes the salt concentration *in the reservoirs*, thereby fixed by the chemical potential of the ions. The value of the Debye length depends on the salt concentration and ranges typically between tens of nanometers (30 nm for $\rho_s = 10^{-4}$ M) down to Angströms (3 Å for $\rho_s = 1$ M). In physical terms, the electrostatic free-energy of an ion dressed by its spherical cloud of counterions with size λ_D is of the order $k_{\mathrm{B}}T$.

The Debye length characterizes the width of the EDL. Interestingly, it is independent of the surface charge and only depends on the bulk ion concentration. However, in the limiting case of a salt free solution (with $\lambda_D \to \infty$), the width of the diffuse layer is merely fixed by the so-called Gouy–Chapman length, to be defined in the next paragraph.

The Debye length plays a central role in nanofluidics for various reasons. First, as quoted above, this is the region in the fluid close to charged surfaces where local charge electro-neutrality is broken. Under an applied electric field, this is thus the region where volume electric forces, $f_e = \rho_e E_e$, will apply, with ρ_e the charge density and E_e an applied electric field. On the other hand these forces vanish in the bulk of the material due to a vanishing charge density $\rho_e \to 0$ far from the surfaces. Therefore tuning the fluid properties inside the nanometric EDL is expected to affect the whole response of the system under an electric field. Accordingly, tuning the dynamics of the fluid in the nanometric EDL, or modifying its structure, will have a macroscopic impact on the fluid dynamics at scales much larger than the Debye length. We will return extensively to this question below, in section IV.B.

Second, a specific behavior is expected when the Debye layers overlap in a nanopore, which occurs when its size is of the order of twice the Debye length. This phenomenon has a strong effect on fluidic transport and has been the object of intense research recently, as reviewed recently by Schoch *et al.*,[40] see section IV. One may cite in particular the phenomena of ion enrichment and exclusion.[2] While these phenomena were known for a long while in membrane technology, they found new applications in the field of nanofluidics. It is at the root of a number of novel fluidic phenomena, such as permselectivity,[2] nanofluidic diodes[6,7] or surface dominated ion transport.[41]

A key practical aspect is that the condition of Debye layer overlap is much less stringent to fulfill with actual nanofabrication techniques. These are now able to produce individual nanopores with at least one dimension in the range of ten nanometers, *i.e.* potentially suited for Debye layer overlap.

c. Length scales associated with surface charge: Gouy-Chapman and Dukhin. Other electrostatic length scales may be constructed on the basis of surface electrostatic properties. We have already mentioned the *Gouy-Chapman length scale*, which shows up for the behavior of a salt solution very close to a charged surface. For a surface charge density Σe, this has an expression $\ell_{GC} = 1/(2\pi\Sigma\ell_B)$. In contrast to the Debye length, the Gouy-Chapman length depends explicitly on the surface density Σ of charges of the confining surface (Σ in units of m^{-2}), but not on the bulk ion concentration. Physically, the Gouy-Chapman length can be defined as the distance from the wall where the electrostatic interaction of a single ion with the wall becomes of the order of the thermal energy. For typical surface charges, say $\Sigma e \sim 50\ \mathrm{mC/m^2}$ ($\approx 0.3\ e/\mathrm{nm}^2$) to fix ideas, then $\ell_{GC} \sim 1$ nm. The Gouy-Chapmann length plays a role for solutions with very small salt concentration. This has an influence on the conductance in nanochannels as we discuss in section IV C.

More interesting is a length which can be defined on the basis of the comparison between the bulk to the surface electric conductance (relating electric current to an applied electric field). This introduces what can be termed as a "Dukhin length", by analogy to the Dukhin number usually introduced for colloids.[155] Indeed the conduction probes the number of free carriers (ions), so that in a channel of width h and surface charge Σ, the equivalent bulk concentration of counterions is $2\Sigma/h$. One may define a Dukhin number $Du = |\Sigma|/(h\rho_s)$, where ρ_s is the concentration of the salt reservoir.[42] A 'Dukhin length' can then be defined as

$$\ell_{Du} = \frac{|\Sigma|}{\rho_s}. \tag{6}$$

To put in numbers, for a surface with a surface charge density $e\Sigma \sim 50\ \mathrm{mC/m^2}$ ($\approx 0.3\ e/\mathrm{nm}^2$), ℓ_{Du} is typically 0.5 nm for $\rho_s = 1$ M, while $\ell_{Du} = 5\ \mu\mathrm{m}$ for $\rho_s = 10^{-4}$ M! The Dukhin length characterizes the channel scale below which surface conduction dominates over the bulk one. In a different context of charge discontinuities at surfaces, it has also been interpreted in terms of an electrokinetic "healing" length.[43]

This length scale can actually be rewritten in terms of the Debye and Gouy-Chapman lengths as $\ell_{Du} \sim \lambda_D^2/\ell_{GC}$. Note also that in the limit where the Debye length is large compared to the Gouy-Chapmann length, $\ell_{GC} < \lambda_D$, the non-linear Poisson–Boltzmann expression for the electrostatic potential[39] allows the above length to be rewritten as $\ell_{Du} \sim \lambda_D \exp[e|V_s|/2k_\mathrm{B}T]$, with V_s the surface potential. The corresponding form for the Dukhin number is more standard in colloid science.[42]

This length plays an important role for the conductance in nanochannels, *cf.* section IV C, where surface effects are shown to strongly affect conductance.[41,44]

3. Slip lengths and surface friction. Up to now we merely considered length scales characterizing the structure of the fluid and its components (ions,...). However the *dynamics* of fluids at interfaces introduce various length scales. This is in particular the case of the so-called *slip length*, b, characterizing the hydrodynamic boundary condition of a fluid at its confining interfaces. The latter is defined according to the Navier boundary condition (BC) as ref. 21

$$b\partial_n v_t = v_t \tag{7}$$

with n,t points to the normal and tangential directions of the surface, v_t is the tangential velocity field and b the slip length. The slip length characterizes the friction of the fluid at the interface and large slip lengths are associated with low liquid–solid friction.

This point has been amply explored experimentally and theoretically: a key result which emerges from this measurements is that the slip length of water at solid surfaces depends crucially on the wettability of the surface.[21,45] We will come back more exhaustively on this point in section III. At this level, one should keep in mind that slip lengths in the range of a few tens of nanometers are typically measured on hydrophobic surfaces, while b is sub-nanometric on hydrophilic surfaces. Note however that very large slip length, in the micron range, may be obtained on nano- and micro-structured interfaces.[46,47]

This offers the possibility to modify the nanofluidics in pores using chemical engineering of the surfaces. As the pore size compares with b a considerable enhancement of fluid transport is accordingly expected. Furthermore one may remark that these values of b also compare with typical Debye lengths. Therefore slippage effects are expected to affect ion transport at charged surfaces. We will discuss these aspects more exhaustively below.

A final remark is that other surface related lengths may be constructed. This is the case in particular of the Kapitza length, which characterizes the boundary condition for thermal transport across an interface, see *e.g.* ref. 21. The Kaptiza length behaves quite similarly to the slip length and strongly depends on the wettability of the interface.[48]

4. Other length scales. The review of length scales above is not exhaustive and many other length scales may enter the game, depending on the problem, geometry and system considered. For example, in the context of transport of macromolecules (colloids, polymers, DNA, RNA, proteins, ...), the typical size of the particle—diameter or radius or gyration—plays of course a central role. More precisely a key quantity is the free energy associated with the confinement of the macromolecule, $\mathcal{F}_c$, which fixes the partitioning of the molecule in the nanopore with respect to the bulk. The probability of the passage of a macromolecule through a pore is expected to scale like $\exp[-\mathcal{F}_c/k_\mathrm{B}T]$, hence fixing its permeability.[49] For example for a polymer chain confined in a pore with size D, the entropic cost of confinement takes the form $\mathcal{F}_c \approx k_\mathrm{B}T(R_g/D)^{1/\nu}$, with $R_g = a \times N^\nu$ the radius of gyration of the polymer (a the monomer size and N the polymer length). For a polymer in good solvent, the Flory exponent is $\nu = 5/3$. Pores the size of which is below the radius of gyration of the molecule thus act as a molecular sieve. This has been explored in various experiments of polymer translocation in nanopores following the pioneering work of Bezrukov *et al.*[50,51] and Kasianowicz,[52] as well as in the context of DNA separation.[53]

C Some practical conclusions

The discussion above is summarized in Fig. 1, where we organized the various scales at play along the scale axis. This points to several conclusions.

1. A number of specific phenomena occur in the nanometric range, which indeed justifies the specificity of the dynamics of fluids in the nanometer range, *i.e.* nanofluidics.

2. An important conclusion is: For water under normal conditions, the Navier–Stokes equation remains valid in nano-channels down to typically 1–2 nanometers. This is good and bad news. Good because *one may safely use NS equations for most nanofluidic phenomena*. This robustness of continuum hydrodynamics is remarkable.

But this is in some sense *bad news* since it means that reaching specific effects associated with the granularity of the fluid requires molecular confinement, *i.e.* confinement below the nanometer. Producing nanochannels with such sizes is technologically difficult to achieve. On the other hand, this probably explains why most biological channels have a molecular entanglement: to reach specificity and avoid the "universality" associated with continuum equations, one should reach the fluid molecular scale.

3. This suggests that the molecular pore is the ultimate scale to reach for nanofluidics, where, as we quoted in the introduction, new solutions and technological breakthrough could emerge as the behavior of matter departs from the common expectations. But at present, there is however still a long way before this scale is technologically accessible.

4. While it remains difficult to tune the bulk behavior of the confined fluid, there is much more room and possibilities to benefit from surface effects. This is clearly apparent in Fig. 1, where surface effects enter the nanofluidic game at much larger length scales. Surface effects enter in particular *via* the Debye length, Dukhin length, slip length, all typically in the ten nanometer range (and even more for the Dukhin length). This suggests that specific effects will show up when one of these lengths will compare with the pore width. Furthermore, particular effects should also occur when two of these lengths become comparable, independently of the confinement. This will be explored in section IV.

Altogether nanofluidics is an incredible playground to play with surface effects!

III. Dynamics at surfaces and the toolbox of nanofluidics

We have discussed above that the bulk laws of hydrodynamics are valid down to very small scales, typically $\sim$1 nm, so that bulk Navier–Stokes equations can be used for most nanofluidic flows. However as the size of the nanochannel decreases, the dynamics at its surface should play an increasingly important role. Navier–Stokes equations require boundary conditions (BC) for the hydrodynamic flow at the device's surface, and a specific knowledge of these BC is a pre-requisite to apprehend flow at the nanoscale.

As quoted in eqn (7), the BC at a solid surface introduces a new length, the so-called Navier length or slip length, which relates the tangential velocity v_t to the shear rate at the wall, Fig. 3:[54]

$$v_t = b\frac{\partial v_t}{\partial z} \tag{8}$$

Obviously the control of the slip length is of major importance for flows in confined geometries, and it can have dramatic consequences on the pressure drop, electrical, and diffusive transport through nanochannels. For instance the relative increase in hydraulic conductance of a cylinder of radius r due to wall slippage is $1 + 8b/r$; it can be much larger for the electrical conductance or electro-osmosis, as we will discuss in the sections below.

As a consequence a sustained interest has been devoted in the last ten years for the investigation of the BC and its dependence on interfacial properties, such as surface topography and liquid–solid interactions. Reviews on this subject can be found in ref. 21, 55 and 56.

A Theoretical expectations for slippage

The theoretical understanding of slippage has been the object of intense research over the last decade, see ref. 21 for a review. The BC has been studied theoretically by Molecular Dynamics simulations, as well as in the context of linear response theory.[45,57–61,63] It is now well understood that the BC on atomically smooth surfaces depends essentially on the structure of the liquid at the interface, itself determined by its interactions, its commensurability with the solid phase, and its global density.[21]

In order to connect the slip length to the interfacial properties, it is useful to interpret the slip length in terms of liquid–solid friction at the interface. The friction force at the liquid–solid interface is linear in slip velocity v_t:

$$F_f = -\mathscr{A}\lambda v_t \tag{9}$$

with λ the liquid–solid friction coefficient and $\mathscr{A}$ the lateral area. By definition the slip length is related to the friction

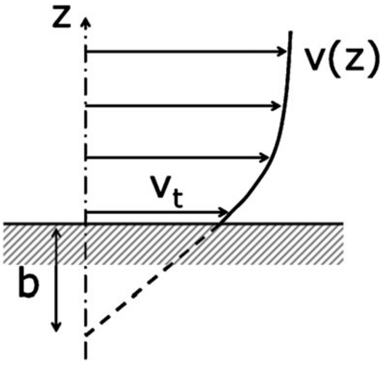


Fig. 3 Hydrodynamic slip and definition of the slip length b: the slip length is the distance in the solid at which the linear extrapolation of the velocity profile vanishes.

coefficient according to $b = \eta/\lambda$, with η the bulk viscosity. As a phenomenological friction coefficient, λ can be expressed in terms of equilibrium properties in the form of a Green–Kubo (GK) relationship:[21,59,62]

$$\lambda = \frac{1}{\mathscr{A}k_B T}\int_0^\infty \langle F_f(t)\cdot F_f(0)\rangle_{\text{equ}}\,\mathrm{d}t \tag{10}$$

where F_f is the total (microscopic) lateral force acting on the surface. In practice, this expression is difficult to estimate. However some useful information on the slip length may be extracted from it.

A rough estimate of slippage effects can be proposed.[63] One may indeed approximate the GK expression as $\lambda \approx \frac{1}{\mathscr{A}k_B T}\langle F_f^2\rangle_{\text{equ}} \times \tau$, with τ a typical relaxation time and $\langle F_f^2\rangle_{\text{equ}}$ the mean-squared lateral surface force. We then write $\tau \sim \sigma^2/D$ where σ is a microscopic characteristic length scale and D the fluid diffusion coefficient, while the rms force is estimated as $\langle F_f^2\rangle_{\text{equ}} \sim C_\perp \rho\sigma(\varepsilon/\sigma)^2$, with ε a typical fluid–solid molecular energy, ρ the fluid density and $C_\perp$ a geometric factor that accounts for roughness effects at the *atomic level* (large $C_\perp$ corresponding to a larger atomic roughness).[45,63] Altogether this provides a microscopic estimate for the slip length as

$$b \sim \frac{k_B T \eta D}{C_\perp \rho \sigma \varepsilon^2} \tag{11}$$

A more systematic derivation leads to a very similar result,[21,62] with eqn (11) multiplied by the inverse of the structure factor of the fluid, $S_w(q_\parallel)$, computed at a characteristic wave-vector of the solid surface $q_\parallel$: $b \propto S_w(q_\parallel)^{-1}$. This term measures a kind of commensurability of the fluid with the underlying solid structure, in full analogy with solid-on-solid friction. Note also that the slip length depends on the product $\eta \times D$, so that according to the Stokes–Einstein relationship, the slip length is not expected to depend on the bulk fluid viscosity (except for specific situations where these two quantities can be decorrelated). The scalings proposed by the above simple estimate are in good agreement with molecular dynamics results.[21,63]

According to eqn (11) significant slippage should occur on very smooth surfaces, in the case of a low density at the wall (which requires both a moderate pressure and unfavorable liquid–solid interactions), or a low value of $S_w(q_\parallel)$, which characterizes liquid/solid commensurability. It also suggests that low energy surfaces, *i.e.* with small liquid–solid interaction ε, should exhibit large slippage. Accordingly, hydrophobic surfaces should exhibit larger slip length than hydrophilic ones.

We plot in Fig. 4 (top) results for the slip length obtained by Molecular Dynamics of a water model over a broad variety of surfaces.[45,63] A 'quasi-universal' dependence of the slip length on the contact angle is measured, with—as expected from the above arguments—an increase of the slip length for more hydrophobic surfaces.

B Experimental results

On the experimental side a race has engaged toward the quantitative characterization of BCs within nanometer resolution, using the most recent developments in optics and scanning probe techniques.[55,56] This experimental challenge has generated considerable progress in experimental tools for nanofluidic measurements, and the values of the slip lengths reported had a tendency to decrease as the resolution and robustness of the techniques has increased. An intrinsic difficulty stems from the fact that the BC is a continuous medium concept involving an hydrodynamic velocity averaged on many molecules in the liquid and extrapolated to the wall: it does not simply reduce to the velocity close to the surface of molecules or tracer particles, which are furthermore submitted to Brownian motion and reversible adsorption. Also, the surface has to be smooth and homogeneous on the scale of the probed area in order to characterize an *intrinsic* BC, and not an *effective* slip length. Thus, the problem of the BC has been a discriminating test of the capacity of instruments to perform quantitative mechanical measurements at the nanoscale. We propose here a brief overview of the current state-of-the-art. This offers the opportunity to quote the various tools which have been developed recently to study flow and fluid dynamics at nanofluidic scales.

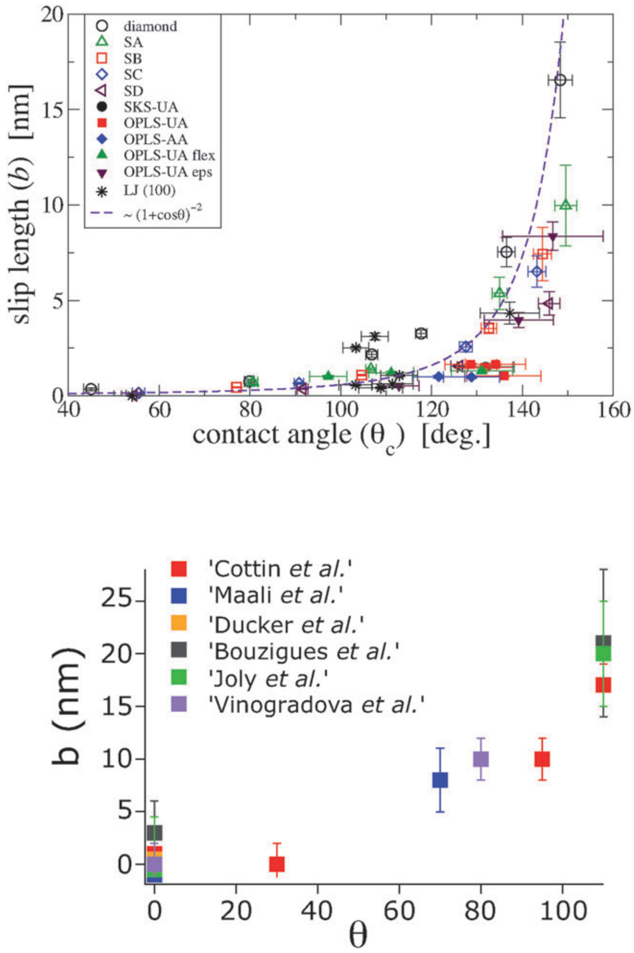


Fig. 4 Slip length b of water as a function of the contact angle on various smooth surfaces. *Top*: Results from Molecular Dynamics simulations by Huang *et al.* for the slip length of water on various surfaces, from ref. 45. The dashed line is a theoretical fit scaling like $b \propto (1 + \cos\theta)^{-2}$. *Bottom* Compilation of various experimental results.

The Surface Force Apparatus provided the first sub-nanometer resolution measurement of BCs, based on the viscous force F_v acting between two crossed cylinders whose distance D is measured within Angström resolution by the so-called FECO interferometric fringes:[64]

$$F_v = \frac{6\pi\eta R^2\dot{D}}{D} f^* \tag{12}$$

$$f_s^*(D/b) = \frac{D}{3b}\left[\left(1+\frac{D}{6b}\right)\ln\left(1+\frac{6b}{D}\right)-1\right] \tag{13}$$

with η the liquid viscosity and R the cylinder radius.[65] In case of a no-slip BC, F_v reduces to the Reynolds force between a sphere and a plane. Chan and Horn[13] investigated the flow of various organic liquids confined between atomically flat mica surfaces and showed that they obeyed the no-slip BC, with the no-slip plane located about one molecular size inside the liquid phase. Their results were confirmed by other experiments running the SFA in a dynamic mode, and extended to water and to various (wetting) surfaces.[14,66] Further investigations focused on non-wetting surfaces and reported the existence of very large slip lengths ($\simeq 1$ μm) associated to a complex behaviour, with slippage appearing only under confinement and above a critical shear stress.[67] However it was shown that elasto-hydrodynamics effects due to non perfectly rigid surfaces (such as glued mica thin films) retrieve this typical behaviour.[68] Experiments on rigid surfaces later reported the existence of substantial slippage of water on smooth homogeneous non-wetting surfaces, with a slip length increasing with the contact angle and reaching a range of 20 nm at contact angles of 110°,[69,70] see Fig. 4.

The AFM with a colloidal probe operates along a similar principle but with a smaller probed area than the SFA. In pioneering investigations Craig *et al.* found a shear-rate dependent BC of sucrose solutions on partially wetting SAMs, with slip lengths reaching 20 nm at high velocities,[71] while Bonacurso *et al.* found a constant (no shear dependent) slip length of 8–9 nm for water on wetting mica and glass surfaces.[72] These dissimilar findings also did not agree with available theoretical results. However Ducker *et al.*, using an independant optical measurement of the probe-distance in the AFM, found a regular no-slip BC on a similar system as the one investigated by Craig *et al.*[73] It was recently shown by Craig *et al.* that the lateral tilt of the cantilever was a significant issue in interpretating hydrodynamic forces.[74] Vinogradova *et al.* also evidenced the important effect of the dissipation due to the cantilever.[75] Using a "snow-man" probe made of two spheres glued on the top of each other, they showed a perfect agreement with the no-slip BC on hydrophilic surfaces, confirming SFA experiments, and a moderate slip length of 10 nm of water onto rough hydrophobic surfaces. Finally dynamic measurements using an oscillatory drive were implemented in the AFM by Maali *et al.* so as to measure the dissipation coefficient with high accuracy.[76] Using this technique they found a no-slip BC of water onto mica surface, and a 8 nm slip length on atomically smooth highly ordered pyrolytic graphite (HOPG) with contact angle 70°.[77]

An alternative way to measure interfacial hydrodynamic dissipation is through the dissipation–fluctuation theorem, for instance by measuring the modification of the Brownian motion of a probe at the vicinity of a surface. In Joly *et al.* experiment[22] the average transverse diffusion coefficient of colloidal particles confined in a liquid slab was measured by Fluorescence Correlation Spectroscopy as a function of the film thickness, and related to the boundary condition on the slab walls. A resolution better than 5 nm is obtained with particles of 200 nm nominal diameter, by averaging the data over 100 000 events crossing the measurement volume. The originality of the method is to give access to the boundary condition without any flow, *i.e.* really at thermodynamic equilibrium. It is thus directly comparable to theoretical predictions. A collateral result is the direct probe of a property of great interest in nanofluidics, *i.e.* the motion of a solute close to a wall.

Besides the dissipative methods, great progress was achieved in the development of optical methods giving direct access to the flow field and its extrapolation to the solid wall with a sub-micrometer resolution.

The Micro-Particle Image Velocity (μ-PIV) technique has been widely used in microfluidics. First attempts to determine the BC yielded slip lengths of micrometer amplitude,[78] difficult to conciliate with theoretical expectations, although the variation of slippage with the liquid–solid interaction were consistent. Gaining better resolution raised many issues. Besides colloidal forces acting on particles close to a wall, such as electro-osmosis induced by the streaming potential, which has to be taken into account, an important source of error is the depletion induced by the colloidal lift.[79] By using the signal of non-moving particles adsorbed to the wall Joseph *et al.*[80] increased the resolution on the slip length to 100 nm in a configuration similar to ref. 78, and found no slippage at this (100 nm) scale whatever the wettability of the surface. A recent investigation using double-focus FCS confirmed the existence of slippage of a few ten nanometers on hydrophobic surfaces.[81] A major improvement toward nanometric resolution is the Total Internal Reflection Fluorescence method which uses an evanescent wave at the interface to be probed, and allows the measurement of the distance of a tracer particle to the wall through the exponential decay of the signal amplitude.[82,83,84,103] Bouzigues *et al.* found a water slip length of 21 nm (±12 nm) on OTS coated glass capillaries,[103] while Lasne *et al.* measured a water slip length of 45 nm (±15 nm) on a similar surface.[84] These results are consistent and the different values of slip length may well be ascribed to slight variations in surface preparation. Note that in such measurements, the size of the probe itself is a major issue in the race toward nanometric resolution: Brownian motion increases as the particle size decreases, not only adding noise to the measurement, but allowing particles to leave the focal plane between two successive images which decreases the locality of the measurement. Some specific statistical treatment has to be performed in order to avoid any bias in the estimated velocity.[84,85,103]

A summary of the results obtained with water on various smooth surfaces is gathered in Fig. 4 plotting the slip length as a function of the contact angle. The various methods show the same trend: water does not slip on hydrophilic surfaces, and develops significant slippage only on strongly hydrophobic surfaces. The highest slip lengths reach the range of 20–30 nm: accordingly hydrodynamic slippage is non-relevant in microfluidic devices, but of major importance in nanochannels. On this basis one expects that water flow in highly confined hydrophobic pores, such as mesoporous media or biological channels (*cf.* the Aquaporin example in the Introduction), should be much less dissipative than ordinary Poiseuille flow. The results obtained with the various methods show a very good qualitative agreement with theory and molecular dynamics simulations performed with water.[45] However the experimental slip lengths show a systematic bias toward higher values than theoretical ones. The discrepancy is actually not understood and might be the effect of gas adsorbed at the wall, possibly condensed in nanobubbles, an active research area recently.[86,87,88]

Finally, the case of organic molecules still raises debates. In the case of perfect wetting, various works on mica surfaces,[13,18] as well as on glass surfaces,[14,70] showed that alkanes and OMCTS have a no-slip velocity either at the wall or at a molecular size inside the liquid phase (negative slip length). However Pit *et al.*, using a fluorescence recovery method with molecular probes in a TIRF configuration, found very large slip lengths ($\simeq 100$ nm) of hexadecane and squalane on atomically smooth saphir surfaces although these liquids wet perfectly the saphir.[89,90] Further work is in progress to understand if these large slip lengths in the case of favorable solid–liquid interactions reflect a particular effect of molecular structure and incommensurability at the saphir/liquid interface. More generally, although the problem of the boundary condition has generated an impressive instrumental progress in the quantitative measurement and control of flows at a nanoscale, a full quantitative confrontation with the theory, *e.g.*eqn (11), is still lacking.

IV. Nanometric diffuse layers and transport therein

The transport of electrolytes within the EDL and the associated phenomena have been extensively studied during the last century due in particular to its central role in colloid science. It has been accordingly discussed exhaustively in textbooks and reviews[42,91] and in the present section we merely point to the main concepts which will be useful for our discussion.

In spite of this long history, new directions and perspectives have been opened very recently in this old domain. Novel insight into the nanometric Debye layer have indeed been reached: new tools now make it possible to investigate the detailed structure and dynamics *inside* the Debye layer. This concerns both the experimental side, see previous section, and numerical side, with the development of ever more powerful Molecular Dynamics tools.

From these more 'molecular' views of the EDL, in contrast to the more standard 'continuum' picture, new concepts and phenomena have emerged. We shall thus merely focus here on the recent progress and new leads opened by nanofluidics in this domain.

A From Debye layer to ion specific effects

An EDL builds up at a charged interface: it is the region of finite width where the surface charge is balanced by a diffuse cloud of counterions.[91] The Debye layer is usually described on the basis of the mean-field Poisson–Boltzmann (PB) theory, see ref. 39 for a review. Ions are described as point charges and only their valency enters the description. Correlations between charges are accordingly neglected. The thermodynamic equilibrium balancing the ions entropy to their electrostatic interaction with the surface (attraction for the counter-ions and repulsion for the co-ions) leads to the introduction of the Debye length, as introduced above in eqn (5) and which we recall here:

$$\lambda_D = (8\pi\ell_B\rho_s)^{-1/2} \tag{14}$$

Getting more into details, the PB description results from the combination of (i) the Poisson equation relating the electrostatic potential V to the charge density $\rho_e = e(\rho_+ - \rho_-)$ according to

$$\Delta V = -\frac{\rho_e}{\varepsilon} \tag{15}$$

with ε the dielectric constant, here identified to its bulk value, and Δ the spatial Laplacian; and (ii) the thermodynamic equilibrium leading to a spatially constant electro-chemical potential for the ions:

$$\mu_\pm = \mu(\rho_\pm) \pm Z_\pm eV \tag{16}$$

with $Z_\pm$ the ion valency, $\rho\pm$ the ion densities and under the above assumptions, $\mu(\rho) = \mu_0 + k_BT \log \rho$ the perfect gas expression for the chemical potential.

Though its crude underlying assumptions, the mean-field PB theory captures most of the physics associated with the EDL and is the usual basis to interpret ion transport. One reason for this success is that ion correlations within the EDL, which are discarded in PB, can generally be neglected. This can be quantified by introducing a coupling parameter, Γ_{cc}, which compares the typical inter-ions electrostatic interaction to thermal energy.[92,93] Following ref. 92 and 93, the latter is estimated as $\Gamma_{cc} \approx e^2/(4\pi\varepsilon\ell)/k_BT = \ell_B/\ell$, with ℓ the mean distance between ions at the surface. This scale is related to the charge density Σ, as $\ell \approx \Sigma^{-1/2}$, so that $\Gamma_{cc} \approx \sqrt{\Sigma\ell_B^2}$.[93] PB is expected to break down when the coupling parameter exceeds unity. For typical surfaces this threshold is not reached: for glass, the surface charge is at most in the 10^{-2} C/m^2 range ($\sim 6 \times 10^{-2} e/\text{nm}^2$), and the coupling parameter Γ_{cc} remains smaller than unity.

Other assumptions inherent to PB could be also questioned. In the PB description, the dielectric constant is assumed to be spatially uniform and identical to its bulk value. This assumption is expected to break down very close to surfaces, in particular hydrophobic, where water ordering induces a local electric dipole pointing outward to the interface.[94] This casts some doubts on the local relation between medium polarization and local electric field. However MD simulations suggest that using the bulk dielectric constant in the Poisson equation is not a critical assumption and can be safely used to describe the EDL.[94]

PB description however misses an important class of effects associated with the specific nature of ions, and which affect the fine structure of the EDL. Ion specificity is intimately connected to Hoffmeister effects, namely that the interactions between charged and neutral objects in aqueous media do depend crucially on the type of ion and not only on electrolyte concentration.[95–97] Such effects occur not only at the air–water interface, where it does affect surface tension depending on ion type, but also at hydrophobic surfaces.[96,98] This effect is evidenced in Fig. 5, where results of molecular dynamics simulations of ions dispersed in water (SPC-E model) from ref. 98 demonstrate that iodide ions do strongly adsorb at air–water and hydrophobic interfaces, while chloride ions do not. Note that an EDL builds up even in the case of a neutral interface, we shall come back to this result and its implications. In general heavier halide ions adsorb at the hydrophobic surface. This effect is mostly absent at a hydrophilic interface. Tools of molecular simulation and new spectroscopic techniques have fully renewed the interest in ion specific phenomena.[97] The origin of the effect involves many aspects, as ion polarizability, image-charge interaction, ion-size (steric) effects, dispersion forces and the mechanisms are still the object of intense research… and also of some fierce debate, such as for the preferred adsorption of hydronium *versus* hydroxyl at interfaces,[99] which fixes the charge of the liquid–vapor interface.

Such ion-specificity effects, which go beyond the traditional PB framework, have profound effects on the ion-transport process within the Debye Layer, which we explore in the next section.

B Interfacial transport: electro-osmosis, streaming currents, and beyond

The EDL is at the origin of various electrokinetic phenomena. Here we shall discuss more particularly electro-osmosis (and -phoresis for colloids) and its symmetric phenomena, streaming currents. Both phenomena take their origin in the ion dynamics within the EDL. Due to the nanometric width of the EDL, these phenomena are '*nanofluidic*' by construction and the various lengths introduced in section II will appear.

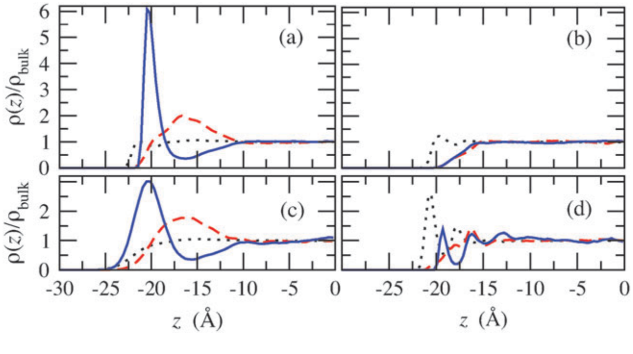


Fig. 5 Simulated density profiles of negative ions (solid lines), positive ions (dashed lines) and water (dotted lines) for roughly 1 M solutions of (a) NaI and (b) NaCl between neutral hydrophobic surfaces, (c) NaI at a vapor–liquid interface and (d) NaI between neutral hydrophilic surfaces. From ref. 98.

Let us first discuss electro-osmosis (EO). EO is the phenomenon by which liquid flow is induced by an electric field. In its simplest form, the fluid velocity v_f beyond the EDL is linearly connected to the applied electric field E_e according to the Smoluchowski formula:[42]

$$v_\infty = -\frac{\varepsilon\zeta}{\eta}E_e \tag{17}$$

with ε the dielectric constant, η the viscosity and ζ the so-called zeta (electrostatic) potential. The common interpretation of ζ is that it is the electrostatic potential at the "shear plane", *i.e.* the position close to the wall where the hydrodynamic velocity vanishes. As we now discuss, this interpretation should be taken with caution.

Let us briefly recall the derivation of eqn (17). Fluid dynamics is described by the stationary Stokes equation:

$$\eta\frac{\partial^2 v_x}{\partial z^2} + \rho_e E_e = 0 \tag{18}$$

Here we use a frame where the electric field E_e is along x and parallel to the planar surface, and the velocity field v_x only depends on the direction z perpendicular to it.

A key remark, already mentioned above, is that the charge density, $\rho_e = e(\rho_+ - \rho_-)$, is non-vanishing only within the EDL, so that the driving force for fluid motion, $F_e = \rho_e E_e$ is limited to that nanometric region, and vanishes otherwise. In order to obtain the velocity profile, the Stokes equation should be integrated twice, and thus requires two boundary conditions. Far from the surface, the velocity profile is plug-like and $\partial_z v|_{z = \infty} = 0$. At the wall surface, the hydrodynamic boundary conditions for the hydrodynamic velocity profile should be specified. We have discussed above that the latter introduces a slip length b, according to eqn (7), as $b\partial v|_{z = 0} = v(z = 0)$. Using Stokes and Poisson equations, one sees immediately that the full velocity profile is related to the electrostatic potential V as

$$v(z) = -\frac{\varepsilon}{\eta}E \times [-V(z) + \zeta] \tag{19}$$

where ζ has the meaning of a zeta potential[156] and takes the expression

$$\zeta = V_0 \times (1 + b\,\kappa_{\rm eff}) \tag{20}$$

where V_0 is the electrostatic potential at the wall and $\kappa_{\rm eff}$ is the surface screening parameter, defined as $\kappa_{\rm eff} = -V'(0)/V_0$.[94,100] For weak potentials this reduces to the inverse Debye length: $\kappa_{\rm eff} \simeq \lambda_D^{-1}$, but its full expression takes into account possible non-linear effects. The fluid velocity far from the surface (*i.e.* outside the EDL) has the Smoluchowski expression above, eqn (17), with the expression of the zeta potential, ζ, given in eqn (20). This prediction was first discussed in ref. 101, and rederived independently in more recent works.[94,100] We emphasize that the above result, eqn (19)–(20), does not make any assumption on the specific model for the EDL, except for a uniform dielectric constant ε and the validity of the Stokes equation, which was discussed above to be valid down to the nanometer scale.

The physical meaning of the above expressions is quite clear. The velocity in the fluid results from a balance between the driving electric force and the viscous friction force on the surface. Per unit surface, this balance writes

$$\eta\frac{v_\infty}{\lambda_D+b}\approx e\Sigma\times E \tag{21}$$

with Σ the surface charge density (in units of m^{-2}) at the wall. The expression for the viscous friction expresses the fact that flow gradients occur on a length $b + \lambda_D$, see Fig. 6. Relating the surface charge to the surface potential V_0 ($e\Sigma = -\varepsilon\partial_z V|_0$), one recovers the above results.

As a side remark, one may note that the slip enhancement does not require any assumption on the electric and fluid behavior. Indeed at a fully general level, the (exact) momentum balance at the surface shows that the velocity at the wall is *exactly* given by $v_{\mathrm{wall}}=\frac{D_{\mathrm{wall}}}{\lambda}E$, with D_{wall} the dielectric displacement computed at the surface, and λ the surface friction coefficient introduced in eqn (9). The latter is related to the slip length according to the definition $\lambda \equiv \eta/b$. This expression for the wall velocity under an applied electric field neither invokes Navier–Stokes equation, nor any assumption on the dielectric behavior. This wall velocity acts as a lower bound to the asymptotic EO velocity, *i.e.* the one measured at infinite distance from the surface. This demonstrates that the enhancement of the EO velocity due to low liquid–solid friction, *i.e.* large slippage, has a fully general validity.

Several conclusions emerge from eqn (20). First, if the wall is characterized by a no-slip BC, with $b = 0$, then the zeta potential ζ identifies with the electrostatic potential V_0 computed at the surface (up to a possible molecular shift due to the precise position of the hydrodynamic BC which defines the so-called shear plane [100]): this is the common assumption, as usually expressed in textbooks.[42] Now, if finite slippage occurs at the surface, then *the zeta potential is much larger than the surface electrostatic potential*, V_0: it is amplified by a factor $1 + b/\lambda_D$ (assuming for the discussion that $\kappa_{\mathrm{eff}} = \lambda_D^{-1}$). This factor may be *very large*, since on bare hydrophobic surfaces b may reach a few tens of nanometers (say $b \sim$ 20–30 nm), while the Debye length typically ranges between 30 nm down to 0.3 nm (for 10^{-4} M to 1 M).

These predictions have been confirmed by molecular Dynamics simulations of electro-osmosis (and streaming current, see below).[94,98,100] Various systems were considered, from a model electrolyte involving ions in a Lennard-Jones solvent, to a more sophisticated SPC-E model of water, with ions of various nature. Altogether simulations confirm the key role of slippage on electrokinetics and results do fully agree with the above description.

On the experimental side, this strong influence of surface dynamics on the electrokinetics at surfaces has not been appreciated and explored in the electrokinetic literature up to now, and very few experimental investigations have been performed on this question.[102,103] One key difficulty is that the above result for the zeta potential, eqn (20), involves a strong entanglement between electrostatics, through surface potential, and fluid dynamics, through hydrodynamic slippage. In order to disentangle the two effects, two independent measurements of these quantities should in principle be performed. To our knowledge the first work on the subject was performed by Churaev *et al.* in ref. 102, and results indeed suggested a slip effect, at the expense however of a rather uncontrolled assumption on the surface potential. More recently this problem was tackled in ref. 103 using the nanoPIV tool discussed above. This set-up allows for two independent measurements of ζ and V_0, leading to an unambiguous confirmation of the above predictions of slippage effect on the zeta potential. Indeed both the velocity profile, which is fitted to eqn (19), and the nanocolloid concentration profile, from which the surface potential is deduced on the basis of the electrostatic repulsion, are measured independently. Velocity profiles from ref. 103 are displayed in Fig. 7. While the two surfaces under consideration (a hydrophilic glass and a silanized hydrophobic surface) had basically the same surface potential, a factor of two is found on the zeta potential under the conditions of the experiment. Results are consistent with a slip length of $\approx$40 nm. The factor of two for ζ/V_0 occurs here due to a Debye length of $\approx$50 nm in the experimental condition of ref. 103 (a large Debye length is required to investigate flow inside the Debye layer at the present spatial resolution, $\sim$20 nm, of the nanoPIV technique). As a side remark, note the large slip velocity at the wall in the slippy case in Fig. 7: this is expected since according to eqn (19), the velocity at the surface is directly proportional to the slip length b, as $v(0)=-\frac{\varepsilon}{\eta}E\times V_0\frac{b}{\lambda_D}$ (assuming, to simplify, that $\kappa_{\mathrm{eff}} \approx \lambda_D^{-1}$).

This result opens up a new perspective to make use of surface physico-chemistry in order to optimize electric-induced transport and control flow by surface properties. It was also argued that such slip-induced optimization would strongly enhance the efficiency of energy conversion devices based on electrokinetic effects.[104,105]

Besides slip effect, ion-specificity also has an important influence on the electro-osmotic transport. As we reported above, the structure of the EDL is affected by the nature of the ions under consideration, especially at hydrophobic walls (and air–water interfaces). This can be best seen by rewriting the expression for the zeta potential, in eqn (20), in a slightly different form:

$$\zeta=-\frac{1}{\varepsilon}\int_0^\infty (z'+b)\rho_e(z')\,\mathrm{d}z' \tag{22}$$

with ρ_e the charge density. This expression *e.g.* comes from the double integration of the Stokes law. As seen in Fig. 5 ion specific effects have a strong influence on the charge distribution, and do indeed strongly modify the value of ζ. This effect was

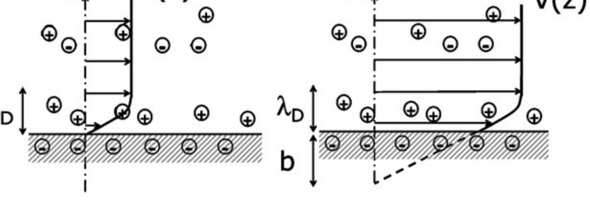


Fig. 6 Sketch of the influence of slippage on the electro-osmotic transport. Slippage reduces the viscous friction in the electric Debye layer, as the hydrodynamic velocity gradient occurs on a length $b + \lambda_D$, instead of λ without slippage. Flow is accordingly enhanced by a factor $1 + b/\lambda_D$.

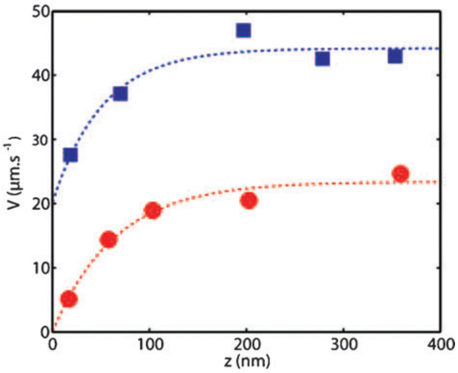


Fig. 7 NanoPIV measurements for the electroosmotic flows close to hydrophilic pyrex ● and hydrophobic OTS-coated surface ■. The deduced zeta potentials are $\zeta = -66 \pm 8$ mV (hydrophilic glass) and $\zeta = -123 \pm 15$ mV (hydrophobic, silanized surface), while the electrostatic potentials were independently measured to be comparable: $V_0 \simeq -69$ mV for the hydrophilic (pyrex) surface and $V_0 \simeq -65$ mV for the hydrophobic (OTS-coated) surface. The salt concentration is 1.5×10^{-4} mol L^{-1} and the driving electric field is $E = 500$ V m^{-1}. Dashed lines are fits to the theoretical predictions using eqn (19), with $\lambda_D = 51 \pm 10$ nm (measured independently) and $b = 0 \pm 10$ nm (bottom) and $b = 38 \pm 6$ nm (top) for the slip lengths. Figure from ref. 103.

studied in various MD simulations of electrokinetics[94,98,106] and several counter-intuitive effects were indeed observed, such as flow reversal (as compared to the expected surface charge),[106] or even the existence of a non-vanishing zeta potential for a *neutral* surface![94,98] These effects can be understood from eqn (22): while for a vanishing surface charge, the system is electro-neutral and $\int_0^\infty \rho_e(z')\,\mathrm{d}z' = 0$, this however does not imply that its first spatial moment, which enters the expression of ζ according to eqn (22), be zero.

More quantitatively, effects of ion-specificity on the zeta potential were rationalized in ref. 94 on the basis of a minimal model, which was subsequently validated on MD simulations. This description is based on the idea that 'big' ions have a larger solvation energy and are therefore attracted to hydrophobic interfaces in order to minimize this cost.[107] An estimate of the solvation free-energy of ions close to the interface allows for analytical predictions for ion-specific effects on the zeta potential.

We finally quote that all of the above discussion on the electro-osmotic transport also applies to the streaming current phenomenon, by which an electric current is induced by the application of a pressure gradient. Again, since the EDL is not electro-neutral, liquid flow will induce an electric current inside the EDL.[42] According to the Onsager symmetry principle, these two phenomena—electro-osmosis and streaming current—are intimately related and the cross coefficients should be identical. The expression for the streaming current I_e under an applied pressure gradient ∇P takes the form:

$$I_e = -\frac{\varepsilon\zeta}{\eta}\mathscr{A}(-\nabla P) \tag{23}$$

where ζ takes the same expression as in eqn (20) and $\mathscr{A}$ is the cross sectional area of the channel.

We end this section by mentioning that the above effects generalize to any interfacial transport phenomena. Indeed the above electrically induced phenomena belong to a more general class of surface induced transport, which also involves phenomena known as diffusio-osmosis or thermo-osmosis and their associated phoretic phenomena for colloidal transport.[21,91] Thermo-osmosis points to fluid motion induced by thermal gradients, while diffusio-omosis corresponds to fluid flow induced by gradients in a solute concentration. Though relatively old, thermophoretic transport has been the object of recent investigations in particular for colloid manipulation.[108–110] Its origin is still the object of an intense debate with recent progress in disentangling the various contributions.[108,109,111] On the other hand, diffusiophoretic and osmotic transport have been less explored[91,112] but their potential in the context of microfluidic applications has been demonstrated recently.[113,114] This opens up the novel possibility of driving and pumping fluids[114]—as well as manipulating colloids—with solute contrasts, which have been barely explored up to now and would deserve further investigations in the context of nanofluidics. A common point to all these phenomena is their "nanofluidic root": as for the electro-osmotic transport discussed above, the driving force for fluid motion is localized within a diffuse layer of nanometric size close to the surfaces. Therefore, this opens up the possibility of strongly amplifying their effect on the basis of slippage effects at the solid interface. As demonstrated theoretically in ref. 115, the amplification is—as for electro-osmosis above—amplified by a factor $1 + b/\lambda$, where λ is the width of the diffuse layer. Similar effects are predicted for thermophoretic transport.[116]

Finally, one may raise the question of interfacial transport on *super-hydrophobic* surfaces. Such surfaces, achieved using nano- or micro-engineering of the surfaces, were shown to considerably enhance the slippage effect and exhibit very large slip length in the micron (or even larger) range.[46,47,117] Accordingly, a naive application of the previous ideas would suggest massive amplification by a factor up to 10^4(!). However, the composite structure of the superhydrophobic interface, involving both solid–gas and liquid–gas interfaces, makes these transport mechanisms far more complex in this case than on a smooth interface. It has been shown that in the regime of a thin Debye layer, no amplification is obtained for electro-osmosis on superhydrophobic surfaces,[114,148] unless a non-vanishing charge exists at the liquid–gas interface. In contrast, a massive amplification is predicted for diffusio-osmosis.[114] This strong prediction has not received an experimental confirmation up to now.

C Surface *versus* bulk: conductance effects

We now discuss the surface conductance effect in nanochannels.

The conductance, K, characterizes by definition the electric current *versus* electric potential drop relationship. As pointed out previously, conductance probes the number of free charge carriers, here ions. In the bulk K is therefore expected to be proportional to the salt concentration ρ_s (ionic strength). Now in the presence of surfaces, the charges brought by the surface lead to a supplementary contribution to the conductance.

Surface conductance is actually a rather classical phenomenon in colloid science,[42] but it was demonstrated quite recently in the context of nanochannel transport by Stein *et al.*[41] and then by other groups.[44]

As we introduced in section II B 2.c, the ratio of bulk to surface charge carriers is characterized by the Dukhin length ℓ_{Du}. This length accordingly describes the competition between bulk and surface conductance. Depending on salt concentration, this length can be much larger than the Debye length. Surface conductance is therefore expected to dominate over the bulk contribution for nanochannels smaller than ℓ_{Du}. The magnitude of the surface conductance effect is amplified in small channels, but a key point is that it does not require Debye layer overlap.

Experimentally, the surface contribution to the conductance shows up as a saturation of the conductance in the limit of small salt concentration, while its expected bulk counterpart is expected to vanish in this limit. The saturation originates in the charge carriers brought by the surface charge, the number of which is independent of the salt concentration. This is illustrated in Fig. 8, from ref. 41, where the conductance is measured in channels of various width, from 1015 nm down to 70 nm.

Let us discuss more specifically this point. In a slit geometry with width h, the general expression for the current (for a unit depth of the slit) takes the form

$$I_e = e\int_{-h/2}^{h/2}\mathrm{d}z[\rho_+(z)u_+(z) - \rho_-(z)u_-(z)] \tag{24}$$

where the averaged velocities of the ions is

$$u_\pm(z) = v(z) \pm e\mu_\pm E_e, \tag{25}$$

with $\mu_\pm$ the ion mobility, defined here as the inverse of the ion friction coefficient (we assume furthermore $\mu_\pm = \mu$). The velocity $v(z)$ is the fluid velocity induced by electro-osmosis under the electric field, as given in eqn (19). The second term is the contribution to the current due to ion electrophoresis.

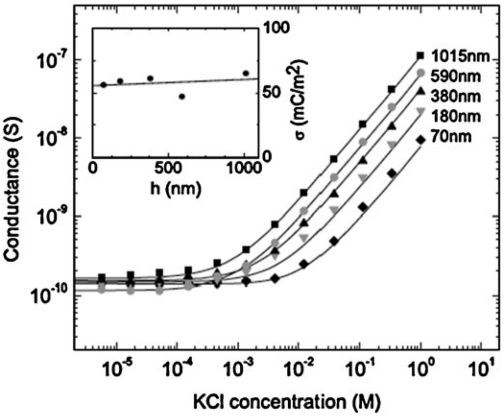


Fig. 8 Channel height dependence of ionic conductance behavior. The conductance of fluidic channels is plotted against salt concentration for various nanochannel width h (h = 1015, 590, 380, 180 and 70 nm). The inset displays the fit values for the surface charge at the nanochannel walls as a function of h. From ref. 41.

Combining the expression of the velocity $v(z)$, as in eqn (19), to eqn (24)–(25), one may remark that the electro-osmotic contribution to the current takes the form of the integral of the charge density times the electrostatic potential: this has therefore the form of an electro-static energy, $\mathscr{E}$. This is confirmed by the full calculation, which leads to a conductance expression in the form:[41,118]

$$I/E_e = 2e^2\mu\rho_s\mathscr{A}(1 + H) \tag{26}$$

with $\mathscr{A}$ the channel cross section ($\mathscr{A} = w \times h$ with w the depth of the channel) and the surface correction H

$$H = \cosh\left(\frac{eV_\mathrm{c}}{k_\mathrm{B}T}\right) - 1 + \frac{\mathscr{E}}{4\rho_s k_\mathrm{B}Th}\left(1 + \frac{1}{2\pi\ell_B\mu\eta}\right) \tag{27}$$

where $V_\mathrm{c} = V(0)$ is the electrostatic potential at the center of the channel, and $\mathscr{E} = \frac{\varepsilon}{2}\int_\mathrm{slit}\mathrm{d}z\left(\frac{\mathrm{d}V}{\mathrm{d}z}\right)^2$ the electrostatic energy (and not the *free* energy). Note that writing $\mu = 3\pi\eta d_i$ with d_i the ion diameter, the last term in eqn (27) is typically of the order d_i/ℓ_B. It originates in the surface-induced—electro-osmotic—contribution to the conductance.

In the low salt concentration regime, $\rho_s \to 0$, the length scales are organized in the order:

$$\ell_{GC} < \lambda_D < \ell_{Du}, \tag{28}$$

since $\lambda_D/\ell_{GC} = \ell_{Du}/\lambda_D \gg 1$ for $\rho_s \to 0$. Saturation will occur as the channel width h is smaller than the Dukhin length, $h < \ell_{Du}$, independently of the order of h and λ_D. Debye layer overlap is accordingly *not* the source of the saturation of the conductance. In this regime, the second term in eqn (27), associated with surface contributions is dominant.

In this limit, the calculation of the electrostatic energy shows that $\mathscr{E} \sim |\Sigma|$. The fact that $\mathscr{E}$ scales linearly in $|\Sigma|$ and not as $|\Sigma|^2$—as would be first guessed—is due to the non-linear contributions to $\mathscr{E}$ in the Poisson–Boltzmann description. These become dominant in the $\rho_s \to 0$ limit where the Debye length becomes larger than the Gouy–Chapman length. Altogether, this shows that for $\rho_s \to 0$, the conductance saturates at a value

$$K_\mathrm{sat} \approx e^2\mu w \times 2|\Sigma|\left(1 + \frac{1}{2\pi\ell_B\mu\eta}\right) \tag{29}$$

with w the depth of the channel. This value is independent of both h and ρ_s. This saturation can be also partly understood in the context of Donnan equilibrium discussed below. In the limit $\rho_s \to 0$, only the counter-ion contributes to the ion concentration in the channel, $\rho_+ + \rho_- \approx 2\Sigma/h$, see eqn (34) below, so that the conductance reduces to $K_\mathrm{sat} \approx e^2\mu w 2|\Sigma|$. This analysis however misses the electro-osmotic contribution to the conductance (second term in the brackets in eqn (29)).

A few remarks are in order. First, the dependence of the conductance on the surface properties, here the surface charge Σ, opens new strategies to tune the nanochannel transport properties, *via* an external control. This has been used by Karnik *et al.* to develop a nanofluidic transistor,[5] allowing the control of the conductance of the nanochannel thanks to a gate voltage.

Another interesting remark is that along the same line as in section IV B, the electro-osmotic contribution to the conductance could be also amplified by slippage effects. This can be readily demonstrated by a direct integration of eqn (24). Accordingly slippage effects add a new contribution to the conductance:

$$K_{\text{slip}} = 2e^2 w \frac{\Sigma^2}{\eta} \times b \tag{30}$$

Assuming Stokes law for ion inverse mobility ($\mu = (3\pi\eta d_{\text{ion}})^{-1}$ with d_{ion} the ion size), then $K_{\text{slip}}/K_{\text{sat}} \sim b \times d_{\text{ion}}|\Sigma| \approx b/\ell_{GC}$: since the Gouy–Chapman length ℓ_{GC} is typically nanometric (or less), this enhancement is therefore very large, even for moderate slip length ($b \sim 30$ nm)!

The channel resistance is accordingly decreased by the same factor in the low salt regime. This amplification of conductance by hydrodynamic slippage opens very interesting perspectives in the context of electrokinetic energy conversion, as was pointed out recently by Pennathur *et al.*[105] and Ren and Stein.[104] Moderate slippage, with slip length of a few tens of nanometers, is predicted to increase the efficiency of the energy conversion up to 40% (and of course even more with larger slip lengths). This attracting result would deserve a thorough experimental confirmation.

D Debye layer overlap and nanofluidic transport

The phenomenon of Debye layer overlap has already been widely explored in the nanofluidic literature. As we pointed out above, various reasons underlie this specific interest: (i) well controlled nanometric pores with size in the range 20–100 nm can be produced using micro-lithography techniques:[40] this does indeed correspond precisely to the range of typical Debye lengths for usual salt concentrations (remember that $\lambda_D = 30$ nm for a salt concentration of 10^{-4} M); (ii) the overlap of Debye length does indeed have a strong influence on ion transport, so that novel transport effects emerge at this scale, with applications for chemical analysis.

The question of Debye layer overlap, and related phenomena, was discussed quite extensively in a recent review by Schoch *et al.*[40] Here we thus focus on the main guiding lines underlying this problem and point to the key physical phenomena at play. In particular we shall discuss an illuminating analogy to transport in semi-conductors.

1. Donnan equilibrium. An important notion underlying the Debye overlap is the so-called "Donnan equilibrium", which is a well know concept in colloid literature. Due to the supplementary charges brought by the nanochannel's surface, a potential drop builds up between the nanochannel's interior and the external reservoir, in order to maintain a spatially uniform chemical potential of the ions. The latter, as introduced above, takes the form (assuming a dilute ion system):

$$\mu(\rho_\pm) = \mu_0 + k_B T \log(\rho_\pm) \pm \text{eV} = \mu_0 + k_B T \log(\rho_s) \tag{31}$$

with ρ_s the (uniform) salt concentration outside the nanochannel, *i.e.* the "reservoir". These equations are completed with the overall electroneutrality over the channel,

$$\int_{-h/2}^{h/2} \mathrm{d}z(\rho_+(z) - \rho_-(z)) = 2\Sigma \tag{32}$$

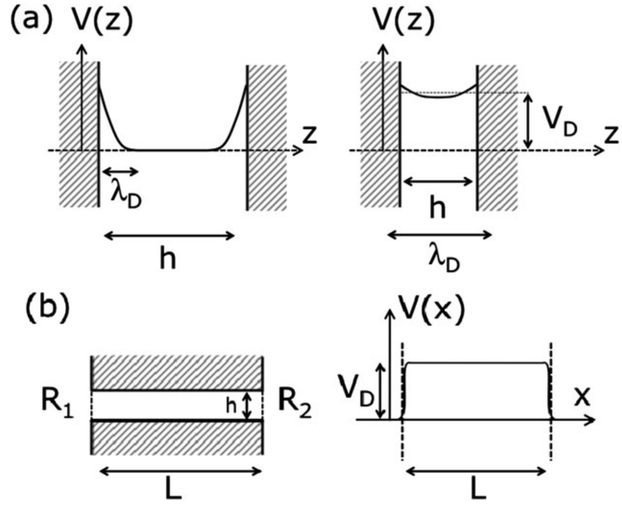


Fig. 9 (a) Sketch of the electrostatic potential in a charged slit with thickness h: (*left*) $h > 2\lambda_D$; and (*right*), $h < 2\lambda_D$, corresponding to Debye layer overlap. (b) A charged nanochannel, with thickness h and length L, connects two reservoirs R_1 and R_2. In the situation of Debye layer overlap, a Donnan potential V_D builds up along the nanochannel.

with z along the direction perpendicular to the nanochannel (here assumed to be a slit), see Fig. 9.

We now focus on the thin pore limit, assuming Debye layer overlap $h \lesssim 2\lambda_D$. In this situation, ion densities and electrostatic potential are approximately spatially uniform over the pore thickness h. Ion densities thus obey:

$$\begin{aligned} \rho_+ \times \rho_- &= \rho_s^2 \\ \rho_+ - \rho_- &= \frac{2\Sigma}{h} \end{aligned} \tag{33}$$

Altogether, this leads to

$$\begin{aligned} \rho_\pm &= \sqrt{\rho_s^2 + \left(\frac{\Sigma}{h}\right)^2} \pm \frac{\Sigma}{h} \\ V_D &= \frac{k_B T}{2e} \log\left[\frac{\rho_-}{\rho_+}\right] \end{aligned} \tag{34}$$

where the mean electrostatic potential V_D is denoted as the Donnan potential.

Note that the amplitude of the Donnan potential V_D is quantified by the ratio ℓ_{Du}/h where $\ell_{Du} = \Sigma/\rho_s$ is the Dukhin length introduced above. To give an order of magnitude $\ell_{Du} \sim 1$ nm for $\rho_s = 1$ M and a typical surface charge.

It is finally of interest to mention the strong analogy between the Donnan equilibrium and the equilibrium of charge carriers in doped semi-conductors (SC). In the latter the electron density ρ_n and the hole density ρ_p obey eqn (33) with the electrolyte density ρ_s replaced by the carrier density in the intrinsic (non-doped) SC. The surface charge in the nano-channel actually acts as impurities delivering additional carriers, with $\Sigma > 0$ corresponding to donors (*N*-doped SC) and $\Sigma < 0$ acceptors (*P*-doped SC). The Donnan potential is then the analogous of the shift in the Fermi energy due to impurities, which determines the voltage difference at equilibrium between regions of different dopage level. The table hereafter, Table 1, summarizes the corresponding quantities in the nano-channel and SC analogy.

Table 1 Equivalence table for corresponding quantities in doped semiconductors and nano-channels

Nanofluidics	Semi-conductors
ρ_- negative ions concentration	ρ_n electron density in the conduction band
ρ_+ positive ions concentration	ρ_p holes density in the conduction band
ρ_s electrolyte concentration	n_i carriers density in the intrinsic SC
Surface charge $2\Sigma/h$	Impurities (doping) concentration
Donnan potential	Shift of the Fermi level

The Donnan potential and Debye layer overlap have many implications on nanofluidic transport.

2. Ion transport and PNP equations. A classical framework to discuss ion transport in narrow pores is the so-called Poisson-Nernst-Planck (PNP) equations. This simplified description of ion transport in strongly confined channels is based on the coupled diffusion-electro-convection of the ions.[119]

As in Fig. 9, the channel thickness is assumed to be small compared to the Debye length, $h \ll \lambda_D$, so that ion concentrations are assumed to be uniform across its thickness. The ion fluxes $J_\pm$ (per unit surface) then takes the form

$$J_\pm = -D_\pm \partial_x \rho_\pm(x) \pm \mu_\pm e \rho_\pm(x)(-\partial_x V)(x) \qquad (35)$$

where $D_\pm$ and $\mu_\pm$ are related by the Einstein relationship, $D_\pm = \mu_\pm k_B T$; x is along the direction of the channel. Note that for simplicity, a single salt and monovalent ions are considered here. A local version of the electroneutrality, eqn (33), may be made along the pore:

$$\rho_+(x) - \rho_-(x) = \frac{2\Sigma(x)}{h} \qquad (36)$$

Finally at the pore entrance and exit, a Donnan electric potential drop builds up, along the description given above, see eqn (34).

To illustrate further this approach, we consider the transport through a single nanochannel, as sketched in Fig. 9(b). At equilibrium the two reservoirs, R_1 and R_2, have the same electric potential and salt concentration, associated with a uniform Donnan potential along the channel. Now if an electrostatic potential drop ΔV or a salt concentration difference $\Delta\rho_s$ is imposed between the reservoirs, ion fluxes will build-up so as to relax towards equilibrium.

In the stationary state, the ion flux is spatially uniform, so that one may solve eqn (35)–(36) for the ion densities and electrostatic potential. This allows the current $I = \mathscr{A}e(J_+ - J_-)$ and total ion flux $\Phi_t = \mathscr{A}(J_+ + J_-)$ to be computed, with $\mathscr{A}$ the cross sectional area. In general, the relationships $I(\Delta V, \Delta\rho_s)$, $\Phi_t(\Delta V, \Delta\rho_s)$ are non-linear.

However, in the limit of small ΔV and $\Delta\rho_s$, a linear relationship can be written in general between the ionic fluxes and corresponding thermodynamic 'forces':[121,122]

$$\begin{bmatrix} I \\ \Phi_t \end{bmatrix} = \frac{\mathscr{A}}{L} \begin{bmatrix} K & \mu_K \\ \mu_K & \mu_{\text{eff}} \end{bmatrix} \begin{bmatrix} -\Delta V \\ -k_B T \Delta[\log \rho_s] \end{bmatrix} \qquad (37)$$

with $\mathscr{A}$ the cross sectional area of the channel and $\Delta[\log \rho_s] = \Delta\rho_s/\rho_s$. Note that due to Onsager (time reversal) symmetry, the non-diagonal coefficients of the matrix are equal.[121,122]

The coefficients of the above matrix can be calculated within the PNP framework. PNP equations are solved analytically for small potential and concentration drops, leading to the following expression for the various coefficients in the matrix, eqn (37):

$$\begin{aligned} K &= 2\mu e^2 \sqrt{\rho_s^2 + \left(\frac{\Sigma}{h}\right)^2} \\ \mu_{\text{eff}} &= 2\mu \times \sqrt{\rho_s^2 + \left(\frac{\Sigma}{h}\right)^2} \\ \mu_K &= e\mu \times \frac{2\Sigma}{h} \end{aligned} \qquad (38)$$

We assumed here that ions have the same mobilites $\mu_\pm = \mu$. The cross effects, associated with the mobility μ_K, originate in the dependence of the Donnan potential on the salt concentration.

Let us conclude this part with a few remarks:

- First one should realize that a number of assumptions are implicitly made in writing PNP equations. In particular the ions are treated as a perfect gas and correlations between ions along the channel are neglected. This may become problematic in strongly confined situations where single file transport (of the solvent) and strong unidirectional electrostatic correlations should build up. As an example, proton transport in single file water has been shown to involve a highly-cooperative mechanism.[120] However these limitations are restricted to single file transport in molecular channels, and should not be a limitation for pore size larger than a nanometer. Furthermore the PNP model is interesting to explore as a guiding line, in the sense that it provides a rather correct physical idea of the (complex) electro-diffusion couplings.
- Furthermore, we may pursue the analogy discussed above between nanochannels and doped SC (see Table 1), and extend the discussion to transport phenomena. Indeed the Poisson–Nernst–Planck equations, eqn (35), are formally identical to the phenomenological transport equations for electrons and holes in SC. Thus, as long as electrical and concentration fluxes *only* are allowed, *nanofluidic devices can in principle reproduce standard SC-based components*, such as diodes and transistors. The nanofluidic diode for instance is based on the properties of a PN junction, *i.e.* the junction between two regions with different doping: according to the equivalence table, Table 1, its nanofluidic equivalent corresponds to two nanochannels with different surface charge Σ/h. We will discuss below (sub-section 5) some recent findings confirming the pertinence of this analogy.
- However one should keep in mind that the transport analogy between fluids and electrons breaks down in the presence of hydrodynamic flow. Actually, PNP equations introduced above do not take into account convective contributions. In a fluidic system, a flow is indeed expected to occur in nanochannels as soon as a voltage drop is applied to its end, due to the body force acting on the mobile charges, but also under a pressure gradient (hydrostatic or osmotic).

This is for example the source of electro-osmotic or streaming current contributions discussed above. In contrast, SCs can not flow—quite obviously—under an applied electric field and may only deform elastically. This is therefore a limitation to the above analogy.

In the case of a strongly confined regime with a large slippage at wall, the hydrodynamic contribution to the ion flux could well become dominant. Such a convection contribution should therefore be added in the form $J^{\pm}_{\rm conv} = v_f \times \rho_{\pm}$, where v_f is the water velocity. The water velocity is in turn coupled to (electric) body forces *via* the momentum conservation equations, in practice the Navier–Stokes equation in its domain of validity. Altogether, the hydrodynamically induced cross phenomena, *i.e.* electro-osmosis, streaming currents, diffusio-osmosis,... (as well as bare osmosis), can be summarized at the linear level in terms of a symmetric transport matrix, in a form similar to eqn (37). This matrix relates linearly the fluxes to thermodynamic forces at work.[122–124] These phenomena have been discussed extensively in the context of membrane transport but generalize here to transport in nanofluidic channels.

More generally the coupling of the ion transport to the fluid transport thus opens the route to new applications in nanofluidic devices, with a richer phenomenology than in SC electronics.

We now illustrate the above concepts in a few practical situations.

3. Permselectivity. As a first example, we consider the permeability of nanopore to ions. It was first shown by Plecis *et al.*[2] that nanochannels exhibit a selective permeability for ion diffusive transport, see Fig. 10(a): ions of the same charge as the nanochannel surface (co-ions) exhibit a lower permeability, while ions of the opposite charge (counter-ions) have a higher permeability through the nanochannel.

This charge specific transport is a direct consequence of a non-vanishing Donnan potential in the nanochannel (and Debye layer overlap). As pointed out above, Fig. 9, counter ions exhibit a higher concentration in the nanochannel as compared to that of a neutral species, $\rho_+ > \rho_n$, while co-ions have correspondingly a lower concentration, $\rho_- < \rho_n$.

The diffusive flux of counter-ion will be accordingly larger than its expectation for neutral species, and *vice versa* for co-ions. This leads therefore to a charge-specific effective diffusion coefficient $D_{\rm eff}$ for the co- and counter- ions.

Following Plecis *et al.*,[2] the effective diffusion coefficient is defined according to the identity

$$J_{\pm} = -D\frac{\Delta\rho_{\pm}}{L} \equiv -D_{\rm eff}\frac{\Delta\rho}{L} \tag{39}$$

taking into account the fact that the local concentration in the nanochannel differs from the one imposed at the two ends of the reservoir. The linearized Poisson–Boltzmann equation is then used to calculate the ion concentration in the slit, $\rho_{\pm}$, as:

$$\begin{aligned} \beta_{\pm} &= \frac{\bar{\rho}_{\pm}(x)}{\rho(x)} = \frac{1}{h}\int_{\rm slit} {\rm d}z \exp[\mp eV(z)/k_{\rm B}T] \\ D_{\rm eff}/D &= \beta_{\pm} \end{aligned} \tag{40}$$

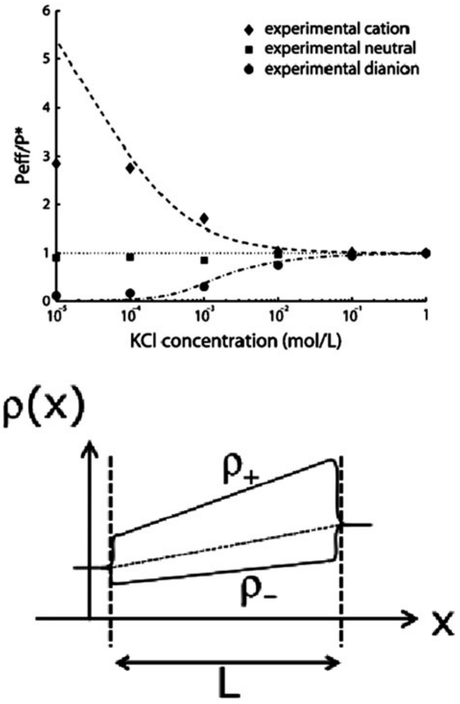


Fig. 10 *Top* Variation of the relative permeability of various probes with different charges, *versus* ionic strength. The experimental results are compared to theoretical fittings obtained with eqn (40) (dotted lines). *Bottom* Concentration profiles of counter- and co-ions, $\rho_{\pm}$, for a given salt concentration drop $\Delta\rho$ between the two ends of the nanochannel. The dotted line represents the linear concentration profile for noninteracting diffusing species, whereas plain lines show the ion profiles in the nanochannel. Adapted from ref. 2.

with $\bar{\rho}_{\pm}(x)$ the (local) ion concentration in the nano-channel, averaged over the channel width h; q the ion charge; $\rho(x)$ the local salt concentration in the absence of electrostatic interactions (fixed by the salt concentration in the reservoirs); and $V(z)$ the electrostatic potential across the channel, for which Plecis *et al.* used a (linearized) Poisson–Boltzmann expression. A very good agreement with experimental results is found, as shown in Fig. 10, showing that this exclusion–enrichment picture does capture the essential ingredients of the ion diffusive transport. Note that using the Donnan description above, one may furthermore approximate the electrostatic potential by its Donnan expression in eqn (34). This leads to an analytical expression for $\beta_{\pm}$ as

$$\beta_{\pm} = \left(\frac{\rho_{\pm}}{\rho_{\mp}}\right)^{1/2} \tag{41}$$

where, as shown in eqn (34), $\rho_{\pm} = \sqrt{\rho_s^2 + \left(\frac{\Sigma}{h}\right)^2} \pm \frac{\Sigma}{h}$. This expression reproduces the results in Fig. 10.

Finally, one should note however that due to the difference in ion permeability, a charge separation will build up between the two ends of the channel. This leads therefore to the creation of a reacting electric field along the channel, which will compensate dynamically for this charge separation.[2] This points to the complex couplings associated with charge transport in nanochannels. More experimental and theoretical work is certainly in order to get further insight in these phenomena.

4. Pre-concentration. Along the same lines, another consequence of Debye layer overlap is the 'pre-concentration' phenomenon, which was first observed by Pu *et al.*,[3] and others.[125] Instead of a concentration drop as above, a voltage drop is applied along the nanochannel. It is then observed that ions enrich at one end and deplete at the other end. This phenomenon is illustrated in Fig. 11.

As above, a key point underlying the phenomenon is that ions of the opposite sign to the channel's surface, counter-ions, are more heavily transported than co-ions. This will create an enriched/depleted zone for both ions at the two ends of the channel.

This phenomenon has attracted a lot of interest, due in particular to its potential applications in the context of chemical analysis, for which it would provide a very interesting way of enhancing the sensitivity of detection methods.

At a basic level, the origin of preconcentration is a "cross effect" as introduced in eqn (37): an electric potential drop ΔV leads to a flux of ions (Φ_t), as quantified by the cross-mobility μ_K. According to the PNP result for μ_K ($\mu_K = e\mu \times 2\Sigma/h$), this cross effect is thus a direct consequence of confinement and the existence of surface charges. Accordingly the magnitude of its effect, as compared *e.g.* to the diffusive flux, depends on the Dukhin number, see eqn (38).

But the detailed mechanisms underlying the preconcentration involve non-linear couplings between the various transport processes of ions in the nanochannel, which are quite complex to rationalize. A systematic description of the phenomenon was proposed by Plecis *et al.* in ref. 126, revealing the existence of various preconcentration regimes. We refer to ref. 40 for a detailed discussion on the mechanisms underlying this process.

5. More complex functionalities: nanofluidic diodes. The analogy with transport in semiconductors quoted above suggests that more complex nanofluid transport phenomena can be obtained in the regime of Debye layer overlap.

We discuss here the analog of a PN junction in SC transport. Using the equivalence table, Table 1, a PN junction corresponds for nanofluidic transport to a nanochannel exhibiting a disymmetric surface charge along its surface.

Such a nanofluidic diode device has been developed by Karnik *et al.*,[6] following a previous work by Siwy *et al.* in a different pore geometry.[7] This is illustrated in Fig. 12 from ref. 6. The surface of a nanochannel is coated with two different surface treatments (half with avidin, half with biotin), leading to a surface charge contrast along the two moieties of the channel, Fig. 12 (left). Accordingly the current *versus* applied electric potential drop characteristics is found to exhibit an asymmetric shape: as in a classical diode, the current passes only in one direction. Similar asymetric *I–V* curves are obtained for pores with asymmetric geometries, like conical pores obtained by track–etch techniques.[7]

At a more quantitative level, Karnik *et al.* discussed the effect on the basis of the PNP equations discussed above, under the assumption of local electroneutrality. In the present disymmetric case, Karnik *et al.* solve numerically these PNP equations for an applied potential drop at the ends of the pore.[6] Solutions are shown in Fig. 12, demonstrating that the PNP framework is indeed able to capture the transport rectification measured experimentally.

This diode behavior can also be discussed in the context of the analogy with semiconductor transport that we put forward

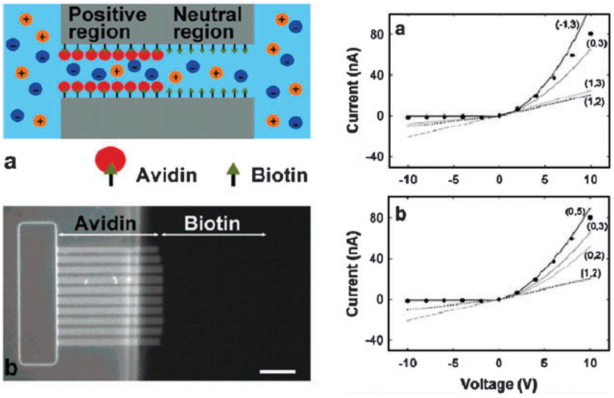


Fig. 12 *Left* (a) A nanofluidic diode is fabricated by patterning a nanochannel with different coating on its two moieties (avidin and biotin here). (b) Epifluorescence image of the fabricated nanofluidic diode, showing fluorescently labeled avidin in half the channel. Scale bar 20 μm. *Right* Current *versus* applied voltage. Symbols are experimental data, while the solid lines are theoretical predictions using a PNP transport model. From ref. 6.

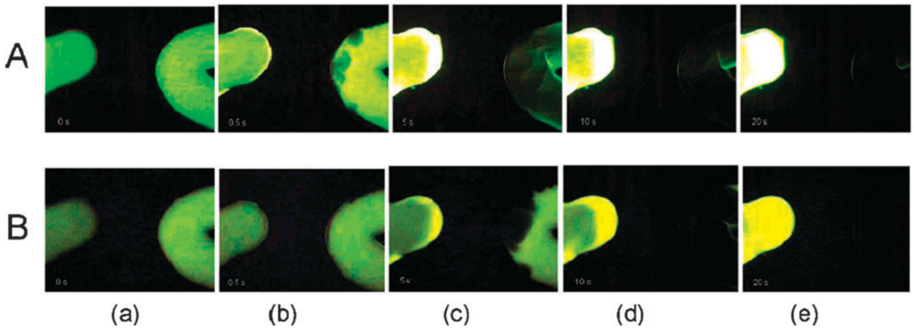


Fig. 11 Two micro-channels filled with fluorescent probes are connnected by a nanochannel (a). After an electric field is applied across the nanochannels, an ion-enrichment and ion-depletion occurs at each end of the nanochannel (b)–(e). Figures (a) to (e) show the evolution of the fluorescence in the microchannel *versus* time (from 0 to 20 s). The probe used was (A) fluorescein, (B) rhodamine 6G. From ref. 3.

above. The geometry described in Fig. 12 is indeed equivalent to that of a classical PN junction, corresponding to two regions with different doping. In the present case of a nanofluidic diode, the doping contrast between the P and N regions is associated with the different surface charge between the two moieties of the nanochannel. The diode effect in PN junctions is classically interpreted in terms of the Space-Charge Zone at the interface between the two regions with different doping. An approximate solution of the PNP equations can be proposed and leads to a rectified *I–V* characteristic which takes the form of the so-called Shockley equation[127]

$$I = I_S\left(\exp\left[\frac{eV}{k_B T}\right] - 1\right) \tag{42}$$

where I_S is the so-called saturation current. This expression is expected to describe well the blocking to non-blocking transition of the junction, *i.e.* for moderate eV/k_BT. However this expression strongly overestimates the current at large eV/k_BT for which a full solution of the PNP equations is needed.

In the Shockley equation, eqn (42), the saturation current is expressed in terms of the donor–acceptor densities, ρ_A and ρ_D, as

$$I_S \approx e\mathscr{A} n_i^2\left(\frac{D_n}{L_n\rho_A} + \frac{D_p}{L_p\rho_D}\right) \tag{43}$$

where n_i is the carrier density, D_n, D_p the diffusion coefficients of electrons and holes, and L_n, L_p the length of the P and N regions. Using the equivalence table, Table 1, one may convert this expression for the ionic transport using $D_n = D_-$, $D_p = D_+$, $L_n = L_p = L/2$, while $\rho_A = 2|\Sigma_A|/h$ and $\rho_D = 2|\Sigma_D|/h$.

This Shockley expression predicts an *I–V* curve in qualitative agreement with the experimental result in Fig. 12. However, quantitatively it does (very strongly) overestimate the current for large voltages *V*, as expected due to the simplifying assumptions underlying the expression in eqn (42).

The analogy is however interesting to capture the underlying physics behind ion rectification in ionic transport. It allows the prediction of a full zoology of fluidic functionalities, in line with their SC analogues.

Finally it is interesting to quote that the disymmetry measured here for *I–V* characteristics generalizes to other transport phenomena. For example, Siwy *et al.* also demonstrated the occurrence of a disymmetric diffusion of ions through conical pores.[128]

Altogether these phenomena demonstrate the richness and complexity of behaviors obtained in the regime of Debye layer overlap. This definitely shows the great potential of nanofluidics, where further phenomena should emerge in the future using the benefit of these couplings.

V. Thermal fluctuations

In this section we raise the question of the role of thermal fluctuations in nanofluidic transport. As a general statistical rule, when the size of the system decreases, fluctuations play an increasingly important role. But under which conditions do they play a role? And which role could they play in nanofluidic transport?

There are quite a few studies on this question in relation to nanofluidic transport. One may cite the experimental study of noise in solid state nanopores by Dekker *et al.*[129] The question of thermal fluctuations was also discussed in the context of numerical molecular dynamics simulations of osmotic flow through carbon nanotubes by Hummer *et al.*[130,131] Simulations show that—in the single file regime—the flow rate is essentially governed by thermal fluctuations rather than hydrodynamics. Hummer *et al.* proposed a 1D random walk to describe the nanofluidic water flow in this stochastic regime. Now, it is fair to remark that fluctuations are indeed expected to play a key role for single file transport, for which dynamics are highly correlated.

Aside from this situation, various indications of the role of thermal fluctuations have been reported, especially in studies involving capillary dynamics. One may cite the noise effects in the breakup of fluid nanojets by Moseler and Landman,[132,133] the noise assisted spreading of drops,[134] and the influence of thermal noise in thin film dewetting[135] in relation to experiments.[136] In these different cases, thermal noise modifies quite strongly the dynamics, even at relatively large length scales (up to typically 100 nm in ref. 135). These different situations can be described within the framework of fluctuating hydrodynamics,[137] in which a random stress tensor is added to the Navier–Stokes terms. While this approach is indeed fruitful in describing the role of thermal fluctuations, a general criterion to quantify the importance of noise is however lacking.

Here, to illustrate the potential importance of fluctuations, we consider a simple situation in which an osmotic flow is driven across a single nanopore. The pore has diameter *d* and length *L* and is *impermeable* to the solute. The pressure drop across the membrane is $\Delta p = k_BT\Delta c$, with *c* the solute concentration. Under the pressure drop Δp, a fluid flow is induced, with mean velocity $\bar{u}$. Assuming NS equations hold gives $\bar{u} \sim \frac{d^2}{\eta}\frac{\Delta p}{L}$. The flow is in the direction of higher solute concentrations.

Now one may ask the question: could fluctuations yield a reverse flow, *i.e.* against the pressure drop, at least for a short time? This would be clearly a (punctual) violation of the second law. And what is the minimal size at which this may occur? A lead to such questions is provided by the so-called Fluctuation Theorem, which quantifies the probability of second law violation by fluctuations. This domain is presently very active, in particular in the context of single molecule spectroscopy.[138] The theorem states that the probability $P(Q_\tau)$ of finding a value Q_τ of the amount of heat dissipated in a time interval τ satisfies in a non-equilibrium state[139–141]

$$P(Q_\tau)/P(-Q_\tau) = \exp(Q_\tau/k_BT) \tag{44}$$

In the nanofluidic situation, this theorem would thus quantify the probability of a flow *against* the osmotic gradient. Let us compute the order of magnitude for the averaged $\bar{Q}_\tau$ in the stationary state: $\bar{Q}_\tau \approx \bar{F} \times \bar{u} \times \tau$, with *F* the frictional force acting on the nanopore surface. One has $F = \pi dL\sigma_w$ with σ_w the stress at the wall given by $\sigma_w \sim d\Delta p/L$. To fix ideas, we compute the amount of heat dissipated over a time needed for a molecule to pass through the whole pore, *i.e.* $\tau = L/\bar{u}$. Altogether this gives

$$\overline{Q}_\tau \sim \Delta p \times \mathscr{V} \qquad (45)$$

with $\mathscr{V} = \frac{\pi}{4}d^2L$ the volume of the pore. According to eqn (44), violations of the second principle are more likely to occur when $\bar{Q}_\tau$ is of the order of the thermal energy. Putting in numbers, if we choose Δc = 1 M (close to physiological conditions), then the condition $Q_\tau \sim k_BT$ is obtained for *nanometric* volumes, *i.e.* for a pore with a size (diameter, length) typically in the nanometer range. This is indeed typically the order of magnitude for the size of biological pores.

The occurrence of such 'counter flow' fluctuations has actually been observed in MD simulations of osmotic flows through carbon nanotube membranes by Hummer *et al.*,[130] and a stochastic model was proposed to account for the water transport across the tube. The above analysis thus fixes the limit where fluctuations starts to be predominant over the mean behavior.

Finally we note that one may perform the same estimate for a different situation of an electric current induced by an electrostatic potential drop ΔV. Assuming a bulk conductance, the dissipated heat is estimated under quite similar terms as $\bar{Q}_\tau \sim e\Delta V \times \rho_S\mathscr{V}$, with ρ_s the salt concentration. Again the condition $Q_\tau \sim k_BT$ is achieved for *nanometric* volumes.

These results point to the crucial role of fluctuations in pores with nanometric size. This is the typical scale of a biological pore: it is then interesting to point out that biological pores are working at the edge of second law violation ($\bar{Q}_\tau \sim k_BT$).

These questions would deserve further experimental and theoretical investigations.

VI. Discussion and perspectives

In this review, we hope to have convincingly shown that nanoscales do indeed play a key role in fluidics. A broad panel of length scales ranging from the molecular to the micrometric scales leads to a rich ensemble of nanofluidic phenomena. While the domain of validity of Navier–Stokes equations was shown to extend down to the nanometer scale, specific transport phenomena show up due to the complex couplings which build up between flow and ionic transport, electrostatics, surface dynamics, *etc*.

We focused in this review on the behavior of fluids at the nanoscale. But in doing so we omitted several important topics in the discussions, however strongly connected to nanofluidics questions.

This concerns in particular the field of transport of (macro)-molecules, polymers, polyelectrolytes or biological molecules (DNA, RNA) through nanopores. This involves either biological nanopores, like the widely studied α-haemolysin, or artificial solid-state nanopores made by ionic drilling of membranes.[142] Starting from the pioneering work by Bezrukov *et al.* and Kasianowicz *et al.*[50–52] there has been a thorough exploration of the translocation mechanisms of macromolecules in tiny pores,[143,144] with important implications for the understanding of the biological translocation process and potentially for single molecule analysis. This domain is now rapidly expanding and we refer to the recent review by Dekker for further reading.[142]

Another aspect that we left aside is the question of nano- and micro-structured surfaces, in line with the recently developed super-hydrophobic surfaces exhibiting the Lotus effect. Superhydrophobic coatings can lead to huge slippage effects, with slip lengths in the micrometer range,[46,47,145,146] as well as other dynamic phenomena, such as hydro-elastic couplings at the superhydrophobic interface.[147] Such surfaces offer the possibility to considerably enhance the efficiency of transport phenomena in particular in the context of slip enhancement discussed above. For example, as we pointed out in the text, a massive amplification, by a factor up to 10^4 (!), is predicted for diffusio-osmosis on superhydrophobic surfaces, which offers the possibility to devise efficient salt pumps.[114] The underlying coupling mechanisms remain however subtle, and this enhancement was shown for example to break down for electro-osmotic transport.[114,148] The implications of nano- and micro-structuration on fluidic transport is still at its infancy and remains to be thoroughly explored. It offers the possibility of coupling fluid dynamics over the scales, from the nano- to the micro-scales, and even to larger macroscopic scales.[149]

Furthermore, beyond these questions, it is interesting to note that nanofluidics offers the possibility of attacking old questions with new points of views, and new control possibilities. We quoted the amazing flow permeability measured by carbon nanotube membranes. While this performance remains to be understood—and possibly investigated at the individual nanotube level—this opens up very promising application in the field of energy, in particular for portable energy sources. Such membranes indeed offer the possibility of very efficient conversion of hydrostatic energy to electrical power at small scales,[104,105] see also ref. 150. As we have discussed above, low (nanoscale) friction at surfaces considerably amplifies electrokinetic processes, such as electro-osmosis or streaming currents.[100,103] The latter phenomenon is able to produce an electrical current from a pressure drop.[151,152] On this basis it was argued[104,105] that the efficiency of this mechano-electric transduction increases considerably for slippery walls: reasonable values suggest an increase of the energy conversion efficiency from a few percent to 30% for slippery surfaces! While the tremendous effect of low surface friction and slippage on charge transport has been recently confirmed experimentally,[103] its expected impact on mechano-electric energy conversion has not received experimental confirmation. Extrapolating with the results obtained with the nanotube membranes, this power conversion efficiency increases to close to 100% (due to the expected extremely weak dissipation) and this amazing prediction would lead to a production of electrical power of several kW m^{-2} for a pressure drop of 1 bar:[104] this is an impressive result, which has not been confirmed up to now. While such predictions have to be taken with care, they strongly suggest exploring the role of nanofluidics in the context of energy conversion. Such nanofluidic energy conversion devices have *a priori* the potential to power larger scale systems (for example in cars or portable devices), due to their high power density and low weight.[1]

An alternative field of application for progress made in nanofluidic transport is desalination. Indeed, as half of humanity has no immediate access to potable water, desalination of seawater has emerged over recent years as an alternative solution to developing fresh water. Reverse osmosis is one of the techniques used in this context, for example in the Ashkelon plant in Israel:[153] seawater is pushed through a membrane impermeable to the salt. Such a process is expensive (typically around 0.5 $\$/\mathrm{m}^3$ for the Ashkelon plant[153]), due—among other factors—to the energy required for this operation. The possibility of reducing this energy by using a considerably more permeable membrane should therefore have a direct impact on the cost of produced water. This is a potential application of membranes made of carbon nanotubes which, as we discussed,[9,35,36] exhibit a permeability 2 to 3 orders of magnitude higher than those with micropores, a fully unexpected and still debated result. Another question for the desalinated water is the mineral composition of the produced water which—if not controlled—could have a deep, negative, impact on agriculture and health. For example the concentration of boron (B), which is toxic to many crops, is high in seawater, and requires to be partially eliminated by desalination post-treatments.[153] Reversely, desalination removes some ions which are essential to plant growth, such as Mg^{2+} or SO_4^{2-}. There is therefore a potential need for better selectivity of filtration.

Achieving selectivity similar to biological channels, such as aquaporins,[8,11,12] is still out of reach, but nanofluidics has a major role to play in this context in order to find the key tuning parameters and architecture. Such applications may still be years away, but the basic principles underlying their operation will develop in the years to come.

Acknowledgements

We thank A. Ajdari, L. Auvray, C. Barentin, A.-L. Biance, J.-L. Barrat, C. Cottin-Bizonne, L. Joly, R. Netz, D. Stein, and C. Ybert for many stimulating discussions. LB thanks in particular D. Stein for illuminating discussions on conductance effects. We acknowledge support from ANR (projects Mikado and Synodos).

References

1 J. C. T. Eijkel and A. van den Berg, *Microfluid. Nanofluid.*, 2005, **1**, 249–267.
2 A. Plecis, R. B. Schoch and P. Renaud, *Nano Lett.*, 2005, **5**, 1147–1155.
3 Q. Pu, J. Yun, H. Temkin and S. Liu, *Nano Lett.*, 2004, **4**, 1099.
4 R. B. M. Schasfoort, S. Schlautmann, J. Hendrikse and A. van den Berg, *Science*, 1999, **286**, 942.
5 R. Karnik, R. Fan, M. Yue, D. Li, P. Yang and A. Majumdar, *Nano Lett.*, 2005, **5**, 943–948.
6 R. Karnik, C. Duan, K. Castelino, H. Daiguji and A. Majumdarn, *Nano Lett.*, 2007, **7**, 547.
7 Z. Siwy and A. Fulinski, *Phys. Rev. Lett.*, 2002, **89**, 198103.
8 H. Sui, B. G. Han, J. K. Lee, P. Walian and B. K. Jap, *Nature*, 2001, **414**, 872.
9 J. K. Holt, H. G. Park, Y. Wang, M. Stadermann, A. B. Artyukhin, C. P. Grigoropoulos, A. Noy and O. Bakajin, *Science*, 2006, **312**, 1034.
10 K. Murata, *et al.*, *Nature*, 2000, **407**, 599.
11 T. Walz, B. L. Smith, M. L. Zeidel, A. Engel and P. Agre, *J. Biol. Chem.*, 1994, **269**, 1583.
12 M. Borgnia and P. Agre, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 2888.
13 D. Y. C. Chan and R. G. Horn, *J. Chem. Phys.*, 1985, **83**, 5311.
14 J.-M. Georges, S. Millot, J.-L. Loubet and A. Tonck, *J. Chem. Phys.*, 1993, **98**, 7345.
15 U. Raviv and J. Klein, *Science*, 2002, **297**, 1540.
16 T.-D. Li, J. Gao, R. Szoszkiewicz, U. Landman and E. Riedo, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2007, **75**, 115415.
17 A. Maali, T. Cohen-Bouhacina, G. Couturier and J.-P. Aimé, *Phys. Rev. Lett.*, 2006, **96**, 086105.
18 T. Becker and F. Mugele, *Phys. Rev. Lett.*, 2003, **91**, 166104.
19 J. A. Thomas and A. J. H. McGaughey, *Phys. Rev. Lett.*, 2009, **102**, 184502.
20 Y. Leng and P. T. Cummings, *Phys. Rev. Lett.*, 2005, **94**, 026101.
21 L. Bocquet and J.-L. Barrat, *Soft Matter*, 2007, **3**, 685.
22 L. Joly, C. Ybert and L. Bocquet, *Phys. Rev. Lett.*, 2006, **96**, 046101.
23 A. Saugey, L. Joly, C. Ybert, J. L. Barrat and L. Bocquet, *J. Phys.: Condens. Matter*, 2005, **17**, S4075–S4090.
24 J. A. Thomas and A. J. H. McGaughey, *Nano Lett.*, 2008, **8**, 2788–2793.
25 Microscopic Derivation of Non-Markovian Thermalization of a Brownian particle: L. Bocquet and J. Piasecki, *J. Stat. Phys.*, 1997, **87**, 1005.
26 E. Helfand, *Phys. Rev.*, 1960, **119**, 1.
27 S. Merabia, S. Shenogin, L. Joly, P. Keblinski and J.-L. Barrat, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 15113–15118.
28 G. Hummer, J. C. Rasalah and J. P. Noworyta, *Nature*, 2001, **414**, 188.
29 T. Chou, *Phys. Rev. Lett.*, 1998, **80**, 85–88.
30 T. Chou and D. Lohse, *Phys. Rev. Lett.*, 1999, **82**, 3552–3555.
31 J. S. Hansen, P. J. Daivis, K. P. Travis and B. D. Todd, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2007, **76**, 041121.
32 J. Goyon, A. Colin, G. Ovarlez, A. Ajdari and L. Bocquet, *Nature*, 2008, **454**, 84–87.
33 J. C. Rasaiah, S. Garde and G. Hummer, *Annu. Rev. Phys. Chem.*, 2008, **59**, 713–740.
34 D. G. Levitt, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1973, **8**, 3050.
35 M. Majumder, *et al.*, *Nature*, 2005, **438**, 44.
36 M. Whitby, L. Cagnon, M. Thanou and N. Quirke, *Nano Lett.*, 2008, **8**, 2632.
37 S. Joseph and N. R. Aluru, *Nano Lett.*, 2008, **8**(2), 452–458.
38 J. Israelachvili and H. Wennerström, *Nature*, 1996, **379**, 219–225.
39 D. Andelman, in *Handbook of Biological Physics*, ed. R. Lipowsky and E. Sackmann, Elsevier Science B.V., 1995, vol. 1.
40 R. B. Schoch, J. Han and P. Renaud, *Rev. Mod. Phys.*, 2008, **80**, 839–883.
41 D. Stein, M. Kruithof and C. Dekker, *Phys. Rev. Lett.*, 2004, **93**, 035901.
42 R. J. Hunter, *Foundations of Colloid Science*, Oxford, 2nd edn, 2001.
43 A. S. Khair and T. M. Squires, *J. Fluid Mech.*, 2008, **615**, 323–334.
44 R. Karnik, K. Castelino, R. Fan, P. Yang and A. Majumdar, *Nano Lett.*, 2005, **5**, 1638–1642.
45 D. Huang, C. Sendner, D. Horinek, R. Netz and L. Bocquet, *Phys. Rev. Lett.*, 2008, **101**, 226101.
46 P. Joseph, C. Cottin-Bizonne, J.-M. Benoît, C. Ybert, C. Journet, P. Tabeling and L. Bocquet, *Phys. Rev. Lett.*, 2006, **97**, 156104.
47 C. Lee, C.-H. Choi and C.-J. Kim, *Phys. Rev. Lett.*, 2008, **101**, 064501.
48 Z. B. Ge, D. G. Cahill and P. V. Braun, *Phys. Rev. Lett.*, 2006, **96**, 186101.
49 P.-G. de Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, 1979.
50 S. M. Bezrukov, I. Vodyanoy and V. A. Parsegian, *Nature*, 1994, **370**, 279–281.
51 S. M. Bezrukov, I. Vodyanoy, R. A. Brutyan and J. J. Kasianowicz, *Macromolecules*, 1996, **29**, 8517–8522.
52 J. J. Kasianowicz, E. Brandin, D. Branton and D. W. Deamer, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 13770.
53 J. Han and H. G. Craighead, *Science*, 2000, **288**, 1026–1029.

54 C. Navier, *Mem. Acad. Sci. Inst. Fr.*, 1823, **6**, 389–416.
55 E. Lauga, M. Brenner and H. Stone, in *Microfluidics: The No-SlipBoundaryCondition*, Handbook of Experimental Fluid Dynamics , ed. C. T. J. Foss and A. Yarin, Springer, New York, 2005, ch. 15.
56 C. Neto, D. Evans, E. Bonaccurso, H.-J. Butt and V. J. Craig, *Rep. Prog. Phys.*, 2005, **68**, 2859–2897.
57 P. Thompson and M. Robbins, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1990, **41**, 6830.
58 L. Bocquet and J.-L. Barrat, *Phys. Rev. Lett.*, 1993, **70**, 2726–2729.
59 J.-L. Barrat and L. Bocquet, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1994, **49**, 3079–3092.
60 M. Cieplak, J. Koplik and J. Banavar, *Phys. Rev. Lett.*, 2001, **86**, 803.
61 N. Priezjev and S. Troian, *Phys. Rev. Lett.*, 2004, **92**, 018302.
62 J.-L. Barrat and L. Bocquet, *Faraday Discuss.*, 1999, **112**, 119.
63 C. Sendner, D. Horinek, L. Bocquet and R. Netz, *Langmuir*, 2009, **25**, 10768–10781.
64 J. Israelachvili and G. Adams, *Nature*, 1976, **262**, 774–6.
65 O. Vinogradova, *Langmuir*, 1995, **11**, 2213–2220.
66 J. Israelachvili, *J. Colloid Interface Sci.*, 1986, **110**(1), 263.
67 Y. Zhu and S. Granick, *Phys. Rev. Lett.*, 2001, **87**, 096105.
68 A. Steinberger, C. C. Cottin-Bizonne, P. Kleimann and E. Charlaix, *Phys. Rev. Lett.*, 2008, **100**, 134501.
69 C. Cottin-Bizonne, B. Cross, A. Steinberger and E. Charlaix, *Phys. Rev. Lett.*, 2005, **94**, 056102.
70 C. Cottin-Bizonne, A. Steinberger, B. Cross, O. Raccurt and E. Charlaix, *Langmuir*, 2008, **24**, 1165–1172.
71 V. Craig, C. Neto and D. Williams, *Phys. Rev. Lett.*, 2001, **87**, 054504–4.
72 E. Bonaccurso, M. Kappl and H.-J. Butt, *Phys. Rev. Lett.*, 2002, **88**, 076103–4.
73 C. D. F. Honig and W. A. Ducker, *Phys. Rev. Lett.*, 2007, **98**, 028305.
74 C. L. Henry and V. J. Craig, *Phys. Chem. Chem. Phys.*, 2009, **11**, 9514–9521.
75 O. Vinogradova, H. Butt, G. Yabukov and F. Feuillebois, *Rev. Sci. Instrum.*, 2001, **72**, 2330.
76 A. Maali, C. Hurth, T. Cohen-Bouhacina, G. Couturier and J.-P. Aimé, *Appl. Phys. Lett.*, 2006, **88**, 163504.
77 A. Maali, T. Cohen-Bouhacina and H. Kellay, *Appl. Phys. Lett.*, 2008, **92**, 053101.
78 D. Tretheway and C. Meinhart, *Phys. Fluids*, 2002, **14**(3), L9–L12.
79 D. Lumma, A. Best, A. Gansen, F. Feuillebois, J. Radler and O. I. Vinogradova, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2003, **67**, 056313.
80 P. Joseph and P. Tabeling, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2005, **71**, 035303.
81 O. I. Vinogradova, K. Koynov, A. Best and F. Feuillebois, *Phys. Rev. Lett.*, 2009, **102**, 118302.
82 P. Huang, J. S. Guasto and K. S. Breuer, *J. Fluid Mech.*, 2006, **566**, 447.
83 R. Sadr, M. Yoda, M. P. Gnanaprakasam and A. T. Conlisk, *Appl. Phys. Lett.*, 2006, **89**, 044103.
84 D. Lasne, A. Maali, Y. Amarouchene, L. Cognet, B. Lounis and H. Kellay, *Phys. Rev. Lett.*, 2008, **100**, 214502.
85 P. Huang, J. Guasto and K. Breuer, *J. Fluid Mech.*, 2009, **637**, 241–265.
86 B. Borkent, S. Dammer, H. Schonherr, G. Vancso and D. Lohse, *Phys. Rev. Lett.*, 2007, **98**, 204502.
87 S. Yang, S. Dammer, N. Bremond, E. Zandvliet, H. J. W. Kooij and D. Lohse, *Langmuir*, 2007, **23**, 7072–7077.
88 W. A. Ducker, *Langmuir*, 2009, **25**, 8907.
89 R. Pit, H. Hervet and L. Léger, *Phys. Rev. Lett.*, 2000, **85**, 980–983.
90 T. Schmadtko, H. Hervet and L. Leger, *Phys. Rev. Lett.*, 2005, **94**, 244501.
91 J. L. Anderson, *Annu. Rev. Fluid Mech.*, 1989, **21**, 61.
92 Y. Levin, E. Trizac and L. Bocquet, *J. Phys.: Condens. Matter*, 2003, **15**, S3523–S3536.
93 I. Rouzina and V. A. Bloomfield, *J. Phys. Chem.*, 1996, **100**, 9977–9989.
94 D. M. Huang, C. Cottin-Bizonne, C. Ybert and L. Bocquet, *Langmuir*, 2008, **24**, 1442–1450.
95 W. Kunz, J. Henle and B. W. Ninham, *Curr. Opin. Colloid Interface Sci.*, 2004, **9**, 19.
96 D. Horinek and R. Netz, *Phys. Rev. Lett.*, 2007, **99**, 226104.
97 P. Jungwirth and D. J. Tobias, *Chem. Rev.*, 2006, **106**, 1259–1281.
98 D. M. Huang, C. Cottin-Bizonne, C. Ybert and L. Bocquet, *Phys. Rev. Lett.*, 2007, **98**, 177801.
99 V. Buch, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 7342.
100 L. Joly, C. Ybert, E. Trizac and L. Bocquet, *Phys. Rev. Lett.*, 2004, **93**, 257805.
101 V. M. Muller, *et al.*, *Colloid J. USSR*, 1986, **48**, 718.
102 N. V. Churaev, J. Ralston, I. P. Sergeeva and V. D. Sobolev, *Adv. Colloid Interface Sci.*, 2002, **96**, 265.
103 C. I. Bouzigues, P. Tabeling and L. Bocquet, *Phys. Rev. Lett.*, 2008, **101**, 114503.
104 Y. Ren and D. Stein, *Nanotechnology*, 2008, **19**, 195707.
105 S. Pennathur, E. Eijkel and A. Ven der Berg, *Lab Chip*, 2007, **7**, 1234.
106 R. Qiao and N. R. Aluru, *Phys. Rev. Lett.*, 2004, **92**, 198301.
107 D. Horinek, A. Herz, L. Vrbka, F. Sedlmeier, S. I. Mamatkulov and R. R. Netz, *Chem. Phys. Lett.*, 2009, **479**, 173–183.
108 R. Piazza, *Soft Matter*, 2008, **4**, 1740–1744.
109 S. Duhr and D. Braun, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 19678–19682.
110 H.-R. Jiang, H. Wada, N. Yoshinaga and M. Sano, *Phys. Rev. Lett.*, 2009, **102**, 208301.
111 A. Wurger, *Phys. Rev. Lett.*, 2008, **101**, 108302.
112 J. P. Ebel, J. L. Anderson and D. C. Prieve, *Langmuir*, 1988, **4**, 396–406.
113 B. Abécassis, C. Cottin-Bizonne, C. Ybert, A. Ajdari and L. Bocquet, *Nat. Mater.*, 2008, **7**, 785.
114 D. M. Huang, C. Cottin-Bizonne, C. Ybert and L. Bocquet, *Phys. Rev. Lett.*, 2008, **101**, 064503.
115 Armand Ajdari and L. Bocquet, *Phys. Rev. Lett.*, 2006, **96**, 186102.
116 J. Morthomas and A. Wurger, *J. Phys.: Condens. Matter*, 2009, **21**, 035103.
117 C. Ybert, C. Barentin, C. Cottin-Bizonne, P. Joseph and L. Bocquet, *Phys. Fluids*, 2007, **19**, 123601.
118 S. Levine, J. R. Marriott and K. Robinson, *J. Chem. Soc., Faraday Trans. 2*, 1975, 711–11.
119 H. Daiguji, P. Yang and A. Majumdar, *Nano Lett.*, 2004, **4**, 137–142.
120 C. Dellago, M. N. Naor and G. Hummer, *Phys. Rev. Lett.*, 2003, **90**, 105902.
121 S. R. De Groot and P. Mazur, *Non-Equilibrium Thermodynamics*, North-Holland, Amsterdam, 1969.
122 E. Brunet and A. Ajdari, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2004, **69**, 016306.
123 O. Kedem and A. Katchalsky, *J. Gen. Physiol.*, 1961, **45**, 143.
124 O. Kedem and A. Katchalsky, *Trans. Faraday Soc.*, 1963, **59**, 1941–1953.
125 Y. C. Wang, A. L. Stevens and J. Han, *Anal. Chem.*, 2005, **77**, 4293–4299.
126 A. Plecis, C. Nanteuil, A.-M. Haghiri-Gosnet and Y. Chen, *Anal. Chem.*, 2008, **80**, 9542–9550.
127 C. Kittel, *Introduction to Solid State Physics*, Wiley, 5th edn, 1976.
128 Z. Siwy, I. D. Kosińska, A. Fuliński and C. R. Martin, *Phys. Rev. Lett.*, 2005, **94**, 048102.
129 R. M. M. Smeets, U. F. Keyser, N. H. Dekker and C. Dekker, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 417–421.
130 A. Kalra, S. Garde and G. Hummer, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 10175.
131 A. Berezhkovskii and G. Hummer, *Phys. Rev. Lett.*, 2002, **89**, 064503.
132 M. Moseler and U. Landman, *Science*, 2000, **289**, 1165.
133 J. Eggers, *Phys. Rev. Lett.*, 2002, **89**, 084502.
134 B. Davidovitch, E. Moro and H. A. Stone, *Phys. Rev. Lett.*, 2005, **95**, 244505.
135 G. Grün, K. Mecke and M. Rauscher, *J. Stat. Phys.*, 2006, **122**, 1261.
136 J. Becker, G. Grün, R. Seemann, H. Mantz, K. Jacobs, K. R. Mecke and R. Blossey, *Nat. Mater.*, 2003, **2**, 59.
137 L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Pergamon Press, London, 1958.

138 F. Ritort, *J. Phys.: Condens. Matter*, 2006, **18**, R531.
139 D. J. Evans, E. G. D. Cohen and G. P. Moriss, *Phys. Rev. Lett.*, 1993, **71**, 2401–2404.
140 R. van Zon and E. G. D. Cohen, *Phys. Rev. Lett.*, 2003, **91**, 110601.
141 R. van Zon, S. Ciliberto and E. G. D. Cohen, *Phys. Rev. Lett.*, 2004, **92**, 130601.
142 C. Dekker, *Nat. Nanotechnol.*, 2007, **2**, 209–215.
143 G. Oukhaled, J. Mathe and A.-L. Biance, *et al.*, *Phys. Rev. Lett.*, 2007, **98**, 158101.
144 S. van Dorp, U. F. Keyser, N. H. Dekker, C. Dekker and S. G. Lemay, *Nat. Phys.*, 2009, **5**, 347–351.
145 C. Cottin-Bizonne, J.-L. Barrat, L. Bocquet and E. Charlaix, *Nat. Mater.*, 2003, **2**, 238.
146 J. Ou and J. P. Rothstein, *Phys. Fluids*, 2005, **17**, 103606.
147 A. Steinberger, C. Cottin-Bizonne, P. Kleimann and E. Charlaix, *Nat. Mater.*, 2007, **6**, 665.
148 T. Squires, *Phys. Fluids*, 2008, **20**, 092105.
149 C. Duez, C. Ybert, C. Clanet and L. Bocquet, *Nat. Phys.*, 2007, **3**, 180–183.
150 Y. Zhao *et al.*, *Adv. Mater.*, 2008, **20**, 1772.
151 F. H. J. van der Heyden, D. J. Bonthuis, D. Stein, C. Meyer and C. Dekker, *Nano Lett.*, 2006, **6**, 2232–2237.
152 F. H. J. van der Heyden, D. J. Bonthuis, D. Stein, C. Meyer and C. Dekker, *Nano Lett.*, 2007, **7**, 1022–1025.
153 U. Yermiyahu *et al.*, *Science*, 2007, **318**, 920.
154 ''Par ma foi ! il y a plus de quarante ans que je dis de la prose sans que j'en susse rien, et je vous suis le plus obligé du monde de m'avoir appris cela.'', which translates into ''By my faith! For more than forty years I have been speaking prose without knowing anything about it and I am much obliged to you for having taught me that''.
155 We specifically thank D. Stein for very interesting discussions concerning this point.
156 We choose here to define the zeta potential as the one appearing in the Smoluchowski formula. According to eqn (20), this may differ from the surface potential, depending on conditions. Another possibility would have been to define the zeta potential as V_0 and consider an amplified electro-osmotic mobility. We preferred to use the former definition.

Chem Soc Rev

REVIEW ARTICLE

View Article Online
View Journal

Cite this: DOI: 10.1039/c8cs00420j

Osmosis, from molecular insights to large-scale applications

Sophie Marbach † and Lydéric Bocquet *

Osmosis is a universal phenomenon occurring in a broad variety of processes and fields. It is the archetype of entropic forces, both trivial in its fundamental expression – the van 't Hoff perfect gas law – and highly subtle in its physical roots. While osmosis is intimately linked with transport across membranes, it also manifests itself as an interfacial transport phenomenon: the so-called diffusio-osmosis and -phoresis, whose consequences are presently actively explored for example for the manipulation of colloidal suspensions or the development of active colloidal swimmers. Here we give a global and unifying view of the phenomenon of osmosis and its consequences with a multi-disciplinary perspective. Pushing the fundamental understanding of osmosis allows one to propose new perspectives for different fields and we highlight a number of examples along these lines, for example introducing the concepts of osmotic diodes, active separation and far from equilibrium osmosis, raising in turn fundamental questions in the thermodynamics of separation. The applications of osmosis are also obviously considerable and span very diverse fields. Here we discuss a selection of phenomena and applications where osmosis shows great promises: osmotic phenomena in membrane science (with recent developments in separation, desalination, reverse osmosis for water purification thanks in particular to the emergence of new nanomaterials); applications in biology and health (in particular discussing the kidney filtration process); osmosis and energy harvesting (in particular, osmotic power and blue energy as well as capacitive mixing); applications in detergency and cleaning, as well as for oil recovery in porous media.

Received 15th February 2019

DOI: 10.1039/c8cs00420j

rsc.li/chem-soc-rev

1 Introduction

From the etymological point of view, osmosis denotes a ''push'' and indeed osmosis is usually associated with the notion of force and pressure. Osmosis is a very old topic, it was first observed centuries ago with reports by Jean-Antoine Nollet in the 18th century. It was rationalized more than one century later by van 't Hoff, who showed that the osmotic pressure took the form of a perfect gas equation of state. In practice, an osmotic pressure is typically expressed across a semi-permeable membrane, *e.g.* a membrane that allows only the solvent to pass while retaining solutes. If two solutions of a liquid containing different solute concentrations are put into contact through such a semi-permeable membrane, the fluid will undergo a driving force pushing it towards the reservoir with the highest solute concentration, see Fig. 1. Reversely, in order to prevent the fluid from passing through the membrane, a pressure has to be applied to the fluid to counteract the flow: the applied pressure is then equal to the osmotic pressure.

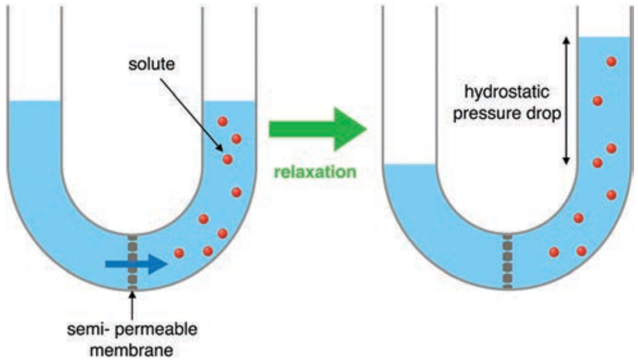


Fig. 1 Key manifestation of osmosis. A semi-permeable membrane allows transport of water upon a solute concentration difference (in red). The flow of water is directed from the fresh water reservoir to the concentrated reservoir.

Osmosis is therefore extremely simple in its expression. Yet it is one of the most subtle physics phenomenon in its roots – it resulted in many debates over years.[1,2] Osmosis also implies subtle phenomena, in particular as a prototypical illustration for the explicit conversion of entropy of mixing into mechanical

---

Laboratoire de Physique de l'Ecole Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université Paris-Diderot, Sorbonne Paris Cité, Paris, France. E-mail: lyderic.bocquet@lps.ens.fr
† Current address: Courant Institute, New York University, New York.

work. In spite of centuries of exploration, osmosis as a field remains very lively, with a number of recent breakthroughs both in its concepts and applications as we shall explore in this review. A simple reason for the importance of osmosis is that it is a very powerful phenomenon: giving just one illustrative number, it is amazing to realize that a concentration difference of ∼1.2 molar, which corresponds roughly to the difference between sea and fresh water (and can be easily achieved in anyone's kitchen), yields an osmotic pressure of ∼30 atmospheres. This is the hydrostatic pressure felt under a 300 m water column! Osmosis has potentially a destructive power, in particular in soft tissues and membranes, with possible fatal consequences.[3] This explains actually why it is also an efficient asset for food preservation (such as fish and meat curing with dry salt).

Osmosis is accordingly also a key and universal phenomenon occurring in many processes, ranging from biological transport in plants, trees and cells, to water filtration, reverse and forward osmosis, energy harvesting and osmotic power, capacitive mixing, oil recovery, detergency and cleaning, active matter, to quote just a few.

The literature on osmosis and its consequences is accordingly absolutely huge,‡ and it may seem hopeless to cover in a single review all aspects of the topic with an exhaustive discussion of all possible applications. Also, such a comprehensive list would probably be useless for readers who want to catch up with the topics related to osmosis. In writing this review, we thus decided to rather present a tutorial and unified perspective of osmosis, obviously with personal views, avoiding exhaustiveness to highlight a number of significant questions discussed in the recent literature. The review will therefore explore the fundamental foundations of osmosis, emphasizing in particular the – sometimes subtle – mechanical balance at play; then report on more recent concepts and applications related to osmosis which – in our opinion – prove promising for future perspectives. We will accordingly put in context phenomena like diffusio-osmosis and -phoresis, as well as "active" (non-equilibrium) counterparts of osmosis, which were realized lately to play a growing role in numerous applications in filtration and energy harvesting.

The review is organized as follows. We start with some basic reminder of the fundamentals of osmosis in terms of equilibrium and non-equilibrium thermodynamics of the underlying process. We further highlight simplistic views clarifying the mechanical aspects of osmosis. We then discuss membrane-less osmosis and the so-called diffusio-osmotic flows. We then show how such phenomena may be harnessed to go beyond the simple views of van 't Hoff. We then explore the transport of particles under solute gradients, diffusio-phoresis, and discuss how this phenomenon can be harnessed to manipulate colloidal assemblies. And we finally illustrate a number of applications for the introduced concepts, from desalination, water treatment, the functioning of the kidney, blue energy harvesting, *etc.* We conclude with some final, brief, perspectives.

2 Osmosis: the van 't Hoff legacy

2.1 A quick history of osmosis

We start this review with a short and non-exhaustive journey through time in order to highlight how a complete understanding of osmosis emerged over time. We refer *e.g.* to ref. 4 for a more detailed historical review. The first occurrence of the term "osmosis" and clear observation of its effects - beyond the seminal work of Nollet – is reported at least as early as in the

---

‡ The word "osmosis" in *Web of Science* results in tens of thousands of referenced papers on this topic.

Sophie Marbach

Sophie Marbach is a post-doctoral fellow at the Courant Institute of New York University. She obtained her PhD from Ecole Normale Supérieure, Paris. Her main interest is to investigate out-of-equilibrium driven processes at the interface of biology, physics and chemistry at the small scales. She uses a number of theoretical and numerical tools with strong connections to experiments to answer open fundamental questions and to help building artificial devices. For example she recently showed how to draw inspiration from the human kidney to build an innovative filtration device.

Lydéric Bocquet

Lydéric Bocquet is director of research at CNRS and joint professor at the Ecole Normale Supérieure, Paris. His research interests are mainly curiosity driven and extend to domains at the interface of fluid dynamics, soft condensed matter and nanoscience. He combines experiments, theory and simulations to explore the intimate mechanisms of fluid interfaces from the macroscopic down to the molecular level. His recent interests aimed at taking benefit of the unexpected fluid transport behavior occurring at the nanoscales to propose new routes for energy harvesting and desalination. He also has a strong interest in every-day life science. He received several awards including the Friedrich Wilhelm Bessel prize of the von Humboldt foundation in 2007 and Advanced Grants of the European Research Council in 2010 and 2018. http://www.phys.ens.fr/∼lbocquet/.

works of Henri Dutrochet in the 1820s.[5,6] He observed swelling events or emptying of pockets driven by the presence of various dissolved components in water (different sugars in plants, sperm in slugs...). In reference to the greek term "osmose" (meaning "impulsion" or "push") he introduced the vocabulary "*endosmose*" and "*exosmose*". Interestingly, Dutrochet served as a pioneer in linking these different topics by claiming that the same physical force could be used to describe all these events,[5] which is indeed a unique and fascinating feature of osmosis. Yet, the mechanisms driving osmotic flow were still unclear, and entangled (or believed to be entangled) with capillary and electrical effects. In 1854 T. Graham introduced the word "osmosis" building on the work of Dutrochet.[7]

Interestingly, the distinction between osmosis and pure diffusion – without a membrane, see Fig. 2 – is not clear from the beginning. The confusion will grow stronger with the work of Adolf Fick in 1855,[8] where he claims that diffusive motion (Fickian diffusion) is the driver for osmotic flow (the water concentration imbalance between the two compartments drives the water flow). The question of finding whether osmotic flow is diffusion-driven or not will be an ongoing debate for a century. That diffusion alone cannot account for osmosis is not widely appreciated. In 1957, the debate is definitely closed by an experimental visualization of water flow, using radioactively labeled water molecules[9] and verified in ref. 10. The flows measured were significantly higher than that expected by pure diffusion.

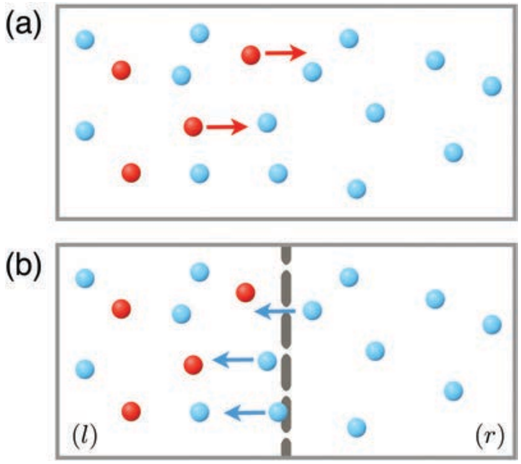


Fig. 2 Osmosis *versus* diffusion. (a) Situation where motion of the red solute particles is governed by diffusion alone (b) osmosis situation, where motion of the blue solvent particles is driven by osmosis; the solute particles being "repelled" by the membrane, the membrane exerts an effective force on the liquid (solute + solvent) that drives solvent flow towards the highly concentrated reservoir.

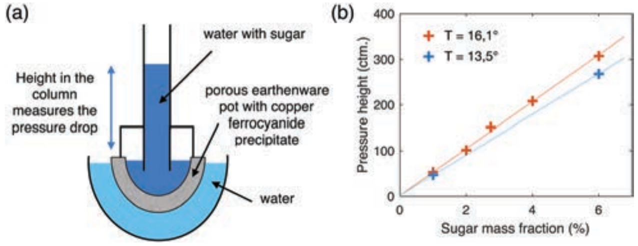


Fig. 3 First measurements of osmotic pressure. (a) Schematic of a Pfeffer cell, the device used by W. Pfeffer to perform the first measurements of osmotic pressure; (b) osmotic pressure as a function of solute concentration at two temperatures, with experimental data points from W. Pfeffer[11] verifying a linear relation (the lines are a guide for the eye).

In 1877, Wilhelm Pfeffer made the first measurements of osmotic pressure,[11] see Fig. 3. At equilibrium, he measured a rise in the concentrated solution, corresponding to a hydraulic pressure drop that is equal to the osmotic pressure. He measured a linear relation between the osmotic pressure and the concentration difference. But also, Pfeffer measured that for each degree rise in temperature, the pressure would go up by 1/270.[12] This fact was reported to Jacobus Henricus van 't Hoff by the botanist Hugo de Vries and van 't Hoff immediately recognized that 270 was an approximation of 273 K. Intrigued by this result, he attempted in 1887 to rationalize this linear dependence[13] and suggested to interpret that the osmotic pressure $\Delta\Pi$ was exerted by the solute particles and equal to the partial pressure that they would have in gas phase (therefore the term "osmotic pressure"):

> *[...] it occurred to me that with the semipermeable barrier all the reversible transformations that so materially ease the application of thermodynamics to gases, become equally available for solutions... That was a ray of light; and led at once to the inescapable conclusion that the osmotic pressure of dilute solutions must vary with temperature entirely as does gas pressure [...].*[12]

then writing

$$\Delta\Pi = k_B T \Delta c_s \qquad (1)$$

with k_B the Boltzmann constant, T temperature and Δc_s the solute imbalance between reservoirs.

Eqn (1) is today referred to as the van 't Hoff law, and gives in practice good agreement for the osmotic pressure measured between two solutions separated by a membrane permeable only to the solvent. For a solute imbalance of $\Delta C_s = 1.2$ mol L^{-1} (corresponding to the ionic strength difference between fresh and sea water, which is twice – two ions for salt – the typical concentration 0.6 mol L^{-1}), we find an osmotic pressure of $\Delta\Pi = k_B T \Delta C_s \mathcal{N}_A \simeq 30$ bar.

At the time, the interpretation of van 't Hoff gave rise to a number of debates.[1,2] In the following decades a great number of theories were invented to describe the osmotic phenomenon and a detailed review of these theories can be found in ref. 14. Among all these theories, two of them caught a lot of attention. One of them was the proof of van 't Hoff's law using the kinetic theory of gas to describe the two solutions[15] (which was later improved for multicomponent systems[16]). The other one is acknowledged today as the common description of osmosis, and makes use of the concept of chemical potentials first introduced by Josiah Willard Gibbs[17] (actually introduced as a physical descriptor required to understand osmosis), that we recall in the next section.

2.2 Thermodynamic equilibrium

We start with the thermodynamic derivation of the osmotic pressure as proposed by Gibbs. We follow here the clear-cut presentation proposed in the textbook by Callen,[18] which we recall here for the purpose of settling properly the foundations. In addition to Callen it is also worth reading the rigorous thermodynamic treatment by Guggenheim.[19] We consider a composite system made of two simple reservoirs (left and right) separated by a rigid wall permeable to component w (usually the solvent) and totally impermeable to all other components (labelled s_1, s_2 and so on). The whole system is in contact with a thermal bath at temperature T. The solvent is in equilibrium over the whole system, *i.e.* over the two reservoirs, while solutes cannot equilibrate between the reservoirs. Due to the imbalance of solute fraction between both reservoirs, the solvent cannot keep a homogeneous pressure across the two reservoirs while ensuring the equality of chemical potential at equilibrium. An (osmotic) pressure drop builds up, which the membrane withstands.

Assuming that the solute concentration is dilute, the Gibbs free energy of a binary system of solvent and dilute solute can be written as

$$G\left(T,p,N_{\mathrm{w}},N_{\mathrm{s}_1}\right) = N_{\mathrm{w}}\mu_{\mathrm{w}}^0(p,T) + N_{\mathrm{s}}\mu_{\mathrm{s}}^0(p,T) + N_{\mathrm{w}}k_{\mathrm{B}}T\ln\frac{N_{\mathrm{w}}}{N_{\mathrm{w}}+N_{\mathrm{s}}} + N_{\mathrm{s}}k_{\mathrm{B}}T\ln\frac{N_{\mathrm{s}}}{N_{\mathrm{w}}+N_{\mathrm{s}}} \tag{2}$$

where $\mu_{\mathrm{w}}^0(p,T)$ and $\mu_{\mathrm{s}}^0(p,T)$ are the chemical potentials of the pure solvent and solute and the last two terms correspond to the entropy of mixing terms. In the dilute regime where $N_{\mathrm{s}} \ll N_{\mathrm{w}}$ the Gibbs free energy simplifies to

$$G(T,p,N_{\mathrm{w}},N_{\mathrm{s}})) = N_{\mathrm{w}}\mu_{\mathrm{w}}^0(p,T) + N_{\mathrm{s}}\mu_{\mathrm{s}}^0(p,T) - k_{\mathrm{B}}TN_{\mathrm{s}} + N_{\mathrm{s}}k_{\mathrm{B}}T\ln\frac{N_{\mathrm{s}}}{N_{\mathrm{w}}} \tag{3}$$

and the chemical potential of the solvent may be obtained as $\mu_{\mathrm{w}} = \partial_{N_{\mathrm{w}}}G$

$$\mu_{\mathrm{w}}(T,p,X) \simeq \mu_{\mathrm{w}}^0(p,T) - k_{\mathrm{B}}TX. \tag{4}$$

with $X \simeq N_{\mathrm{s}}/N_{\mathrm{w}}$ the solute molar fraction. The chemical potential balance, $\mu_{\mathrm{w}}^{(\mathrm{l})} = \mu_{\mathrm{w}}^{(\mathrm{r})}$, thus writes

$$\mu_{\mathrm{w}}\left(T,p^{(\mathrm{l})},0\right) - k_{\mathrm{B}}T\frac{N_{\mathrm{s}}^{(\mathrm{l})}}{N_{\mathrm{w}}^{(\mathrm{l})}} = \mu_{\mathrm{w}}\left(T,p^{(\mathrm{r})},0\right) - k_{\mathrm{B}}T\frac{N_{\mathrm{s}}^{(\mathrm{r})}}{N_{\mathrm{w}}^{(\mathrm{r})}}. \tag{5}$$

Noting then that for small pressure drops, $\mu_{\mathrm{w}}(T,p^{(\mathrm{r})},0) \simeq \mu_{\mathrm{w}}(T,p^{(\mathrm{l})},0) + (p^{(\mathrm{r})} - p^{(\mathrm{l})})v_{\mathrm{w}}$, with $v_{\mathrm{w}} = \partial_p\mu_{\mathrm{w}}(T,p,0)$ the molecular volume, one deduces finally

$$\Delta\Pi = p^{(\mathrm{r})} - p^{(\mathrm{l})} = k_{\mathrm{B}}T\left[\frac{N_{\mathrm{s}}^{(\mathrm{r})}}{V^{(\mathrm{r})}} - \frac{N_{\mathrm{s}}^{(\mathrm{l})}}{V^{(\mathrm{l})}}\right] \tag{6}$$

Introducing the concentration as $c_{\mathrm{s}} = N_{\mathrm{s}}/V$, one thus recovers the result of van 't Hoff

$$\Delta\Pi = k_{\mathrm{B}}T\Delta c_{\mathrm{s}}. \tag{7}$$

In the case of several dilute solutes, this generalizes simply to

$$\Delta\Pi = k_{\mathrm{B}}T\left[\frac{N_{\mathrm{s}_1}^{(\mathrm{r})} + N_{\mathrm{s}_2}^{(\mathrm{r})} + \ldots}{V^{(\mathrm{r})}} - \frac{N_{\mathrm{s}_1}^{(\mathrm{l})} + N_{\mathrm{s}_2}^{(\mathrm{l})} + \ldots}{V^{(\mathrm{l})}}\right]. \tag{8}$$

The derivation above is limited to dilute solutes. For arbitrary molar fractions X of solute/solvent mixtures, the osmotic pressure is given in terms of the general expression for the pressure, namely[20]

$$\Pi(X) = X\frac{\partial f}{\partial X} - f[X] + f[X=0], \tag{9}$$

with $f(X) = F/V$ the Helmholtz free energy density calculated for a solute molar fraction X. Deviations from ideality are for example measured for polymers, where the range of validity of the van 't Hoff law decreases with increasing molecular weight.[20] Deviations are also expected for highly concentrated brines or solvent mixtures, *e.g.* in the context of solvophoresis, see below ref. 21.

An interesting, and quite counter-intuitive remark is that – provided it is semi-permeable – the membrane characteristics do not appear in this thermodynamic expression for the osmotic pressure. Another puzzling remark is that the osmotic pressure is a colligative property, *i.e.* it does not depend on the nature of the solute (nor that of the membrane), but only on the concentration of the solute. This is relevant when the membrane is completely impermeable to the solute, but when the membrane is only partially impermeable, or when there are different solutes with different permeation properties, there may be both a solvent and a solute flux driven by the solute concentration imbalance (in opposite directions).[22–26] The osmotic pressure is then usually assumed to be reduced by a so-called (dimensionless) reflection factor, σ, which depends on the specific properties of solvent–membrane interactions and transport. This requires to go beyond the thermodynamic equilibrium and consider the detailed mass and solute transport across the membrane, as we now explore.

2.3 Osmotic fluxes and thermodynamic forces

Following the work of Staverman,[27] Kedem and Katchalsky derived a relation between solute and solvent flows through a porous membrane and the corresponding thermodynamic forces,[28] based on Onsager's framework of irreversible processes.[29,30]

As in the previous section, we consider a composite system made of two simple reservoirs (left and right), containing a solvent w and a solute s. The reservoirs are separated by a rigid wall, which is now permeable to all components, but with a differential permeability between the solute and the solvent. Obviously, the objective of the membrane is somehow to reject the solute but the rejection is incomplete here. The whole system is put in contact with a thermal bath at temperature T.

The entropy production (per unit membrane area $\mathscr{A}$) is accordingly written as:

$$\Phi = \frac{T}{\mathscr{A}}\frac{\mathrm{d}S}{\mathrm{d}t} = -\left(\mu_{\mathrm{w}}^{(\mathrm{r})} - \mu_{\mathrm{w}}^{(\mathrm{l})}\right)\frac{\mathrm{d}N_{\mathrm{w}}^{(\mathrm{r})}}{\mathrm{d}t} - \left(\mu_{\mathrm{s}}^{(\mathrm{r})} - \mu_{\mathrm{s}}^{(\mathrm{l})}\right)\frac{\mathrm{d}N_{\mathrm{s}}^{(\mathrm{r})}}{\mathrm{d}t} \tag{10}$$

with $\frac{\mathrm{d}N_{\mathrm{i}}^{(\mathrm{r})}}{\mathrm{d}t}$ the flux of molecules of component i per unit area. The dissipation function of eqn (10) is a product of fluxes $\frac{\mathrm{d}N_{\mathrm{i}}^{(\mathrm{r})}}{\mathrm{d}t}$ and the corresponding thermodynamic forces, here the differences in chemical potentials.

Now, restricting ourselves to ideal solutions for simplicity, one may write the chemical potential difference as $\mu_i^{(r)} - \mu_i^{(l)} = \nu_i\Delta p + k_BT\Delta\ln X_i$ where X_i is the molar fraction of component i and $\nu_i = (\partial\mu_i/\partial p)$ the molar volume of i. Accordingly, $\mu_s^{(r)} - \mu_s^{(l)} = \nu_s\Delta p + k_BT\dfrac{\Delta c_s}{c_s}$ for the solute and $\mu_w^{(r)} - \mu_w^{(l)} = \nu_w\Delta p - k_BT\dfrac{\Delta c_s}{c_w}$ for the solvent (where we used $c_s \ll c_w$). Eqn (10) then rewrites:

$$\Phi = -\left.\quad \nu_w\frac{dN_w^{(r)}}{dt} + \nu_s\frac{dN_s^{(r)}}{dt}\right)\Delta p - \left.\quad \frac{1}{c_s}\frac{dN_s^{(r)}}{dt} - \frac{1}{c_w}\frac{dN_w^{(r)}}{dt}\right)k_BT\Delta c_s \tag{11}$$

From the dissipation function in eqn (11), we may thus identify a new set of forces and fluxes: new forces are $-\Delta p$ and $-k_BT\Delta c_s$, respectively the hydrostatic pressure and solute concentration imbalance; new flows are (a) the total volume flow through the membrane (sum of all flows):

$$Q = \nu_w\frac{dN_w^{(r)}}{dt} + \nu_s\frac{dN_s^{(r)}}{dt} \tag{12}$$

and (b) the excess solute flow (as compared to the solute flow carried by the solvent) or the exchange flow:

$$J_e = \frac{1}{c_s}\frac{dN_s^{(r)}}{dt} - \frac{1}{c_w}\frac{dN_w^{(r)}}{dt} \tag{13}$$

Under the assumption that the concentration of solute is small $c_s \ll c_w$, one may thus rewrite $J_e \simeq J_s/c_s - Q$ where J_s is the solute flow.

The framework of irreversible processes assumes a linear relation between fluxes and forces,[30] hereby taking the form

$$\left.\begin{matrix} Q \\ J_s - c_sQ \end{matrix}\right) = \mathbb{L} \times \left.\begin{matrix} -\Delta p \\ -k_BT\Delta\log c_s \end{matrix}\right). \tag{14}$$

where $\mathbb{L}$ is the transport matrix. Importantly, as we discuss below and in Section 3.2.2, this matrix is symmetric according to Onsager's principle – due to microscopic time reversibility – and definite positive – due to the second principle of thermodynamics.

The question then amounts to characterizing the transport coefficients of this matrix. By identifying limiting regimes, Kedem and Kachalsky rewrote these transport equations in a more explicit form as[28,31–34]

$$Q = -\mathscr{L}_{hyd}(\Delta p - \sigma k_BT\Delta c_s), \tag{15}$$

$$J_s = -\mathscr{L}_D\omega_s\Delta c_s + c_s(1 - \sigma)Q, \tag{16}$$

where $\mathscr{L}_{hyd} = \kappa_{hyd}\mathscr{A}/(\eta L)$ is the solvent permeance through the membrane with κ_{hyd} the permeability (with units of a length squared), $\mathscr{A}$ the membrane area, η the fluid viscosity, and L the membrane thickness; $\mathscr{L}_D = \mathscr{A}D_s/L$ is the solute permeability with D_s the diffusion coefficient of the solute. Eqn (15) is often referred to as the Starling equation in the physiology literature,[35] see *e.g.* ref. 36 and 37. The osmotic pressure generated by the large scale molecules involved in the body (complex proteins such as albumin and more) is referred to as the oncotic pressure. These equations introduce two dimensionless (numerical) factors: σ is the so-called reflection or selectivity coefficient and ω_s is a solute "mobility" across the membrane – both of which we discuss in details below.

The Onsager symmetry relations for eqn (14) can be verified by exploring two limiting cases: (1) the situation where $\Delta p = 0$ yields osmotic flow only as $Q = \sigma\mathscr{L}_{hyd}c_s\Delta\mu$ (using $\Delta\mu = k_BT\Delta c_s/c_s$ in the dilute case); (2) and the situation where $\Delta\mu = 0$ yields $J_s - c_sQ = \sigma\mathscr{L}_{hyd}c_s\Delta p$. One obtains therefore $[Q/\Delta\mu]_{\Delta p=0} = [(J_s - c_sQ)/\Delta p]_{\Delta\mu=0}$ and the symmetry of the transport matrix is indeed verified.

The reflection coefficient and the solute mobility. The Kedem–Katchalsky equations introduce the reflection coefficient σ mentioned previously and first described by Staverman.[27] This coefficient *a priori* depends on the relative interactions of the membrane with the solute and solvent.[22,23,38] The Kedem–Katchalsky framework also introduces the permeability of the solute through the membrane *via* the combination $\mathscr{L}_D\omega_s$. A fully semi-permeable membrane corresponds to the case where $\sigma = 1$ and $\omega_s = 0$: the solute flux vanishes $J_s = 0$ and the pressure driving the fluid identifies with the van 't Hoff result $\Delta\Pi = k_BT\Delta c_s$. Reversely, a "transparent" membrane which is fully permeable to both solute and solvent correspond to $\sigma = 0$ and $\omega_s = 1$: no osmotic pressure is expressed and the solute flux reduces to Fick's law.

In the intermediate case, the membrane is partially permeable to the solute and we expect $0 < \sigma < 1$, see Fig. 4. As an example, in a pure Nafion membrane about 18 μm thick, the reflection coefficient between water and KCl salt was measured as $\sigma = 0.82$ (at concentration 0.25 mol L^{-1}).[39]

Interestingly, cases with negative reflection coefficient, $\sigma < 0$, were reported. This situation is often termed anomalous osmosis[40,41] and it corresponds to situations where the solute is more permeable than the solvent. We will discuss in Section 4 various examples where such a situation with reversed osmosis occurs.

The specificity of the membrane and its interaction with the solute molecules actually come into play into this reflection coefficient σ. A number of models have tried to rationalize the dependence of σ on the chemical and physical properties of the components. The first models took into account steric effects (similar to Fig. 4), where in fact the volume accessible to the solute inside the pore would differ (because of its typically larger size) than that accessible to the solvent.[42–44] The next

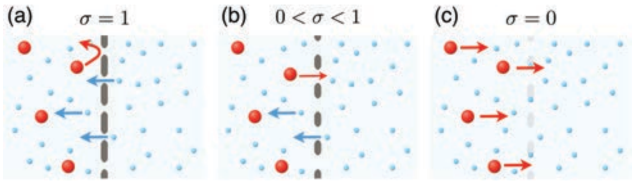


Fig. 4 Examples of reflection coefficient based on steric exclusion. In (a) the blue solvent only may traverse the pores; while in (b) the red solute particles may also traverse; their permeability through the pores is however small since the accessible volume in the pore for the red solute particles is smaller than that for the blue solvent. (c) The membrane is now fully permeable to all species, and therefore diffusion dominates and solute particles move towards the low concentration side.

generation of models sought to include as well hydrodynamic interactions, investigating how friction induced by the proximity of the solute to the pore walls would reduce permeability.[45,46] Anderson also studied interactions with the pore walls and adsorption of the solute in the pore.[38] Similarly, the "mobility" coefficient ω_s entering the transport equations will depend both on the solute–membrane interactions and transport parameters. A first, naive, estimate is to identify this coefficient with the partition coefficient of the solute between the membrane and the reservoirs at equilibrium, $K_s = c_s^m/c_s^{bulk}$, so that $\omega_s = K_s$.[31] But this estimate does not account for the complex transport processes occurring within the membrane. Interestingly, the non-dimensional coefficients ω_s and σ are expected to be linearly related,[31] as $1 - \sigma \propto \omega_s$, a result that we will recover below in a specific case.

Altogether, a complete determination of the reflection and mobility coefficients requires to implement a microscopic description of the membrane–fluid interactions. We will explore below and in Section 3 various situations highlighting how playing with interactions may lead to advanced osmotic transport behavior.

2.4 Mechanical views of osmosis: a tutorial perspective

Beyond the general formalism introduced above, it is interesting to get further fundamental insights into the microscopic mechanisms which underlie osmosis. In particular it is of interest to get some intuition on the mechanical force balance associated with the osmotic pressure. To do so, we will reduce the microscopic ingredients of osmosis to their minimal function and this description has merely a tutorial purpose. Still it is very enlightening in order to understand how the connection between "microscopic" parameters and thermodynamic forces builds up. Such mechanistic views of osmosis also allow to envision advanced osmotic phenomena, beyond the van 't Hoff perspective. Alternative approaches with similar illustrative objectives were proposed for one-dimensional single file channels, see ref. 47 and 48.

We pointed out above that the van 't Hoff law for the osmotic pressure does not involve the membrane properties *per se*, provided that it is semi-permeable. So it is tempting to replace the membrane by a crude equivalent, namely an energy barrier acting on the solute only, say $\mathcal{U}(x)$ (assuming for simplicity a unidimensional geometry) – see Fig. 5. This approach, which captures the minimal ingredients at play in osmosis, was first introduced by Manning[49] in the low concentration regime, and generalized more recently to explore the osmotic transport across perm-selective charged nanochannels[50] or in non-linear regimes at high solute concentrations.[51] One may note that such a potential barrier can also be physically achieved; for example, it may be generated from a nonuniform electric field acting on a polar solute in a nonpolar solvent,[52] or it can represent the nonequivalent interactions of solute and of solvent particles with a permeable membrane, *e.g.*, charge interactions.[50,53]

Let us first consider the ideal case where the barrier's maximum is high, *i.e.* $\mathcal{U}_{max} \gg k_BT$, so that the solute cannot cross the barrier: this is the perfectly semi-permeable case. In both reservoirs the solute is at equilibrium and the solute profile follows accordingly the Boltzmann relation

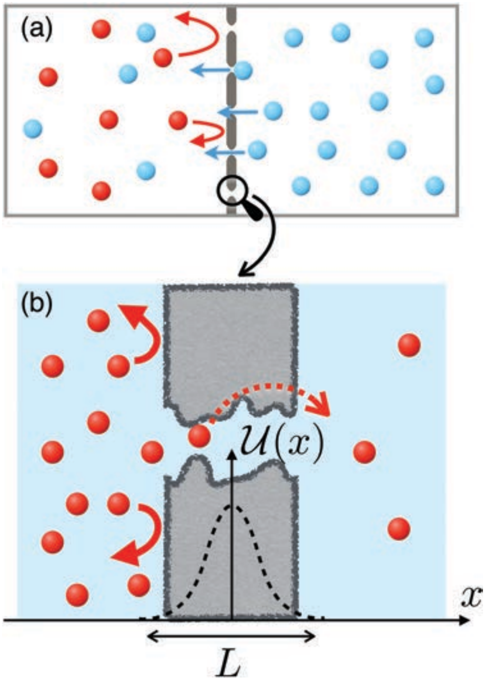


Fig. 5 The physical force driving osmosis is interaction of the solute with the membrane. (a) The membrane exerts a repulsive force (red arrows) on the solute particles (red) that creates a pressure gradient (or a void on the immediate left hand side of the membrane) that drives the flow (blue arrows of blue solvent particles). (b) Mechanical view of osmosis: the partially permeable membrane may be viewed as an energy barrier for the solute molecules (in red) that they have to overcome in order to traverse the membrane.

$$c_s^{(r)/(l)}(x) = c_s^{(r)/(l)} \times e^{\frac{-\mathcal{U}(x)}{k_BT}} \tag{17}$$

Now, a key remark is that the force on a fluid element of volume $d\tau$ (consisting of the solvent and solute mixture) will write

$$df(x) = c_s^{(r)/(l)}(x) \times (-\partial_x\mathcal{U}(x))d\tau. \tag{18}$$

with $d\tau = \mathcal{A}dx$; $\mathcal{A}$ is the membrane area. The total force per unit area acting on the fluid is accordingly integrated over x

$$\frac{F_T}{\mathcal{A}} = \int_0^{\infty} dx\, c_s^{(r)} e^{-\frac{\mathcal{U}(x)}{k_BT}} \times (-\partial_x\mathcal{U}(x)) + \int_{-\infty}^{0} dx\, c_s^{(l)} e^{-\frac{\mathcal{U}(x)}{k_BT}} \times (-\partial_x\mathcal{U}(x)) \tag{19}$$

(where we arbitrarily put $x = 0$ at the position of the maximum of the energy barrier), leading immediately to

$$\frac{F_T}{\mathcal{A}} = k_BT \times [c_s^{(r)} - c_s^{(l)}] \equiv \Delta\Pi \tag{20}$$

where we neglected terms behaving as $\exp[-\mathcal{U}_{max}/k_BT]$. Altogether this simple approach allows one to retrieve the van 't Hoff law. It highlights the mechanical origin of osmosis: as is transparent from the previous derivation, the osmotic pressure results from the fact that the reservoir containing more solute particle will generate a higher repelling force on the fluid than from the other reservoir: accordingly a fluid flow will be generated from the low to the high concentrations, hence diluting the more concentrated reservoir.

While the above approach is instrinsically at equilibrium, it can be easily generalized to a non-equilibrium situation by releasing the assumption of an (infinitely) high energy barrier: in this case the solute can cross the "membrane" between the two reservoirs at a finite rate, see Fig. 5, generating a solute flux. We further assume that the membrane is fully permeable to the solvent (no energy barrier acting on it), with a permeance $\mathcal{L}_{\text{hyd}}$ relating the fluid flux Q to the pressure drop Δp in the absence of a concentration difference: $Q = \mathcal{L}_{\text{hyd}}(-\Delta p)$.

The stationary dynamics of the system is described by the coupled set of equations for the solute diffusive dynamics – Smoluchowski equation – and fluid transport – Navier–Stokes equation. In the 1D geometry described above, the stationary solute concentration $c_s(x)$ obeys a Smoluchowski equation:

$$0 = \partial_t c_s = -\partial_x j_s = -\partial_x(-D_s\partial_x c_s + \lambda_s c_s(-\partial_x\mathcal{U}) + v_x c_s), \qquad (21)$$

where $j_s = J_s/\mathcal{A}$ is the solute flux per unit surface, D_s is the solute diffusion coefficient, $\lambda_s = D_s/k_BT$ the mobility and v_x the local fluid velocity. We will further assume a low Péclet number, $\text{Pe} = v_xL/D_s \ll 1$, such that the convective term of eqn (21) is negligible. This is valid for low permeability (nanoporous) membranes. The full derivation including the convective term was considered in ref. 49. Since the solute flux across the membrane J_s is constant in time and spatially uniform, eqn (21) is explicitly solved with respect to the concentration as:

$$c_s(x) = c_s^{(r)} - \Delta c_s e^{-\beta\mathcal{U}(x)} \frac{\int_x^{L/2} dx' \exp[+\beta\mathcal{U}(x')]}{\int_{-L/2}^{L/2} dx' \exp[+\beta\mathcal{U}(x')]}, \qquad (22)$$

where $\beta = 1/k_BT$. The solute concentration difference between the two volumes is $\Delta c_s = c_s^{(r)} - c_s^{(l)}$. For simplicity we assumed that the barrier has an extension L.

Now turning to the momentum conservation equation for the fluid (solvent + solute), the flow field $\boldsymbol{v}$ of the fluid obeys a Stokes equation (neglecting inertial terms)

$$0 = -\boldsymbol{\nabla} p + \eta\boldsymbol{\nabla}^2\boldsymbol{v} + \boldsymbol{f}_{\text{ext}}, \qquad (23)$$

where p is the fluid pressure and $\boldsymbol{f}_{\text{ext}}$ represents the total volume forces acting on the system, *e.g.* the forces acting on the solvent and on the solute, here

$$\boldsymbol{f}_{\text{ext}} = c_s(-\boldsymbol{\nabla}\mathcal{U}). \qquad (24)$$

The driving force inducing the solvent flow along the x axis is accordingly written in terms of an apparent pressure drop, $-\partial_x\mathcal{P} = -\partial_x p + c_s(x)(-\partial_x\mathcal{U})$. The membrane, *via* its potential $\mathcal{U}$, will therefore create an average force on the fluid, which writes per unit surface

$$-\Delta\mathcal{P} = -\Delta p + \int_{-L/2}^{L/2} dx c_s(-\partial_x\mathcal{U}) \equiv -\Delta p + \sigma\Delta\Pi, \qquad (25)$$

where Δ means the difference of a quantity between the two sides. The second term of eqn (25) can be interpreted as the osmotic contribution. Using the expression for the concentration profile given in eqn (22), one recovers the classical van 't Hoff law of the osmotic pressure, $\Delta\Pi = k_BT\Delta c_s$, and furthermore obtains an expression for the reflection coefficient σ as

$$\sigma = 1 - \frac{L}{\int_{-L/2}^{L/2} dx' \exp[+\beta\mathcal{U}(x')]}. \qquad (26)$$

The above result correctly recovers the case of a completely semi-permeable membrane (no solute flux across the membrane), *i.e.*, $\beta\mathcal{U} \gg 1$ and $\sigma \to 1$, yielding $-\Delta\mathcal{P} = -\Delta[p - \Pi]$. In the intermediate cases, although the membrane is permeable, a flow arises due to the solute concentration gradient even in the absence of an imposed pressure gradient. When the potential is repulsive and small $\mathcal{U} \sim k_BT$, then $0 < \sigma < 1$; the flow is in the direction of increasing concentration.

Integrating eqn (23) over the membrane area ($\mathcal{A}$) and thickness (L) allows the total flux Q to be expressed as

$$Q = -\mathcal{L}_{\text{hyd}}(\Delta p - \sigma k_BT\Delta c_s). \qquad (27)$$

Here the permeance $\mathcal{L}_{\text{hyd}}$ can be expressed in terms of the permeability, κ_{hyd}, as $\mathcal{L}_{\text{hyd}} = \dfrac{\mathcal{A}}{L}\dfrac{\kappa_{\text{hyd}}}{\eta}$. The permeability κ_{hyd} is defined formally in terms of the flow as $\kappa_{\text{hyd}}{}^{-1} = \langle -\nabla^2 v\rangle/\langle v\rangle$, where $\langle\cdot\rangle = \mathcal{V}^{-1}\iint dx d\mathcal{A}(\cdot)$ denotes an average over the pore volume, here $\mathcal{V} = \mathcal{A}L$. These parameters, κ_{hyd} and $\mathcal{L}_{\text{hyd}}$, take into account the detailed geometry of the pores in the membrane (pore cross section, length, *etc.*). Overall eqn (27) agrees with the Kedem–Kachalsky result in eqn (15). While this approach is derived here in the dilute regime for the solute, it can be generalized to arbitrary concentrations, see ref. 51.

As a last remark, it is interesting to note that the mechanistic approach highlights an underlying fundamental symmetry in the transport phenomenon. Indeed eqn (25) introduces the osmotic pressure as the driving force on the fluid: $\int_{-L/2}^{L/2} dx c_s(-\partial_x\mathcal{U}) = \sigma\Delta\Pi$. Now the Smoluchowski equation for the solute – integrated over the membrane thickness L, eqn (21) – contains the very same term and one may accordingly rewrite the solute flux as

$$J_s = -\frac{D_s}{L}\mathcal{A}\left[\Delta c_s - \sigma\frac{\Delta\Pi}{k_BT}\right] \qquad (28)$$

The solute flux is therefore intimately related to the osmotic pressure. As is transparent from this equation, the van 't Hoff osmotic pressure is fully expressed, *i.e.* $\Delta\Pi = k_BT\Delta c_s$, only when the solute flux vanishes $J_s = 0$ ($\sigma = 1$ and $\omega_s = 0$). Reversely for a fully permeable membrane $J_s = -\dfrac{D_s\mathcal{A}}{L}\Delta c_s$, and there is no osmotic pressure ($\sigma = 0$ and $\omega_s = 1$). Finally this equation can be rewritten as $J_s = -D_s\mathcal{A}\ (1 - \sigma)\Delta c_s/L$, so that the "mobility" coefficient ω_s is related here to the reflection coefficient as $\omega_s = 1 - \sigma$.

In this first part we have reviewed the basic understanding of osmosis, from the historical discovery of the phenomenon to the precise understanding of the effect in terms of thermodynamic forces. Although simplistic, the previous mechanical/kinetic approach provides a fruitful and complementary perspective on osmotic transport, which suggests a number of generalizations – that we will discuss below. It also reveals that the key aspect of osmosis is not really the membrane itself, but the existence of differential forces acting separately on the solvent and the solute. This is crucial to understand a number of phenomena related to osmosis that we discuss below.

3 Osmosis without a membrane

Situations where differential forces act on the solvent and the solute occur naturally, especially at interfaces: for example a charged surface does act specifically on dissolved ions, repelling co-ions and attracting counter-ions; or a neutral hard wall will repel polymers *via* excluded volume. As we now discuss in the following sections, these specific forces may be harnessed to induce interfacially-driven osmotic flows.

The geometry we will consider here involves a solid surface along which a solute gradient, or more generally a thermodynamic force – an electric field, a temperature gradient… – is established, as sketched in Fig. 6 and 7. Under an electric field, the net electric forces occurring within the diffuse interface close to the solid will push the fluid and generate a so-called electro-osmosis flow for the solvent. But as we will show below, a solute gradient ∇c_∞ parallel to the surface can also generate fluid motion whose amplitude is proportional to ∇c_∞:

$$v_{\mathrm{DO}} \propto \nabla c_\infty. \tag{29}$$

This latter phenomenon is usually coined as *diffusio-osmosis*. The phenomenon bears some fundamental analogy with Marangoni effects where a gradient of surface tension at an interface may drive fluid (or reversely particle) motion as $v_\mathrm{f} \propto \nabla\gamma_{\mathrm{LV}}$.[54] Now extending Marangoni flows to solid–fluid interfaces is definitely not obvious, but it was recognized by Derjaguin and collaborators[55,56] that the diffuse nature of the interface may allow the fluid to "slip" over the solid surface under a concentration gradient. Diffusio-osmosis is accordingly an interfacially driven flow, and takes its origin in the interfacial structure of the solute close to the solid surface, within the first few nanometers close to the surface.

3.1 From electro- to diffusio-osmosis

3.1.1 From electro-osmosis…. Let us start with the canonical example of electro-osmosis, *i.e.* the fluid flow close to a solid surface generated under an applied electric field. A solution containing ions will build up a so-called electric double layer (EDL) close to any charged surface: counter-ions are attracted by the surface charge, while co-ions are repelled. The surface charge, say Σ, is balanced in the fluid by a density of charge $\rho_\mathrm{e} = e(c_+ - c_-)$, defined as the difference between the density of positive and negative ions (assuming monovalent ions here for illustrative purposes). The resulting double layer is diffuse and extends over a finite width, see Fig. 6. The structure of the EDL was amply discussed in many textbooks and reviews, and we refer in particular to ref. 57–59 for further insights. As a rule of thumb, the extension of the EDL is typically given by the Debye screening length,[58,60] defined as

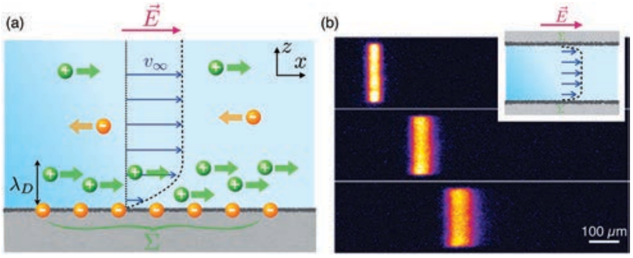


Fig. 6 Electro-osmosis. (a) In the presence of a surface charge Σ, an electrical double layer forms extending typically over a distance fixed by the Debye length λ_D. Under an applied electric field $\vec{E}$ parallel to the surface, a net electric force builds up due to the unbalanced charge within the electric double layer. It drives a solvent flow parallel to the surface, extending as a flat flow profile into the bulk. (b) Visualization of the electro-osmotic flow in a capillary, *via* the displacement of (neutral) tagged molecules. In contrast to the parabolic Poiseuille flow, the electro-osmotic flow profile takes the form of a plug flow, which barely disperses the dye. Reproduced from ref. 61 with permission from Springer Nature, copyright 2005.

$$\lambda_\mathrm{D} = \frac{1}{\sqrt{8\pi\ell_\mathrm{B} c_\mathrm{s}}} \tag{30}$$

where c_s is the (bulk) salt concentration in the bulk and $\ell_\mathrm{B} = e^2/4\pi\varepsilon k_\mathrm{B}T$ is the Bjerrum length (ε is the dielectric permittivity of water). Typically for water at room temperature, ℓ_B = 0.7 nm and the Debye length ranges between 30 nm for a salt concentration of 10^{-4} mol l^{-1} to 0.3 nm for a 1 mol l^{-1} salt concentration.

Within the EDL, there is a net charge density in the fluid, and whenever an external electric field is applied to the fluid (parallel to the surface), this will generate a net bulk force $\rho_\mathrm{e}E$. The Stokes equation for the fluid velocity writes accordingly in the direction x (parallel to the solid interface)

$$\eta\partial_{zz}v + \rho_\mathrm{e}E = 0 \tag{31}$$

where z is the direction perpendicular to the interface. The pressure-gradient term vanishes for the shear-flow considered here. Using the Poisson equation $\rho_\mathrm{e} = -\varepsilon\partial_{zz}V_\mathrm{e}$, relating the charge density to the electric potential V_e in the fluid, one can integrate twice eqn (31) to obtain the velocity profile

$$v(z) = -\frac{\varepsilon E}{\eta}[V_\mathrm{e}(z=0) - V_\mathrm{e}(z)] \tag{32}$$

where a no-slip boundary condition was assumed here. The electrical potential at the interface is usually identified as the zeta potential $V_\mathrm{e}(z = 0) = \zeta$. The electro-osmotic velocity is constant beyond the EDL and reaches its asymptotic value

$$v_\infty = \mu_{\mathrm{EO}}E \tag{33}$$

where $\mu_{\mathrm{EO}} = -\dfrac{\varepsilon\zeta}{\eta}$ is the electro-osmotic mobility. In the presence of hydrodynamic slippage on the surface, the electro-osmotic mobility is typically enhanced by a factor $1 + b/\lambda_\mathrm{D}$, where b is the slip length, see ref. 62–64 for more details. We finally note that the ζ-potential may be rewritten as a function of the electrical concentration ρ_e (by integrating twice eqn (31))

$$\zeta = -\frac{1}{\varepsilon}\int_0^\infty z\rho_e(z)\mathrm{d}z. \quad (34)$$

From a physical point of view, electro-osmosis may be seen as a force balance between the viscous friction force at the interface and the electrostatic driving force within the EDL. The velocity field is expected to establish over the Debye length λ_D and thus the fluid friction force is typically $\sim\eta v_\infty/\lambda_D$. Now the body electrical force within the EDL is simply $\Sigma \times E$ where Σ is the surface charge. From Gauss' electrostatic boundary condition, we have $\Sigma = -\varepsilon\partial_z V_e|_{z=0} \approx -\varepsilon\zeta/\lambda_D$. Altogether the force balance thus takes the form

$$\eta\frac{v_\infty}{\lambda_D} \approx \Sigma \times E \approx -\frac{\varepsilon\zeta}{\lambda_D} \times E \quad (35)$$

and this leads accordingly to the expression in eqn (33) for the electro-osmotic mobility. A simple extension of this argument highlights immediately the potential role of hydrodynamic slippage: with a slip length b, the viscous friction force will reduce to $\sim\eta v_\infty/(\lambda_D + b)$ while keeping the body force identical, so that the electro-osmotic velocity will be increased by a factor $1 + b/\lambda_D$. Altogether, the electro-osmotic flow thus takes its origin within the very few nanometers close to the boundary and can be therefore strongly affected by molecular details: hydrodynamic slippage,[62] nanoscale roughness,[65] contamination,[66] dielectric inhomogeneities,[67] *etc.* This makes the underlying physics of interfacial transport both complex and very rich.

3.1.2 . . .To diffusio-osmosis. . .. While electro-osmosis corresponds to interfacially driven fluid motion under an external electric field, diffusio-osmotic motion occurs under the gradient of a solute, $\partial_x c_\infty$, in the vicinity of a solid surface – see Fig. 7. Similarly to electro-osmosis, a key ingredient is the specific interaction of the solute with the surface, which occurs within a diffuse layer of finite thickness. Reflecting the discussion of osmosis across a model potential barrier in Section 2.4, the solute will be assumed to interact *via* an external potential $\mathcal{U}(z)$ with the solid surface. One noticeable difference to the previous membrane case though is that this potential now acts perpendicular to the solid surface and solute gradient (*i.e.* depending on z but not on x), see Fig. 7.

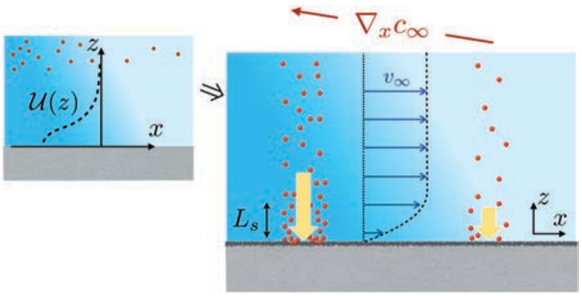


Fig. 7 Diffusio-osmosis. A gradient of solute imposed far from the surface induces a fluid flow. Here the solute is assumed to interact with the surface *via* an interaction potential $\mathcal{U}(z)$ (an adsorbing profile on the figure). The solute interaction with the surface induces a force on the fluid, here towards the surface, which is higher in the more concentrated area. This normal force converts into a parallel pressure drop which generates a fluid flow from high to low concentrations (and reversely for a repelling interaction).

Diffusio-osmosis with neutral solutes. We first consider the case of neutral solutes. The fluid velocity and solute density obey the coupled Stokes and Smoluchowski equations, which write in the stationary state as:

$$\begin{aligned} 0 &= -\nabla p + \eta\nabla^2 \boldsymbol{v} + (-\nabla\mathcal{U}), \\ 0 &= -\nabla\cdot[-D_s\nabla c_s + \lambda_s c_s(-\nabla\mathcal{U}) + \boldsymbol{v}c_s] \end{aligned} \quad (36)$$

At infinity, we assume a fixed gradient $\partial_x c_\infty$ along x for the solute concentration.

These coupled equations are strongly entangled. However in the limit of a thin interfacial layer – corresponding to a range for the potential $\mathcal{U}(z)$ which is small compared to the lateral variations of the solute gradient, one expects the concentration profile to relax quickly to a local equilibrium across the diffuse layer $c_s(x,z) \simeq c_\infty(x)\exp(-\mathcal{U}(z)/k_BT)$.

Turning now to the fluid transport equation, the Stokes equation projected along the z direction writes simply

$$0 = -\partial_z p + c_s(-\partial_z\mathcal{U}) \quad (37)$$

because the z component of fluid velocity is expected to be negligible for thin layers. We can integrate this pressure balance to obtain

$$p(x,z) - p_\infty = k_BTc_s(x,z) - k_BTc_\infty(x) \quad (38)$$

which can be seen as an osmotic equilibrium across the diffuse layer.[68] In simple terms, the existence of a specific solute-wall interaction allows the membrane to "express" the solute osmotic pressure $\Pi(x,z) = k_BTc_s(x,z)$ within the interfacial layer. However the effects of the latter disappear in the bulk ($z \to \infty$) and there is no bulk osmotic pressure gradient.

Now inserting the pressure from eqn (38) into the Stokes equation projected along x, see eqn (36), leads to

$$\eta\partial_z{}^2 v_x - \partial_x[p(x,z) - p_\infty] = 0. \quad (39)$$

Following the same steps as for the electro-osmosis, one obtains the fluid velocity along the x coordinate in the bulk fluid as

$$v_\infty = \mu_{DO} \times (-k_BT\nabla_x c_\infty) \quad (40)$$

with the diffusio-osmotic mobility μ_{DO} given by

$$\begin{aligned} \mu_{DO} &= \frac{1}{\eta}\int_0^\infty z\left(\frac{c_s(x,z)}{c_\infty} - 1\right)\mathrm{d}z \\ &= \frac{1}{\eta}\int_0^\infty z\left(\exp\left(\frac{-\mathcal{U}(z)}{k_BT}\right) - 1\right)\mathrm{d}z. \end{aligned} \quad (41)$$

This expression is similar to eqn (34) for the electro-osmotic mobility. The effect of hydrodynamic slippage on the surface can also be taken into account, along the same lines as in ref. 69 and 70 and leads to an enhancement factor of the diffusio-osmotic mobility scaling as $(1 + b/\lambda)$, where b is the slip length and λ is the typical width of the diffuse interface. The amplification effect is expected to be massive on superhydrophobic surfaces[70] and amplification by orders of magnitude are predicted. Interestingly for strongly hydrophobic surfaces where the liquid–vapor interface dominates, the diffusio-osmotic velocity takes the physically transparent expression $v_{\mathrm{DO}} = \frac{b_{\mathrm{eff}}}{\eta}\nabla\gamma_{\mathrm{LV}}$, where b_{eff} is the effective slip length on the superhydrophobic surface and γ_{LV} is the (solute concentration dependent) surface tension of the liquid–vapor interface.

Similarly as in electro-osmosis, diffusio-osmosis can be interpreted in terms of a simple force balance within the diffuse layer. A first integration of eqn (39) indeed shows that diffusio-osmotic flow results from the balance between the viscous stress on the surface and an osmotic pressure gradient integrated over the diffuse layer:

$$0 = \eta\partial_x v_x|_{\mathrm{wall}} + \int_0^\infty \mathrm{d}z\partial_x[\Pi(x,z) - \Pi_\infty(x)] \tag{42}$$

Simple estimates of the various terms lead to a more qualitative version of this force balance as

$$\eta\frac{v_\infty}{\lambda} \sim \pm\lambda \times (-k_{\mathrm{B}}T\nabla_x c_\infty) \tag{43}$$

where λ is defined here as the range of the potential $\mathscr{U}$, and the $\pm$ sign depends on whether the solute is attracted or depleted by the surface. This leads to $v_\infty \sim \pm\frac{\lambda^2}{\eta} \times (-k_{\mathrm{B}}T\nabla_x c_\infty)$ in full agreement with eqn (40) and (41).

Diffusio-osmosis is definitely an osmotic flow, *e.g.* a flow driven by an osmotic pressure gradient located within the diffuse layer. However the direction of the diffusio-osmotic flow can be along or against the gradient of the solute, in strong contrast to bare osmosis which induces a flow towards the highest solute concentration. That is highlighted in the expression of the diffusio-osmotic mobility, eqn (41), which can be positive or negative depending on the attractive or repulsive nature of the interaction potential $\mathscr{U}(z)$. As a rule of thumb, the sign of the mobility will be dominantly determined by the adsorption $\Gamma = \int_0^\infty \mathrm{d}z(c(x,z)/c_\infty(x) - 1)$. If there is a surface excess ($\Gamma > 0$ or $\mathscr{U}(z) < 0$), the solvent flow goes towards the low concentrated area ($\mu_{\mathrm{DO}} > 0$). That may appear as surprising because it amounts to concentrating even more the already concentrated solution; we shall discuss this apparent paradox in Section 3.2.2. Reversely a surface depletion resulting from a repulsion of the solute from the wall ($\Gamma < 0$ or $\mathscr{U}(z) > 0$) reverses the direction of the solvent flow towards the high concentrated zone ($\mu_{\mathrm{DO}} < 0$). An interesting limiting case for this behavior is exemplified by a solute interacting with the wall *via* steric effect, *i.e.* hard-core excluded volumes. For a solute particle with radius R, the mobility in eqn (39) reduces to

$$\mu_{\mathrm{DO}}^{\mathrm{steric}} = -\frac{R^2}{2\eta}. \tag{44}$$

This behavior was measured in particular in ref. 25 for the diffusio-osmotic flow under a neutral polymer concentration gradient, see Fig. 8 for an illustration. A final remark is that this simple rule for the correlation between adsorption and the sign of diffusio-osmosis is not exact and may fail for more complex interactions between the solute and the wall, for instance with an oscillatory spatial dependence of the concentration profile due to layering. The sign of μ_{DO} may then be expected to differ from the sign of the adsorption Γ. In this case, no obvious conclusion can be made for the direction of the diffusio-osmotic velocity and a full calculation has to be made, see for example ref. 26.

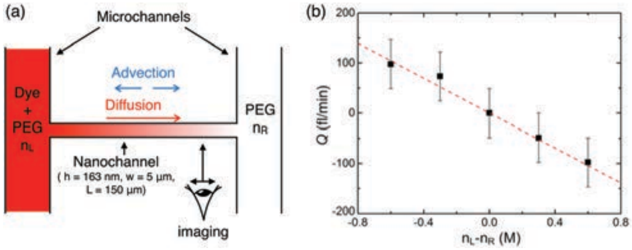


Fig. 8 Experimental evidence for diffusio-osmosis. (a) A gradient of polyethylene glycol polymer PEG is maintained along a nanochannel thanks to lateral microchannels acting as reservoirs. The nanochannel is 160 nm in thickness and is fully permeable to PEG. Under a PEG concentration gradient, a diffusio-osmotic flow arises: water moves towards higher concentrations of PEG. The flow rate Q is measured *via* the concentration profile of a dye. (b) Measured diffusio-osmotic flux Q as a function of the PEG concentration difference, showing a velocity proportional to the PEG concentration difference. This behavior and the sign of the effect are consistent with a steric exclusion of PEG on the surfaces, as predicted in eqn (44). (a) and (b) are reproduced and adapted from ref. 25 with permission from the American Physical Society (APS), copyright 2014.

Diffusio-osmosis with electrolytes. We now discuss specifically the case of diffusio-osmosis under salinity gradients. Here, as for electro-osmosis, the diffuse layer corresponds to the electric double layer created close to a charged surface, see Fig. 9. The derivation follows similar steps as above, from eqn (36)–(41), except that one has to take into account the spatial distribution

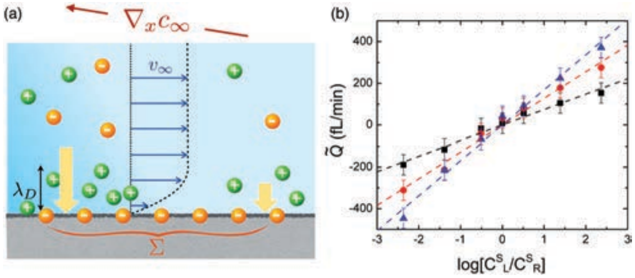


Fig. 9 Diffusio-osmosis with charged electrolytes. (a) Geometry; an electrolyte concentration difference is imposed far from the charged interface. The electrical interaction with the surface induces an (attractive) electrostatic force – sketched with arrows – on the fluid which is larger where the salt concentration is larger, hereby inducing a net flow towards the low salinity region. (b) Measurement of the diffusio-osmotic flux as a function of the difference of the logarithm of the salt concentration between two reservoirs, in a similar way as in Fig. 8. Note the reversal of the sign as compared to Fig. 8b. Reproduced from ref. 25 with permission from the APS, copyright 2014.

of both the counter- and co-ions in the EDL that follow a Poisson–Boltzmann distribution, see ref. 71. In this case the diffusio-osmotic velocity is shown to take the form

$$v_\infty = D_{\rm DO}(-\nabla \log c_\infty) \qquad (45)$$

where we introduced a mobility $D_{\rm DO}$ which has now the units of a diffusion coefficient. It takes the expression[54]

$$D_{\rm DO} = \frac{k_{\rm B}T}{2\pi\eta\ell_{\rm B}} \times \log\left(\cosh^2\frac{\Phi_0}{4}\right) \qquad (46)$$

where $\Phi_0 = eV_0/k_{\rm B}T$ is the dimensionless surface potential V_0 (usually identified with the zeta potential). Note that for an electrolyte with unequal diffusion coefficients for the anions and cations ($D_+ \neq D_-$), a diffusion electric field builds up under the gradient of the salt concentration (if no current exists in the bulk). This takes the form $E_{\rm diff} = \frac{k_{\rm B}T}{e}\delta\nabla\log c_\infty$ with $\delta = (D_+ - D_-)/(D_+ + D_-)$ and adds a supplementary electro-osmotic contribution to the diffusio-osmotic velocity as $v_{\rm diff} = -\frac{\varepsilon\zeta}{\eta} \times E_{\rm diff}$. Accordingly this leads to a supplementary contribution to the mobility as:

$$D_{\rm DO}^{\rm diff} = -\frac{\varepsilon\zeta}{\eta} \times \frac{k_{\rm B}T}{e}\delta \qquad (47)$$

An important remark is that for electrolytes the velocity is proportional to the gradient of the logarithm of salt concentration, in contrast to solutes where it is basically linear in the gradient, see eqn (40). We will refer to this dependence as "log-sensing" by analogy to behaviors occurring for the chemotaxis of biological entities (*e.g.* bacteria). Such a dependence may be understood on the basis of the simple scaling argument based on the force balance above, see eqn (43). Indeed the thickness of the diffuse layer is now given by the Debye length, and $v_\infty \simeq \frac{{\lambda_{\rm D}}^2}{\eta} \times (-k_{\rm B}T\nabla_x c_\infty)$. Since the Debye length depends on the salt concentration as ${\lambda_{\rm D}}^2 = (8\pi l_{\rm B} c_\infty)^{-1}$, one obtains:

$$v_\infty \approx -\frac{k_{\rm B}T}{8\pi\eta l_{\rm B}}\nabla_x \log c_\infty. \qquad (48)$$

which is qualitatively similar to the exact results in eqn (46) and predicts log-sensing for diffusio-osmosis with electrolytes.

This behavior is confirmed by experimental investigations of water flows under salinity gradients in nanofluidic circuits,[25] see Fig. 9. Diffusio-osmotic flow of water under salinity gradients was also evidenced across carbon nanotube membranes,[72] confirming further that diffusio-osmosis was acting against bare osmosis. In an alternative configuration, diffusio-osmosis was also shown to induce very large ionic currents under salinity gradients.[73–75] We will come back to such cross effects associated with diffusio-osmosis in Section 3.2.1, as well as in the section dedicated to blue energy harvesting, Section 6.3. In a very different field, diffusio-osmotic flows were also shown to strongly impact and shape the reactive fluid flows occurring in the solid Earth.[76] Log-sensing has also many counter-intuitive consequences and a variety of applications,[77,78] which we will discuss more specifically in the context of diffusio-phoresis in Section 5.

Solvo-osmosis and diffusio-osmosis with mixtures. Up to now we considered merely dilute solute solutions, but all previous results can be generalized to mixtures of liquids with any molar fraction of its constituents. The key ingredient remains that the two constituents interact differently with the solid substrate. As shown in ref. 51, the diffusio-osmotic velocity now takes the expression

$$v_\infty = \mu_{\rm DO}(-\nabla_x \Pi[X_\infty(x)]), \qquad (49)$$

where Π is the generalized osmotic pressure defined in eqn (9), calculated for the molar fraction X_∞, hence generalizing the expression in eqn (40). The diffusio-osmotic mobility $\mu_{\rm DO}$ is still given by the initial expression $\mu_{\rm DO} = \frac{1}{\eta}\int_0^\infty {\rm d}z' z' \left(\frac{c_{\rm s}(x,z')}{c_\infty(x)} - 1\right)$. However, for a solute-substrate interaction potential $\mathscr{U}$, the concentration profile $c_{\rm s}(x,z)$ is now implicitly related to the value in the bulk $c_\infty(x)$ *via* the local equilibrium condition $\mu[c_{\rm s}(x,z)] + \mathscr{U}(z) \simeq \mu[c_\infty(x)]$.

Diffusio-osmosis with ethanol–water mixtures was investigated recently in ref. 26. But the majority of existing experimental investigations merely explored the reverse configuration of phoretic transport of particles under gradients of liquid composition, denoted as "solvo-phoresis".[21,79] Interestingly in ref. 21, the phoretic transport of colloidal (polystyrene) particles in ethanol–water mixtures resulted in a "log-sensing" behavior of the particle diffusio-phoretic velocity, obeying $V = D_{\rm SP}\nabla\log X$, with here X the ethanol mole fraction.

3.1.3 … And electro-chemical equivalence.

In the case of electrolytes, the two previous transport phenomena, electro- and diffusio- osmosis, are fundamentally intertwined. Indeed, from the thermodynamic point of view, the chemical potential and the electric potential contributions merge into the electrochemical potential: $\mu_{\rm el} = \mu + qV$ (with q the ion charge and V the electric potential). There is accordingly a deep analogy when driving the system under gradients of chemical potential (diffusio-osmosis) or driving under gradients of electric potential (electro-osmosis). An illuminating discussion on this point and the corresponding force balance is provided by T. Squires in ref. 59 and 80 and we reproduce the essentials of the argumentation here.

Let us consider in full generality that a gradient of the electrochemical potential is applied in the bulk far from the boundary, $\nabla\mu_{\rm el,i}^{\rm B}$ (along the direction of the solid surface, say x); the index i runs over the various ion species in the solution. As we discussed above for both electro- and diffusio-osmosis, this will generate net thermodynamic forces on individual ion specie i, which may be written as $f_i(x,z) = -\nabla\mu_{{\rm el},i}(x,z)$. A key remark is that the electrochemical potential is approximately constant across the EDL, *i.e.* $\mu_{{\rm el},i}(x,z) \simeq \mu_{{\rm el},i}^{\rm B}(x)$, so that the individual force rewrites $f_i(z) \simeq -\nabla\mu_{{\rm el},i}^{\rm B}$. The interfacial motion results from the forces in excess to the bulk, so that the corresponding total force acting on the fluid rewrites

$$f_{\rm T} = \sum_{i=1,n} \Delta c_i \times f_i(z) = \sum_{i=1,n} \Delta c_i \times \left(-\nabla\mu_{{\rm el},i}^{\rm B}\right) \qquad (50)$$

where the sum runs over n ion species and $\Delta c_i = c_i(x,z) - c_i^{\rm B}(x)$ is the excess ion concentration in the boundary layer, as compared to the bulk. This driving force will generate a flow according to the Stokes equation $\eta\partial_z{}^2 v_x + f_T = 0$ and following the same steps as above, one obtains the far field slip velocity as

$$v_\infty = -\sum_{i=1,n}\frac{1}{\eta}\nabla\mu^{\mathrm{B}}_{\mathrm{el},i}\int_0^\infty \mathrm{d}z\, z\Delta c_i(z) \equiv \sum_i M_i\left(-\nabla\mu^{\mathrm{B}}_{\mathrm{el},i}\right), \qquad (51)$$

where the mobility M_i takes the expression $M_i = \frac{1}{\eta}\int_0^\infty \mathrm{d}z\, z\Delta c_i(z)$. For symmetric and monovalent ions, these mobilities can be exactly calculated using Poisson–Boltzmann framework, leading to

$$M_\pm = \frac{\varepsilon}{e\eta}\left[\mp\frac{\zeta}{2} + \frac{k_\mathrm{B}T}{e}\times\log\left(\cosh^2\frac{\Phi_0}{4}\right)\right] \qquad (52)$$

with $\Phi_0 = eV_0/k_\mathrm{B}T$ the dimensionless surface potential and here $\zeta \equiv V_0$ the zeta potential.

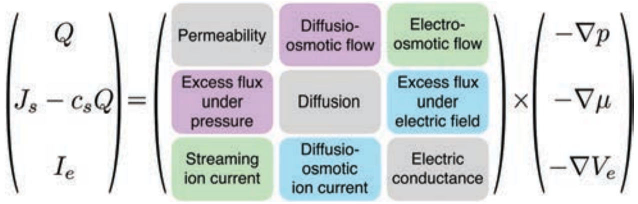
Under a constant electric field $\nabla\mu^{\mathrm{B}}_\pm = \mp eE_0$ and the electro-osmotic mobility is predicted as $\mu_\mathrm{EO} = \frac{1}{e}(M_+ - M_-)$, in full agreement with the previous result in eqn (33) and (34). Under an imposed ionic strength gradient in the bulk, then $\nabla\mu^{\mathrm{B}}_\pm = k_\mathrm{B}T\nabla\log c_\pm$ are identical and $M_\mathrm{DO} = M_+ + M_- \equiv D_\mathrm{DO}$, again in full agreement with the previous result in eqn (46).

3.2 Transport matrix and symmetry considerations

3.2.1 Transport matrix and cross fluxes. As introduced in Section 2.3, the framework of irreversible processes allows one to write a linear relation between thermodynamic forces and fluxes.[30] Adding the electric forces to the set of forces, one may generalize the results in eqn (14) to obtain linear transport equations now relating the solvent flux Q, excess solute flux $J_\mathrm{s} - c_\mathrm{s}Q$ and electric current I_e to the pressure gradient $-\nabla p$, chemical potential gradient $-\nabla\mu$ and the applied electric field $-\nabla V_\mathrm{e}$, and summarized as

$$\begin{pmatrix} Q \\ J_\mathrm{s} - c_\mathrm{s}Q \\ I_\mathrm{e} \end{pmatrix} = \mathbb{L}\times\begin{pmatrix} -\nabla p \\ -\nabla\mu \\ -\nabla V_\mathrm{e} \end{pmatrix}, \qquad (53)$$

Due to Onsager principle, this matrix is symmetric and positive definite.[30] Each term of this matrix corresponds to a specific transport phenomenon. Diagonal terms are associated respectively with permeability (characterizing solvent flux under a pressure drop), diffusion (characterizing solute flux under an applied solute gradient) and electrical conductance (characterizing ionic current under an applied electric field). The off-diagonal terms correspond to cross effects. We detail below the cross effects that are all recapitulated in Fig. 10.



$$\begin{pmatrix} Q \\ J_s - c_sQ \\ I_e \end{pmatrix} = \begin{pmatrix} \text{Permeability} & \text{Diffusio-osmotic flow} & \text{Electro-osmotic flow} \\ \text{Excess flux under pressure} & \text{Diffusion} & \text{Excess flux under electric field} \\ \text{Streaming ion current} & \text{Diffusio-osmotic ion current} & \text{Electric conductance} \end{pmatrix} \times \begin{pmatrix} -\nabla p \\ -\nabla\mu \\ -\nabla V_e \end{pmatrix}$$

Fig. 10 Transport matrix. Explicit transport matrix $\mathbb{L}$ as presented in eqn (53), with colors indicating symmetric terms.

In the first row of the matrix, electro-osmosis and diffusio-osmosis – explored so far – correspond to the terms relating the solvent flux Q to a chemical gradient $-\nabla\mu$ and an electric field $-\nabla V_\mathrm{e}$ respectively. A key consequence of the symmetry of the matrix is that the same mobilities characterize symmetric transport phenomena. For example consider the first column of the matrix $\mathbb{L}$, one finds that the electro-osmotic mobility and diffusio-osmotic mobility also describe respectively the electric current and excess solute flux generated under a pressure drop, as

$$\begin{aligned} I_e &= \mathscr{A}\mu_\mathrm{EO}\times(-\nabla p) \\ J_\mathrm{s} - c_\mathrm{s}Q &= \mathscr{A}\mu_\mathrm{DO}\times(-\nabla p) \end{aligned} \qquad (54)$$

where $\mathscr{A}$ is the channel cross section. The first corresponds to the so-called streaming current and takes its origin in the motion of mobile ions in the EDL which are carried by the pressure-driven flow; the pressure-driven excess solute flux has a similar physical origin.

Streaming currents are commonly measured in experiments,[60,62] even down to single carbon and boron-nitride nanotubes.[73] To our knowledge, no experimental measurement of pressure-driven excess solute flux has been performed up to now. However this is not the case in molecular dynamics simulations where it is far easier to measure the diffusio-osmotic mobility *via* the pressure-driven excess flux[26,69] – see details in Section 3.5.

Now, the transport matrix suggests that an electric current can be generated under an osmotic gradient, which we term here the diffusio-osmotic ion current, following

$$I_\mathrm{DO} = K_\mathrm{osm}\times(-\nabla\log c_\infty). \qquad (55)$$

Let us consider a channel in the form of a slit of width w and height h (with $w \gg h$ to simplify). Using Poisson–Boltzmann to describe the EDL, one can calculate the corresponding osmotic electric current[64,73,81] and the mobility takes the form

$$K_\mathrm{osm} = \alpha\times(-\Sigma)\frac{k_\mathrm{B}T}{2\pi\eta\ell_\mathrm{B}}\left(1 - \frac{\sinh^{-1}\chi}{\chi}\right) \qquad (56)$$

where $\alpha \simeq 2w$ is the perimeter of the channel cross section. In this expression we introduced $\chi = \sinh\frac{|\Phi_0|}{2}$ with $\Phi_0 = eV_0/k_\mathrm{B}T$ the dimensionless surface potential V_0. In the Poisson–Boltzmann framework χ is related to the surface charge Σ according to $\chi = 2\pi\ell_\mathrm{B}\lambda_\mathrm{D}|\Sigma|/e$ with λ_D the Debye length, so that $\chi \propto |\Sigma|$. This formula can be extended to take slippage on the surface into account, as well as mobile surface charges.[64] More precisely the diffusio-osmotic ion current takes its origin in the motion of ions in the EDL which are carried by the diffusio-osmotic flow. As a simple estimate we may write that $I_\mathrm{DO} \approx (-\Sigma)\times v_\mathrm{DO}$, where v_DO is the diffusio-osmotic velocity: using the expression eqn (40) for v_DO, one indeed recovers eqn (56). However the prediction of eqn (56) reports a more complex dependence, since the linear dependence in Σ is only valid for large enough Σ, while for low Σ, one finds that $K_\mathrm{osm} \propto \Sigma^3$, *i.e.* vanishingly small. Such osmotically driven currents have been measured experimentally in various systems, nanochannels, single nanotubes, single nanopores – see ref. 73, 74, 82 and 83 – to cite a few. This effect finds important applications in the context of blue energy harvesting,[75] that we will explore in detail in Section 6.3.

3.2.2 **Entropy production with diffusio-osmosis.** We pointed out above that the sign of the diffusio-osmostic mobility, $\mu_{\rm DO}$, can be either positive or negative, so that the corresponding flux can be along or against the concentration gradient. A negative $\mu_{\rm DO}$ may appear at first sight striking since the direction of the solvent flow corresponds to that of an increase in salt concentration, thus leading to an apparent violation of the second principle. This is however not the case, as it can be verified from a calculation taking into account all relevant fluxes. To highlight this situation, let us consider a membrane separating two reservoirs with fixed volumes; the concentration on the left/right reservoir is $c_{\rm s}(t) = c_0 \mp \dfrac{\Delta c_{\rm s}(t)}{2}$. The pore size is assumed here to be larger than the solute diameter so that the membrane is permeable to the solute and there is no bare osmotic pressure. A salinity gradient however generates a diffusio-osmotic flow on the pore surface. Based on the transport matrix formulation, eqn (14), one may write the solvent and (excess) solute fluxes as a function of the solute concentration and pressure gradients according to:

$$\begin{aligned} Q &= \frac{\mathscr{A}_{\rm p}}{L}[\kappa(-\Delta p) + \mu_{\rm DO}(-k_{\rm B}T\Delta c_{\rm s})] \\ J_{\rm s} - c_0 Q &= \frac{\mathscr{A}_{\rm p}}{L}[\mu_{\rm DO}c_0(-\Delta p) + \lambda_{\rm s}(-k_{\rm B}T\Delta c_{\rm s})] \end{aligned} \tag{57}$$

with $\mathscr{A}_{\rm p}$ the total (open) pore area of the membrane, L its thickness and $\lambda_{\rm s} = D_{\rm s}/k_{\rm B}T$ the diffusive mobility of the solute across the membrane, defined in terms of the solute diffusion coefficient; κ is defined in terms of the permeance as $\mathscr{L}_{\rm hyd} \equiv \kappa\mathscr{A}_{\rm p}/L$ (note that $\kappa = \kappa_{\rm hyd}/\eta$ where $\kappa_{\rm hyd}$ is the permeability introduced above). The second principle imposes that the transport matrix in eqn (14) and (57) should be definite positive. Accordingly, the determinant $\det(\mathbb{L}) \propto \kappa\lambda_{\rm s} - \mu_{\rm DO}{}^2 c_0$ must be strictly positive.

On the other hand, since the volume is fixed, the flux vanishes, $Q = 0$, and the solute flux writes

$$J_{\rm s} = \frac{\mathscr{A}_{\rm p}}{L}\frac{1}{\kappa}\left[\lambda_{\rm s}\kappa - \mu_{\rm DO}{}^2 c_0\right](-k_{\rm B}T\Delta c_{\rm s}) \tag{58}$$

we find that the term in brackets is proportional to the determinant $\det(\mathbb{L})$, and therefore is constrained by the second principle to be positive. Accordingly, whatever the sign of the diffusio-osmotic mobility $\mu_{\rm DO}$ and the corresponding diffusio-osmotic solvent flux, the total solute flux will go down the solute gradient, as expected from the second principle.

3.3 The peculiarity of diffusio-osmosis across an orifice

In the previous sections, we implicitly considered (diffusio-osmotic) transport across long channels, so that fluid flow is translationally invariant along the channel's length. However transport across thin membrane pores[84–86] raises the question of the specificity of these geometries in which the channel length L may decrease down to molecular lengths, in particular with the advent of 2D materials such as graphene, h-BN and $\rm MoS_2$ as membranes for fluidic transport.[74,87,88] For example, recent measurements across nanopores in $\rm MoS_2$ membranes reported huge diffusio-osmotic ion currents under salinity gradients.[74] In another experiment, gradients of salts were shown to strongly increase the capture rate of DNA molecules across solid-state nanopores.[84]

For long channels the driving force for fluid transport, *e.g.* the gradient of the chemical potential, is expected to scale as its inverse length, $\nabla\mu = \Delta\mu/L$. This would suggest that the driving force diverges as $1/L$ in the limit of nanopores where $L \to 0$. However entrance effects level off this diverging behavior to a value typically fixed by the lateral size of the pore, say a its radius – see Fig. 11. As a rule of thumb, one may expect that $\nabla\mu \approx \Delta\mu/a$ (see for example ref. 89 for the conductance of ion channels). However the flow in and out of the pore is expected to be strongly disturbed, as shown for example for electro-osmosis across nanopores in thin membranes.[90–92] Similar effects are accordingly expected to apply to diffusio-osmotic transport.

The diffusio-osmotic flow across a nanopore with vanishing thickness was recently calculated analytically by Rankin *et al.*[93] The calculation is best performed in oblate-spheroidal coordinates, in line with a similar calculation for electro-osmosis in ref. 90. The averaged diffusio-osmotic velocity $v_{\rm DO}$ across the pore, which is defined in terms of the diffusio-osmotic flux $Q = \pi a^2 v_{\rm DO}$, is proportional to $\Delta c_{\rm s}$ the difference (and not the gradient as in eqn (40)) of the solute concentration between the two sides of the membrane: $v_{\rm DO} = \mu_{\rm DO}^{\rm pore}(-k_{\rm B}T\Delta c_{\rm s})$. The general expression for the mobility derived in ref. 93 takes the form

$$\mu_{\rm DO}^{\rm pore} = -\frac{2a}{\pi^2\eta}\int_0^1 {\rm d}\xi\,\xi^2\int_0^\infty {\rm d}\nu\,\frac{{\rm e}^{-\mathscr{U}/k_{\rm B}T} - 1}{1+\nu^2} \tag{59}$$

where (ν,ξ) are the oblate spheroidal coordinates (iso-ν and iso-ξ curves are respectively oblate spheroids and hyperboloids of revolution). This expression involves a complex spatial

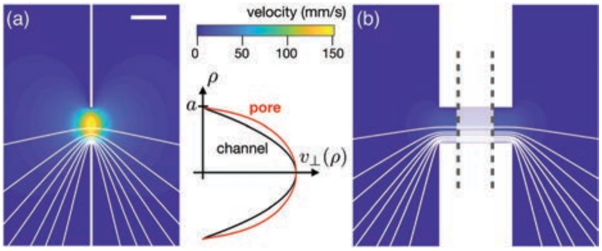


Fig. 11 Peculiarity of pressure flow across an orifice. (a) Simulated flow velocity and streamlines across a pore with, say, radius a = 10 nm and thickness $L = a/10$ under a pressure drop Δp = 1 bar. The scale bar is 10 nm. The streamlines are spaced equally in magnitude at the center of the pore. (b) Simulated flow velocity and streamlines across a channel with same radius a and thickness $L = 10a$ under the same pressure drop. The scales (velocity and geometry) are the same as for (a). For readability the whole channel length is not plotted. (center) Normalized velocity profile (perpendicular to the membrane) at the center of the membrane, comparing the channel and pore cases.

average of the Boltzmann weight $\mathrm{e}^{-\mathscr{U}/k_\mathrm{B}T} - 1$, which should be compared to the corresponding simple expression in eqn (41) for the planar case.

The above result can be simplified for certain functional forms of the potential $\mathscr{U}$. For example, assuming that the interaction potential $\mathscr{U}$ depends only on variable ξ allows the mobility to be rewritten in terms of the two-dimensional interaction within the pore only as $\mu_\mathrm{DO}^\mathrm{pore} = \frac{1}{\pi a^2 \eta}\int_0^a \mathrm{d}\rho\rho\sqrt{a^2 - \rho^2}\left[\mathrm{e}^{-\mathscr{U}(\rho)/k_\mathrm{B}T} - 1\right]$ with ρ the axi-symmetric distance to the center of the pore.[93] This expression for the mobility can be recovered thanks to the symmetry of the transport matrix eqn (53). Indeed $\mu_\mathrm{DO}^\mathrm{pore}$ can also be calculated in terms of the excess solute flux under a pressure driven flow: in this case the velocity profile was shown to be semicircular (and not parabolic),[94–96] see Fig. 11, and the excess solute flux conveyed by the circular flow reduces to the above expression.

The complexity associated with diffusio-osmosis across an orifice is also highlighted by the predicted dependence of the mobility on the pore size a and the range λ of the interaction $\mathscr{U}$. Let us focus the discussion for the thin diffuse layer case, where $\lambda \ll a$ (we refer to ref. 93 for a full discussion). As a reference, the diffusio-osmotic mobility across a long channel with length L was shown previously to scale as $\mu_\mathrm{DO} \sim \frac{\lambda^2}{\eta L}$, see *e.g.* eqn (43). However, for an orifice in a thin membrane, Rankin *et al.* showed on the basis of eqn (59) that the mobility exhibits a variety of non-trivial scalings, with $\mu_\mathrm{DO}^\mathrm{pore} \sim \lambda^\gamma a^{1-\gamma}$, and an exponent γ that depends on the details of the interaction potential $\mathscr{U}$. For example, for the potential discussed above, which assumes a dependence as $\mathscr{U}(\xi)$, one finds $\gamma = 3/2$; but for a potential depending on the distance to the membrane or to the edge of the pore, then $\gamma = 2$.[93] In the latter case, $\gamma = 2$, the diffusio-osmotic mobility scales as $\mu_\mathrm{DO}^\mathrm{pore} \sim \lambda^2/a$, which corresponds to the long channel result with the length L replaced by a. But for other values of the exponent γ this simple rule of thumb does not apply, making the diffusio-osmotic transport across the orifice quite peculiar.

As a last comment, it is possible to extend qualitatively these results to electrolyte solutions, by assuming that the potential range λ identifies with the Debye length. This suggests an anomalous salinity dependence for the diffusio-osmotic mobility $D_\mathrm{DO} = v_\mathrm{DO}/[-k_\mathrm{B}T\Delta \log c_\mathrm{s}] \propto {c_\mathrm{s}}^{1-\gamma/2}$, in contrast to long channels where $D_\mathrm{DO} \propto c_\mathrm{s}^0$. The nanopore geometry may thus depart from the log-sensing behavior of diffusio-osmotic transport under salinity gradients. These results remain however to be fully assessed experimentally.

3.4 Alternative interfacial transport: thermo-osmosis

Extending on electro- and diffusio-osmosis, thermo-osmosis corresponds to fluid motion under gradients of temperature; see Fig. 12a. Such effects were reported as early as in the 1900s.[97,98] Thermo-osmosis was first rationalized in terms of thermodynamic forces by Derjaguin *et al.*[54,99,100] Similarly as for diffusio-osmosis in eqn (41) and electro-osmosis in eqn (34), the net velocity generated far from the surface is predicted as[54]

$$v_\infty = \left(\frac{-2}{\eta}\int_0^\infty z\delta h(z)\mathrm{d}z\right)\nabla \log T_\infty \qquad (60)$$

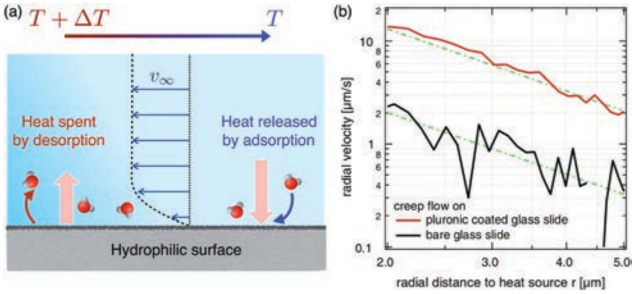


Fig. 12 Thermo-osmosis near an interface. (a) Geometry; a temperature gradient is imposed far from a hydrophilic surface. A thermal flux from the hot to the cold region is therefore installed. The interaction of water with the surface induces a force (light red arrows) that varies along the surface due to the thermal gradient, inducing a net flow. (b) Thermo-osmotic flow measured on two different surfaces (going towards the higher temperatures) as a function of the distance to the heat source. Reproduced from ref. 101 with permission from the APS, copyright 2016.

where T_∞ is the temperature far from the surface and $\delta h(z)$ is the excess specific enthalpy in the interfacial layer as compared to the bulk liquid. If the solid surface is *e.g.* hydrophilic, then $\delta h(z) \lesssim 0$ and the flow of water is directed toward higher temperatures, see Fig. 12 and ref. 101 and 102. An interpretation of thermo-osmosis (and -phoresis) in terms of interfacial surface tension modification, and therefore Marangoni-like flow generation, has also been suggested[103] and formalized.[104,105] The transport of fluids or particles under thermal forces led to strong debates between the interfacial approach discussed above and an ''energetic'' approach,[106–108] which attempts to write the net driving force acting on a particle as the gradient of a thermodynamic quantity.[106] The resulting Soret coefficient – defined as the ratio between the thermophoretic mobility and particle diffusion coefficient – highlights a different dependence on the particle size as compared to the interfacial framework discussed above. Although attractive, the energetic approach was then extensively criticized.[107,108]

As for electro- and diffusio-osmosis, the details of the interfacial dynamics, for example slippage at the interface, is expected to strongly affect thermo-osmotic flows. This has been evidenced for example in molecular dynamics where a huge enhancement of thermo-osmosis was measured with slip,[109,110] although the exact dependence of thermo-osmosis on the interfacial properties was measured to be substantially complex.[110] We refer to the ref. 102 and 107 for more in-depth discussion on thermo-osmosis and -phoresis.

Lately thermo-osmosis has gained growing attention in terms of applications and we briefly comment here on this aspect. Many applications of the phenomenon are done in the context of thermo-phoresis, or displacement of colloidal particles under thermal gradients (in a similar way to diffusio-phoresis, see Fig. 18a). This phenomenon was harvested to manipulate colloids and build structures[107,111–113] with advanced applications in microfluidics[101] or towards DNA detection.[114] Among other phenomena, it was shown that couplings between thermo- and diffusio-phoretic drivings allow to finely manipulate colloidal structures.[115] Also, thermophoresis of molecules can provide detailed information about particles and molecules (size, charge and hydration shell) and this provide very efficient analytical tools to probe protein in biological liquids.[116,117] In a different context, applications of thermo-osmosis were suggested for the recovery of water from organic waste-water,[118] as well as for energy harvesting from thermal differences (and waste heat).[110,119,120]

3.5 Numerical simulations of (diffusio-)osmotic transport: methodologies and results

Molecular simulations have now become a highly efficient tool to explore the fundamental properties of fluids and materials. Molecular dynamics simulate the many-body dynamics of particles and molecules, either at equilibrium or far from equilibrium, submitted to various thermodynamic forces. They provide detailed information on the molecular processes at play. In the present context of studying osmotic forces and related fluxes, this represents a key opportunity to understand the fundamental and subtle origins underlying interfacial transport and how these can be affected by the microscopic details of the interface.

Simulating electro-osmosis is relatively straightforward in the sense that the effect of the electric field converts directly into an electric force acting on the suspended ions. This has led to numerous molecular dynamics studies of electro-osmosis, as well as of streaming currents, allowing to decipher a wealth of phenomena associated with transport within the electric double layer.[62,121] Now, simulating diffusio- and thermo-osmosis is by far more difficult and subtle. Indeed such transport occurs under thermodynamic forces associated with the gradient of concentration or temperature, and these can not obviously be represented in terms of mechanical forces acting on the simulated particles. We discuss in this section recent developments in the numerical methodologies allowing to perform simulations of osmotic transport.

For bare osmosis, direct simulations can be performed using two explicit reservoirs with difference of solute concentration. For example such implementation was used by Kalra *et al.* in the study of osmosis across carbon nanotubes.[122] This configuration has the drawback that osmosis occurs in the transient regime since the reservoirs empty/fill during the osmotic process and this limits statistics. Osmosis was later rationalized in more simple terms by simplifying the explicit membrane description to reduce it to a confining potential acting on the solute only.[123–125] This is the numerical pendant to the mechanical views of osmosis described in Section 2.4.

The numerical implementation of diffusio-osmosis in molecular dynamics is far more complex since one should be able to represent the chemical gradient in terms of a microscopic force acting on the particles. Various methods to investigate diffusio-osmotic transport were proposed in the recent literature and we discuss them now.

Using symmetry relations to infer transport coefficients. It turns out that it is far easier to calculate the diffusio-osmotic mobility by exploiting the symmetry of the transport matrix. Recalling the general relation between fluxes and forces,

$$\begin{pmatrix} Q \\ J_\mathrm{s} - c_\mathrm{s}Q \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix} \times \begin{pmatrix} -\Delta p \\ -k_\mathrm{B}T\Delta\log c_\mathrm{s} \end{pmatrix}. \qquad (61)$$

the Onsager symmetry for the transport matrix implies that $L_{21} = L_{12}$. Accordingly, calculating the diffusio-osmotic mobility as a water flux under a concentration gradient, here $L_{12} = Q/(-k_\mathrm{B}T\Delta\log c_\mathrm{s})$, is therefore equivalent to calculating the excess solute flux under a pressure gradient, here $L_{21} = (J_\mathrm{s} - c_\mathrm{s}Q)/(-\Delta p)$ – see Fig. 13a. The latter is far easier to implement numerically in non-equilibrium molecular dynamics (NEMD) since it requires only to generate a pressure-driven flow and measure the integrated solute flux (or locally the velocity and solute concentration profile). This can be performed with periodic boundary conditions along the flow, so that the resulting diffusio-osmotic mobility is indeed characteristic of the liquid–solid interface under scrunity, and does not depend on *e.g.* entrance effects into the pore. This methodology was successfully applied to quantify the diffusio-osmotic mobility on a variety of interfaces, including superhydrophobic surfaces, graphene, and with various liquids.[26,69,70,110] We discuss below some results of the simulations.

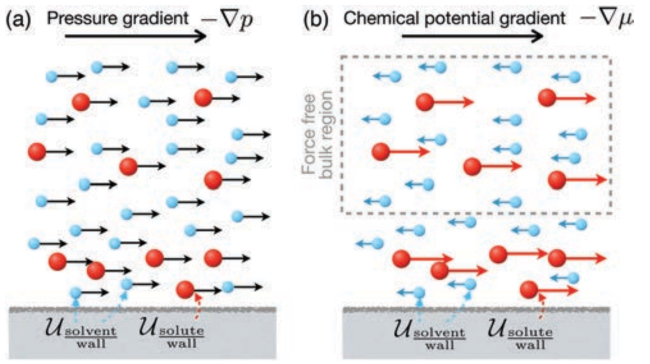


Fig. 13 Non-equilibrium molecular dynamics of osmotic interfacial transport. Inspired from ref. 125. (a) Excess flux under pressure gradient. The pressure gradient is obtained by applying a force acting on each particle and the solute flux in excess to the bulk $J_\mathrm{s} - c_\mathrm{s}Q$ is measured. (b) NEMD method to simulate interfacial transport under chemical potential gradients. The chemical potential gradient is modeled as a forward force per solute particle (red) and a properly defined counter force per solvent particle (blue) such that the total force on the fluid in the bulk is zero. The local diffusio-osmotic velocity profile is directly measured.

Equilibrium fluctuations for linear response coefficients. Transport coefficients may also be inferred from equilibrium fluctuations by making use of Green–Kubo (GK) relations for the various mobilities. The transport coefficients introduced in the transport matrix $\mathbb{L}$ can indeed be written in terms of a time-correlation function of the fluctuating fluxes Q_i at thermal equilibrium. Such formal relations are obtained thanks to linear-response theory and the fluctuation–dissipation theorem.[126–128] They provide generic expressions for the non-equilibrium mobilities in terms of equilibrium correlation functions in the form

$$L_{ij} = \frac{\mathscr{V}}{k_\mathrm{B}T}\int_0^\infty \mathrm{d}t\langle Q_i(t)Q_j(0)\rangle \qquad (62)$$

where $\mathscr{V}$ is the system volume and $\{Q_i\}$ are the fluxes under scrutiny. The symmetry of the transport matrix originates in the time-symmetry of the underlying microscopic dynamics.[20] The simplest route to obtain the Green–Kubo formula for the diffusio-osmotic mobility is to consider the solute excess flux generated under a pressure drop since the latter is equivalent to a body force applied to all system particles. The linear-response formalism then immediately leads to[125]

$$L_{21} = L_{21} = \frac{\mathscr{V}}{k_{\mathrm{B}}T}\int_0^\infty \langle (J_{\mathrm{s}} - c_\infty Q)(t)Q(0)\rangle \mathrm{d}t. \tag{63}$$

In the case of a channel with length L and cross area $\mathscr{A}$, one has $L_{21} = L_{21} \equiv \frac{\mathscr{A}}{L}\mu_{\mathrm{DO}}$. We refer to ref. 125 and 129 for detailed derivations of these GK equations.

These GK formula allow to calculate numerically the diffusio-osmotic mobility, as well as any off-diagonal terms of eqn (53), by estimating the correlation functions in eqn (63) in equilibrium simulations. This approach was followed in ref. 125, 129 and 130 and the resulting mobilities were successfully compared with results of NEMD simulations, as discussed below.

Non-equilibrium molecular dynamics and mechanical representation of chemical gradients. While the equilibrium approach provides proper foundations to calculate diffusio-osmosis, non-equilibrium simulations proves usually more practical to calculate transport coefficients, *e.g.* in terms of statistics. However, as we emphasized above, this requires to build a proper numerical scheme to implement a mechanical equivalent of the chemical potential gradient. One interesting route was suggested by Yoshida *et al.* in ref. 129 and then applied to electro- and diffusio-osmotic transport of electrolytes: the authors ran different simulations where forces f_j are applied separately to each individual specie, here {solvent, anions, cations}, allowing to calculate the corresponding individual fluxes Q_i and deduce the mobilities for the individual species $M_{i,j} = Q_i/f_j$; the electro- and diffusio-osmotic mobilities are then calculated by proper linear combinations of the mobilities of individual species, in order to deduce the electro- and diffusio-osmosis. This approach echoes directly the discussion in Section 3.1.3, in which the electro-osmotic and diffusio-osmotic mobilities are deduced from the individual ion mobilities, defined above as $M_\pm$.[129]

It is however relevant to develop numerical methods to simulate explicitly the diffusio-osmotic flows. Such a numerical scheme was recently proposed in ref. 125, in which a proper mechanical set of driving forces is applied to the system to mimick the chemical potential gradient of the solute. To do so, the scheme applies differential forces on the solute and on the solvent, see Fig. 13b: (i) an external force F_μ on each solute particle in the whole system; (ii) a counter force $-[N_{\mathrm{s}}^{\mathrm{B}}/(N^{\mathrm{B}} - N_{\mathrm{s}}^{\mathrm{B}})] \times F_\mu$, acting on each solvent particle. Here $N_{\mathrm{s}}^{\mathrm{B}}$ and N^{B} are respectively the number of solute particles and the total number of particles in a properly defined bulk region (''sufficiently'' far from the surface). The counter force is set to ensure a force free balance in the bulk volume. Most important, it can be verified that applying linear-response theory to the system with this set of forces allows one to show that the resulting diffusio-osmotic mobility does identify with the GK relation in eqn (63): this therefore fully validates the theoretical foundations of the proposed numerical scheme. The corresponding effective chemical potential is then related to the applied external force F_μ *via*

$$-\nabla_x \mu = F_\mu \frac{N^{\mathrm{B}}}{N^{\mathrm{B}} - N_{\mathrm{s}}^{\mathrm{B}}}. \tag{64}$$

This approach leads as expected to a velocity profile exhibiting a strong gradient within the interfacial layer, and then a plug flow far from the surface. The deduced diffusio-osmotic mobility obtained from the NEMD scheme was checked to be identical to both the equilibrium GK results and those obtained from the excess flux under pressure-driven flow introduced above.[125]

Some difficulties with the microscopic stress tensor. The continuum approach, as described above, allows one to predict diffusio-osmotic transport in terms of a surface pressure gradient. In a different approach, it is accordingly tempting to obtain the diffusio-osmotic flow by a direct numerical calculation of the local microscopic pressure in the fluid. However, as was demonstrated by Frenkel and collaborators in a series of papers,[131–134] a major difficulty in this approach is that there is no unique expression for the local microscopic pressure tensor (*e.g.* in terms of the position and velocities of individual particles and the microscopic forces acting on them). Accordingly various microscopic definitions of the pressure tensor lead to different numerical results. Such difficulty was evidenced for diffusio-osmotic flows,[132] but also for thermo-osmotic flows.[133,134]

***S*ome results of simulations.** Simulations have allowed to gain much insights into diffusio-osmotic transport. Various fluids, *e.g.* Lennard-Jones fluids, but also electrolytes and water-ethanol mixtures, and various interfaces were considered, hydrophilic or hydrophobic surfaces, graphene, superhydrophobic surfaces, *etc.* Among highlighted effects one may quote the impact of hydrodynamic slippage of the fluid at the surface, which does boost considerably the diffusio-osmotic mobility on hydrophobic[69] and graphene surfaces,[125] and even more on super-hydrophobic surfaces.[70] The enhancement of the diffusio-osmotic mobility scales typically like the ratio between the (effective) slip length and the interfacial length, as mentioned in the previous sections.

Simulations also give some insights on the local diffusio-osmotic velocity profile and its relation to the concentration profiles. Within the continuum framework discussed in the previous section, the velocity profile is obtained simply by integration of the Stokes equation of motion in eqn (39), with the pressure expression given in eqn (38). Simulations actually show usually a very good agreement between the continuum prediction and the velocity profiles measured in the NEMD simulations, giving strong support to the continuum description. Such an agreement may be considered as surprising in view of the strong velocity gradients occurring on length scales in the range of a few molecular size. However the Stokes equation is known to be surprisingly robust down to molecular lengthscales[62] and this explains its success in predicting diffusio-osmotic flows.

Finally, the continuum framework allows one to relate the diffusio-osmotic mobility to the concentration profile, and more particularly to its first spatial moment, see eqn (41). It is accordingly tempting to relate – as for Marangoni effects – the amplitude of diffusio-osmotic transport to the adsorbed quantity, defined as $\Gamma = \int dz(c_s(z) - c_\infty)$. The latter is directly connected to the surface tension *via* the Gibbs–Duhem relation. The adsorbed quantity Γ provides in most cases a good estimate for the diffusio-osmotic mobility and its sign. However – as mentioned earlier – for complex concentration profiles the relation was found to be more complex than this simple rule of thumb (for example for water-ethanol mixture at interfaces[26]).

4 Osmosis beyond van 't Hoff

4.1 Advanced osmosis and nanofluidics

The previous section highlighted molecular insights into osmotic phenomena, unveiling the underlying driving forces at play. However such perspectives also suggest possible extensions to obtain more advanced osmotic transport beyond the linear framework presented before. In this section we discuss osmotic transport across channels with more complex geometries involving symmetry breaking, or active parts. Our objective in this section is to show that it is possible to extend simple osmosis beyond the van 't Hoff paradigm and design advanced functionalities resulting in non-linear and active transport.

In this context it is interesting to observe that in biological species (bacteria, archaea, fungi,...) many membrane channels do achieve advanced functionalities in order to regulate osmosis: for example rectified osmosis[135] – *e.g.* an osmotic flow with a non-linear dependence on the concentration gradient-, or gated osmosis to prevent lysis and survive osmotic shocks in mechanosensitive channels[3] (with diffusio-osmosis identified as a potential mechanism for the gating mechanism in deformable structures[136]). These few examples highlight the possibility of going far beyond the van 't Hoff paradigm, thanks to properly designed (active) nanochannels.

We believe that the advent of nanofluidics has a key role to play in this regard, in order to identify new types of behaviors which could be scaled-up to macroscopic membranes. The new opportunities brought by nanofluidics in terms of the variety of nanoscale geometries and materials, combined with state-of-the-art experimental instrumentation, allows one to fabricate and investigate fundamentally the transport in ever smaller channels, with ever more complex and rich behaviors. Carbon nanotubes, down to nanometric sizes[73,138–140] can now be manipulated and inserted in devices were water is flown through – see Fig. 14a. Single nanopores can be carved or etched in membranes that are only an atomic layer thick[74] and may be accordingly functionalized,[141] see Fig. 14b. It is also now possible to fabricate Ångström scale slits using graphene sheets as spacers, reaching confinement thicknesses down to ~3 Å[142,143] – see Fig. 14c.

Such nanofluidic technologies offer new possibilities in the context of osmotic transport. They allow nearly molecular scale designs, leading to various nanofluidic-specific effects which may be key assets for new separation techniques and water filtration: from specific ion exclusion effects[87,138,140,143,144] with a number of anomalous ionic effects to be investigated,[145] to extremely fast permeation of water, in particular through carbon nanotubes.[139,140,146–148] Also new types of nanoscale membranes have also emerged recently, offering new designs as compared to traditional membranes: for example, with dedicated patterns of hydrophilic and hydrophobic regions;[149] or tailor-designed DNA origami channels,[86,150] and – last but not least – the multilayer membranes of graphene (so-called graphene oxide membranes).[75,151,152]

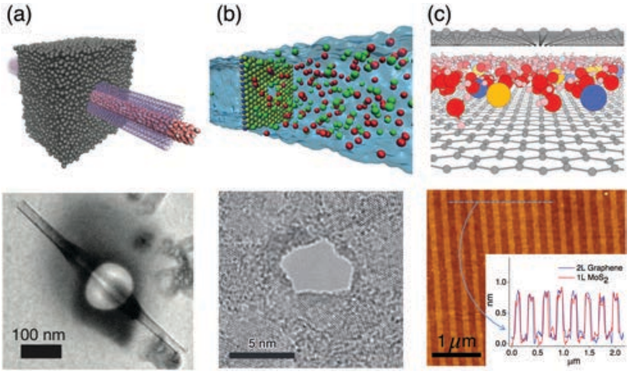


Fig. 14 From nanoscale to Ångström scale pores. (a) Reproduced from ref. 73. (top) Molecular dynamics representation of water flowing through a transmembrane multi-wall boron-nitride nanotube and (bottom) transmission electron microscope (TEM) picture of its experimental counterpart. (b) (top) Molecular dynamics representation of a nanopore in a mono-layer MoS_2 membrane (in blue and yellow) and the salt (green and red) in solution and (bottom) TEM picture of its experimental counterpart, a 5 nm pore. Reproduced from ref. 74 with permission from Springer Nature, copyright 2016. (c) (top) Molecular dynamics representation of water and ions (orange and blue) flowing through a graphene slit with 7 Ångström spacing. (Courtesy from Timothée Mouterde) (bottom) AFM image of bilayer graphene spacers on top of the bottom graphite layer. Inset: Height profiles yield an estimate of 7 Å for the thickness of spacers made from 2 layers of graphene or 1 layer of MoS_2 (the blue line shows the scan position for the corresponding trace in the inset). Reproduced from ref. 137 with permission from AAAS, copyright 2017.

This constitutes a new and exciting playground, in which osmotic phenomena may (and should) flourish in various forms. We discuss in the next paragraphs two such examples: the development of osmotic diodes, and an active counterpart of osmosis, which both lead to tunable osmotic driving beyond van 't Hoff.

4.2 Osmotic diodes and osmotic pressure rectification

One of the successes of nanofluidics was to demonstrate the possibility to design diodes for ionic transport, in full analogy with their electronic counterpart.[62,85,153] This takes the form of a non-linear and rectified response for the ionic current *versus* the applied voltage. Typically an ionic diode behavior manifests itself in channels with an asymmetric design, *e.g.* an asymmetric surface charge or an asymmetric geometry. Such behavior is expected to occur in the regime where the Dukhin number is of order one and asymmetric along the channel:[154] the Dukhin number is defined here as $Du = \Sigma/c_s h$, where Σ is the surface charge density, c_s the bulk salt concentration and h a characteristic channel dimension. It quantifies the importance of surface *versus* bulk electric conduction. As such ionic diodes may find interesting applications to boost osmotic power generation under salinity gradients, see ref. 75 and Section 6.3.

Now, coming back to osmosis, the asymmetry of the design may be expected to yield an asymmetric interaction of the membrane with the electrolyte, hence an asymmetric push on the liquid. It was shown in ref. 50 that such asymmetric geometry – depicted in Fig. 15a – results in an osmotic diode, with a rectified osmotic pressure *versus* the concentration gradient (non linear dependence), furthermore tunable by the applied electric field.

The description builds on the previous mechanical views of osmosis, in Section 2.4. The Stokes equation for fluid motion writes

$$0 = -\boldsymbol{\nabla} p + \eta \boldsymbol{\nabla}^2 \boldsymbol{u} + \rho_e(-\boldsymbol{\nabla} V_e), \qquad (65)$$

with $\rho_e = e(c_+ - c_-)$ the charge density, $c_\pm$ is the concentration of positive and negative ions (assumed here as monovalent for simplicity) and V_e is the electric potential. Following the same steps as in Section 2.4 to integrate the fluid equations of motion in the channel, the general relation between flow and pressure takes the expression

$$Q = \mathscr{L}_{\rm hyd} \times -\Delta[p - \Pi_{\rm app}] \qquad (66)$$

where $\mathscr{L}_{\rm hyd}$ is the channel permeance introduced above. The apparent osmotic pressure between the two sides of the channel is accordingly defined as

$$\Delta\Pi_{\rm app} = \frac{1}{\mathscr{A}} \iint {\rm d}\mathscr{A} {\rm d}x \rho_e \times (-\nabla V_e) \qquad (67)$$

with $\mathscr{A}$ the cross section of the pore, L its length. The ion concentration profiles obey the Poisson–Nernst–Planck equations, coupling the diffusive dynamics to the applied electric forces. In spite of the expected non-linear dependence of the osmotic pressure in terms of driving forces, the symmetry in the force balance and ionic transport equations, which was highlighted in Section 2.4 and eqn (28) for the simplest geometry, still holds. There is accordingly a linear relation between the apparent osmotic pressure in eqn (67) and the total surface ion flux j_s:

$$\Delta\Pi_{\rm app} = k_B T \Delta c_s + j_s \times \frac{L}{D}. \qquad (68)$$

It is therefore expected that the rectifying behavior in the ion flux, akin to the current diode, thus translates into a rectifying osmotic pressure.

Solving the full equations in the geometry presented in Fig. 15a yields a fluid flux:[50]

$$Q = \mathscr{L}_{\rm hyd}[\sigma k_B T \Delta c_s - \Delta p] + Q_S[\Delta c_s]\left(\exp\left(\frac{e\Delta V_e}{k_B T}\right) - 1\right) \qquad (69)$$

where the reflection coefficient σ is now a non-linear function of the concentrations in both reservoirs and Q_S plays the role of a "limiting fluid flux".[50] The apparent osmotic pressure $\Delta\Pi_{\rm app} = Q/\mathscr{L}_{\rm hyd}$ is plotted in Fig. 15. The rectification and diode behavior *versus* concentration is weak for zero voltage but strongly enhanced for higher applied voltage bias.

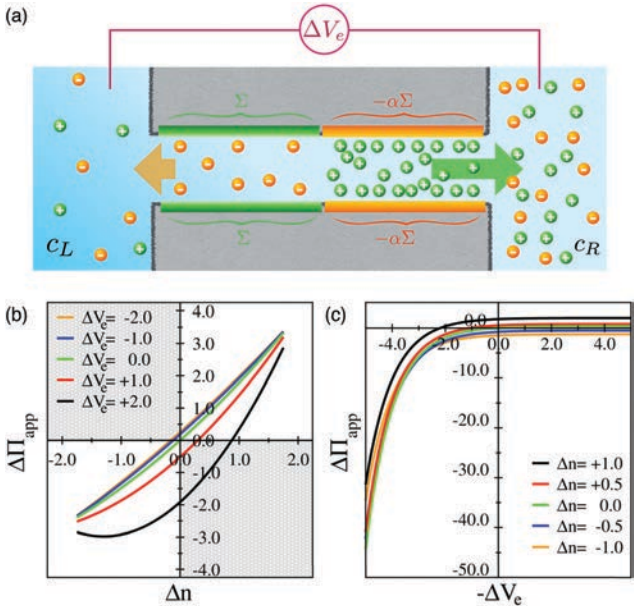


Fig. 15 Osmotic rectification in an osmotic diode. (a) A nanochannel presents an asymmetric surface charge with $\Sigma > 0$ and $-\alpha\Sigma$ on the other side with $\alpha \neq 1$. The ions are therefore submitted to an asymmetric force (between one side and the other, in colored arrows) that drives the osmotic flow. Inspired from ref. 50. Apparent osmotic pressure $\Delta\Pi_{\rm app}$ *versus* (b) salinity gradient $\Delta n = n_R - n_L$ (where $n_i = c_i/c_0$ is a normalized concentration) and *versus* (c) applied voltage drop $\Delta V_e = V_R - V_L$ (normalized by k_BT/e) as obtained from an analytical solution for the flows of all species in the nanochannel. (b) and (c) are reproduced from ref. 50 with permission from the APS, copyright 2013. More information can be found in ref. 50.

Examples of permeability rectification are actually numerous in the biological world. They are harnessed *e.g.* in plant cells[135] or in animal cells.[155–157] Surprisingly in all the studies that we are aware of, entering flows are notably larger than outer flows, and up to 10 times higher in some mammalian fibroplasts.[157] It is fascinating to see how most cells are therefore adapted to fill in faster than they would swell under the same conditions, probably with a link to survival strategies. We highlight that rectified osmotic flows could be used in a variety of fields. In fact, Fig. 15b and the results reported in ref. 50, demonstrate that water may be flown against the natural osmotic gradient, with water flowing to the high salinity reservoir, depending on the voltage applied. Furthermore, under an oscillating electric field, with proper conditions, this induced water flow against the natural osmotic gradient will be maintained. This opens new perspectives *e.g.* for advanced water purification strategies and active filtration with oscillating fields, as we discuss later.

4.3 Towards active osmosis

We discuss now a second class of examples of osmotic phenomena that goes beyond the van 't Hoff paradigm. As we exhaustively discussed, the idea of osmosis is closely related to semi-permeability and sieving – with the membrane playing the role of a simple colander. However one may consider how the

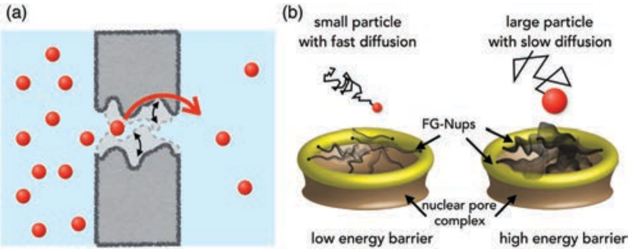


Fig. 16 Osmosis through out-of-equilibrium pores. (a) Illustration of a pore with a time-dependent shape, where the inner pore size may be either smaller either larger than the typical solute size. (b) Reproduced and adapted from ref. 160 with permission from Springer Nature, copyright 2016. Representation of the spatiotemporal motion of grafted phenylalanine–glycine nucleoporins (FG-Nups) inside the selectivity filter of the nuclear pore complex. Small (resp. large) particles that diffuse fast (resp. slow) see effectively the FG-Nups as static leaving small openings (resp. moving and everywhere) and therefore encounter only a "small" energy barrier for translocation (resp. large).

osmotic pressure builds up in membranes with time-dependent pores: a pore which opens and closes over time – see Fig. 16 – will exhibit a time-dependent size exclusion and sieving is thus expected to generate an intermittent osmotic push on the fluid. The resulting osmotic pressure is expected to be some time-average of the push, which remains to be properly defined. But injecting energy at the pore scale – here *via* the time-dependent opening of the pore – may also lead to far-from-equilibrium behaviors, allowing possibly to bypass the entropy bottleneck. Osmosis across dynamically stimulated pores is therefore a subtle problem, which requires proper microscopic foundations. Beyond the fundamental question, adding the sieving frequency as a new tuning parameter may improve separation and selectivity properties of the membranes.[158,159]

The question of dynamic osmosis actually arises naturally in biological pores, *e.g.* ion channels, since their shape is affected by thermal fluctuations or demonstrates out-of-equilibrium motion.[161] The question of out-of-equilibrium osmosis was discussed in the literature in the 1980s[162] and a few molecular dynamics studies were pursued,[161,163,164] usually with a focus on the specificities of the biological channels under scrutiny. In fact, temporal dynamics of biochannels are strongly believed to be connected to the selectivity properties of the channel. For example, the fluctuations of the refined structure of the selectivity filter of the KcsA channel is believed to be a key factor for extremely refined passage of the potassium ion.[161] Further, such temporal dynamics of the structure may provide an efficient alternative to simple steric sieving for selectivity. This was noticed in the nuclear pore complex,[160] where particles see an effective translocation barrier which is dependent on their diffusion properties (see Fig. 16b).

To our knowledge, there is no general framework discussing the concept of "dynamic osmosis". Several simple models were considered recently by the authors in ref. 158 and 159. Here we illustrate a few basic concepts underlying this active osmosis process and the opportunities that it offers.

In line with our previous discussion of osmotic phenomena, it is particularly fruitful to address the question under the perspective of the mechanical insights, where the pore with fluctuating shape is modeled as a time-dependent energy barrier, say $\mathscr{U}(x,t)$, using similar notations as previously. The average osmotic force acting on the fluid is again obtained in terms of the force acting on the fluid integrated over the channel size and averaged over time. It writes within this framework

$$\Delta\Pi_{\mathrm{app}} = \left\langle \int \mathrm{d}x c_{\mathrm{s}}(x,t) \times (-\nabla_x \mathscr{U}(x,t)) \right\rangle_t. \tag{70}$$

where here $\langle\cdot\rangle_t$ denotes a time average. As for the static (passive) case, the Smoluchowski equation for the solute allows one to rewrite the apparent osmotic pressure in terms of the solute flux across the fluctuating barrier:

$$\Delta\Pi_{\mathrm{app}} = k_{\mathrm{B}}T\Delta c_{\mathrm{s}} + \langle j_{\mathrm{s}}\rangle_t \times \frac{L}{D_{\mathrm{s}}}. \tag{71}$$

It is interesting to note that the concept of osmotic force $\Delta\Pi_{\mathrm{app}}$ connects directly to the question of translocation of solute molecules across a fluctuating barrier – *via* the solute surface flux j_{s}. That specific question was actually the topic of an exhaustive literature in the context of ratchets, molecular motors, or stochastic resonance.[165,166] Numerous counter-intuitive consequences were highlighted, both theoretically and experimentally, like directed motion, "uphill" transport against gradients, enhanced translocation, *etc.* Accordingly the previous symmetry relation eqn (71) shows that the existence of a finite flux $\langle j_{\mathrm{s}}\rangle_t$, with possibly unconventional dependencies on the concentrations in the reservoirs, will have consequences on osmotic transport, *i.e.* leading to flow of the suspending fluid itself and not only solute motion.

To illustrate this behavior, it is instructive to consider a simple example, made of an asymmetric membrane as in Fig. 17, which oscillates in time as an "on/off" process over a time interval $\tau/2 = \pi/\omega$. When "on", the barrier height is considered as much larger than the thermal energy. This process bares similarities with the ratchet process in ref. 167 and subsequent references, where solute pumping was demonstrated. In elementary terms, when the barrier is "off", solute molecules from both sides diffuse freely (see Fig. 17b). Now, when the barrier is back "on", solute that has crossed the maximum point of the barrier will slide down to the opposite side (see Fig. 17c). This process leads to a finite flux of solute averaged over a period, $\langle j_{\mathrm{s}}\rangle_t$, which can be exactly calculated in the simple model considered, see ref. 168. For example, in the quasi-static (low frequency) regime, the averaged flux reduces in this simple system to

$$\langle j_{\mathrm{s}}\rangle_t \underset{\omega\to 0}{\sim} L\omega \times [C_2 \times \delta_0 - C_1 \times (1-\delta_0)], \tag{72}$$

with C_1 and C_2 the solute concentrations in both reservoirs and δ_0 the potential asymmetry, see Fig. 17. Beyond the specific expression (restricted to this specific regime and model), this result highlights the possibility of uphill solute transport (pumping against concentration gradients) or enhanced solute flux, depending on the direction of the concentration gradient *versus* the pore asymmetry. This behavior is summarized in Fig. 17d and e.

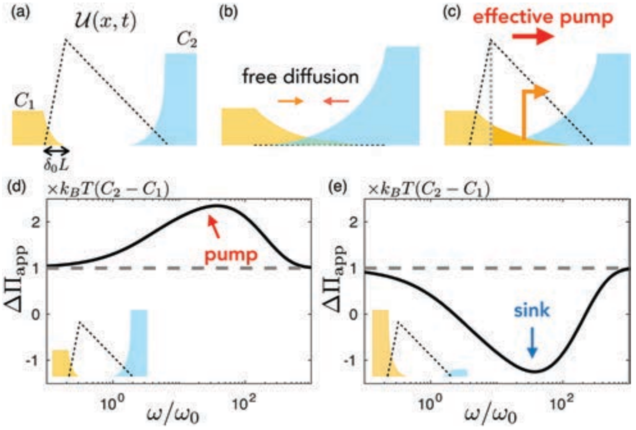


Fig. 17 Active membrane as a pump (or sink). (a) An asymmetric potential barrier (representing the membrane acting on the solute) separates two solute reservoirs with different concentrations. (b) As the barrier is temporarily lowered, the solute may diffuse inwards from both sides. (c) If the barrier is risen back, the solute that has diffused beyond the maximum point of the energy barrier will be carried to the other side. In the example shown here, more solute from the lower concentration side has traversed beyond the maximal point. This solute is then transported to the highly concentrated side. The active membrane therefore acts as a pump. (d and e) Apparent rejection coefficient as obtained from the on/off energy barrier model described in the text.[168] The normalizing frequency is $\omega_0 = D_s/L^2$, and the asymmetry parameter is here $\delta_0 = 0.1$. Insets indicate the solute concentration on both sides; $\Delta c_s = C_2 - C_1$; for (d), $C_1 = 0.4C_0$ and $C_2 = C_0$ where C_0 is some arbitrary concentration and for (e) $C_1 = C_0$ and $C_2 = 0.1C_0$.

Now inserting this result for the flux in eqn (71), one therefore predicts highly counter-intuitive behaviors for the osmotic pressure, *i.e.* the driving force acting on the fluid. In Fig. 17d and e, we have introduced and plotted the apparent osmotic pressure

$$\Delta\Pi_{\mathrm{app}} = k_{\mathrm{B}}T(C_2 - C_1) + \frac{k_{\mathrm{B}}TL}{D_{\mathrm{s}}} \times \langle j_{\mathrm{s}}\rangle_t(C_1, C_2, \omega) \qquad (73)$$

based on the full solution of the simplistic previous model. Notably this plot highlights the possibility of "resonant osmosis" for a characteristic frequency (in the form of an extremum of σ_{app}), but this simple model also suggests that – depending on the direction of the potential asymmetry against the concentration gradient – the rejection may be larger than unity (pumping regime) or even decrease towards negative values.[168]

This points to a wealth of intriguing behaviors for osmotic phenomena, which were barely explored up to now. As emphasized above, the recent development of nanofluidics suggests many routes to develop such active pores in artificial channels, *e.g* using voltage-gated nanochannels[169–172] or UV light[173] or stimulated surface chemical reactivity.[174,175] Other externally controlled existing devices include thermally responsive nanochannels.[176,177] The challenge awaiting is to achieve such active control in yet smaller devices to significantly impact water or ion transport.

The foundations of active osmosis remain therefore to be properly investigated. The present discussion is merely an appetizer to illustrate the abundance of "exotic" behaviors which could be unveiled in this context.

5 From diffusio-phoresis of particles to active matter

The previous sections showed how gradients of solutes induce fluid motion in the presence of an interface *via* the diffusio-osmotic phenomenon. Symmetrically when a (solid) particle is suspended in a quiescent fluid, gradients of solute will induce motion of the particle. This motion, called "diffusio-phoresis", relies on the very same osmotic driving forces, occurring within the interfacial layer at the particle boundary. The idea to transport large particles harnessing osmotic forces appeared first in the Russian literature with the works of Derjaguin and Dukhin[55,56,99] and was more thoroughly investigated in the 1990s.[79,178,179] We refer to the review by Anderson in ref. 54 for a dedicated discussion of the underlying transport mechanisms and some of its subtle features.

Diffusio-phoresis and its consequences have gained renewed interest for the last decade, highlighting an increasing number of situations where this phenomenon is shown to play a role, as well as dedicated applications in various domains. Basically diffusio-phoresis occurs whenever there is a gradient of solute or of a mixture of solutes and such situations are ubiquitous.[180,181] Here we summarize the main elements of the phenomenon and focus on a number of elementary implications. More explicit applications will be discussed in the next section, Section 6.

5.1 *E pur si muove*: from diffusio-osmosis to diffusio-phoretic motion

The diffusio-phoretic velocity of a particle under a (dilute) solute gradient writes as[54]

$$\boldsymbol{v}_{\mathrm{DP}} = \mu_{\mathrm{DP}} \times (-k_{\mathrm{B}}T\boldsymbol{\nabla}c_{\infty}) \qquad (74)$$

Physically the phenomenon at stake is sketched in Fig. 18: the solute gradient at the solute surface induces a diffusio-osmotic slip velocity of the fluid (relative to the solid particle) beyond the interfacial diffuse layer; the particle is put in motion in order to precisely negate the corresponding velocity. For spherical particles, the value of the mobility μ_{DP} defined in eqn (74) identifies with minus the corresponding diffusio-osmotic mobility, as given previously in Section 3.1:

$$\mu_{\mathrm{DP}} = -\mu_{\mathrm{DO}}. \qquad (75)$$

For example, for a solute interacting *via* a potential $\mathscr{U}$ with the particle, the diffusio-phoretic mobility is minus the mobility in eqn (41):

$$\mu_{\mathrm{DP}} = -\mu_{\mathrm{DO}} = -\frac{1}{\eta}\int_0^{\infty} z\left(\exp\left(\frac{-\mathscr{U}(z)}{k_0 T}\right) - 1\right)\mathrm{d}z. \qquad (76)$$

Interestingly, provided the value for the diffusio-osmotic mobility is constant over the particle's surface, it was shown by Morrison that this result holds for any particle shape (the argument is valid for any interfacially driven transport[54,183]).

5.1.1 Phoresis in the thin layer limit. Summarizing briefly the derivation, eqn (76) is obtained by separating the

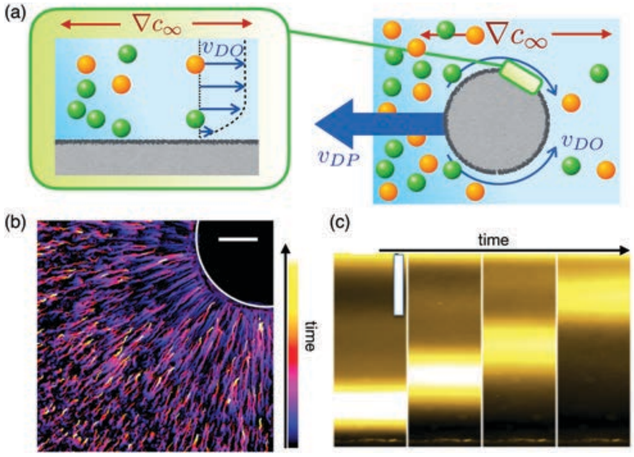


Fig. 18 From osmosis to phoresis. (a) Under a concentration gradient, a particle is put into motion *via* diffusio-osmosis occurring at its surface. (b) Time-stamped stream lines of decane droplet migration towards a hydrogel beacon initially loaded with sodium dodecyl sulfate (SDS), acting as a long-range solute source. Adapted from ref. 182 with permission from the United States National Academy of Science, copyright 2016. The scale bar is 100 μm. (c) Diffusio-phoretic transport of fluorescent λ-DNA under a LiCl gradient (Δc_s = 100 mM over a range of 800 μm, highest concentration being up), scale bar is 100 μm. Images at 100, 150, 200 and 300 s. Adapted from ref. 77.

diffusio-osmotic driving, which occurs at the particle surface within the diffuse layer of thickness λ, and the far-field flow occurring beyond the diffuse layer. Following the prediction in Section 3.1, the near-field diffusio-osmotic flow results in a diffusio-osmotic "slip" velocity of the fluid relative to the particle, with amplitude $\mu_{\mathrm{DO}} \times (-k_{\mathrm{B}}T\boldsymbol{\nabla}c_{\mathrm{s}})_t$, where the index t refers to the tangential component along the particle surface. The concentration gradient in the vicinity of the surface $(\boldsymbol{\nabla}c_{\mathrm{s}})_t$ is related to the far concentration field c_{s}. It obeys Ficks's equation $\boldsymbol{\nabla}^2 c_{\mathrm{s}} = 0$ together with the boundary condition at infinity fixing the concentration gradient, $c_{\mathrm{s}}(r \to \infty) \simeq z\nabla c_\infty$, with z the coordinate along the direction of the gradient. This gives (in spherical coordinates):

$$c_{\mathrm{s}}(r,\theta) = R\nabla c_\infty \left(\frac{r}{R} + \frac{1}{2}\left(\frac{R}{r}\right)^2\right) \times \cos\theta \qquad (77)$$

outside of the diffuse layer; with r, θ the spherical coordinates. Back to the flow, the velocity field outside the diffuse layer obeys the Stokes equation

$$\eta\boldsymbol{\nabla}\boldsymbol{v} - \boldsymbol{\nabla}p = 0, \qquad (78)$$

with the boundary conditions on the particle given by the tangential slip velocity and at infinity given by the uniform flow field

$$\boldsymbol{v}(r = R^+) = -\frac{3\sin\theta}{2}v_{\mathrm{slip}}\boldsymbol{t}, \ \& \ \boldsymbol{v}(r \to \infty) = -\boldsymbol{v}_{\mathrm{DP}} \qquad (79)$$

in the particle frame of reference; $v_{\mathrm{slip}} = \mu_{\mathrm{DO}} \times (-k_{\mathrm{B}}T\nabla c_\infty)$, $R^+ \approx R + \lambda$ denotes the position on the particle surface but located beyond the diffuse layer (here considered as infinitesimal); $\boldsymbol{t}$ is the tangential vector on the particle's surface. To lowest order, the solution to the previous equations results in a flow dominated by a Stokeslet

$$\begin{aligned} v_r &= -v_{\mathrm{DP}}\cos\theta + \frac{F}{4\pi\eta}\frac{\cos\theta}{r} + \mathcal{O}\left(\frac{1}{r^3}\right) \\ v_\theta &= v_{\mathrm{DP}}\sin\theta - \frac{F}{8\pi\eta}\frac{\sin\theta}{r} + \mathcal{O}\left(\frac{1}{r^3}\right) \end{aligned} \qquad (80)$$

where $F = 6\pi\eta R(v_{\mathrm{DP}} + v_{\mathrm{slip}})$. As can be easily verified, F identifies with (minus) the force acting on the particle. At steady-state, the particle is moving with a constant velocity v_{DP} and no force acts on it. Accordingly, the diffusio-phoretic velocity $\boldsymbol{v}_{\mathrm{DP}}$ is fixed by imposing $F = 0$ and this results in eqn (74). This shows implicitly that the far velocity profile scales like $1/r^3$ and can be rewritten as in ref. 54

$$\boldsymbol{v}_{\mathrm{l}}(\boldsymbol{r}) = \frac{1}{2}\left(\frac{R}{r}\right)^3\left[3\frac{\boldsymbol{rr}}{r^2} - \boldsymbol{I}\right]\cdot\boldsymbol{v}_{\mathrm{DP}}. \qquad (81)$$

where we came back to the lab frame of reference, $\boldsymbol{v}_{\mathrm{l}}(\boldsymbol{r}) = \boldsymbol{v}(\boldsymbol{r}) - \boldsymbol{v}_{\mathrm{DP}}$. Accordingly, the hydrodynamic interaction between particles undergoing diffusio-phoretic transport is weak, in contrast to *e.g.* gravity driven transport where the fluid velocity scales like $1/r$ far from the particle. This has important consequences for the phoretic transport in confinement.[54,184,185]

5.1.2 Osmotic force balance on particles. Let us come back to the force balance underlying diffusio-phoresis. We emphasized above that diffusio-phoresis, like any interfacially driven transport, is a force-free motion: the particle moves without any force acting on it, *i.e.* the global resulting force acting on the particle vanishes.[54] This has counter-intuitive consequences and led to various mis-interpretations and debates concerning osmotically-driven transport of particles,[186–191] in particular in the context of phoretic self-propulsion (see Section 5.3 and ref. 192). We thus take the proper space here to discuss the osmotic force balance.

A naive interpretation of diffusio-phoresis is that the particle velocity v_{DP} under a solute gradient results from the balance of Stokes' viscous force $F_{\mathrm{v}} = 6\pi\eta R v_{\mathrm{DP}}$ and the osmotic force resulting from the gradient of the osmotic pressure integrated over the particle surface, hypothetically scaling as $F_{\mathrm{osm}} \sim R^2 \times R\nabla\Pi$. Balancing the two forces one finds a phoretic velocity behaving as $v_{\mathrm{DP}} \sim R^2\frac{k_{\mathrm{B}}T}{\eta}\nabla c_\infty$. Looking at the expression for the diffusio-phoretic mobility in the thin layer limit, eqn (74) and (76), the latter argument does not match the previous estimate by a factor of order $(R/\lambda)^2$, where λ is the range of the potential of interaction between the solute and the particle.

The difference between the two scalings originates in the fact that for interfacially driven motion, the velocity gradients occur mostly over the thickness λ of the diffuse layer, and not on the particle size R, as *e.g.* for the Stokes flow. More fundamentally, this raises the question of how osmotic pressure is expressed: the existence of a difference of solute concentration between the two sides of the colloidal particle does not obviously imply the existence of a corresponding osmotic pressure and this belief led to much confusion. The argument above based on the global force balance is globally flawed and needs to be properly clarified.

In his exhaustive work in ref. 191 following the debate of ref. 186–190, Brady tackled the question based on a "micro-mechanical" analysis of the solute and solvent transport in the presence of the colloidal particles.

However, in order to properly solve the riddle and reconcile the various approaches, one needs to go into the details of the force distribution and write properly the force balance on the particle undergoing diffusio-phoretic transport. It is accordingly interesting to relax the hypothesis of a thin diffuse layer, and consider more explicitly the transport inside the diffuse layer, as was explored by various authors, using *e.g.* controlled asymptoptic expansions.[193–195]

General results for the hydrodynamic flow and mobility can be obtained without assuming a thin diffuse layer. We consider that the interaction between the solute and the particle occurs *via* a radially symmetric potential $\mathscr{U}(r)$, so that the Stokes equations now writes

$$\eta\boldsymbol{\nabla}^2\boldsymbol{v} - \boldsymbol{\nabla}p + c_\mathrm{s}(\boldsymbol{r})(-\boldsymbol{\nabla}\boldsymbol{\mathscr{U}}) = 0. \tag{82}$$

The boundary conditions are now replaced by the no-slip boundary condition on the particle's surface, as well as the prescribed velocity at infinity (in the frame of reference of the particle):

$$\boldsymbol{v}(r = R) = \boldsymbol{0} \text{ and } \boldsymbol{v}(r \to \infty) = -\boldsymbol{v}_\mathrm{DP} \tag{83}$$

The concentration profile obeys a Smoluchowski equation in the presence of the external potential $\mathscr{U}(r)$, in the form

$$0 = -\boldsymbol{\nabla}\cdot[-D_\mathrm{s}\boldsymbol{\nabla}c_\mathrm{s} + \lambda_\mathrm{s}c_\mathrm{s}(-\boldsymbol{\nabla}\mathscr{U})] \tag{84}$$

with the boundary condition at infinity accounting for a constant solute gradient $c_\mathrm{s}(r \to \infty) \simeq r\cos\theta\nabla c_\infty$ (convective transport is neglected here). Given the symmetry of the problem, the solution takes the general form $c_\mathrm{s}(r,\theta) = c_0(r)\cos\theta$, with c_0 scaling with the gradient at infinity as $c_0 \propto R\nabla c_\infty$. Altogether this is a self-consistent equation for the solute concentration field. It should therefore be considered as a source term for the fluid transport eqn (82). For large distances to the particle ($r \gg \lambda$), it reduces to the previous result in eqn (77).

Interestingly, the solution of eqn (82) for the velocity profile can be calculated exactly for any radially symmetric potential $\mathscr{U}(r)$, by extending textbook techniques for the Stokes problem in ref. 96; see also ref. 196 for a related calculation in the context of electro-phoresis. It can be demonstrated that the solution for $\boldsymbol{v}(\boldsymbol{r})$ still takes the same form as in eqn (80), but the force along the axis of the gradient appearing in the Stokeslet term ($v \sim F/r$) term now takes the expression

$$F = 6\pi R\eta v_\mathrm{DP} - \pi R^2\int_R^\infty c_0(r)(-\partial_r\mathscr{U})(r) \times \varphi(r)\mathrm{d}r \tag{85}$$

with $\varphi(r) = \frac{2}{3}\left(3\frac{r}{R} - 2\left(\frac{r}{R}\right)^2 - \frac{R}{r}\right)$ a dimensionless function, the factor $\frac{2}{3}$ originating from the angular average. The diffusio-phoretic velocity results from the force-free condition, $F = 0$, and therefore it writes

$$v_\mathrm{DP} = \frac{\pi R^2}{6\pi\eta R}\int_R^\infty c_0(r)(-\partial_r\mathscr{U})(r) \times \varphi(r)\mathrm{d}r \tag{86}$$

Remembering that $c_0(r) \propto R\nabla c_\infty$, this equation generalizes the previous result obtained in the thin layer limit.

At first sight, eqn (85) and (86) appear as a force balance between the Stokes friction $6\pi R\eta v_\mathrm{DP}$ and the osmotic force, here written in terms of the local force $c_0(r)(-\partial_r\mathscr{U})(r)$ integrated over the particle surface (and potential range). The latter represents the push of the solute molecules on the particle. Actually, eqn (86) is very similar to eqn (2.7) in ref. 191, with the r-dependent term $\pi R^2 \times \varphi(r)$ replaced in ref. 191 by the prefactor $L(R)$. However the integrated "osmotic push" is weighted here by the local factor $\varphi(r)$ (in contrast to ref. 191) and this detail actually changes the whole scaling for the mobility.

Indeed in the thin diffuse layer limit, with $r - R \sim \lambda \ll R$, then one may expand $\varphi(r) \simeq -2(r - R)^2/R^2$, while the concentration profile $c_0(r)$ can be approximated as

$$c_0(r) \simeq R\nabla c_\infty \times \left[\frac{r}{R} + \frac{1}{2}\left(\frac{R}{r}\right)^2\right]\exp[-\mathscr{U}(r)/k_\mathrm{B}T]. \tag{87}$$

One may then verify that the above eqn (86) indeed reduces to the results in eqn (74) and (76) predicted by the thin layer approach. In other words, the weight $\varphi(r) \sim \lambda^2/R^2$ is a signature of the fact that the velocity gradients occur on the potential width λ and not on the particle size R. An osmotic pressure is indeed expressed at the particle's surface and yields diffusio-phoretic transport, but in a very subtle way which does not reduce to considering only the direct solute force. This corrects the naive argument suggested at the beginning of the section.

The exact calculation above also allows one to gain key insight into the local force acting on the particle. The latter is the sum of the hydrodynamic shear force, normal pressure and direct interaction with the solute. Using the exact results for the velocity profile in the thin layer regime, $\lambda \ll R$, one predicts

$$\begin{aligned} f_r &= 3L_\mathrm{s}R^2k_\mathrm{B}T\nabla c_\infty\cos\theta \\ f_\theta &= \frac{3}{2}L_\mathrm{s}R^2k_\mathrm{B}T\nabla c_\infty\sin\theta \end{aligned} \tag{88}$$

where $L_\mathrm{s} = \int_R^\infty\left(\mathrm{e}^{-\beta\mathscr{U}(x)} - 1\right)\mathrm{d}x$ has the dimension of a length and quantifies the excess adsorption on the interface. Eqn (88) can be recovered easily with a simplistic argument: one expects this osmotic force to scale as $\mathscr{V}_\mathrm{int}\nabla\Pi = \mathscr{V}_\mathrm{int}\nabla(k_\mathrm{B}Tc_\infty)$ where $\mathscr{V}_\mathrm{int}$ is the interaction volume. Writing L_s the typical interaction lengthscale we have $\mathscr{V}_\mathrm{int} \approx 4\pi R^2L_\mathrm{s}$, leading accordingly to eqn (88). While the integrated total force does vanish as expected, the osmotic gradients do generate an inhomogeneous local tension on the surface of the particle, as plotted in Fig. 19a. Accordingly, if one considers that the particle is elastically deformable, such tensions would generate a deformation of the particle following the shape sketched in Fig. 19b.

The situation is very different for electro-phoretic transport. As was first discussed in ref. 197, for electro-phoresis there is a local force balance between the direct electric force acting on the particle and the hydrodynamic shear acting on its surface: accordingly the local force simply vanishes identically. In physical terms, this is due to the fact that the electric force acting on the colloid particle exactly balances the electrical

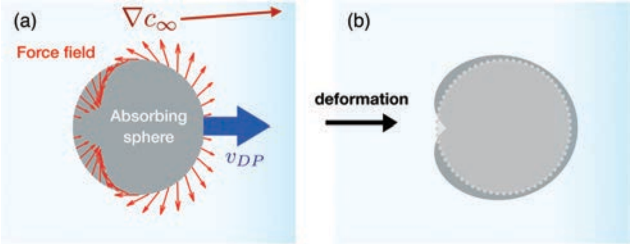


Fig. 19 Local force acting on a diffusio-phoretic sphere. (a) Local force field acting on a sphere experiencing diffusio-phoresis with absorption at its surface in a solute gradient. The local force is plotted with an arbitrary factor amplitude factor (the same for each vector) and projected in the 2D plane; (b) potential resulting deformation of the sphere, axisymmetric view, when the deformation is assumed to be proportional to the local force.

force acting on the electric double layer because of local electroneutrality (the charge in the electric double layer being exactly opposite to the surface charge). This can be actually verified explicitly by extending the previous calculations to electro-phoresis. This can be performed for weak electrostatic potential along the lines in ref. 196, and it predicts indeed a vanishing local force.

Accordingly, particles undergoing electro-phoresis are not expected to deform, in contrast to diffusio-phoresis which leads to local deformations. Such results remain to be experimentally studied in order to observe the modification of a particle conformation undergoing diffusio-phoretic drift. We note however that in the context of thermo-phoresis, DNA was reported to stretch under a temperature gradient.[198] Such effects could have interesting applications in the context of separation of particles, since their shape will differ depending on their size.

As a last comment, the previous discussion neglected surface transport at the surface of the particle: this involves convection of solute in the interfacial region, but also fluid slippage at the particle surface, which will both affect the steady-state concentration field of the solute around the particle. This leads to corrections to the mobility as a function of a (properly defined) Péclet number, as introduced in ref. 68, 69, 193 and 199.

5.1.3 The diffusio-phoretic mobility. Let us now focus on the mobility. As for diffusio-osmosis, the diffusio-phoretic mobility scales as $\mu_{\mathrm{DP}} \approx \pm\lambda^2/\eta$ where λ is the thickness of the interfacial layer. For electrolytes, the latter identifies with the Debye layer thickness and one expects accordingly that $\mu_{\mathrm{DP}} \sim 1/c_{\mathrm{s}}$ so that one usually writes the diffusio-phoretic velocity under salt gradients as:

$$\boldsymbol{v}_{\mathrm{DP}} = D_{\mathrm{DP}} \boldsymbol{\nabla} \log c_{\mathrm{s}} \tag{89}$$

The diffusio-phoretic mobility D_{DP} has now the dimension of a diffusion coefficient. According to the previous estimates, one expects for electrolytes that $D_{\mathrm{DP}} \approx k_{\mathrm{B}}T/(8\pi\eta\ell_{\mathrm{B}})$ with ℓ_{B} the Bjerrum length, so that the value for D_{DP} is expected to be in the range – though slightly smaller – of diffusion coefficients of molecules (thus far larger than any colloid diffusion coefficient): experimentally typical values for D_{DP} are in the range $D_{\mathrm{DP}} \sim 0.1\text{–}1 \times 10^{-10}\ \mathrm{m^2\ s^{-1}}$.[78,200] Note that, as for diffusio-osmosis, diffusio-phoresis under electrolyte gradients with unequal diffusion coefficients for the anions and cations ($D_+ \neq D_-$) has an electro-phoretic contribution similar to eqn (47) which can become quantitatively predominant; see also ref. 54. The expression in eqn (89) highlights a "log-sensing" behavior, similar to that observed in bacteria, *e.g.* in *E. coli*.[201] It is at the basis of various unconventional transport phenomena which we discuss below.

Aside the case of electrolytes, other classes of relevant interactions involve steric exclusions – *e.g.* for neutral polymers – for which the mobility is expected to scale as $\mu_{\mathrm{DP}} = R_{\mathrm{p}}{}^2/\eta$ with R_{p} the excluded particle diameter of the solute.[54,115,202]

On the experimental side, diffusio-phoresis has been investigated in numerous studies. First measurements were performed by the Russian school,[56] and later by Prieve, Anderson and collaborators in the 90's.[178,179] However the development of microfluidics over the last two decades has allowed to develop dedicated systems in which concentration gradients are perfectly controlled and tunable. It was then possible to measure diffusio-phoretic motion and obtain further insights into the phenomenon and its consequences.[21,77–79,200,203–207]

While the above discussion assumed implicitly a dilute solute, the phenomenon is expected to occur under gradients of mixtures, and is denoted as solvo-phoresis.[79] This was for example investigated in a recent study by Paustian *et al.*,[21] who showed that polystyrene colloids undergo motion in gradients of water-ethanol mixtures. The velocity of the particles was shown experimentally to obey

$$\boldsymbol{v}_{\mathrm{DP}} = D_{\mathrm{DP}} \boldsymbol{\nabla} \log X \tag{90}$$

where X is the ethanol molar fraction, thus pointing to nonideality effects. It would be interesting to disentangle the contribution of the dependence of the interfacial thickness with the molar fraction. As a final remark, a slightly different phenomenon is the so-called osmo-phoresis, which is obtained for permeable particles in which their interface plays the role of a semi-permeable membrane and reported in ref. 208.

5.2 Harnessing diffusio-phoresis: membrane less separation, log-sensing and localization

In this section we highlight a number of chosen examples to illustrate the impact and the applications of diffusio-phoresis in diverse physical situations. An interesting feature of diffusio-phoresis is that complex patterns of solute gradients can be rather easily achieved – in relative contrast to electric fields as driving forces – so that this phenomenon can induce particle motion in quite subtle ways leading to a wealth of counter-intuitive behaviors. Such solute patterns may occur naturally, for example due to evaporation leading then to drying film stratification,[202] in membrane fouling,[207] or at dead-end pores, allowing for boosted extraction of particles in porous media,[209,210] as well as in hydrothermal pores with steep pH gradients.[181] Alternatively static or dynamic patterns of solutes can be designed thanks to dedicated microfluidic devices.[78,203] An illuminating example was reported recently in ref. 182, showing that "chemical" beacons emitting solutes may allow to engineer ultra-long range nonequilibrium interactions between particles, up to millimeters – see Fig. 18b.

Instead of being exhaustive, we discuss here several ‘‘elementary mechanisms’’, which serve the purpose of highlighting the versatile manipulation of particle assemblies *via* diffusiophoretic motion.

Boosting migration. As a first example, we highlight how diffusio-phoretic transport leads to strongly enhanced migration of particles, with the fast solute ‘‘towing’’ the large, slow, particles. This effect was demonstrated in particular in coflow geometries, such as in Fig. 20, which was considered in various papers.[200,205,206,211] This geometry is quite generic in microfluidics in the context of mixing and serves here the purpose of highlighting consequences of diffusio-phoretic motion. Colloids have a low diffusion constant and therefore barely mix in such a geometry, see Fig. 20a. Adding tiny amounts of salt, typically millimolars, drastically boosts colloid dispersion – see Fig. 20b and c – with an observed effective diffusion coefficient of the colloids which is 10 to 100 times larger than the equilibrium diffusion coefficient. As mentioned earlier, this is a consequence of diffusio-phoretic motion of the colloids under the salinity gradients present across the various parts of the channel. This can be rationalized on the basis of simple arguments. The growth rate for the width of the colloid suspension writes as $\mathrm{d}w/\mathrm{d}t = 2v_{\mathrm{DP}} = 2D_{\mathrm{DP}}\nabla\log c_{\mathrm{s}}$, with c_{s} the inhomogeneous salt concentration. The latter relaxes *via* Fick's diffusion and $\nabla\log c_{\mathrm{s}} \approx \pm 1\big/\sqrt{D_{\mathrm{s}}t}$ (the sign depending on the salt gradient direction), so that

$$\frac{\mathrm{d}w}{\mathrm{d}t} \approx \pm\frac{2D_{\mathrm{DP}}}{\sqrt{D_{\mathrm{s}}t}}. \qquad (91)$$

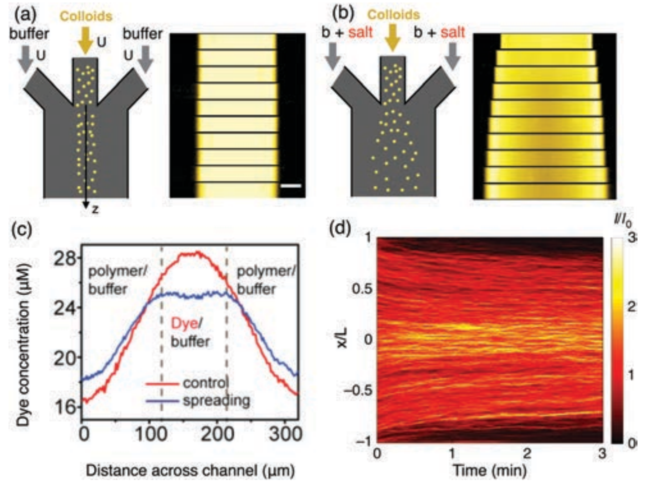


Fig. 20 Harnessing diffusio-phoresis to transport particles. (a and b) Adapted from ref. 200. Fluorescent colloids are injected in the central branch of a microfluidic channel with three branches, with the same inlet velocity. They are imaged at different positions along the channel (side). In case (a) all channels have the same buffer composition (control), whereas in (b) salt (typically 10 mM of NaCl, LiCl or KCl) is added to the side channels. Although little dispersion is seen in case (a) due to the low diffusivity of colloids, a strong migration towards the high salinity is observed in (b). The horizontal scale bar is 50 μm. (c) Dispersion of a dye (Rh6G, 50 μM) at a cross-section of a microfluidic system with three branches similar to (a). The dye enters the central channel and the side channels are filled with polymer (5 wt% Ficoll 400 K) for the spreading experiment (control, without polymer). Adapted from ref. 206 with permission from Springer Nature, copyright 2017. (d) Spatio-temporal evolution of particles upon exposure to CO$_2$ gradients (CO$_2$ is flown above and below, in $x = \pm L$). The particles are polystyrene, diameter 0.5 μm, dispersed in a liquid buffer, and L = 400 μm. Adapted from ref. 205, image under Creative Commons Attribution 4.0 International License.

Note that in the experiments of Fig. 20, the effective time is related to the position z along the channel as $t = z/U$ (U the average flow velocity). Integrating this equation yields immediately the observed diffusive like behavior,

$$w(t) - w_0 = \pm\sqrt{2D_{\mathrm{eff}}t}. \qquad (92)$$

with a diffusion coefficient $D_{\mathrm{eff}} \simeq D_{\mathrm{DP}}{}^2/D_{\mathrm{s}}$. Quantitatively D_{eff} is of the order of a (fraction of) salt coefficient D_{s} so that $D_{\mathrm{eff}} \gg D_0$ (the colloid diffusion coefficient) and colloids ‘‘diffuse’’ much faster in the presence of (even minute) concentration gradients. Similar behaviors in equivalent geometries have been reported under CO$_2$ gradients[205,211] or polymer gradients.[206]

Localization. As a second example, diffusio-phoretic motion can be harnessed to manipulate and localize particle assemblies. Interestingly in the biological world, bacteria are capable of using solute contrasts to localize proteins.[212] Localization is then used as an information for further vital processes, for instance localization of the ring of the FtsZ protein at midcell is used for cellular division.[213,214] Similar features can be obtained on the basis of diffusio-phoresis under salt gradients, harnessing the log-sensing feature discussed above in eqn (89), and leading to particle accumulation.[77,78] Indeed, under a linear salt concentration gradient, the diffusio-osmotic velocity is not uniform and will be larger in regions of lower salt concentration – see Fig. 21a. Alternating the gradient over time leads to rectification of diffusio-phoretic motion and accumulation of the colloids towards *e.g.* the center of the cell, as highlighted in Fig. 21b. It is interesting on this elementary example to formalize the rectification process. The colloid and salt equations of transport obey the coupled Smoluchowski and diffusion equations

$$\begin{aligned} \partial_t\rho &= -\boldsymbol{\nabla}\cdot(-D_0\boldsymbol{\nabla}\rho + D_{\mathrm{DP}}\boldsymbol{\nabla}[\log c_{\mathrm{s}}]\times\rho) \\ \partial_t c_{\mathrm{s}} &= D_{\mathrm{s}}\boldsymbol{\nabla}^2 c_{\mathrm{s}} \end{aligned} \qquad (93)$$

with ρ the colloid density and c_{s} the salt concentration; D_0 and D_{s} are the particle and salt diffusion constants and D_{DP} the particle diffusio-phoretic mobility. Let us simplify the geometry to fix ideas and consider a one-dimensional channel. We consider an oscillating salt concentration profile, $\nabla c_{\mathrm{s}}(x,t) = f(t)\Delta c_{\mathrm{s}}/\ell$, with ℓ a characteristic length scale and $f(t)$ an oscillating function of time with zero average. Averaging over the rapid salt concentration oscillations, the mean diffusio-phoretic velocity which enters the Smoluchowski equation simplifies to $\bar{v}_{\mathrm{DP}} = D_{\mathrm{DP}}\langle\nabla[\log c_{\mathrm{s}}]\rangle_t \approx -D_{\mathrm{DP}}(\langle f^2\rangle_t/\ell^2)\times x$, where x is the distance to the center of the cell and $\langle\cdot\rangle_t$ an average over time. It can be rewritten in terms of an effective potential *via*

$$\bar{v}_{\mathrm{DP}} \equiv \mu_0 \times -\partial_x\mathscr{U}_{\mathrm{eff}}, \qquad (94)$$

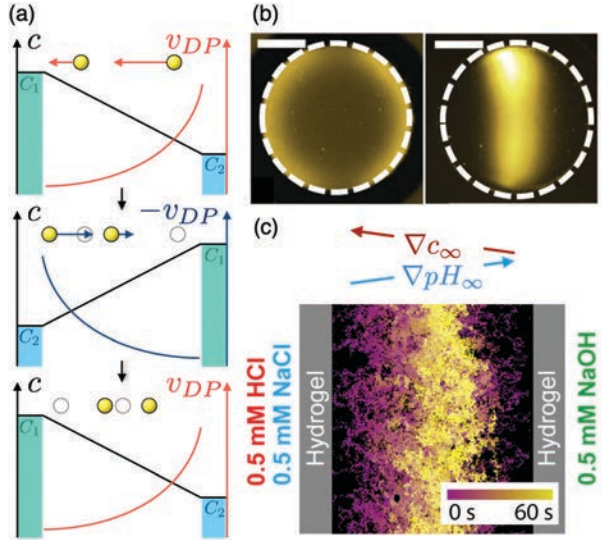


Fig. 21 Focusing particles with diffusio-phoretic transport. (a) Trajectories of two (yellow) particles starting at different lateral positions upon an alternating concentration gradient. As the diffusio-phoretic velocity scales logarithmically with solute concentration, the particles closer to lower concentrations move faster, resulting in localization of the particles at the center. (b) Alternating fluids are flushed on both sides of a circular microfluidic well containing fluorescent colloids. The concentration of LiCl in the two side channels alternates (left/right) with period 480 s, between buffer alone and 100 mM. The scale bar is 200 μm. (left) Initial particle distribution in the well and (right) stationary colloid distribution under the alternating concentration gradient. Reproduced from ref. 77. (c) Diffusio-phoresis under combined steady pH and salt gradients. Flowing NaOH and HCl solutions separately in two reservoir channels establishes a gradient in pH from 3.3 (left) to 10.7 (right), within which diffusio-phoretic particles proceed monotonically to the right. A NaCl gradient is superimposed on the pH gradient inducing diffusio-phoresis to the left. Streamline images reveal unexpected focusing at a location within the channel. Adapted from ref. 80 with permission from the APS copyright 2016.

with $\mu_0 = D_0/k_BT$ the colloid mobility and

$$\mathscr{U}_{\mathrm{eff}}(x) = \frac{k_BT}{2}\langle f^2\rangle_t \frac{x^2}{\sigma_\ell^2}, \tag{95}$$

with $\sigma_\ell = \frac{\ell}{2}\sqrt{\frac{D_0}{D_{\mathrm{DP}}}} \ll \ell$. This illustrates that the rectified diffusio-phoresis of the colloids can be interpreted in terms of a harmonic trapping potential towards the central node (x = 0) of the solute concentration oscillation pattern. This allows one to manipulate the colloidal population *via* time-dependent solute gradients.

Alternative routes for focusing colloidal populations were proposed using diffusio-phoretic transport without the requirement of time-dependent fields[80] – see Fig. 21c. These make use of combined steady gradients of salt and pH which are shown to yield localization of the particles. The interpretation of this subtle phenomenon incidentally highlights that the effects of gradients cannot be simply superimposed for diffusio-phoresis and a new formulation of coupled diffusio-phoretic transport was required, rewriting the driving solute gradients in terms of the corresponding ion fluxes. Such combination of gradients of pH and salts was suggested to occur in hydrothermal pores, with potential consequences on the emergence of an ion-gradient-driven early protometabolism and the origin of life.[181] Finally focusing of colloidal particles was demonstrated in dead-end pore geometries,[210] with potential applications to preconcentration, separation, and sorting of particles.

Osmotic shock. As a last example, we discuss a very striking and counter-intuitive behavior stemming from log-sensing, coined as osmotic shock, which was discussed in ref. 77. It illustrates that diffusio-phoresis keeps a long-lasting "memory" of solute gradients, even when they would be expected to be already homogenized. Consider a situation where the colloids are spread in a reservoir with lateral size ℓ, with initially a uniform solute concentration c_0. Then at time t = 0, solute is flushed at the boundaries, $c_s(x = \pm\ell/2, t) = 0$ for $t > 0$ (simplifying to 1D, and x is the coordinate from the center of the reservoir). After a short transient, the solute concentration profile will decay to zero according to $c_s(x,t) \simeq c_0 \exp[-t/\tau]\cos(\pi x/\ell)$, with $\tau = \ell^2/D_s$ the diffusion timescale of the solute. The diffusio-phoretic velocity of the colloids then writes $v_{\mathrm{DP}} = D_{\mathrm{DP}} \times \nabla \log c(x,t)$, so that

$$v_{\mathrm{DP}} = D_{\mathrm{DP}} \times \frac{\pi}{\ell}\tan\left(\frac{\pi x}{\ell}\right) \approx D_{\mathrm{DP}} \times \frac{\pi^2}{\ell^2} \times x, \tag{96}$$

pointing towards the center of the reservoir, $x = 0$, hence gathering the colloid population towards this position. From eqn (96) it is clear that the DP velocity is therefore independent of time! This leads to the counter-intuitive result that diffusio-phoretic motion occurs on far longer timescales than the solute diffusion timescale. This behavior was highlighted experimentally in ref. 77, where diffusio-phoretic motion of colloid particles was observed on timescales ten times longer than the naive diffusive timescale for the salt. Log-sensing is an efficient approach to localize colloids, but its application for trapping of other types of particles, *e.g.* polymers, remains to be explored. This could possibly be harnessed to improve sensors or traps for high throughput chemical reactions.[215] Applications to information storage and retrieval could also be explored.[216] Log-sensing also helps remove particles or fluids trapped in dead-end pores, as we will discuss in Section 6.4.

As a final word on this section, the role of diffusio-phoretic transport in biological context has yet to be readily explored and quantified. In the toolbox of living systems, concentration gradients play a versatile role, readily exploited in many aspects of the biological machinery, such as energy reservoirs, but also serving more sophisticated functionalities, associated with spatial signaling, localization and pattern formation at the various scales involved in the biological processes. One may cite *e.g.* enzyme transport,[217] protein localization in bacteria[212] or more generally in spatial cell biology the use of concentration patterns for positional information,[218–220] to quote a few. Obviously chemotaxis in biological organisms under solute gradients is a highly complex phenomenon stemming from the interplay of complex signalling pathways, quite far from the simple diffusio-phoretic transport discussed here. But one may reversely remark that the consequences of diffusio-phoresis as a physical phenomenon cannot be overlooked in biological materials, especially in the presence of ubiquitous gradients. This has been barely explored.[221]

5.3 From self-propulsion to self-assembly

Beyond the idea of passive diffusio-phoresis, where particles move under externally imposed solute gradients, arose the idea that the solute concentration gradients could be generated on the particle themselves, *e.g. via* chemical reactions occurring at their surfaces. For an asymmetric chemical reactivity, this self-diffusio-phoresis process thus generates self-propulsion of the particles, fueled by the chemical reactions.[192,222,223] Together with other phenomena leading to self-propulsion, this triggered the emergence of the field of active matter, which has exploded over the last decade. It is not our purpose to review this field, as this goes beyond the scope of our focus on osmotic forces and we refer to some recent reviews on this topic.[223–225] We however highlight here a few phenomena where osmosis, *via* diffusio-phoresis and related mechanisms, is explicitly at play.

Self-diffusio-phoresis. On the experimental side, the phenomenon of self-diffusio-phoresis was pioneered by Paxton and coworkers,[226] who showed self propulsion of Platinum/Gold nanorods in hydrogen peroxide. Hydrogen peroxide is chemically transformed differentially on both metals, either forming or being depleted on each side of the rod, and this creates a gradient of the reacting specie (here hydrogen peroxide) driving diffusio-phoresis, see Fig. 22. For such bimetallic particles with redox reactions on both sides of the particle, self-electro-phoresis may actually contribute to the driving force, *via* motion of charges (electrons and ions) within and outside the particle. Self-diffusio-phoresis was further demonstrated in colloidal janus particles of various materials, see ref. 227–236, and ref. 224 and 237 for a more exhaustive literature on this aspect.

On the theoretical side, the mechanisms by which the creation or removal of species on the particle's surface generate an osmotic pressure gradient and motion are not obvious and there has been some initial debate on this question, see ref. 186–190. This echoes directly our discussion about the diffusio-phoretic force balance in Section 5.1.2. Actually the question was pioneered by Lammert *et al.* on the putative self-electro-phoresis of biological cells or vesicles driven by non-uniform ion pumping across the bounding membrane.[238] Echoing this situation, an illuminating model for self-propulsion *via* asymmetric osmotic driving force was introduced by Golestanian *et al.*[111,239] They considered a particle exhibiting a non-uniform chemical reactivity on its surface, as defined by the corresponding solute flux on its surface $\alpha(\boldsymbol{r}) = -D_r\nabla_\perp C_{\text{react}}$ (corresponding to the generation or consumption of the solute by the chemical reaction), D_r being the diffusion constant of the reactant. This boundary term is coupled to the diffusive dynamics of the solute concentration in the bulk. The resulting concentration gradient induces a diffusio-osmotic slip velocity v_{DO} at the surface, depending on the position, and accordingly particle motion. In the case of a janus sphere, exhibiting a contrasting chemical reactivity on its two moieties, the self-diffusio-phoretic velocity V takes the simple expression

$$V = \langle v_{\text{DO}}\rangle_{\text{sphere}} = \frac{1}{8D_r}(\alpha_- - \alpha_+)(\mu_+ + \mu_-) \qquad (97)$$

where $\alpha_\pm$ is the chemical reactivity on the two sides and $\mu_\pm$ the local surface phoretic mobility. We emphasize though that on the experimental side, other mechanisms also contribute to the motion of catalytically self-propelled particles, like self-electro-phoresis.[192]

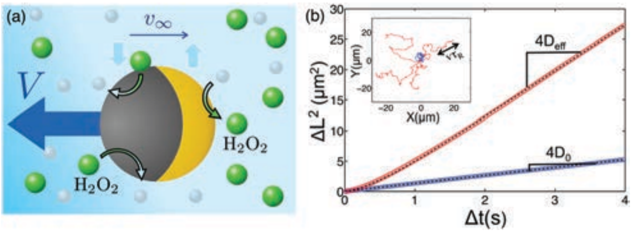


Fig. 22 Self-propelled particles. (a) Chemical reactions occur differentially at the front and at the rear of a reactive colloidal particle, thereby inducing a chemical concentration gradient. This leads to diffusio-osmotic driving at the surface, hereby displacing the particle. (b) Experimental mean squared displacements $\Delta L^2(\Delta t)$ and 2D trajectories (inset) for bare (blue) and active colloids (red) in 7.5% H_2O_2 solution. Bare colloids (bottom) show standard diffusion (ΔL^2 linear in time), while the mean squared displacement of active colloids shows a ballistic motion at small timescales and a diffusive motion at longer timescales. The measured diffusion coefficients are D_0 = 0.33 μm² s^{-1} for bare and D_{eff} = 1.9 μm² s^{-1} for active colloids. Reproduced from ref. 229 with permission from the APS, copyright 2010.

To some extent, such "active particles" mimick self-propelled biological organisms. The possibility to fabricate artificially these systems constitutes a playground to study far-from-equilibrium behaviors.[224] Because active particles consume energy at a local scale, their collective behavior is *a priori* not constrained – at least to some extent – by thermodynamics and may possibly allow to break the bottleneck of the second principle; *cf.* the beautiful example of the rotating Feynman ratchet with active materials in ref. 240.

Active suspensions. Such active particles have fascinating behaviors, and we focus on a few examples. First, self propulsion leads to ballistic motion on short timescales, but orientational random motion leads to diffusive behavior on long timescales, in a way similar to the so-called "run and tumble" motion of bacteria. The effective diffusion coefficient is however far larger than the bare diffusion coefficient based on the Stokes–Einstein estimates D_0,[228,229] see Fig. 22b. As a rule of thumb the effective diffusion coefficient is typically $D_{\text{eff}} \approx V^2 \times \tau_R$, where V is the self-propelling velocity and τ_R is the timescale for rotational Brownian motion: $\tau_R \sim D_{\text{rot}}^{-1} \sim R^2/D_0$, where D_{rot} is the rotational diffusion coefficient, R the particle size. As a consequence the particles behave as a "hot" bath, with a high effective temperature defined from a "fluctuation–dissipation" – like relation as $k_BT_{\text{eff}} = D_{\text{eff}}/\mu_0$, where $\mu_0 = D_0/k_BT$ the bare particle mobility, so that

$$k_BT_{\text{eff}} \approx k_BT \times \frac{V^2\tau_R}{D_0} \qquad (98)$$

(up to numerical prefactors). Altogether this predicts that $T_{\text{eff}}/T \sim \text{Pe}^2$ where the Péclet number is defined in terms of the self-phoretic velocity as Pe = VR/D_0. This prediction was further confirmed experimentally.[229]

Osmotic pressure of active suspensions. The question of the osmotic pressure created by active particles was raised in a number of theoretical and experimental works.[234,241–244] As we introduced above, the osmotic pressure acting on the fluid can be defined mechanically *via* the average force exerted by the active particles on a semi-permeable wall. In the case where the active particles exhibit a Boltzmann-like equilibrium in the presence of the wall (say, represented by an external potential), as shown *e.g.* for sedimentation profiles,[229] then the osmotic pressure reduces to the van 't Hoff law, except that the temperature is replaced by the effective temperature of the suspension:

$$\Delta\Pi \simeq k_{\mathrm{B}}T_{\mathrm{eff}} \times \Delta\rho \tag{99}$$

where ρ is here the concentration of active particles, and the effective temperature T_{eff} was introduced above. This matches the equation of state measured using sedimentation profiles.[234]

However Boltzmann-like equilibrium is expected to fail in some limiting situations for active particles. In particular it is commonly observed that self-propelled particles do hit ''compulsively'' hard surfaces, similarly to a fly on a window:[240] in such cases, the particle remains stuck at the membrane's surface until it reorients, and there is a non-Boltzmannian accumulation of particles at the membrane. This is nicely exemplified in the run-and-tumble model under an external field, where strong deviations from the Boltzmann profile is predicted when the typical drift velocity $V_{\mathrm{d}} = \mu_0\mathscr{F}_{\mathrm{ext}}$ (with $\mathscr{F}_{\mathrm{ext}}$ the maximum external force, say, due to the separating membrane), is larger than the particle self-propulsion speed V.[245] This situation occurs for steep potentials. When such accumulation occurs, the corresponding osmotic push will differ from the simple van 't Hoff law, see ref. 243 and 244. In particular, the osmotic pressure depends on the properties of the membrane itself and its interaction with the particles – typically *via* the ratio between the typical membrane characteristic steepness and the particle mean-free path,[244] in strong contrast to the van 't Hoff ''universal'' relation. Similar deviations from the van 't Hoff law also occurs when the interaction between the active particles and the membrane involves wall-induced rotational torques.[242]

Towards self-organization and self-assembly. Another interesting feature of particles propelling *via* self-diffusio-phoresis is that they interact *via chemical signaling*. Propelled particles act as a beacon – similarly to the situation considered in ref. 182 – and leave a ''trace'' of their passage in the form of a diffusing cloud of chemicals which will be felt by other particles, see Fig. 23. Accordingly other active particles will be reoriented towards or against[246,247] the active particle *via* diffusio-phoretic motion (on top of their self-driving motion). Indeed the surface creation or consumption of solutes generates long-distance distortion of the solute concentration profile, typically relaxing spatially as a monopole $\delta c_{\mathrm{s}} \sim 1/r$ (the scaling deriving from Fick's equation with a sink). This long-range interaction is for example highlighted in Fig. 23b, showing the diffusio-phoretic attractive motion induced by a single beacon, from ref. 231. At shorter range, the interaction may become more complex and requires detailed investigation of the chemical drivings.[247]

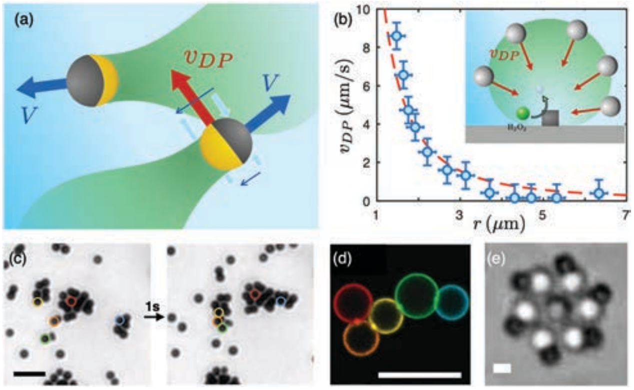


Fig. 23 Out-of-equilibrium self-assembly. (a) Sketch of the diffusing chemical trace left behind active particles, and which modifies the local chemical gradient. Another active particle approaching this chemical gradient will therefore sense a different driving velocity along its edge, changing its trajectory (here in an attractive configuration). (b) Courtesy from Jeremie Palacci, data from ref. 231. Passive colloid diffusio-phoretic speed as a function of the distance to a hematite cube that can be used - with blue light - to catalyse the dissociation of H_2O_2. (c) Spontaneous self-assembly of active particles. The particles form various cluster sizes and shapes. The scale bar is 10 μm. Reproduced from ref. 230 with permission from the APS, copyright 2012. (d) Sequential self-assembly of DNA-grafted droplets, the different colors represent different functionalizations. The scale bar is 10 μm. Reproduced from ref. 252, image under Creative Commons Attribution 4.0 International License. (e) Targeted assembly of phototactic swimmers into nanogears. The scale bar is 1 μm. Reproduced from ref. 235 with permission from Springer Nature, copyright 2018.

This osmotic-induced chemical interaction is at the origin of many advanced collective properties of active particles, such as clustering,[224,230,232,248] or self-assembly,[231,235] see Fig. 23c and d. Out-of-equilibrium self-assembly has raised enormous interest, since activity leads to unexpected structures, with the hope of designing novel and smart materials.[224]

It is interesting to formalize the basics of the phenomenon at stake by writing the coupled diffusion-reaction equation for the colloid population and solute concentration. For the purpose of illustration, we only consider here a single neutral chemically generated specie which acts as a chemo-attractant to the colloids. These dynamical equations actually identify with the so-called Keller–Segel equation, which were written to describe the chemotactic aggregation of a slime mold (amoebae) under the perspective of a dynamical instability:[230,249,250]

$$\begin{aligned} \partial_t\rho &= -\boldsymbol{\nabla}.(-D_{\mathrm{eff}}\boldsymbol{\nabla}\rho + (\mu_{\mathrm{DP}}\boldsymbol{\nabla}c_{\mathrm{s}} \times \rho)) \\ \partial_t c_{\mathrm{s}} &= D_{\mathrm{s}}\boldsymbol{\nabla}^2 c_{\mathrm{s}} + \alpha\rho \simeq 0 \end{aligned} \tag{100}$$

with D_{eff} the effective diffusion coefficient of the active colloids, D_{s} the diffusion coefficient of the ''chemo-attractant'' specie and α the chemical rate of the powering chemical reaction occurring at the surface of each colloid; we assume here that the solute dynamics are fast. By analogy to electrostatics, the second equation for the solute allows one to obtain the solute concentration as a function of the colloid density, as $c_{\mathrm{s}}(\boldsymbol{r}) \simeq \alpha/D_{\mathrm{s}}\int d\boldsymbol{r}'\rho(\boldsymbol{r}')/4\pi|\boldsymbol{r} - \boldsymbol{r}'|$.

When introduced in the first equation in eqn (100), this shows that the diffusio-phoretic attraction acts as a self-consistent effective potential such that

$$\mathscr{U}_{\text{eff}}(\boldsymbol{r},\{\rho\}) = -\frac{\mu_{\text{DP}}}{\mu_0}\frac{\alpha}{D_{\text{s}}}\int \mathrm{d}\boldsymbol{r}'\frac{\rho(\boldsymbol{r}')}{4\pi|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{101}$$

with μ_0 the bare colloid mobility. Equivalently, the dynamics of the colloid population can be formally derived from an effective free energy functional of the colloid system which takes the simple form

$$\begin{aligned}\mathscr{F}_{\text{eff}} &= k_{\text{B}}T_{\text{eff}}\int \mathrm{d}\boldsymbol{r}[\rho(\boldsymbol{r})\log\rho(\boldsymbol{r}) - \rho(\boldsymbol{r})] \\ &\quad - \frac{1}{2}\left(\frac{\mu_{\text{DP}}}{\mu_0}\frac{\alpha}{D_{\text{s}}}\right)\int \mathrm{d}\boldsymbol{r}'\mathrm{d}\boldsymbol{r}''\frac{\rho(\boldsymbol{r}')\rho(\boldsymbol{r}'')}{4\pi|\boldsymbol{r}'-\boldsymbol{r}''|}\end{aligned} \tag{102}$$

This highlights that the "osmotic interaction" *via* the diffusio-phoretic motion induced by the solute traces leads to long range non-equilibrium interactions. This is expected to lead to strong collective effects, such as the clusterization observed experimentally. For attractive systems, $\mu_{\text{DP}} > 0$, these equations are formally analogous to a (non-inertial) gravitational system. Keller–Segel and subsequent works have shown that the above equations predicts an aggregation mechanism, similar to a Chandrasekhar gravitational collapse.[249,250] Similar conclusions were predicted for thermally active colloids[251] where the threshold for collapse was rederived.

However the clusters observed in the experiments, *e.g.* in ref. 230, 231 and 248, do not correspond to full collapse and are rather dynamic, with clusters reaching a finite size and continuously rearranging over time, see Fig. 23c. As shown in ref. 253 and 254, this behavior can be reproduced by Keller–Segel-like dynamics provided both translational and rotational phoretic conditions are properly taken into account in the kinetics. Using this framework, it is furthermore possible to predict the condition in which dynamic clusterization occurs,[253,254] in good agreement with the experiments.

Beyond clusterization, the self-diffusio-phoretic motion of particles and their osmotic interactions were shown lately to lead to the self-assembly of active particles into higher levels of structure organization. This is highlighted in Fig. 23e, from ref. 235, where self-spinning microgears are built on the basis of these non-equilibrium interactions. Beyond, more "on demand" structures are possible, like structures assembled through DNA-grafted interfaces[252,255,256] (see Fig. 23d).

6 Osmosis, towards applications

From food processing in biological organisms,[257–259] to reverse osmosis for desalination and energy generation from salinity differences,[31,73,260,261] osmotic forces are harvested in a considerable number of applications in very different domains. In this section we review more specifically a variety of such applications based on (recently) elucidated transport mechanisms relying on osmotic forces.

6.1 Water treatment and membrane separation

6.1.1 Reverse and forward osmosis and their limitations. Access to clean water and cleaning water from industrial waste is a great challenge:[262] in 2015 still 663 million people worldwide lacked access to drinkable water,[263] and cleaning waste water is becoming a major challenge in oil and gas industries;[264,265] going further, some new regulations may appear to enforce a zero liquid discharge for industrial waste, thus requiring complete recycling of water resources.[266] On a day-to-day basis, humanity consumes the equivalent of 10–100 cubic kilometers of fresh water[267] for all purposes (agriculture, industry, domestic). Because fresh water is not directly accessible everywhere, and in order to cover the growing need for freshwater, desalination of sea water and cleaning of waste water have become essential.

Lately, membrane based technology has established itself for water purification. Reverse osmosis is the most broadly used technique (representing 62–65% of the installed capacity in 2015 for desalination,[268,269] 24% being thermal based technologies). Reverse osmosis relies on the very simple principle of applying an external large hydrostatic pressure to counterbalance the osmotic pressure difference and induce a flow of water towards the low concentration side – see Fig. 24a. In particular, one can therefore extract water from seawater by concentrating seawater even more, or extract water from waste water (in a simplistic view). For desalination the pressures involved are typically of 30–50 bars in order to exceed the osmotic pressure.

In a different approach, forward osmosis (combined with thermal methods for desalination) makes use of draw solutions to counterbalance the salinity induced osmotic pressure[270–272] – see Fig. 24b. Generating a high osmotic pressure, typically above the 30 bars of pressure between sea and fresh water, requires draw solutes which are highly soluble in water, and also with a sufficiently small size (hence low molecular weight). Indeed, as a rule of thumb, for a solute with elementary volume $v_0 \sim r^3$ with r the solute size, the maximum osmotic pressure which can be achieved is typically $\Delta\Pi \sim \frac{k_{\text{B}}T}{v_0}$. This would suggest that solutes with size above 1 nm are not able to achieve a sufficiently high osmotic pressure for desalination

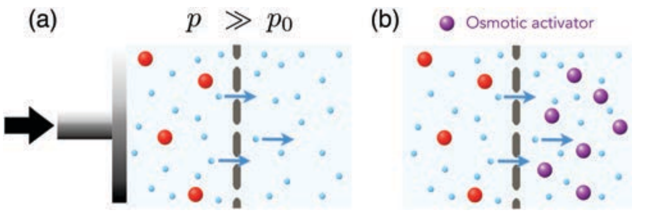


Fig. 24 Reverse osmosis and forward osmosis. (a) Schematic explaining reverse osmosis, occurring *via* a piston applying a large hydrostatic pressure p such that the difference to the atmospheric pressure (pressure of the other compartment) is larger than the osmotic pressure $p - p_0 \geq \Delta\Pi$. (b) Schematic explaining forward osmosis, occurring *via* the addition of some soluble species (here the purple solute) that increases the osmotic pressure on the drought side and therefore "attracts" water from the brine side.

(note that this argument forgets departures from ideality, which could increase more strongly the osmotic pressure). In a more subtle way, a solute with a strong affinity towards water may also decrease the water chemical potential and modify accordingly the pressure. This echoes the huge pressure drops measured with hydrogel structures.[273]

A number of other membrane based techniques with similar geometry are used or being developed, from electrodialysis (based on electric potential driving of salts)[274,275] to capacitive deionization[379–381] but also shock electrodialysis based on the idea of combining salt recovery with a porous charged material,[276] concentration polarization,[277] and other techniques harnessing chemical phenomena like adsorption desalination[278] or biodesalination[279,280] and bio-water treatment.[281]

Membrane-based technologies suffer from a number of limitations. First, they have a high-propensity to fouling by molecules which are larger than the critical molecule size allowed to pass;[282] also due to the large pressures applied during reverse osmosis. Second, because they are passive membranes – essentially discriminating particles upon their size – they can not be at the same time very selective and highly permeable. This was formalized for ultrafiltration membranes in ref. 283 and 284. In fact increasing the permeability of a membrane (and therefore the energy required to recover a given amount of cleaned water) requires essentially to broaden the size of the pores, as water flow within pores is limited by friction on the pore walls. However this leads inevitably to a decrease in the selectivity or separation properties of the membrane; and reciprocally. This is called the selectivity-permeability trade-off. For nanofiltration membranes (used for reverse osmosis and so on) the same trade-off exists although the proper establishment of a limiting regime is still empirical[261] – see Fig. 25. Finally, one challenging progress route for membrane separation is the ability to perform molecular scale design[285] and therefore to ensure the best selectivity properties to eliminate *e.g.* micropollutants, some of which are of great concern for health.[286] Overall, it should be realized that the main current challenge in desalination and water purification is not really the permeability of the membrane, but rather achieving a well-controlled selectivity to retain/reject specific species.

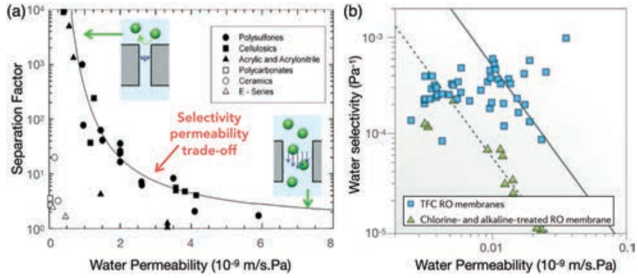


Fig. 25 Selectivity permeability trade-off. (a) Adapted from ref. 283. Selectivity *versus* permeability values for ultrafiltration membranes (used for separation of larger molecules than salt ions). Bovine serum albumine is the model molecule for selectivity. The line indicates the standard selectivity permeability model (with a log normal distribution of pores, Poiseuille flow and selectivity given by Zeman and Wales exclusion rules[287]). (b) Adapted from ref. 261 with permission from Springer Nature, copyright 2016. Selectivity *versus* permeability values for reverse osmosis membranes using salt as the model specie for selectivity. The lines correspond to the empirical models inspired from ref. 288 to relate maximal selectivity and permeability in reverse osmosis membranes.

6.1.2 What can we expect from new nanomaterials and nanofluidic devices? It may appear that developing the ''ideal membrane'', which is both highly selective and highly permeable, is like squaring the circle. However, Nature has achieved this tour-de-force, with water porins like aquaporins exhibiting unrivalled performances in terms of selectivity and permeability.[289] This requires to develop new artificial materials with properly decorated nanopores allowing for such exquisite design, for example self-assembled artificial water channels,[289–291] or tailor-made DNA origami channels.[86,150] This is actually a challenge that nanoscale science may be up to.[88,152,289,292]

The development of new nanomaterials has indeed allowed the emergence of new avenues for membrane separation. Graphitic materials of various forms and geometries, such as carbon nanotubes, graphene and lately graphene oxides membranes, have raised considerable promises, see ref. 88 and 152 for reviews on this topic. Carbon materials were consistently shown to exhibit ultra-low water friction and high permeability, and this represents a key asset to minimize the viscous loss in separation processes. Furthermore advanced functionalization allows one to decorate the nanotubes improving selectivity.[72,293,294] Membranes made of nanopores in *e.g.* 2D graphene sheets have a molecular thickness, while keeping high mechanical strength: this accordingly increases the driving forces for transport (which scale like the inverse thickness) by orders of magnitude, hence all transport coefficients and the overall efficiency of the process.[88]

Still, these graphitic systems – carbon nanotubes, graphene slits and more generally 2D materials – remain difficult to fabricate as large-scale membranes. Upscaling towards industrial applications is a considerable challenge. The advent of graphene oxide membranes and their derivatives may change the story. These are constituted of graphene flakes, which organize into parallel stacks of graphene layers, having nanoslits in a staggered alignment and an interlayer distance which is typically below the nanometer. In spite of the complex labyrinthine flow across multiple graphene layers,[295] the membranes demonstrate large permeability.[296,297] Last but not least, they are relatively easy to fabricate at large scale. Such systems therefore appear as ideal membranes for ionic separation[298] and may well revolutionize the domain of filtration.

Now, beyond materials themselves, it should be realized that nanoscales allow for many new ''exotic'' transport phenomena, the consequences of which have – up to now – been barely harnessed. One may quote for example the membranes made of hydrophobic nanopores, making use of nanobubbles as a semipermeable sieve for osmotic phenomena;[149] or the ionic and osmotic diodes, allowing for rectified transport in membranes, or the active osmotic phenomenon, as we discussed above; or in a different context, the specific adsorption properties of graphene oxide membranes allowing for water-ethanol separation in membranes,[299] which are far more efficient than standard distillation. Such routes would deserve a proper exploration to go beyond the relatively basic sieving principles underlying membrane science. These should offer alternative routes for filtration and separation which still need to be invented.

6.2 Osmosis in biological systems: aquaporins, ion pumps and the kidney

Osmotic forces are harvested in the biological world in a considerable variety of phenomena and contexts: to store energy, induce mechanical motion, control ejection and absorption of compounds, *etc.* Osmotic pressure was much studied at first in plants,[5] and one may in fact assess that plant life depends on osmotic forces. Indeed plants can not rely on muscle for force generation, yet they are able to achieve tensile and compressive stresses on a much wider range.[300] To produce motion or growth, they rely on an underlying hydraulic machinery driven by osmotic or humidity gradients. For example, phloem§ flow harnesses osmotic driving to transport sugar over long distances.[258,259,301–303] Osmotic transport is critical to regulate size in *e.g.* conifer leaf.[304] The opening and closing of stomata on leaves¶ [305] and the circadian motion of various plants and flowers[306,307] is regulated by swelling or shrinking driven by water flows. Those flows are generally actuated by active transport of solutes through specialized pumps.[308,309] Even biofilms harvest osmotic pressure gradients in the extracellular space for surface motility.[310]

In animals and human beings, a number of processes involve osmotic flows for water or volume regulation and transport: from the kidney[311] to the liver[312] and the intestine,[313] not forgetting salivary secretion.[314] Cells harvest osmotic forces in a variety of ways, most obviously to control expansion and regulate size, *e.g.* in cysts,[315] and also to regulate absorption[316] or ejection of genetic material[317] *via* small capsules. A number of processes also harvest more subtle forces in a fascinating way and we cite a few to engage the curious reader. Osmotic pressure changes may affect frequency of miniature end-plate potentials in neuromuscular junctions,[318,319] but also drive oscillatory flows for cell regeneration.[320] Electro-osmosis is harvested for uphill transport of water by insects in draught areas[321] but also more generally for epithelial transport.[322,323]

The list of examples of osmotic transport in biological systems is nearly infinite, and occurs at all possible scales from individual molecules to organs and tissues. It is pointless to attempt a thorough review. Rather, we discuss below in more detail three specific biological phenomena related to osmosis. These examples raise in particular the question of whether such phenomena may be mimicked artificially to achieve advanced osmotic transport in artificial devices.

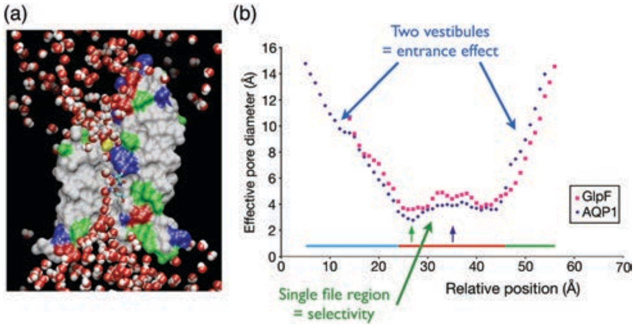


Fig. 26 The aquaporin AQP1 water channel. (a) Molecular dynamics simulation of water transversing an aquaporin channel. Snapshot from movie in ref. 333 under Creative Commons license, in complement to ref. 334. (b) Effective pore diameter of the AQP1 and GlpF channels. Pore diameters were determined with AMBER-based van der Waals radii and analysed using the program HOLE38. Reproduced and adapted from ref. 335 with permission from Springer Nature, copyright 2001.

6.2.1 Aquaporins: the ideal semi-permeable membrane. A decisive turnpoint in the study of nanoscale systems was triggered notably by the discovery of nanoscale channels in biology. One of the most famous of these channel families is the aquaporin family (the most common being AQP1 or CHIP-28, see Fig. 26a).[324,325] An aquaporin is a water-specific channel; aquaporins are present in many organs in living systems, animals, but also plants:[326] they play a central role in the human kidney (see below), are also key role players in red blood cells and many other organs,[327] and regulate water uptake in plants.[328] The striking specificity of aquaporins is that they are both highly selective to water and highly permeable. The permeability of an aquaporin was measured notably by P. Agre *et al.* to be in the range of $p_f = 11.7 \times 10^{14}$ cm^3 s^{-1} at 37 °C‖ [329,330] (with p_f here defined as $Q = p_f v_w \Delta p / k_B T$; Q is the water flux and v_w the bulk water molecular volume). This corresponds to ≈ 3 million water molecules translocating per second per bar (p_f being related to the particle flux dN/dt according to $dN/dt = (p_f/k_B T)\Delta p$). The value of the permeability of the aquaporin is much larger than that for other channels, see ref. 289 for a comparison, or what would predicted by continuum dynamics at these scales.[62,332]

Aquaporins present several intriguing features: surprisingly they are hydrophobic channels[335] and they are extremely constricted[335] – only 3 Å in diameter at the narrowest point that allows for this selectivity. An aquaporin-based membrane constitutes therefore a somewhat ideal semipermeable membrane. All of its exceptional transport properties are intimately connected to its nanoscale (and even Ångström-scale) structure – thus hinting to the striking and appealing properties of fluid flow at the nanoscale. It is thus natural to look for artificial solutions for semi-permeable membranes harvesting properly designed nanoscale structures.[290,291] For example, aquaporins present a sophisticated hourglass shape, that is believed to enhance the water permeability,[336] see Fig. 26b. Such a geometry could be readily mimicked using *e.g.* pore coatings[337] to enhance permeability of membranes.

6.2.2 Kidney: an ultra-efficient and unconventional osmotic exchanger. As a second example, we discuss the separation

---

§ Phloem is a living tissue that transports soluble compounds in particular sugar in plants.

¶ Stomata are small pores at the leaves surface that control leaf transpiration.

‖ A more recent measurement in ref. 331 suggests that this value may actually have been underestimated by a factor 5.

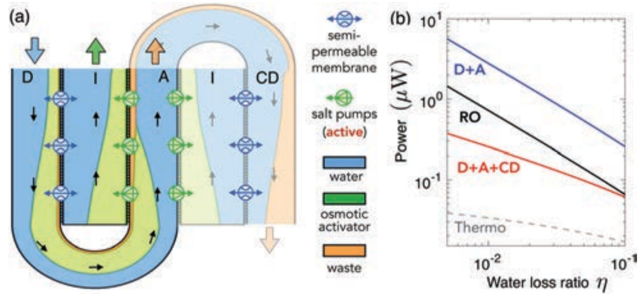


Fig. 27 The osmotic exchanger principle of the kidney. (a) Inspired from ref. 311. Water, salt and urea molar fractions are represented in various colors along the U tube (descending, D and ascending A) limbs and the interstitium (I). For visibility, the water molar fraction was divided by 100. Black arrows represent the direction of flow. A semi-permeable membrane (containing aquaporins) separates the descending limb and the interstitium, while the ascending limb contains salt pumps transporting actively the salt to the interstitium. A third limb, the collecting duct (CD, in lighter collers) also exchanges with the interstitium *via* a semipermeable membrane. The latter is crucial for the overall efficiency of the separation process. (b) Adapted from ref. 311. Power required for the functioning of the separation process as a function of the targeted water loss ratio: for the simple loop geometry (A + D), for the full serpentine (A + D + CD), as compared to the equivalent reverse osmosis process under a pressure gradient (RO). (a) and (b) are under Creative Commons Attribution 3.0 License.

process occurring in the kidneys. As we highlight, the efficiency of the kidney filtration process takes its root in a very unconventional osmotic process, and could be inspirational for future separation technologies.

Per day, the human kidney is capable of recycling about 200 L of water and 1.5 kg of salt, separating urea from water and salt at the low cost of 0.5 kJ L^{-1}[338] while readsorbing $\simeq$99% of the water input. The core of the kidney separation process lies in the millions of parallel filtration substructures called nephrons.[338] A striking feature is that the nephrons of all mammals present a precise loop geometry, the so-called Loop of Henle – in the shape of the letter ''U'' – see Fig. 27a. This loop plays a key role in the urinary concentrating mechanism and has been extensively studied from a biological and physiological point of view.[338–342]

As put forward in ref. 311, the U loop acts as an osmotic exchanger, similar in concept to a thermal exchanger – see Fig. 27a. A mixture of water, salt and urea (or any other compound to be separated) enters and flows through the tubular U loop. Water and ions may be exchanged through the tube walls with a common interconnecting media, called the interstitium. On the descending side (D), aquaporins allow for water permeation across the walls. On the ascending side (A), salt is actively pumped, using an external source of energy (in the case of the kidneys, the dissociation energy of Adenosine Tri-Phosphate, ATP). This pumped salt results in an increased salt concentration in the interstitium, higher than the concentration of salt and urea in the descending tube (D). The osmotic pressure is therefore inverted and drives water from the U tube to the interstitium across the aquaporin channels. As a result, urea is highly concentrated in the U loop, while salt and water are redirected from the interstitium towards the blood circulatory network. This U-shape geometrical design is key to the efficient operation of the separation. Note that the third limb following the U-tube plays a crucial role in enhancing the separation efficiency.[311]

One may actually estimate the working efficiency of this osmotic exchanger in a simple way, providing a lower bound on the separation ratio. It is quantified in terms of the amount of lost water $\eta = c_{\text{w}}^{\text{A,top}} v_{\text{A,top}} / c_{\text{w}}^{\text{D,top}} v_{\text{D,top}}$, where v is the flow velocity calculated at the top of the ascending (A) or descending (D) branch, and c_{w} is the concentration of water. For the system to work, water has to flow from the descending branch towards the interstitium and this requires that chemical activities obey $a_{\text{D,top}}^{\text{Water}} \geq a_{\text{I,top}}^{\text{Water}}$. The latter can be expressed simply (in the low concentration regime) in terms of molar fractions and one obtains

$$\frac{c_{\text{w}}^{\text{D,top}}}{c_{\text{w}}^{\text{D,top}} + c_{\text{s}}^{\text{D,top}} + c_{\text{waste}}^{\text{D,top}}} \geq \frac{c_{\text{w}}^{\text{I,top}}}{c_{\text{w}}^{\text{I,top}} + c_{\text{s}}^{\text{I,top}}} \tag{103}$$

where c_{w}, c_{s} and c_{waste} are respectively the concentrations of water, osmotic activator (salt) and waste. Assuming that all the osmotic activator has been reabsorbed in the upper branch yields $c_{\text{s}}^{\text{I,top}} = \frac{v_{\text{D,top}}}{v_{\text{I,top}}} c_{\text{s}}^{\text{D,top}}$. Water flow is conserved and thus $c_{\text{w}}^{\text{D,top}} v_{\text{D,top}} = c_{\text{w}}^{\text{I,top}} v_{\text{I,top}} + c_{\text{w}}^{\text{A,top}} v_{\text{A,top}}$. A lower bound for the fraction of lost water η_{lost} can then be simply deduced from eqn (103) as

$$\eta_{\text{lost}} \geq \left. \frac{c_{\text{waste}}^{\text{D,top}}}{c_{\text{s}}^{\text{D,top}} + c_{\text{waste}}^{\text{D,top}}} \right)^{n}. \tag{104}$$

with $n = 1$. For the geometry including a third reabsorbing branch, the collecting duct, see Fig. 27a, a similar reasoning yields the same result with $n = 2$. The square exponent thus leads to much smaller lost water fraction η_{lost} showing that this third branch is essential in the overall efficiency of the kidney separation. Using physiological values for the concentration, this estimate provides a prediction for water reabsorption, and thus urea separation, in the range of $\eta_{\text{lost}} \sim 1\%$, which is in excellent agreement with every-day life experience; see ref. 311. To some extent, note that the osmotic exchanger of the kidney may be compared to a forward osmosis process. However the key difference is the geometry with 3 limbs that allows for a more efficient reabsorption of water.

In fact, energy wise, this system is also shown to be far more efficient than standard reverse osmosis principles, as can be estimated within the above model, see Fig. 27b. In living systems, the nephron operates the separation of urea from water near the thermodynamic limit, $\simeq$0.2 kJ L^{-1}.[311] Yet, standard dialytic filtration systems, which are based on reverse osmosis and passive equilibration with a dialysate, require more than two orders of magnitude more energy.[343]

Some attempts to build artificial devices mimicking the nephron were reported in the literature, but they rely on biological tissues or cell mediated transport, and cannot be easily scaled up and transferred to other separation devices.[344–346]

None of the approaches so far rely on the specific geometry of the U-loop to improve the filtration process. Mimicking the separation process occurring in the kidney based on the physical perspective described above can now be foreseen using microfluidic elementary building blocks.

6.2.3 Proton pumps, chemi-osmosis and advanced ionic machinery. As a last example, we discuss proton pumps and channels, which are compelling illustrations of how Nature harvests osmotic forces to drive mechanical parts. Biological systems have developed a fascinating artillery of devices to passively and actively transport ions, namely ionic channels and ion pumps. Among these, proton pumps are canonical examples. We detail a few examples below.

Proton pumps to build proton gradients. There exists a great variety of ways to actively transport protons in biology, from combined proton-electron transfer in cytochrome oxidase (crucial for respiration[347]) to proton pumps implying the participation of ATP – the latter are called H^+-ATPases.[348] ATP-ases play a key role in bio-energetics and are ubiquitous in many forms of life and plants.[349] They include three types. The P-type ATP-ases include in particular the plasma membrane H^+-ATPase, that uses the dissociation energy of ATP to form gradients of protons. These gradients are crucial for plant movement (from phloem loading, to size regulation in the stomatal aperture, to tip growing systems** [350]). The V-type ATP-ases also use the dissociation energy of ATP to form gradients of protons. Interestingly this chemical reaction is accompanied by a rotary motion of the protein. It is central to many processes in animals,[351] from acid base balance in the kidney, pH maintenance in mechanosensory hair cells, bone resorption, tumour metastasis, sperm motility and maturation *etc.* The last type, the F-type, can work similarly to the V-type[352,353] and consumes ATP to form gradients of protons depending on aerobic conditions.[354] However it most commonly works the reverse way, *e.g.* consuming the proton gradient and synthesizing ATP, and we discuss that below.

Proton gradients harvested for energy vectorization and locomotion. The idea that osmotic gradients could be harvested for advanced functionalities was introduced as early as in the 1960s, by the seminal work of Mitchell in ref. 355. He introduced the concept of chemi-osmotic coupling, namely that a chemical reaction may be powered by the directed channeling of a specie. In the case of the F-type ATP-ase, directed motion of protons (note that the full reaction does imply the production of water on one side of the membrane) catalyzes the synthesis of ATP through a rotary motion[353] – see Fig. 28a. As the F-type catalyzes the formation of ATP (that is the vector for energy in all living systems), it is central to all forms of life.[356] The rotary motion occurring during synthesis can be harvested for artificial locomotion of inorganic devices.[357] Harvesting proton gradients for locomotion is more commonly performed not by the F-type ATP-ase but by the bacterial flagellar motor[358] – see Fig. 28b. The bacterial flagellar motor is an impressive 45 nm[358] ionic machinery at the root of bacterial locomotion *via* flagellar rotation notably in *E. coli*.[359] A proton gradient induces spontaneous transport of protons through stator parts (MotA/MotB).[360] As the transport is gated through these channels, it induces a ratchet-like motion of the rotor part of the motor.[361] The flagellum is attached to the rotor and therefore rotates. Around 1200 ions translocating per rotation generate a force at the base of the flagellar motor of about 200 pN.[358] The flagellum rotates at about 100 Hz[358] allowing *E. coli* to swim at more than 10 body lengths per second!

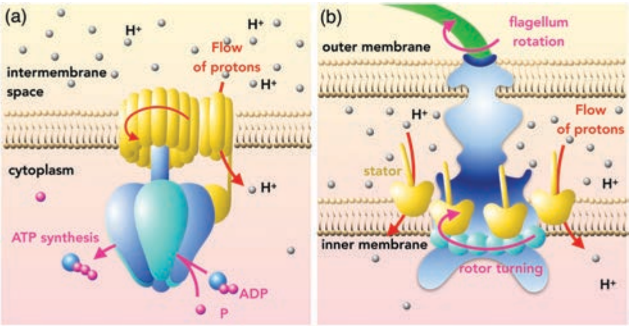


Fig. 28 Harvesting proton gradients: energy vectorization and locomotion. (a) Simplistic view of an F-type H^+-ATPase, here working as an ATP synthesis enzyme. A proton gradient is maintained between the intermembrane space and the cytoplasm by the respiratory cycle. Protons thus naturally flow inwards through the proton channel of the ATPase (in yellow). This triggers a mechanical rotation of the central element of the ATPase that in turn catalyzes the synthesis of ATP from adenosine diphosphate (ADP) and phosphate (P). (b) Simplistic view of the bacterial flagellar motor (in blue). The proton gradient transverses here the stator parts of the motor (in yellow), namely the MotA/MotB complexes. These are responsible for turning the basal rotor of the flagellar motor. As the flagellum is attached to the motor, this induces rotation of the flagellum and allows for bacterial locomotion.

Nanoscale ionic machinery. The proper function of the F-type enzyme is dependent on a subtle balance of osmotic and chemical potentials for proper function[362] and the detailed mechanisms involving motion and electric field coupling to the proton flux are still investigated.[363] Further physical insight on the detailed flows in the proton pump but more broadly on ionic channels is required to establish biomimetic principles to construct similar ionic machines with artificial material. Such physical insight is also dependent on better modeling of ion transport at the ultimate scales, with strong charge interactions, breakdown of hydrodynamics, *etc.*

6.3 Blue energy harvesting: osmotic power and capacitive mixing

As we have seen, filtration and separation of molecules requires energy input to counteract the entropy of mixing. Reversely, entropic energy harvesting may be possible by mixing molecules. The energy harvested from differences in salinity, *e.g.* by mixing sea water and fresh river water, is called blue energy. The maximal entropic energy collected by mixing volumes of sea and river water is typically 0.8 kW h m^{-3}, see ref. 364. Over the earth, counting the natural potential resources where rivers

---

** Tip-growing systems, such as pollen tubes or root hairs, continuously grow in one direction.

flow in the ocean such as the Amazonian river, a total of around 1 TW of power could be harvested, amounting to 8500 TW h in a year.[260] This is to be compared with the actual production of other renewable energies: in 2015, hydraulic energy production is ~4000 TW h, the nuclear energy around 2600 TW h, and wind and solar 1100 TW h altogether.[365] In the global energy balance, blue energy, as a renewable and non-intermittent source of energy, has thus a great potential. Here we focus on some energetic and osmotic aspects of blue energy and refer to ref. 75 for a more detailed review of the current status of blue energy harvesting.

The current attempts to harvest blue energy have essentially relied on two techniques, as sketched in Fig. 29. Pressure-retarded osmosis (PRO) harvests the natural osmotic force between sea water and river water when they are separated by a semipermeable membrane to activate a turbine to generate electricity. Reverse electro-dialysis (RED) uses diffusion gradients of salts between sea water and river water to directly generate (ionic) electric currents by separating the corresponding ion fluxes using a multi-stack of cation and anion selective membranes.[366] Both strategies rely on separation of water from ions or ions from water, and therefore require subnanoporous structures which impede the water fluxes and diminish energetic efficiency. Current PRO technologies are only able to produce up to 3 W m^{-2}, less than the critical 5 W m^{-2} for economic viability.[367] The reasons for such a low performance can be readily understood: while the osmotic pressure at the interface between sea water and fresh water is considerable and reaches 30 bars, the permeability of the semi-permeable membrane is extremely small since its pore structure is in the sub-nanometer scale to sieve ions: the power, which is the product of flow rate and pressure drop is accordingly small.

On the other hand, state-of-the-art RED achieved up to 8 W m^{-2} in controlled environment,[82] and there is an industrial hope for blue energy harvesting which is currently explored with the REDStack project in the Netherlands.[75,366]

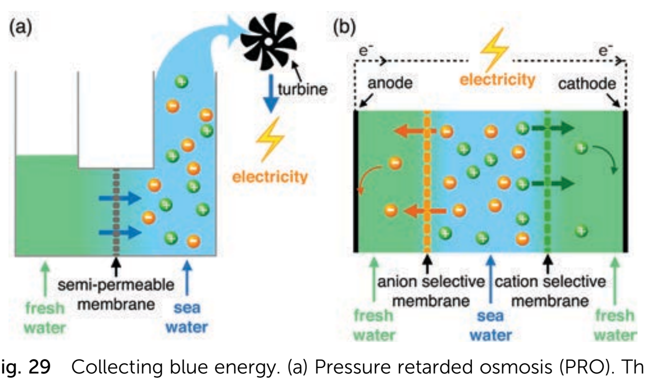


Fig. 29 Collecting blue energy. (a) Pressure retarded osmosis (PRO). The mixing of sea water and fresh water across a semi-permeable membrane drives a water flow that turns a turbine generating energy. (b) Reverse electro-dialysis (RED). Fresh and sea water are separated by stacks of alternating cation and anion selective membranes. Spontaneous diffusion induces fluxes of ions through the selective membranes, which is captured at the boundaries by reactive electrodes producing an electric current. Usually RED is performed by alternating fresh and sea water a dozen times, although only three layers are represented on the figure.

Still we note that the above power figures should not be considered as negligible, because membrane systems are quite compact and hundreds of square meters of membranes can be packed over a single ground square meter. Such performances should be compared to the 2.5 W per square meter of ground field required for a Windmill farm,[368] due to the very large required distance between windmills to prevent flow interactions. This illustrates that blue energy is actually already competitive as such in spite of the poor performances of PRO and RED.

Beyond PRO and RED, it was shown recently that new nanomaterials and nanofluidic transport constitute key assets that allow to boost considerably these performances.[73–75,87,369] Experiments across nanotubes of boron-nitride (BN), and subsequently across MoS$_2$ nanoporous membranes, reported huge ionic currents. A puzzling remark is that the BN nanotubes in the experiments of ref. 73 or the MoS$_2$ nanopores of ref. 74 are permeable to ions, in contrast to the canonical views of RED involving cation and anion selective membranes. The origin of the osmotic current was then shown to be the diffusio-osmotic ionic currents taking place at the surface of the materials, coupled to the considerable surface charge exhibited by these systems – see Fig. 30. We reported in the previous section the corresponding ionic current in eqn (55) and (56), and for a membrane constituted of N tubes of radius R, length L and surface charge Σ, the ionic current can be estimated as

$$I_{\rm osm} \approx N2\pi R\Sigma \times v_{\rm DO} \approx N\frac{2\pi R}{L}\Sigma D_{\rm DO} \times \Delta \log c_{\rm s} \qquad (105)$$

where $v_{\rm DO}$ the diffusio-osmotic water flow speed and $D_{\rm DO}$ is the diffusio-osmotic mobility, typically $D_{\rm DO} \sim k_{\rm B}T/(8\pi\eta\ell_{\rm B})$. This prediction was fully confirmed experimentally in ref. 73.

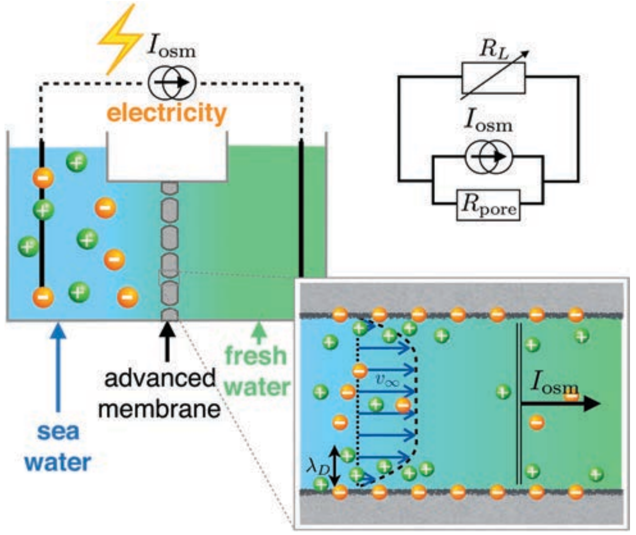


Fig. 30 Blue energy with diffusio-osmosis. A porous membrane with large and charged pores (zoom) induces a diffusio-osmotic plug-like flow with center velocity v_∞ upon a salt concentration difference ((*e.g.* here between sea and fresh water) as seen in Fig. 9). This flow drives excess charges in the electric double layer producing a net ionic current $I_{\rm osm}$ that can be harvested in a load resistance $R_{\rm L}$ – top right electric schematic.

Therefore – and this is a key asset – the blue energy does not require full selectivity of the membrane, in contrast to RED standards. Connecting the membrane to a load resistance R_L – Fig. 30, the maximum osmotic power which can be harvested is easily found to be

$$\mathscr{P} = \frac{1}{4}R_{\text{pore}}{I_{\text{osm}}}^2 \tag{106}$$

where R_{pore} is the pore or membrane resistance (that can be obtained from standard conductance measurements). Osmotic power reaches thousands of Watts per square meter in BN nanotubes, and even up to 10^6 W m^{-2} for the 2D MoS_2 due to its molecular thickness (leading to huge gradients). This estimate actually suggests to couple diffusio-osmotic current generation with an asymmetric pore geometry leading to ionic diode behavior:[75] blocking the ionic backflow thanks to the diode property allows one to boost the output power by reducing Joule losses (see details in ref. 75 and Fig. 30). Asymmetric channels were indeed shown to improve energy harvesting.[370,371]

This methodology can be readily generalized to other materials which are better suited for upscaling as compared to BN nanotubes. Key progress has been made recently in this direction.[75] Using diffusio-osmotic currents thus constitutes a promising route for improved blue energy harvesting, making it possibly relevant to industrial scale.

Beyond these membrane-based routes, the so-called "capacitive mixing" methodology is an alternative approach to harvest osmotic energy.[372] The principle is to charge and discharge an ionic capacitor by alternating flows of salty and fresh water. Capacitor plates are connected to current collectors. First (step A on Fig. 31a) salty water is flushed in, charging the capacitor plates, resulting in a closed circuit current in the load resistance. Then (step B), salty water is replaced by fresh water. When the circuit is closed again on the load resistance (step C), the capacitor plates discharge into the bulk as fresh water is less salty, resulting in a current in the opposite direction. The circuit is opened and fresh water is replaced by salty water (step D) and the cycle may start again.

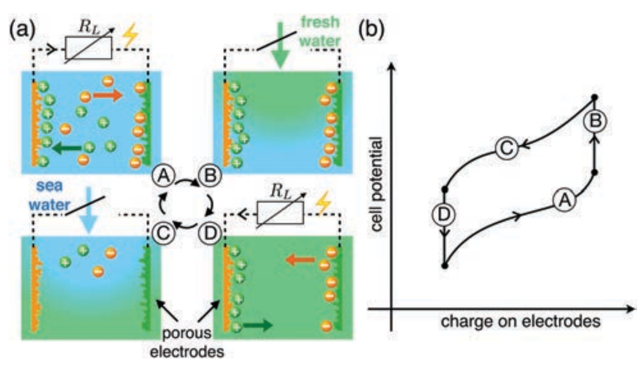


Fig. 31 Capacitive Mixing to collect blue energy. (a) Capacitive mixing cycle. Two electrodes with functionalized surfaces (such that one is positively charged in surface (green) and the other negatively charged in surface (orange)) are embedded in a fluidic device where salty water and fresh water are alternatively flushed in a cycle. (b) Associated voltage *versus* charge cycle. The cycle is described further in the text.

The power generated may be computed from the area of the cycle in the voltage/charge plane – see Fig. 31b. Typically, over 1 cycle (about 20 h[373]), 1 J per gram of carbon electrode may be collected. To compare with previous results, we estimate that 1 carbon plate of 6 × 6 cm^2 is about 1 g, such that one may recover around 0.2 W m^{-2} with capacitive mixing. Capacitive mixing therefore requires significant progress in optimizing the cell setup and the nanoporous structure to enhance performances.[374,375]

6.4 Dead-end pores: detergency, particle and liquid osmotic extraction

We have demonstrated in the previous sections how efficient diffusio-phoresis is to boost migration of particles. Combined with the ability to generate gradients of solute (in particular of salts) at small scales, it proves a method of choice in various applications to extract particles or liquids from dead-end pores. We discuss shortly two examples where diffusio-osmotic forces are harnessed.

A nice application of diffusio-osmosis was highlighted recently in the context of cleaning and the significance of rinsing in laundry detergency.[376] The question at stake here is how to extract particles which are stuck in dead-end pores in the porous matrix constituting the fabric. A simple flow resulting from mechanical action may not be able to perform this task, especially since particles buried in small pores in the interyarn pore space may not be recovered by advection because flow is channelled by larger pores (see Fig. 32a and c). Experiments then showed that rinsing with fresh water generates

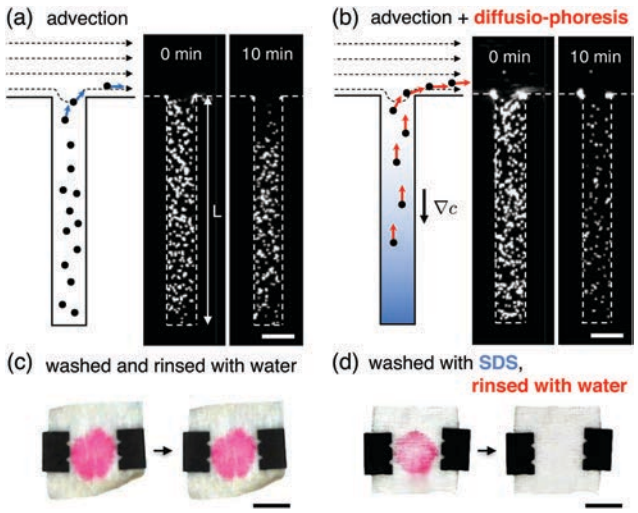


Fig. 32 Particle removal with diffusio-phoresis. Reproduced from ref. 376 with permission from the APS, copyright 2018. (a) Fluorescence image sequence showing particles in a dead-end microfluidic pore, upon advection in the main conduct. The solutions are composed of SDS at 10 mM. (b) Same as (a) with a solute gradient, where the inner pore solute concentration is 10 mM and the outer (main channel) is 0.1 mM. All scale bars are 50 μm. (c and d) A piece of cotton fabric is stained with colored colloidal particles (polystyrene latex). The piece of fabric is washed and rinsed in water (c) or washed in 10 mM SDS and rinsed with water (d) then photographed immediately after rinsing (left) and 120 s afterwards (right). All scale bars are 1 cm.

surfactant gradients at the scale of the fabric fibres and this in turn leads to diffusio-phoretic motion of the particles inside the dead-end pores. This flushing geometry echoes the osmotic shock discussed above. As highlighted in Fig. 32b and d, the gradient-induced motion allows one to extract particles from the intertwined network of pores. This suggests that, after laundering with any kind of detergent, rinsing with fresh water will allow a diffusio-phoretic push to wash out dirt and stains. It is worthwhile noting that detergency within this prospect also benefits from the log-sensing and osmotic shock effect discussed in Section 5.2: particle removal by this mechanism is effective on significantly long time scales, allowing for proper removal of the particles.

This type of mechanism based on diffusio-phoretic migration is versatile and applies to any flushing geometry. Various recent experiments considered extraction of particles – colloidal particles and oil emulsions – from dead-end pores.[209,210] Such results have also obvious applications in a different context, in geology for example, where dissolution and recrystallization at the mineral-fluid interface leads to ubiquitous salt gradients at the root of diffusio-phoretic and -osmotic transport.[76,377] In fact, flushing by fresh water was shown to enhance considerably oil recovery, a method coined as "Low salinity enhanced oil recovery".[378] While the very origin of this phenomenon is still debated, it is quite clear that diffusio-osmotic flows will play a key role in recovering biphasic mixtures using salinity gradients. Consider oil in a porous structure with typical pore radius a, as sketched in Fig. 33, where oil is trapped in dead-end pores. After a flush with fresh water, a diffusio-osmotic flow may be generated at the surface of the porous material, with velocity $v_{\mathrm{DO}} = -D_{\mathrm{DO}}\nabla \log c_{\mathrm{s}}$. Assuming first that oil is blocked, this generates a counter-balancing pressure gradient, such that the total flux is vanishing, leading to a pressure drop

$$\Delta p_{\mathrm{DO}} = -\frac{8\eta D_{\mathrm{DO}}}{a^2}\Delta[\log c_{\mathrm{s}}] \tag{107}$$

along the dead-end channel (and independent of the channel length). Putting in numbers, with a strong salinity gradient between salty water at 1 M and fresh water at 0.1 mM to fix ideas, we find Δp_{DO} = 0.07 bar for a = 100 nm and up to Δp_{DO} = 30 bars for a = 5 nm. This has to be compared to the oil–water capillary pressure expressed as $\Delta P_{\mathrm{cap}} = \frac{\gamma}{a}$ with $\gamma \simeq$ 10–20 mN m^{-1} a typical surface tension at the oil–water interface (possibly decorated with injected surfactants). While Δp_{cap} = 2 bar $> \Delta p_{\mathrm{DO}}$ for a = 100 nm, it is in the same range for a = 5 nm with $\Delta p_{\mathrm{cap}} \sim \Delta p_{\mathrm{DO}} \sim$ 40 bar. For very small pores, the pressure induced by diffusio-osmosis – which scales as $1/a^2$ – is thus able to bypass the capillary pressure, scaling as $1/a$. These simplistic estimates are made for illustration only and would deserve more detailed experimental investigations. They highlight the efficiency of diffusio-osmotic effects to extract liquids which are deeply confined within nanometric dead-end porosity.

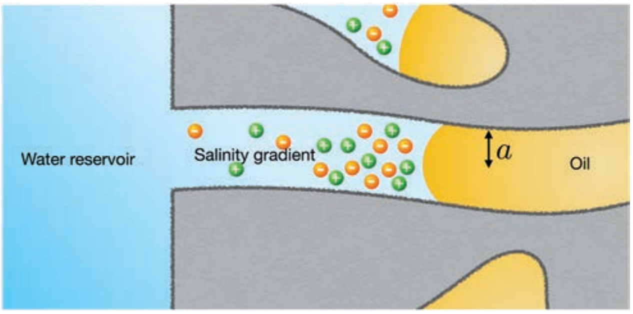


Fig. 33 Diffusio-osmotic effects for oil recovery under salinity gradients. Enhanced oil recovery is traditionally preformed by injecting sea water in the reservoir to push the oil. However flushing with fresh water slugs is known to boost the process. Gradients of salinity within dead end pores may help bypassing the capillary forces blocking the oil within the porosity.

7 Concluding remarks and perspectives

As is clear from our discussion in the previous section, osmosis is ubiquitous and crucial to an impressive number of processes, with extremely diverse manifestations. In spite of this diversity, a key and universal aspect of osmosis is that it may be interpreted as a driving force, exerted by the membrane (or a surface, or a particle's surface, and so on) on the solute particles. As we have seen in many situations in detail, we typically expect the apparent osmotic pressure to write generically as

$$\Delta\Pi_{\mathrm{app}} \simeq \langle c_{\mathrm{s}}(-\nabla\mathcal{U}_{\mathrm{eff}})\rangle$$

with $\langle\cdot\rangle$ some specific average and $\mathcal{U}_{\mathrm{eff}}$ the effective interaction potential. This mechanical perspective allows one to interpret most osmotic related phenomena (diffusio-osmosis, diffusio-phoresis, active osmosis, *etc.*). Beyond this generic description, a proper description of the forces at play is required in more specific examples, as we showed on the subtle example of the force balance in diffusio-phoresis.

Our understanding of osmotic related phenomena is still blurred by a number of open riddles. Non-equilibrium osmotic flows should be investigated, in particular to harvest non-equilibrium forces for advanced transport of species, which offer a number of promising avenues. Introducing more reliable descriptions and understanding for ionic transport at the smallest scales should also open the way to build advanced ionic detectors and ionic-powered machinery. At micrometric scales, a number of processes could be improved, harvesting the properties of specific geometries – as in the kidney – together with a clever mix of osmotic forces – as diffusio-phoresis for detergency.

Overall we still have a lot to learn from Nature and how it harvests osmosis in many forms, for separation purposes, energy storage and harvesting, *etc.* Today osmosis is usually harnessed in its most basic form, for example as the prototypical example of osmotic pressure across a semi-permeable membrane. Yet Nature has developed far more clever and far more complex examples. Mimicking the natural wonders with artificial systems is a great challenge but it opens new avenues for many outstanding societal questions that are worth the journey.

8 List of symbols

We report below symbols that are used frequently throughout the review.

a	Pore radius
$\mathcal{A}$	Membrane or pore area
b	Slip length of the surface
β	$= 1/k_\mathrm{B}T$
c_s	Solute concentration
c_w	Water or solvent concentration
$c_{+/-}$	Concentration of positive or negative ions
D_DO	Diffusio-osmotic "diffusion" coefficient
D_DP	Diffusio-phoretic "diffusion" coefficient
D_0	Colloid diffusion coefficient
D_s	Diffusion coefficient of the solute
e	Elementary charge
E	Electric field
ε	Dielectric permittivity of the fluid
η	Solvent viscosity
I_DO	Diffusio-osmotic ion current
I_e	Electric current
J_e	Exchange or Excess solute flow
J_s	Solute flow
j_s	$= J_\mathrm{s}/\mathcal{A}$ solute flow per unit area
k_B	Boltzmann's constant
K_osm	Osmotic electric mobility
L	Thickness of the membrane or length of the pore
κ_hyd	Permeance of the membrane or pore
ℓ_B	$= e^2/4\pi\varepsilon k_\mathrm{B}T$ Bjerrum length
$\mathbb{L}$	Transport matrix (or a part of the full matrix)
$\mathscr{L}_\mathrm{D}$	$= \dfrac{D_\mathrm{s}\mathcal{A}}{\eta L}$, solute permeability of the membrane or pore
$\mathscr{L}_\mathrm{hyd}$	$= \dfrac{\kappa_\mathrm{hyd}\mathcal{A}}{\eta L}$, hydrodynamic permeability of the membrane or pore
λ	Range of potential interactions
λ_D	$= 1/\sqrt{8\pi\ell_\mathrm{B}c_\mathrm{s}}$ Debye length
λ_s	$= \dfrac{D_\mathrm{s}}{k_\mathrm{B}T}$ mobility of the solute
μ_DO	Diffusio-osmotic mobility
μ_DP	Diffusio-phoretic mobility
μ_EO	Electro-osmotic mobility
μ_i^0	Chemical potential of the pure specie i
μ_i	Chemical potential of specie i
N_i	Number of molecules of specie i
ω_s	Solute "mobility" across the membrane
p	Pressure
Π	Osmotic pressure
Q	Volume flow
R	Particle size
ρ_e	Charge density
σ	Reflection or selectivity coefficient
Σ	Surface charge
T	Temperature
$\mathcal{U}(x)$	Potential barrier representing the membrane
$\boldsymbol{v}$	Velocity field of the fluid
v_DO	Diffusio-osmotic velocity
v_DP	Diffusio-phoretic velocity
v_EO	Electro-osmotic velocity
v_w	Molar volume of water
V_e	Electric potential
X	Solute molar fraction
ζ	Zeta potential

Conflicts of interest

There are no conflicts to declare

Acknowledgements

The authors are grateful for the numerous discussions they enjoyed with Marie-Laure Bocquet, Daan Frenkel, David Huang, Jérémie Palacci, Benjamin Rotenberg, Alessandro Siria, Todd Squires, Emmanuel Trizac, Patrick Warren. Authors also acknowledge pertinent feedback from Maarten Biesheuvel and Alan Kay. L. B. acknowledges funding from the European Union's H2020 Framework Programme/ERC Advanced Grant 785911-*Shadoks* and European Union's H2020 Framework Programme/FET *NanoPhlow*. S. M. and L. B. acknowledge funding from ANR project *Neptune*.

Notes and references

1 M. L. Meyer and J. van 't Hoff, *Recl. Trav. Chim. Pays-Bas*, 1890, **9**, 157–161.
2 J. van 't Hoff, *Z. Phys. Chem.*, 1890, **5**, 174–176.
3 C. Kung, B. Martinac and S. Sukharev, *Annu. Rev. Microbiol.*, 2010, **64**, 313–329.
4 D. C. Guell and H. Brenner, *Ind. Eng. Chem. Res.*, 1996, **35**, 3004–3014.
5 R. J. H. Dutrochet, *L'agent immédiat du mouvement vital dévoilé dans sa nature et dans son mode d'action, chez les vegétaux et chez les animaux*, Baillière, 1826.
6 R. H. Dutrochet, *J. Membr. Sci.*, 1995, **100**, 5–7.
7 T. Graham, *et al.*, *Philos. Trans. R. Soc. London*, 1854, **144**, 177–228.
8 A. Fick, *London Edinburgh Philos. Mag. J. Sci.*, 1855, **10**, 30–39.
9 A. Mauro, *Science*, 1957, **126**, 252–253.
10 E. Robbins and A. Mauro, *J. Gen. Physiol.*, 1960, **43**, 523–532.
11 W. Pfeffer, *Osmotische untersuchungen: studien zur zellmechanik*, W. Engelmann, 1877.
12 G. Wald, *J. Chem. Educ.*, 1986, **63**, 658.
13 J. H. van 't Hoff, *Z. Phys. Chem.*, 1887, **1**, 481–508.
14 F. G. Borg, arXiv preprint physics/0305011, 2003.

15 A. Einstein, *Ann. Phys.*, 1905, **17**, 1.

16 W. G. McMillan Jr and J. E. Mayer, *J. Chem. Phys.*, 1945, **13**, 276–305.

17 J. W. Gibbs, *Nature*, 1897, **55**, 461.

18 H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 1998.

19 E. A. Guggenheim, *Thermodynamics: An Advanced Treatment for Chemists and Physicists*, Amsterdam, North-Holland, 5th edn, 1967, ch. 4, p. 170.

20 J.-L. Barrat and J.-P. Hansen, *Basic concepts for simple and complex liquids*, Cambridge University Press, 2003.

21 J. S. Paustian, C. D. Angulo, R. Nery-Azevedo, N. Shi, A. I. Abdel-Fattah and T. M. Squires, *Langmuir*, 2015, **31**, 4402–4410.

22 J. Talen and A. Staverman, *Trans. Faraday Soc.*, 1965, **61**, 2800–2804.

23 J. Talen and A. Staverman, *Trans. Faraday Soc.*, 1965, **61**, 2794–2799.

24 J. N. Weinstein and S. R. Caplan, *Science*, 1968, **161**, 70–72.

25 C. Lee, C. Cottin-Bizonne, A.-L. Biance, P. Joseph, L. Bocquet and C. Ybert, *Phys. Rev. Lett.*, 2014, **112**, 244501.

26 C. Lee, C. Cottin-Bizonne, R. Fulcrand, L. Joly and C. Ybert, *J. Phys. Chem. Lett.*, 2017, **8**, 478–483.

27 A. Staverman, *Recl. Trav. Chim. Pays-Bas*, 1951, **70**, 344–352.

28 O. Kedem and A. Katchalsky, *Biochim. Biophys. Acta*, 1958, **27**, 229–246.

29 L. Onsager, *Phys. Rev.*, 1931, **37**, 405.

30 S. R. De Groot and P. Mazur, *Non-equilibrium thermodynamics*, Courier Corporation, 2013.

31 O. Kedem and A. Katchalsky, *J. Gen. Physiol.*, 1961, **45**, 143–179.

32 O. Kedem and A. Katchalsky, *Trans. Faraday Soc.*, 1963, **59**, 1918–1930.

33 O. Kedem and A. Katchalsky, *Trans. Faraday Soc.*, 1963, **59**, 1931–1940.

34 O. Kedem and A. Katchalsky, *Trans. Faraday Soc.*, 1963, **59**, 1941–1953.

35 E. H. Starling, *J. Phys.*, 1896, **19**, 312–326.

36 J. R. Pappenheimer, *Phys. Rev.*, 1953, **33**, 387–423.

37 R. Adamson, J. Lenz, X. Zhang, G. Adamson, S. Weinbaum and F. Curry, *J. Phys.*, 2004, **557**, 889–907.

38 J. L. Anderson and D. M. Malone, *Biophys. J.*, 1974, **14**, 957–982.

39 A. Yamauchi, Y. Shin, M. Shinozaki and M. Kawabe, *J. Membr. Sci.*, 2000, **170**, 1–7.

40 H. Fujita and Y. Kobatake, *J. Colloid Interface Sci.*, 1968, **27**, 609–615.

41 A. Hill, *Q. Rev. Biophys.*, 1979, **12**, 67–99.

42 J. D. Ferry, *Chem. Rev.*, 1936, **18**, 373–455.

43 P. M. Ray, *Plant Physiol.*, 1960, **35**, 783.

44 J. C. Giddings, E. Kucera, C. P. Russell and M. N. Myers, *J. Phys. Chem. B*, 1968, **72**, 4397–4408.

45 E. M. Renkin, *J. Gen. Physiol.*, 1954, **38**, 225–243.

46 J. L. Anderson and J. A. Quinn, *Biophys. J.*, 1974, **14**, 130.

47 T. Chou, *Phys. Rev. Lett.*, 1998, **80**, 85.

48 T. Chou, *J. Chem. Phys.*, 1999, **110**, 606–615.

49 G. S. Manning, *J. Chem. Phys.*, 1968, **49**, 2668–2675.

50 C. B. Picallo, S. Gravelle, L. Joly, E. Charlaix and L. Bocquet, *Phys. Rev. Lett.*, 2013, **111**, 244501.

51 S. Marbach, H. Yoshida and L. Bocquet, *J. Chem. Phys.*, 2017, **146**, 194701.

52 P. J. W. Debye, P. Debye, B. Eckstein, W. Barber and G. Arquette, *Equilibrium and sedimentation of uncharged particles in inhomogeneous electric fields*, Academic Press, 1954, pp. 273–285.

53 E. Grim and K. Sollner, *J. Gen. Physiol.*, 1957, **40**, 887–899.

54 J. L. Anderson, *Annu. Rev. Fluid Mech.*, 1989, **21**, 61–99.

55 B. Derjaguin, G. Sidorenkov, E. Zubashchenko and E. Kiseleva, *Prog. Surf. Sci.*, 1993, **43**, 138–152.

56 B. Derjaguin, S. Dukhin and A. Korotkova, *Prog. Surf. Sci.*, 1993, **43**, 153–158.

57 R. J. Hunter, *Foundations of colloid science*, Oxford university press, 2001.

58 D. Andelman, *Handbook of biological physics*, Elsevier, 1995, vol. 1, pp. 603–642.

59 T. M. Squires, *Fluids, Colloids and Soft Materials: An Introduction to Soft Matter Physics*, 2016, 59–79.

60 R. B. Schoch, J. Han and P. Renaud, *Rev. Mod. Phys.*, 2008, **80**, 839.

61 S. Devasenathipathy and J. Santiago, *Microscale Diagnostic Techniques*, Springer, 2005, pp. 113–154.

62 L. Bocquet and E. Charlaix, *Chem. Soc. Rev.*, 2010, **39**, 1073–1095.

63 A. S. Khair and T. M. Squires, *Phys. Fluids*, 2009, **21**, 042001.

64 T. Mouterde and L. Bocquet, *Eur. Phys. J. E: Soft Matter Biol. Phys.*, 2018, **41**, 148.

65 R. Messinger and T. Squires, *Phys. Rev. Lett.*, 2010, **105**, 144503.

66 A. J. Pascall and T. M. Squires, *Phys. Rev. Lett.*, 2010, **104**, 088301.

67 D. J. Bonthuis and R. R. Netz, *Langmuir*, 2012, **28**, 16049–16059.

68 J. L. Anderson and D. C. Prieve, *Langmuir*, 1991, **7**, 403–406.

69 A. Ajdari and L. Bocquet, *Phys. Rev. Lett.*, 2006, **96**, 186102.

70 D. M. Huang, C. Cottin-Bizonne, C. Ybert and L. Bocquet, *Phys. Rev. Lett.*, 2008, **101**, 064503.

71 D. Prieve, J. Anderson, J. Ebel and M. Lowell, *J. Fluid Mech.*, 1984, **148**, 247–269.

72 M. Lokesh, S. K. Youn and H. G. Park, *Nano Lett.*, 2018, **18**, 6679–6685.

73 A. Siria, P. Poncharal, A.-L. Biance, R. Fulcrand, X. Blase, S. T. Purcell and L. Bocquet, *Nature*, 2013, **494**, 455.

74 J. Feng, M. Graf, K. Liu, D. Ovchinnikov, D. Dumcenco, M. Heiranian, V. Nandigana, N. R. Aluru, A. Kis and A. Radenovic, *Nature*, 2016, **536**, 197.

75 A. Siria, M.-L. Bocquet and L. Bocquet, *Nat. Rev. Chem.*, 2017, **1**, 0091.

76 O. Plümper, A. Botan, C. Los, Y. Liu, A. Malthe-Sørenssen and B. Jamtveit, *Nat. Geosci.*, 2017, **10**, 685.

77 J. Palacci, C. Cottin-Bizonne, C. Ybert and L. Bocquet, *Soft Matter*, 2012, **8**, 980–994.

78 J. Palacci, B. Abécassis, C. Cottin-Bizonne, C. Ybert and L. Bocquet, *Phys. Rev. Lett.*, 2010, **104**, 138302.

79 M. Kosmulski and E. Matuevi, *J. Colloid Interface Sci.*, 1992, **150**, 291–294.

80 N. Shi, R. Nery-Azevedo, A. I. Abdel-Fattah and T. M. Squires, *Phys. Rev. Lett.*, 2016, **117**, 258001.

81 J. Fair and J. Osterle, *J. Chem. Phys.*, 1971, **54**, 3307–3316.

82 D.-K. Kim, C. Duan, Y.-F. Chen and A. Majumdar, *Microfluid. Nanofluid.*, 2010, **9**, 1215–1224.

83 Z. Zhang, X. Sui, P. Li, G. Xie, X.-Y. Kong, K. Xiao, L. Gao, L. Wen and L. Jiang, *J. Am. Chem. Soc.*, 2017, **139**, 8905–8914.

84 M. Wanunu, W. Morrison, Y. Rabin, A. Y. Grosberg and A. Meller, *Nat. Nanotechnol.*, 2010, **5**, 160.

85 Z. S. Siwy, *Adv. Funct. Mater.*, 2006, **16**, 735–746.

86 N. A. Bell, C. R. Engst, M. Ablay, G. Divitini, C. Ducati, T. Liedl and U. F. Keyser, *Nano Lett.*, 2011, **12**, 512–517.

87 M. I. Walker, K. Ubych, V. Saraswat, E. A. Chalklen, P. Braeuninger-Weimer, S. Caneva, R. S. Weatherup, S. Hofmann and U. F. Keyser, *ACS Nano*, 2017, **11**, 1340–1346.

88 L. Wang, M. S. Boutilier, P. R. Kidambi, D. Jang, N. G. Hadjiconstantinou and R. Karnik, *Nat. Nanotechnol.*, 2017, **12**, 509.

89 J. E. Hall, *J. Gen. Physiol.*, 1975, **66**, 531–532.

90 M. Mao, J. Sherwood and S. Ghosal, *J. Fluid Mech.*, 2014, **749**, 167–183.

91 J. D. Sherwood, M. Mao and S. Ghosal, *Langmuir*, 2014, **30**, 9261–9272.

92 D. V. Melnikov, Z. K. Hulings and M. E. Gracheva, *Phys. Rev. E*, 2017, **95**, 063105.

93 D. J. Rankin, L. Bocquet and D. M. Huang, 2019, arXiv.org/abs/1904.10636.

94 R. A. Sampson, *Philos. Trans. R. Soc., A*, 1891, **182**, 449–518.

95 Z. Dagan, S. Weinbaum and R. Pfeffer, *J. Fluid Mech.*, 1982, **115**, 505–523.

96 J. Happel and H. Brenner, *Low Reynolds number hydrodynamics: with special applications to particulate media*, Springer Science & Business Media, 2012, vol. 1.

97 G. Lippmann, *Compt. rend*, 1907, **145**, 104–105.

98 M. Aubert, *Ann. Chim. Phys.*, 1912, **26**, 165.

99 B. V. Derjaguin, *Surface Forces and Surfactant Systems*, Springer, 1987, pp. 17–30.

100 B. Derjaguin, N. Churaev and V. Muller, *Surface Forces*, Springer, 1987, pp. 390–409.

101 A. P. Bregulla, A. Würger, K. Günther, M. Mertig and F. Cichos, *Phys. Rev. Lett.*, 2016, **116**, 188303.

102 V. M. Barragán and S. Kjelstrup, *J. Non-Equilib. Thermodyn.*, 2017, **42**, 217–236.

103 E. Ruckenstein, *J. Colloid Interface Sci.*, 1981, **83**, 77–81.

104 R. Piazza, *J. Phys.: Condens. Matter*, 2004, **16**, S4195.

105 A. Parola and R. Piazza, *Eur. Phys. J. E: Soft Matter Biol. Phys.*, 2004, **15**, 255–263.

106 S. Duhr and D. Braun, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 19678–19682.

107 R. Piazza, *Soft Matter*, 2008, **4**, 1740–1744.

108 R. Piazza and A. Parola, *J. Phys.: Condens. Matter*, 2008, **20**, 153102.

109 J. Morthomas and A. Würger, *J. Phys.: Condens. Matter*, 2008, **21**, 035103.

110 L. Fu, S. Merabia and L. Joly, *Phys. Rev. Lett.*, 2017, **119**, 214501.

111 R. Golestanian, T. Liverpool and A. Ajdari, *New J. Phys.*, 2007, **9**, 126.

112 R. Di Leonardo, F. Ianni and G. Ruocco, *Langmuir*, 2009, **25**, 4247–4250.

113 A. Würger, *Rep. Prog. Phys.*, 2010, **73**, 126601.

114 L.-H. Yu and Y.-F. Chen, *Anal. Chem.*, 2015, **87**, 2845–2851.

115 H.-R. Jiang, H. Wada, N. Yoshinaga and M. Sano, *Phys. Rev. Lett.*, 2009, **102**, 208301.

116 C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun and S. Duhr, *Nat. Commun.*, 2010, **1**(100), 1–7.

117 T. Dau, E. Edeleva, S. Seidel, R. Stockley, D. Braun and D. E. Jenne, *Sci. Rep.*, 2016, **6**, 35413.

118 A. F. Al-Alawy and R. M. Al-Alawy, *Iraqi Journal of Chemical and Petroleum Engineering*, 2016, **17**, 53–68.

119 N. Kuipers, J. H. Hanemaaijer, H. Brouwer, J. van Medevoort, A. Jansen, F. Altena, P. van der Vleuten and H. Bak, *Desalin. Water Treat.*, 2015, **55**, 2766–2776.

120 A. P. Straub, N. Y. Yip, S. Lin, J. Lee and M. Elimelech, *Nat. Energy*, 2016, **1**, 16090.

121 B. Rotenberg and I. Pagonabarraga, *Mol. Phys.*, 2013, **111**, 827–842.

122 A. Kalra, S. Garde and G. Hummer, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 10175–10180.

123 Y. Luo and B. Roux, *J. Phys. Chem. Lett.*, 2009, **1**, 183–189.

124 T. W. Lion and R. J. Allen, *J. Chem. Phys.*, 2012, **137**, 244911.

125 H. Yoshida, S. Marbach and L. Bocquet, *J. Chem. Phys.*, 2017, **146**, 194702.

126 I. McDonald and J. Hansen, *Theory of simple liquids*, Academic Press, London, 1986, vol. 2, p. 179.

127 V. Marry, J.-F. Dufrêche, M. Jardat and P. Turq, *Mol. Phys.*, 2003, **101**, 3111–3119.

128 L. Bocquet and J.-L. Barrat, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1994, **49**, 3079.

129 H. Yoshida, H. Mizuno, T. Kinjo, H. Washizu and J.-L. Barrat, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2014, **90**, 052113.

130 H. Yoshida, H. Mizuno, T. Kinjo, H. Washizu and J.-L. Barrat, *J. Chem. Phys.*, 2014, **140**, 214701.

131 Y. Liu, R. Ganti, H. G. Burton, X. Zhang, W. Wang and D. Frenkel, *Phys. Rev. Lett.*, 2017, **119**, 224502.

132 Y. Liu, R. Ganti and D. Frenkel, *J. Phys.: Condens. Matter*, 2018, **30**, 205002.

133 R. Ganti, Y. Liu and D. Frenkel, *Phys. Rev. Lett.*, 2017, **119**, 038002.

134 R. Ganti, Y. Liu and D. Frenkel, *Phys. Rev. Lett.*, 2018, **121**, 068002.

135 K. Kiyosawa and M. Tazawa, *Protoplasma*, 1973, **78**, 203–214.

136 D. J. Bonthuis and R. Golestanian, *Phys. Rev. Lett.*, 2014, **113**, 148101.

137 A. Esfandiar, B. Radha, F. Wang, Q. Yang, S. Hu, S. Garaj, R. Nair, A. Geim and K. Gopinadhan, *Science*, 2017, **358**, 511–513.

138 W. Choi, Z. W. Ulissi, S. F. Shimizu, D. O. Bellisario, M. D. Ellison and M. S. Strano, *Nat. Commun.*, 2013, **4**, 2397.

139 E. Secchi, S. Marbach, A. Niguès, D. Stein, A. Siria and L. Bocquet, *Nature*, 2016, **537**, 210.

140 R. H. Tunuguntla, R. Y. Henley, Y.-C. Yao, T. A. Pham, M. Wanunu and A. Noy, *Science*, 2017, **357**, 792–796.

141 K. A. Mahmoud, B. Mansoor, A. Mansour and M. Khraisheh, *Desalination*, 2015, **356**, 208–225.

142 R. Nair, H. Wu, P. Jayaram, I. Grigorieva and A. Geim, *Science*, 2012, **335**, 442–444.

143 B. Radha, A. Esfandiar, F. Wang, A. Rooney, K. Gopinadhan, A. Keerthi, A. Mishchenko, A. Janardanan, P. Blake and L. Fumagalli, *et al.*, *Nature*, 2016, **538**, 222.

144 F. Fornasiero, H. G. Park, J. K. Holt, M. Stadermann, C. P. Grigoropoulos, A. Noy and O. Bakajin, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 17250–17255.

145 J. Feng, K. Liu, M. Graf, D. Dumcenco, A. Kis, M. Di Ventra and A. Radenovic, *Nat. Mater.*, 2016, **15**, 850.

146 M. Majumder, N. Chopra, R. Andrews and B. J. Hinds, *Nature*, 2005, **438**, 44.

147 J. K. Holt, H. G. Park, Y. Wang, M. Stadermann, A. B. Artyukhin, C. P. Grigoropoulos, A. Noy and O. Bakajin, *Science*, 2006, **312**, 1034–1037.

148 M. Whitby, L. Cagnon, M. Thanou and N. Quirke, *Nano Lett.*, 2008, **8**, 2632–2637.

149 J. Lee, T. Laoui and R. Karnik, *Nat. Nanotechnol.*, 2014, **9**, 317.

150 M. Langecker, V. Arnaut, T. G. Martin, J. List, S. Renner, M. Mayer, H. Dietz and F. C. Simmel, *Science*, 2012, **338**, 932–936.

151 R. Joshi, P. Carbone, F. C. Wang, V. G. Kravets, Y. Su, I. V. Grigorieva, H. Wu, A. K. Geim and R. R. Nair, *Science*, 2014, **343**, 752–754.

152 M. Majumder, A. Siria and L. Bocquet, *MRS Bull.*, 2017, **42**, 278–282.

153 R. Karnik, C. Duan, K. Castelino, H. Daiguji and A. Majumdar, *Nano Lett.*, 2007, **7**, 547–551.

154 A. Poggioli, A. Siria and L. Bocquet, *J. Phys. Chem. B*, 2019, **123**(5), 1171–1185.

155 R. E. Farmer and R. I. Macey, *Biochim. Biophys. Acta*, 1970, **196**, 53–65.

156 C. Toupin, M. Le Maguer and L. McGann, *Cryobiology*, 1989, **26**, 431–444.

157 D. B. Peckys, F. Kleinhans and P. Mazur, *PLoS One*, 2011, **6**, e23643.

158 S. Marbach and L. Bocquet, *J. Chem. Phys.*, 2017, **147**, 154701.

159 S. Marbach, D. S. Dean and L. Bocquet, *Nat. Phys.*, 2018, **14**, 1108–1113.

160 Y. Sakiyama, A. Mazur, L. E. Kapinos and R. Y. Lim, *Nat. Nanotechnol.*, 2016, **11**, 719.

161 S. Y. Noskov, S. Berneche and B. Roux, *Nature*, 2004, **431**, 830.

162 G. Eisenman and R. Horn, *J. Membr. Biol.*, 1983, **76**, 197–225.

163 P. Läuger, W. Stephan and E. Frehland, *Biochim. Biophys. Acta*, 1980, **602**, 167–180.

164 H. Schröder, *J. Chem. Phys.*, 1983, **79**, 1997–2005.

165 L. Gammaitoni, P. Hänggi, P. Jung and F. Marchesoni, *Rev. Mod. Phys.*, 1998, **70**, 223.

166 P. Reimann and P. Hänggi, *Appl. Phys. A*, 2002, **75**, 169–178.

167 J. Rousselet, L. Salome, A. Ajdari and J. Prostt, *Nature*, 1994, **370**, 446.

168 S. Marbach, N. Kavokine and L. Bocquet, *J. Chem. Phys.*, 2019, submitted.

169 R. Karnik, R. Fan, M. Yue, D. Li, P. Yang and A. Majumdar, *Nano Lett.*, 2005, **5**, 943–948.

170 E. B. Kalman, O. Sudre, I. Vlassiouk and Z. S. Siwy, *Anal. Bioanal. Chem.*, 2009, **394**, 413–419.

171 W. Guan, R. Fan and M. A. Reed, *Nat. Commun.*, 2011, **2**, 506.

172 K.-G. Zhou, K. Vasu, C. Cherian, M. Neek-Amal, J. C. Zhang, H. Ghorbanfekr-Kalashami, K. Huang, O. Marshall, V. Kravets and J. Abraham, *et al.*, arXiv preprint arXiv:1805.06390, 2018.

173 J. Moorthy, C. Khoury, J. S. Moore and D. J. Beebe, *Sens. Actuators, B*, 2001, **75**, 223–229.

174 H. Zhang, X. Hou, L. Zeng, F. Yang, L. Li, D. Yan, Y. Tian and L. Jiang, *J. Am. Chem. Soc.*, 2013, **135**, 16102–16110.

175 P. Liu, G. Xie, P. Li, Z. Zhang, L. Yang, Y. Zhao, C. Zhu, X.-Y. Kong, L. Jiang and L. Wen, *NPG Asia Mater.*, 2018, **10**, 849–857.

176 J. Liu, N. Wang, L.-J. Yu, A. Karton, W. Li, W. Zhang, F. Guo, L. Hou, Q. Cheng and L. Jiang, *et al.*, *Nat. Commun.*, 2017, **8**, 2011.

177 L. Hu, S. Gao, X. Ding, D. Wang, J. Jiang, J. Jin and L. Jiang, *ACS Nano*, 2015, **9**, 4835–4842.

178 D. C. Prieve and R. Roman, *J. Chem. Soc., Faraday Trans. 2*, 1987, **83**, 1287–1306.

179 J. Ebel, J. L. Anderson and D. Prieve, *Langmuir*, 1988, **4**, 396–406.

180 D. Velegol, A. Garg, R. Guha, A. Kar and M. Kumar, *Soft Matter*, 2016, **12**, 4686–4703.

181 F. M. Möller, F. Kriegel, M. Kieß, V. Sojo and D. Braun, *Angew. Chem., Int. Ed.*, 2017, **56**, 2340–2344.

182 A. Banerjee, I. Williams, R. N. Azevedo, M. E. Helgeson and T. M. Squires, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 8612–8617.

183 F. Morrison Jr, *J. Colloid Interface Sci.*, 1970, **34**, 210–214.

184 B. Rallabandi, F. Yang and H. A. Stone, arXiv preprint arXiv:1901.04311, 2019.

185 A. Chamolly, T. Ishikawa and E. Lauga, *New J. Phys.*, 2017, **19**, 115001.

186 U. M. Córdova-Figueroa and J. F. Brady, *Phys. Rev. Lett.*, 2008, **100**, 158303.

187 F. Jülicher and J. Prost, *Phys. Rev. Lett.*, 2009, **103**, 079801.

188 T. M. Fischer and P. Dhar, *Phys. Rev. Lett.*, 2009, **102**, 159801.

189 U. M. Córdova-Figueroa and J. F. Brady, *Phys. Rev. Lett.*, 2009, **103**, 079802.

190 U. M. Córdova-Figueroa and J. F. Brady, *Phys. Rev. Lett.*, 2009, **102**, 159802.

191 J. F. Brady, *J. Fluid Mech.*, 2011, **667**, 216–259.

192 J. L. Moran and J. D. Posner, *Annu. Rev. Fluid Mech.*, 2017, **49**, 511–540.

193 B. Sabass and U. Seifert, *J. Chem. Phys.*, 2012, **136**, 064508.

194 N. Sharifi-Mood, J. Koplik and C. Maldarelli, *Phys. Fluids*, 2013, **25**, 012001.

195 U. Córdova-Figueroa, J. Brady and S. Shklyaev, *Soft Matter*, 2013, **9**, 6382–6390.

196 H. Ohshima, T. W. Healy and L. R. White, *J. Chem. Soc., Faraday Trans. 2*, 1983, **79**, 1613–1628.

197 D. Long, J.-L. Viovy and A. Ajdari, *Phys. Rev. Lett.*, 1996, **76**, 3858.

198 H.-R. Jiang and M. Sano, *Appl. Phys. Lett.*, 2007, **91**, 154104.

199 S. Michelin and E. Lauga, *J. Fluid Mech.*, 2014, **747**, 572–604.

200 B. Abécassis, C. Cottin-Bizonne, C. Ybert, A. Ajdari and L. Bocquet, *Nat. Mater.*, 2008, **7**, 785.

201 Y. V. Kalinin, L. Jiang, Y. Tu and M. Wu, *Biophys. J.*, 2009, **96**, 2439–2448.

202 R. P. Sear and P. B. Warren, *Phys. Rev. E*, 2017, **96**, 062602.

203 J. S. Paustian, R. N. Azevedo, S.-T. B. Lundin, M. J. Gilkey and T. M. Squires, *Phys. Rev. X*, 2013, **3**, 041010.

204 R. Nery-Azevedo, A. Banerjee and T. M. Squires, *Langmuir*, 2017, **33**, 9694–9702.

205 S. Shin, O. Shardt, P. B. Warren and H. A. Stone, *Nat. Commun.*, 2017, **8**, 15181.

206 R. Guha, F. Mohajerani, M. Collins, S. Ghosh, A. Sen and D. Velegol, *J. Am. Chem. Soc.*, 2017, **139**, 15588–15591.

207 D. Florea, S. Musa, J. M. Huyghe and H. M. Wyss, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, 201322857.

208 J. Nardi, R. Bruinsma and E. Sackmann, *Phys. Rev. Lett.*, 1999, **82**, 5168.

209 A. Kar, T.-Y. Chiang, I. Ortiz Rivera, A. Sen and D. Velegol, *ACS Nano*, 2015, **9**, 746–753.

210 S. Shin, E. Um, B. Sabass, J. T. Ault, M. Rahimi, P. B. Warren and H. A. Stone, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 257–261.

211 S. Shin, J. T. Ault, P. B. Warren and H. A. Stone, *Phys. Rev. X*, 2017, **7**, 041038.

212 L. Shapiro, H. H. McAdams and R. Losick, *Science*, 2009, **326**, 1225–1228.

213 M. Osawa, D. E. Anderson and H. P. Erickson, *Science*, 2008, **320**, 792–794.

214 M. Loose, E. Fischer-Friedrich, J. Ries, K. Kruse and P. Schwille, *Science*, 2008, **320**, 789–792.

215 M. Krishnan, N. Mojarad, P. Kukura and V. Sandoghdar, *Nature*, 2010, **467**, 692.

216 C. J. Myers, M. Celebrano and M. Krishnan, *Nat. Nanotechnol.*, 2015, **10**, 886.

217 J. Agudo-Canalejo, T. Adeleke-Larodo, P. Illien and R. Golestanian, *Acc. Chem. Res.*, 2018, **51**, 2365–2372.

218 J. Lutkenhaus, *Science*, 2008, **320**, 755–756.

219 L. Rothfield, A. Taghbalout and Y.-L. Shih, *Nat. Rev. Microbiol.*, 2005, **3**, 959.

220 A. Eldar, R. Dorfman, D. Weiss, H. Ashe, B.-Z. Shilo and N. Barkai, *Nature*, 2002, **419**, 304.

221 R. Sear, arXiv:1901.00802, 2019.

222 A. Aubret, S. Ramananarivo and J. Palacci, *Curr. Opin. Colloid Interface Sci.*, 2017, **30**, 81–89.

223 P. Illien, R. Golestanian and A. Sen, *Chem. Soc. Rev.*, 2017, **46**, 5508–5518.

224 C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe and G. Volpe, *Rev. Mod. Phys.*, 2016, **88**, 045006.

225 P. H. Colberg, S. Y. Reigh, B. Robertson and R. Kapral, *Acc. Chem. Res.*, 2014, **47**, 3504–3511.

226 W. F. Paxton, K. C. Kistler, C. C. Olmeda, A. Sen, S. K. S. Angelo, Y. Cao, T. E. Mallouk, P. E. Lammert and V. H. Crespi, *J. Am. Chem. Soc.*, 2004, **126**, 13424–13431.

227 N. Mano and A. Heller, *J. Am. Chem. Soc.*, 2005, **127**, 11574–11575.

228 J. R. Howse, R. A. Jones, A. J. Ryan, T. Gough, R. Vafabakhsh and R. Golestanian, *Phys. Rev. Lett.*, 2007, **99**, 048102.

229 J. Palacci, C. Cottin-Bizonne, C. Ybert and L. Bocquet, *Phys. Rev. Lett.*, 2010, **105**, 088304.

230 I. Theurkauff, C. Cottin-Bizonne, J. Palacci, C. Ybert and L. Bocquet, *Phys. Rev. Lett.*, 2012, **108**, 268303.

231 J. Palacci, S. Sacanna, A. P. Steinberg, D. J. Pine and P. M. Chaikin, *Science*, 2013, 1230020.

232 I. Buttinoni, J. Bialké, F. Kümmel, H. Löwen, C. Bechinger and T. Speck, *Phys. Rev. Lett.*, 2013, **110**, 238301.

233 G. Mino, T. E. Mallouk, T. Darnige, M. Hoyos, J. Dauchet, J. Dunstan, R. Soto, Y. Wang, A. Rousselet and E. Clement, *Phys. Rev. Lett.*, 2011, **106**, 048102.

234 F. Ginot, I. Theurkauff, D. Levis, C. Ybert, L. Bocquet, L. Berthier and C. Cottin-Bizonne, *Phys. Rev. X*, 2015, **5**, 011004.

235 A. Aubret, M. Youssef, S. Sacanna and J. Palacci, *Nat. Phys.*, 2018, **14**, 1114–1118.

236 A. Brown and W. Poon, *Soft Matter*, 2014, **10**, 4016–4027.

237 A. T. Brown, W. C. Poon, C. Holm and J. de Graaf, *Soft Matter*, 2017, **13**, 1200–1222.

238 P. Lammert, J. Prost and R. Bruinsma, *J. Theor. Biol.*, 1996, **178**, 387–391.

239 R. Golestanian, T. B. Liverpool and A. Ajdari, *Phys. Rev. Lett.*, 2005, **94**, 220801.

240 R. Di Leonardo, L. Angelani, D. Dell'Arciprete, G. Ruocco, V. Iebba, S. Schippa, M. Conte, F. Mecarini, F. De Angelis and E. Di Fabrizio, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**(21), 9541–9545.

241 T. W. Lion and R. J. Allen, *EPL*, 2014, **106**, 34003.

242 A. P. Solon, Y. Fily, A. Baskaran, M. E. Cates, Y. Kafri, M. Kardar and J. Tailleur, *Nat. Phys.*, 2015, **11**, 673.

243 J. Rodenburg, M. Dijkstra and R. van Roij, *Soft Matter*, 2017, **13**, 8957–8963.

244 S. C. Takatori, R. De Dier, J. Vermant and J. F. Brady, *Nat. Commun.*, 2016, **7**, 10694.

245 J. Tailleur and M. Cates, *Phys. Rev. Lett.*, 2008, **100**, 218103.

246 P. G. Moerman, H. W. Moyses, E. B. Van Der Wee, D. G. Grier, A. Van Blaaderen, W. K. Kegel, J. Groenewold and J. Brujic, *Phys. Rev. E*, 2017, **96**, 032607.

247 S. Y. Reigh, P. Chuphal, S. Thakur and R. Kapral, *Soft Matter*, 2018, **14**, 6043–6057.

248 F. Ginot, I. Theurkauff, F. Detcheverry, C. Ybert and C. Cottin-Bizonne, *Nat. Commun.*, 2018, **9**, 696.

249 E. F. Keller and L. A. Segel, *J. Theor. Biol.*, 1970, **26**, 399–415.

250 M. P. Brenner, L. S. Levitov and E. O. Budrene, *Biophys. J.*, 1998, **74**, 1677–1693.

251 R. Golestanian, *Phys. Rev. Lett.*, 2012, **108**, 038303.

252 Y. Zhang, A. McMullen, L.-L. Pontani, X. He, R. Sha, N. C. Seeman, J. Brujic and P. M. Chaikin, *Nat. Commun.*, 2017, **8**, 21.

253 O. Pohl and H. Stark, *Phys. Rev. Lett.*, 2014, **112**, 238303.

254 H. Stark, *Acc. Chem. Res.*, 2018, **51**, 2681–2688.

255 L. Feng, R. Dreyfus, R. Sha, N. C. Seeman and P. M. Chaikin, *Adv. Mater.*, 2013, **25**, 2779–2783.

256 A. McMullen, M. Holmes-Cerfon, F. Sciortino, A. Y. Grosberg and J. Brujic, *Phys. Rev. Lett.*, 2018, **121**, 138002.

257 P. Sheeler and D. E. Bianchi, *Cell and molecular biology*, Wiley, New York, 1987.

258 K. H. Jensen, E. Rio, R. Hansen, C. Clanet and T. Bohr, *J. Fluid Mech.*, 2009, **636**, 371–396.

259 J. Comtet, K. H. Jensen, R. Turgeon, A. D. Stroock and A. Hosoi, *Nat. Plants*, 2017, **3**, 17032.

260 B. E. Logan and M. Elimelech, *Nature*, 2012, **488**, 313.

261 J. R. Werber, C. O. Osuji and M. Elimelech, *Nat. Rev. Mater.*, 2016, **1**, 16018.

262 E. Jones, M. Qadir, M. T. van Vliet, V. Smakhtin and S.-M. Kang, *Sci. Total Environ.*, 2019, **657**, 1343–1356.

263 W. H. Organization, W. J. W. Supply and S. M. Programme, *Progress on sanitation and drinking water: 2015 update and MDG assessment*, World Health Organization, 2015.

264 R. D. Vidic, S. L. Brantley, J. M. Vandenbossche, D. Yoxtheimer and J. D. Abad, *Science*, 2013, **340**, 1235009.

265 K. B. Gregory, R. D. Vidic and D. A. Dzombak, *Elements*, 2011, **7**, 181–186.

266 V. C. Onishi, A. Carrero-Parreno, J. A. Reyes-Labarta, E. S. Fraga and J. A. Caballero, *J. Cleaner Prod.*, 2017, **140**, 1399–1414.

267 A. Y. Hoekstra and M. M. Mekonnen, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 3232–3237.

268 A. Alkaisi, R. Mossad and A. Sharifian-Barforoush, *Energy Procedia*, 2017, **110**, 268–274.

269 G. Amy, N. Ghaffour, Z. Li, L. Francis, R. V. Linares, T. Missimer and S. Lattemann, *Desalination*, 2017, **401**, 16–21.

270 J. R. McCutcheon, R. L. McGinnis and M. Elimelech, *Desalination*, 2005, **174**, 1–11.

271 R. V. Linares, Z. Li, V. Yangali-Quintanilla, N. Ghaffour, G. Amy, T. Leiknes and J. S. Vrouwenvelder, *Water Res.*, 2016, **88**, 225–234.

272 Q. Chen, Q. Ge, W. Xu and W. Pan, *J. Membr. Sci.*, 2019, **574**, 10–16.

273 T. D. Wheeler and A. D. Stroock, *Nature*, 2008, **455**, 208.

274 R. Semiat, *Environmental Sci. Technol.*, 2008, **42**, 8193–8201.

275 Y.-M. Chao and T. Liang, *Desalination*, 2008, **221**, 433–439.

276 D. Deng, E. V. Dydek, J.-H. Han, S. Schlumpberger, A. Mani, B. Zaltzman and M. Z. Bazant, *Langmuir*, 2013, **29**, 16167–16177.

277 S. J. Kim, S. H. Ko, K. H. Kang and J. Han, *Nat. Nanotechnol.*, 2010, **5**, 297.

278 M. W. Shahzad, M. Burhan, L. Ang and K. C. Ng, *Emerging Technologies for Sustainable Desalination Handbook*, Elsevier, 2018, pp. 3–34.

279 I. V. Arámburo-Miranda and E. H. Ruelas-Ramrez, *Environ. Sci. Pollut. Res.*, 2017, **24**, 25676–25681.

280 K. Minas, E. Karunakaran, T. Bond, C. Gandy, A. Honsbein, M. Madsen, J. Amezaga, A. Amtmann, M. Templeton and C. Biggs, *et al.*, *Desalin. Water Treat.*, 2015, **55**, 2647–2668.

281 H. Liu, R. Ramnarayanan and B. E. Logan, *Environmental Sci. Technol.*, 2004, **38**, 2281–2285.

282 M. Elimelech and W. A. Phillip, *Science*, 2011, **333**, 712–717.

283 A. Mehta and A. L. Zydney, *J. Membr. Sci.*, 2005, **249**, 245–249.

284 H. B. Park, J. Kamcev, L. M. Robeson, M. Elimelech and B. D. Freeman, *Science*, 2017, **356**, eaab0530.

285 J. R. Werber, A. Deshmukh and M. Elimelech, *Environ. Sci. Technol. Lett.*, 2016, **3**, 112–120.

286 L. W. Stephens, R. Lindsay, M. Milne, A. Klein, V. Fuchs and L. Rodenburg, *Proceedings of the Water Environment Federation*, 2016, **2016**, 5423–5429.

287 L. Zeman and M. Wales, *Synthetic membranes*, 1981, **2**, 411–434.

288 G. M. Geise, H. B. Park, A. C. Sagle, B. D. Freeman and J. E. McGrath, *J. Membr. Sci.*, 2011, **369**, 130–138.

289 M. Barboiu, *Chem. Commun.*, 2016, **52**, 5657–5665.

290 Y.-X. Shen, W. C. Song, D. R. Barden, T. Ren, C. Lang, H. Feroz, C. B. Henderson, P. O. Saboe, D. Tsai and H. Yan, *et al.*, *Nat. Commun.*, 2018, **9**, 2294.

291 Z. Sun, I. Kocsis, Y. Li, Y.-M. Legrand and M. Barboiu, *Faraday Discuss.*, 2018, **209**, 113–124.

292 M. S. Mauter, I. Zucker, F. Perreault, J. R. Werber, J.-H. Kim and M. Elimelech, *Nat. Sustainability*, 2018, **1**, 166.

293 P. Nednoor, V. G. Gavalas, N. Chopra, B. J. Hinds and L. G. Bachas, *J. Mater. Chem.*, 2007, **17**, 1755–1757.

294 W.-F. Chan, E. Marand and S. M. Martin, *J. Membr. Sci.*, 2016, **509**, 125–137.

295 H. Yoshida and L. Bocquet, *J. Chem. Phys.*, 2016, **144**, 234701.

296 J. Abraham, K. S. Vasu, C. D. Williams, K. Gopinadhan, Y. Su, C. T. Cherian, J. Dix, E. Prestat, S. J. Haigh and I. V. Grigorieva, *et al.*, *Nat. Nanotechnol.*, 2017, **12**, 546.

297 K. H. Thebo, X. Qian, Q. Zhang, L. Chen, H.-M. Cheng and W. Ren, *Nat. Commun.*, 2018, **9**, 1486.

298 B. Mi, *Science*, 2014, **343**, 740–742.

299 S. Gravelle, H. Yoshida, L. Joly, C. Ybert and L. Bocquet, *J. Chem. Phys.*, 2016, **145**, 124708.

300 J. Dumais and Y. Forterre, *Annu. Rev. Fluid Mech.*, 2012, **44**, 453–478.

301 T. Broyer, *Bot. Rev.*, 1947, **13**, 1–58.

302 M. V. Thompson and N. M. Holbrook, *Plant, Cell Environ.*, 2003, **26**, 1561–1577.

303 K. H. Jensen, J. Lee, T. Bohr, H. Bruus, N. M. Holbrook and M. A. Zwieniecki, *J. R. Soc., Interface*, 2011, **8**, 1155–1165.

304 H. Rademaker, M. A. Zwieniecki, T. Bohr and K. H. Jensen, *Phys. Rev. E*, 2017, **95**, 042402.

305 T. Mansfield, A. Hetherington and C. Atkinson, *Annu. Rev. Plant Biol.*, 1990, **41**, 55–75.

306 R. L. Satter and A. W. Galston, *Annu. Rev. Plant Physiol.*, 1981, **32**, 83–110.

307 W. G. van Doorn and U. van Meeteren, *J. Exp. Bot.*, 2003, **54**, 1801–1812.

308 R. Hedrich and J. I. Schroeder, *Annu. Rev. Plant Biol.*, 1989, **40**, 539–569.

309 M. Irving, S. Ritter, A. Tomos and D. Koller, *Bot. Acta*, 1997, **110**, 118–126.

310 A. Seminara, T. E. Angelini, J. N. Wilking, H. Vlamakis, S. Ebrahim, R. Kolter, D. A. Weitz and M. P. Brenner, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 1116–1121.

311 S. Marbach and L. Bocquet, *Phys. Rev. X*, 2016, **6**, 031008.

312 K. Meyer, O. Ostrenko, G. Bourantas, H. Morales-Navarrete, N. Porat-Shliom, F. Segovia-Miranda, H. Nonaka, A. Ghaemi, J.-M. Verbavatz and L. Brusch, *et al.*, *Cell Syst.*, 2017, **4**, 277–290.

313 J. R. Thiagarajah and A. S. Verkman, *Physiology of the Gastrointestinal Tract*, Elsevier, 6th edn, 2018, pp. 1249–1272.

314 R. J. Turner and H. Sugiya, *Oral diseases*, 2002, **8**, 3–11.

315 E. Gin, E. M. Tanaka and L. Brusch, *J. Theor. Biol.*, 2010, **264**, 1077–1088.

316 C. Rauch and E. Farge, *Biophys. J.*, 2000, **78**, 3036–3047.

317 A. Evilevitch, L. Lavelle, C. M. Knobler, E. Raspaud and W. M. Gelbart, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 9292–9295.

318 J. Hubbard, S. Jones and E. Landau, *J. Phys.*, 1968, **197**, 639–657.

319 E. Furshpan, *J. Phys.*, 1956, **134**, 689–697.

320 M. Kücken, J. Soriano, P. A. Pullarkat, A. Ott and E. M. Nicola, *Biophys. J.*, 2008, **95**, 978–985.

321 J. Küppers, A. Plagemann and U. Thurm, *J. Membr. Biol.*, 1986, **91**, 107–119.

322 J. Fischbarg, J. A. Hernandez, A. A. Rubashkin, P. Iserovich, V. I. Cacace and C. F. Kusnier, *J. Membr. Biol.*, 2017, **250**, 327–333.

323 M. Dvoriashyna, A. J. Foss, E. A. Gaffney, O. E. Jensen and R. Repetto, *J. Theor. Biol.*, 2018, **456**, 233–248.

324 G. M. Preston, T. P. Carroll, W. B. Guggino and P. Agre, *Science*, 1992, **256**, 385–387.

325 K. Murata, K. Mitsuoka, T. Hirai, T. Walz, P. Agre, J. B. Heymann, A. Engel and Y. Fujiyoshi, *Nature*, 2000, **407**, 599.

326 C. Maurel, L. Verdoucq, D.-T. Luu and V. Santoni, *Annu. Rev. Plant Biol.*, 2008, **59**, 595–624.

327 P. Agre, *Proc. Am. Thorac. Soc.*, 2006, **3**, 5–13.

328 H. Javot and C. Maurel, *Ann. Bot.*, 2002, **90**, 301–313.

329 M. L. Zeidel, S. V. Ambudkar, B. L. Smith and P. Agre, *Biochemistry*, 1992, **31**, 7436–7440.

330 B. Roux, J.-Y. Lapointe and D. G. Bichet, *Med. Sci.*, 2001, **17**, 115–116.

331 A. Horner, F. Zocher, J. Preiner, N. Ollinger, C. Siligan, S. A. Akimov and P. Pohl, *Sci. Adv.*, 2015, **1**, e1400083.

332 T. Walz, B. L. Smith, M. L. Zeidel, A. Engel and P. Agre, *J. Biol. Chem.*, 1994, **269**, 1583–1586.

333 T. UIUC, Water channels in cell membranes, YouTube, https://www.youtube.com/watch?v=GSi5-y6NHjY.

334 E. Tajkhorshid, P. Nollert, M. Ø. Jensen, L. J. Miercke, J. O'connell, R. M. Stroud and K. Schulten, *Science*, 2002, **296**, 525–530.

335 H. Sui, B.-G. Han, J. K. Lee, P. Walian and B. K. Jap, *Nature*, 2001, **414**, 872.

336 S. Gravelle, L. Joly, F. Detcheverry, C. Ybert, C. Cottin-Bizonne and L. Bocquet, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, 201306447.

337 G. Pérez-Mitta, J. S. Tuninetti, W. Knoll, C. Trautmann, M. E. Toimil-Molares and O. Azzaroni, *J. Am. Chem. Soc.*, 2015, **137**, 6011–6017.

338 R. Greger and U. Windhorst, *From Cellular Mechanisms to Integration*, 1996, vol. 2.

339 P. H. Baylis, *Baillieres Clin Endocrinol. Metab.*, 1989, **3**(2), 229–578.

340 J. L. Stephenson, *Kidney Int.*, 1972, **2**, 85–94.

341 A. T. Layton, *Am. J. Physiol.*, 2010, **300**, F356–F371.

342 A. Edwards, *Am. J. Physiol.*, 2009, **298**, F475–F484.

343 Gambro, Phoenix X 36 brochure, 2009.

344 J. T. Borenstein, E. J. Weinberg, B. K. Orrick, C. Sundback, M. R. Kaazempur-Mofrad and J. P. Vacanti, *Tissue Eng.*, 2007, **13**, 1837–1844.

345 J. C. Kim, F. Garzotto, F. Nalesso, D. Cruz, J. H. Kim, E. Kang, H. C. Kim and C. Ronco, *Expert Rev. Med. Devices*, 2011, **8**, 567–579.

346 P. Armignacco, F. Garzotto, M. Neri, A. Lorenzin and C. Ronco, *Blood Purif.*, 2015, **39**, 110–114.

347 V. R. Kaila, M. I. Verkhovsky and M. Wikström, *Chem. Rev.*, 2010, **110**, 7062–7081.

348 P. L. Pedersen and E. Carafoli, *Trends Biochem. Sci.*, 1987, **12**, 146–150.

349 A. Y. Mulkidjanian, K. S. Makarova, M. Y. Galperin and E. V. Koonin, *Nat. Rev. Microbiol.*, 2007, **5**, 892.

350 M. G. Palmgren, *Annu. Rev. Plant Biol.*, 2001, **52**, 817–845.

351 T. Nishi and M. Forgac, *Nat. Rev. Mol. Cell Biol.*, 2002, **3**, 94.

352 T. Murata, I. Yamato, Y. Kakinuma, A. G. Leslie and J. E. Walker, *Science*, 2005, **308**, 654–659.

353 T. Meier, P. Polzer, K. Diederichs, W. Welte and P. Dimroth, *Science*, 2005, **308**, 659–662.

354 R. A. Capaldi and R. Aggeler, *Trends Biochem. Sci.*, 2002, **27**, 154–160.

355 P. Mitchell, *Nature*, 1961, **191**, 144–148.

356 W. Junge and N. Nelson, *Annu. Rev. Biochem.*, 2015, **84**, 631–657.

357 R. K. Soong, G. D. Bachand, H. P. Neves, A. G. Olkhovets, H. G. Craighead and C. D. Montemagno, *Science*, 2000, **290**, 1555–1558.

358 Y. Sowa and R. M. Berry, *Q. Rev. Biophys.*, 2008, **41**, 103–132.

359 M. Silverman and M. Simon, *Nature*, 1974, **249**, 73.

360 S. Kojima and D. F. Blair, *Biochemistry*, 2004, **43**, 26–34.

361 Y. Nishihara and A. Kitao, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 7737–7742.

362 R. Cereijo-Santaló, *Arch. Biochem. Biophys.*, 1972, **152**, 78–82.

363 J. H. Miller Jr, K. I. Rajapakshe, H. L. Infante and J. R. Claycomb, *PLoS One*, 2013, **8**, e74978.

364 N. Y. Yip and M. Elimelech, *Environ. Sci. Technol.*, 2012, **46**, 5230–5239.

365 I. Statistics, International Energy Agency, 2017.

366 R. A. Tufa, S. Pawlowski, J. Veerman, K. Bouzek, E. Fontananova, G. di Profio, S. Velizarov, J. G. Crespo, K. Nijmeijer and E. Curcio, *Appl. Energy*, 2018, **225**, 290–331.

367 S. E. Skilhagen, *Desalin. Water Treat.*, 2010, **15**, 271–278.

368 D. MacKay, *Sustainable Energy-without the hot air*, UIT Cambridge, 2008.

369 D. J. Rankin and D. M. Huang, *Langmuir*, 2016, **32**, 3420–3432.

370 X. Zhu, J. Hao, B. Bao, Y. Zhou, H. Zhang, J. Pang, Z. Jiang and L. Jiang, *Sci. Adv.*, 2018, **4**, eaau1665.

371 C.-Y. Lin, C. Combs, Y.-S. Su, L.-H. Yeh and Z. S. Siwy, *J. Am. Chem. Soc.*, 2019, **141**(8), 3691–3698.

372 D. Brogioli, *Phys. Rev. Lett.*, 2009, **103**, 058501.

373 D. Brogioli, R. Zhao and P. Biesheuvel, *Energy Environ. Sci.*, 2011, **4**, 772–777.

374 M. Simoncelli, N. Ganfoud, A. Sene, M. Haefele, B. Daffos, P.-L. Taberna, M. Salanne, P. Simon and B. Rotenberg, *Phys. Rev. X*, 2018, **8**, 021024.

375 M. Marino, O. Kozynchenko, S. Tennison and D. Brogioli, *J. Phys.: Condens. Matter*, 2016, **28**, 114004.

376 S. Shin, P. B. Warren and H. A. Stone, *Phys. Rev. Appl.*, 2018, **9**, 034012.

377 A. Putnis, *Science*, 2014, **343**, 1441–1442.

378 A. Lager, K. J. Webb, I. R. Collins and D. M. Richmond, *et al.*, SPE Symposium on Improved Oil Recovery, 2008.

379 Y. Oren, *Desalination*, 2008, **228**, 10–29.

380 M. E. Suss, T. F. Baumann, W. L. Bourcier, C. M. Spadaccini, K. A. Rose, J. G. Santiago and M. Stadermann, *Energy Environ. Sci.*, 2012, **5**, 9511–9519.

381 T. Kim, C. A. Gorski and B. E. Logan, *Environ. Sci. Technol. Lett.*, 2017, **4**, 444–449.

Phoretic active matter

Ramin Golestanian

- What is Diffusiophoresis
- Microscopic Theory of Diffusiophoresis
- Self-diffusiophoresis
- Stochastic Dynamics of Phoretically Active Particles
- Experiments on Self-phoresis
- Mixtures of Apolar Active Colloids: Active Molecules
- Mixtures of Apolar Active Colloids: Stability of Suspensions
- Polar Active Colloids: Moment Expansion
- Non-equilibrium Dynamics of Active Enzymes
- Phoresis on the Slow Lane: Trail-following Bacteria
- Chemotaxis and Cell Division

- Monday : 11h00-12h30
- Tuesday : 08h30-10h00
- Tuesday : 10h30-12h00

arXiv:1909.03747v1 [cond-mat.soft] 9 Sep 2019

PHORETIC ACTIVE MATTER

Ramin Golestanian

Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany
Rudolf Peierls Centre for Theoretical Physics, University of Oxford, UK

(Dated: September 10, 2019)

`These notes are an account of a series of lectures I gave at the Les Houches Summer School ‘‘Active Matter and Non-equilibrium Statistical Physics’’ during August and September 2018.`

Contents

I. WHAT IS DIFFUSIOPHORESIS

To understand diffusiophoresis we need to take a proper account of the effect of the boundaries on the solvent and the solute (Anderson, 1989; Anderson *et al.*, 1982; Anderson and Prieve, 1984; Derjaguin, 1987; Derjaguin *et al.*, 1947; Prieve *et al.*, 1982). Let us zoom in near a boundary and assume that the boundary is interacting with the solute particles with an interaction potential $W(\boldsymbol{r})$, leading to the force of $\boldsymbol{F} = -\nabla W(\boldsymbol{r})$ acting on individual particles see Fig. I.1. The continuity equation for the solute concentration $\rho(\boldsymbol{r}, t)$ can be written as

$$\partial_t \rho + \nabla \cdot \boldsymbol{J} = 0, \tag{I.1}$$

where the current density $\boldsymbol{J}(\boldsymbol{r}, t)$ is defined as

$$\boldsymbol{J} = -D\nabla\rho + \beta D\rho(-\nabla W) + \rho\boldsymbol{v}, \tag{I.2}$$

where $\beta = 1/(k_\mathrm{B}T)$ and Einstein relation is assumed between the mobility and diffusion coefficient. The corresponding governing equation equation for the solvent is given by Stokes equation

$$-\eta\nabla^2\boldsymbol{v} = -\nabla p + \boldsymbol{f}, \tag{I.3}$$

that is complemented by the incompressibility constraint

$$\nabla \cdot \boldsymbol{v} = 0. \tag{I.4}$$

In Eq. (I.3), η is the viscosity, p is the pressure, and $\boldsymbol{f}$ represents the body force density acting on the solvent. Noting that the force acting on the solute particles will be transmitted to the solvent by way of force balance for each solute particle, we can write

$$\boldsymbol{f} = \rho\boldsymbol{F} = \rho(-\nabla W). \tag{I.5}$$

This relation closes the system of equations for the solvent and the solute that should be simultaneously solved for the concentration and the velocity profiles.

In the stationary situation, the equation for the concentration reads

$$-D\nabla^2\rho + \boldsymbol{v}\cdot\nabla\rho + \beta D\nabla\cdot\boldsymbol{f} = 0. \tag{I.6}$$

Using the incompressibility constraint, we can also find an equation for the pressure, which reads

$$-\nabla^2 p + \nabla\cdot\boldsymbol{f} = 0. \tag{I.7}$$

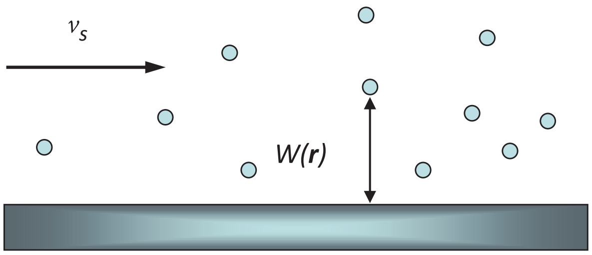


FIG. I.1 Fluid moves past a solid surface in the presence of a solute concentration gradient due to the interaction $W(\boldsymbol{r})$ between the solute particles and the solid surface. The so-called fluid slip velocity is as shown when the interaction is predominantly repulsive.

Combining Eqs. (I.6) and (I.7) yields

$$\nabla^2 \left[p - k_{\mathrm{B}}T\rho\right] + \frac{1}{\beta D}\,\boldsymbol{v}\cdot\nabla\rho = 0, \tag{I.8}$$

which does not involve the body force that incorporates the molecular interactions.

The potential W is expected to have a very short range, say σ, through which it starts from infinity—representing the impenetrability or the excluded volume effect of the surface—and decays to zero. If the wall is impenetrable to the solute particles, the normal current should be negligibly small in the vicinity of the wall, $J_\perp \simeq 0$. If the wall is impenetrable to the fluid, the normal fluid velocity should also be negligibly small near the wall, $v_\perp \simeq 0$. Then, Eq. (I.2) requires that the singular contribution in that neighborhood due to ∇W is balanced by the gradient of concentration, namely

$$-k_{\mathrm{B}}T\partial_\perp\rho + \rho\left(-\partial_\perp W\right) \simeq 0, \tag{I.9}$$

within a distance σ from the wall. This can be solved to give

$$\rho^s(\boldsymbol{r}) = \rho_{\mathrm{out}}\,\mathrm{e}^{-W(\boldsymbol{r})/k_{\mathrm{B}}T}, \tag{I.10}$$

in the "slip" region, where ρ_{out} is the concentration of the solute immediately after the wall potential has died off. Note that Eq. (I.10) implies a strong depletion of the solute particles near the wall ($\rho\,|_{\mathrm{wall}} = 0$).

The presence of the body force in the above equations implies that both the concentration and the pressure have singular behaviors in the vicinity of the wall. However, Eq. (I.8) suggests that the combination $p - k_{\mathrm{B}}T\rho$ is a smooth function through the domain of action of the wall potential, and does not entail any singular terms despite both p and ρ having singular behaviour near the wall. In fact, to be consistent with the above approximation scheme, the velocity term that represents advection should be neglected in this equation in the vicinity of the wall, or the so-called slip region.

Using this smoothness property, we can relate the pressure and concentration profiles in the slip region to the those in the outer region, as follows

$$p^s = p_{\mathrm{out}} + k_{\mathrm{B}}T(\rho^s - \rho_{\mathrm{out}}) = p_{\mathrm{out}} + k_{\mathrm{B}}T\rho_{\mathrm{out}}\left[\mathrm{e}^{-W(\boldsymbol{r})/k_{\mathrm{B}}T} - 1\right], \tag{I.11}$$

where p_{out} is the pressure just outside the slip domain, and Eq. (I.10) is used to arrive at the second form.

The above framework suggests that we can separate the slip region (where the interaction is at work) from the outer region, see Fig. I.2, work out the velocity of the solvent at the boundary between the two regions, and use it as a boundary condition for the outer problem.

A. Slip Velocity near Surfaces

The flow field inside the slip region can be determined within our approximation using the Stokes equation that ensures force balance for the component of the velocity that is parallel to the surface

$$-\eta\nabla^2 v_\parallel = -\partial_\parallel p, \tag{I.12}$$

where the body force is neglected because it is assumed to be in the perpendicular direction ($f_\parallel = 0$). Since we are interested in propulsion that is driven by concentration gradient and not external pressure

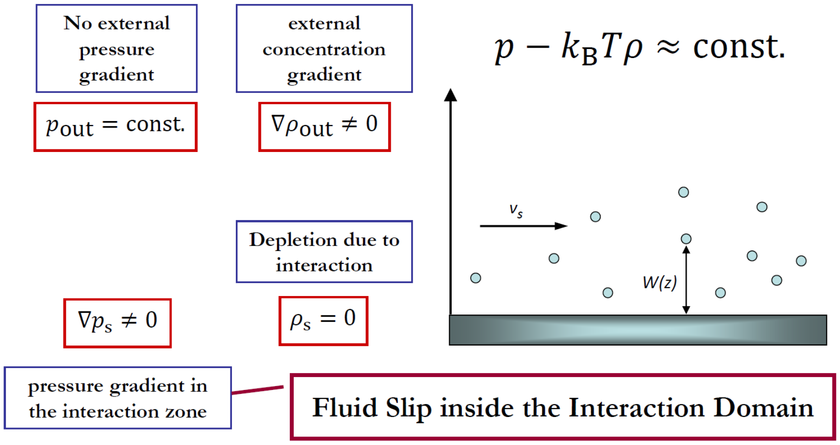


FIG. I.2 We can divide space into the slip region near the surface, where the interaction $W(\boldsymbol{r})$ is strong, and the outer region far from the surface. Whereas the pressure p and the solute concentration ρ behave very differently in the two regions, the combination $p - k_\mathrm{B}T\rho$ behaves smoothly and allows us to relate the values in the two regions, see Eq. (I.11). The pressure gradient in the slip region then generates a fluid slip velocity, which has the direction shown when the interaction is predominantly repulsive.

gradient, we assume $\partial_\| p_\mathrm{out} = 0$. Using Eq. (I.11), Eq. (I.12) gives

$$-\nabla^2 v_\| = \frac{k_\mathrm{B}T}{\eta}(\partial_\| \rho_\mathrm{out})\left[1 - \mathrm{e}^{-W(\boldsymbol{r})/k_\mathrm{B}T}\right]. \tag{I.13}$$

We note that variations in the (normal) z-direction are considerably stronger than variations in the parallel direction, we can implement a lubrication-like approximation $\nabla^2 \approx \partial_z^2$. The relevant boundary conditions are $\eta\partial_z v_\|\big|_{z\to\infty} = 0$ (zero externally applied shear rate) and $v_\|\big|_{z=0} = 0$ (no-slip condition at the wall). Integrating the first moment of Eq. (I.13) with respect to the normal coordinate z subject to the above boundary conditions, we find $-\int_0^\infty \mathrm{d}z\, z\partial_z^2 v_\| = v_\|\big|_{z\to\infty} \equiv v_s$ on the left hand side, where the slip velocity v_s is defined as the asymptotic value of the parallel velocity. We thus obtain

$$v_s = \mu\,\partial_\| \rho_\mathrm{out}, \tag{I.14}$$

where

$$\mu = \frac{k_\mathrm{B}T}{\eta}\int_0^\infty \mathrm{d}z\, z\left[1 - \mathrm{e}^{-W(z)/k_\mathrm{B}T}\right]. \tag{I.15}$$

is the *phoretic mobility* of the system. The surface slip velocity outside of the slip layer can be used together with the Stokes equation to solve for the flow field. Defining the Derjaguin length via

$$\lambda_\mathrm{D}^2 = \int_0^\infty \mathrm{d}z\, z\left[1 - \mathrm{e}^{-W(z)/k_\mathrm{B}T}\right]. \tag{I.16}$$

we can write the phoretic mobility as

$$\mu = \frac{k_\mathrm{B}T}{\eta}\lambda_\mathrm{D}^2. \tag{I.17}$$

Note that $\lambda_\mathrm{D}^2 > 0$ corresponds to cases where W is predominantly repulsive, whereas $\lambda_\mathrm{D}^2 < 0$ corresponds to cases where W is predominantly attractive. When $\lambda_\mathrm{D} \ll R$, where R is the characteristic radius of curvature of the surface, the slip boundary condition on the fluid velocity can be used.

B. Phoretic Drift Velocity of Colloidal Particles

We can now make use of the reciprocal theorem of Lorentz (Happel and Brenner, 1981; Lorentz, 1896; Stone and Samuel, 1996) to calculate the propulsion velocity directly from the surface slip. To do this, we start from the definition of stress tensor for an incompressible flow

$$\sigma_{ij} = -p\delta_{ij} + \eta(\partial_i v_j + \partial_j v_i), \tag{I.18}$$

using which, we can write the governing equations of the Stokes flow as

$$\begin{aligned} \nabla \cdot \boldsymbol{v} &= 0, && \text{(I.19)} \\ \nabla \cdot \boldsymbol{\sigma} &= 0. && \text{(I.20)} \end{aligned}$$

If we have two solutions of the above equations in the same domain $\mathcal{D}(t)$, namely, $(\boldsymbol{v}^{(1)}, \boldsymbol{\sigma}^{(1)})$ and $(\boldsymbol{v}^{(2)}, \boldsymbol{\sigma}^{(2)})$, then we know from Green's theorem that the following relation holds between them

$$\int_{\mathcal{S}} \mathrm{d}S \; \boldsymbol{n} \cdot \boldsymbol{\sigma}^{(2)} \cdot \boldsymbol{v}^{(1)} = \int_{\mathcal{S}} \mathrm{d}S \; \boldsymbol{n} \cdot \boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{v}^{(2)}, \tag{I.21}$$

where $\boldsymbol{n}$ is the normal unit vector perpendicular to the surface $\mathcal{S}$ that defines the boundary of $\mathcal{D}$. Let us now choose $(\boldsymbol{v}^{(1)}, \boldsymbol{\sigma}^{(1)})$ to be the force-free and torque-free motion of an object with a surface slip velocity boundary condition, and $(\boldsymbol{v}^{(2)}, \boldsymbol{\sigma}^{(2)})$ to describe the motion of the same object when dragged through the viscous fluid by an external force $\boldsymbol{F}^{(2)}$ with velocity $\boldsymbol{V}^{(2)}$. Since $\boldsymbol{v}^{(2)}|_{\mathcal{S}} = \boldsymbol{V}^{(2)}$ and solution (1) is force-free, then the right hand side of Eq. (I.21) vanishes. We can split the velocity of solution (1) as $\boldsymbol{v}^{(1)}|_{\mathcal{S}} = \boldsymbol{V}^{(1)} + \boldsymbol{v}_s$, where $\boldsymbol{V}^{(1)}$ is a net drift velocity for the particle and the relative velocity component is given by the surface slip velocity $\boldsymbol{v}_s$. With this composition, Eq. (I.21) gives $\boldsymbol{F}^{(2)} \cdot \boldsymbol{V}^{(1)} = -\int_{\mathcal{S}(t)} \mathrm{d}S \; \boldsymbol{n} \cdot \boldsymbol{\sigma}^{(2)} \cdot \boldsymbol{v}_s$. Considering that for a sphere of radius a we have $\boldsymbol{n} \cdot \boldsymbol{\sigma}^{(2)} = \frac{1}{4\pi R^2} \boldsymbol{F}^{(2)}$, we find the drift velocity of the force-free and torque-free sphere as

$$\boldsymbol{V} = -\frac{1}{4\pi R^2} \int_{\mathcal{S}} \mathrm{d}S \; \boldsymbol{v}_s, \tag{I.22}$$

where we have dropped the superscript (1).

For a diffusiophoretic sphere, we find

$$\boldsymbol{V} = -\frac{1}{4\pi R^2} \int_{\mathcal{S}} \mathrm{d}S \; \mu \nabla_{\parallel} \rho_{\mathrm{out}} = -\mu \nabla_{\parallel} \rho_{\mathrm{out}}^{\infty}, \tag{I.23}$$

where a solution for diffusion equation with vanishing normal flux boundary condition around a sphere has been used to perform the integration.

In a similar manner, we can show that the angular velocity of a spherical particle is given as

$$\boldsymbol{\Omega} = -\frac{3}{8\pi R^3} \int_{\mathcal{S}} \mathrm{d}S \; \hat{\boldsymbol{e}}_r \times \boldsymbol{v}_s. \tag{I.24}$$

Consider a patterned particle with axial symmetry about a given axis $\boldsymbol{n}$, with a mobility pattern that can expanded in the basis of Legendre polynomials as $\mu(\theta) = \sum_{\ell} \mu_{\ell} P_{\ell}(\cos\theta)$. Then, Eq. (I.24) yields

$$\boldsymbol{\Omega} = -\frac{3}{4} \frac{\mu_1}{R} \Big(\boldsymbol{n} \times \nabla_{\parallel} \rho_{\mathrm{out}}^{\infty} \Big), \tag{I.25}$$

which shows that the mobility pattern of the particle should have a non-vanishing first harmonic in order for diffusiophoresis to lead to an angular velocity in a concentration gradient.

II. MICROSCOPIC THEORY OF DIFFUSIOPHORESIS

The subtle interactions between solute molecules, colloids, and the incompressible solvent that lead to diffusiophoretic drift can be understood more easily from a microscopic perspective. Let us start from a reduced two-body description of the problem where we consider a colloidal particle A at position $\boldsymbol{R}$ and a solute molecule B at position $\boldsymbol{X}$, which interact via a the potential $W^{\mathrm{AB}}(\boldsymbol{R}-\boldsymbol{X})$ within the framework of the Fokker-Planck equation. The relevant governing equation for the two-body distribution reads (Agudo-Canalejo *et al.*, 2018b)

$$\begin{aligned}
\partial_t \rho_{\mathrm{AB}}(\boldsymbol{R},\boldsymbol{X},t) &= \nabla_{\boldsymbol{R}}\cdot\boldsymbol{\mu}^{\mathrm{AA}}\cdot\left[k_{\mathrm{B}}T\nabla_{\boldsymbol{R}}\rho_{\mathrm{AB}}+\left(\nabla_{\boldsymbol{R}}W^{\mathrm{AB}}\right)\rho_{\mathrm{AB}}\right]\\
&+ \nabla_{\boldsymbol{R}}\cdot\boldsymbol{\mu}^{\mathrm{AB}}\cdot\left[k_{\mathrm{B}}T\nabla_{\boldsymbol{X}}\rho_{\mathrm{AB}}+\left(\nabla_{\boldsymbol{X}}W^{\mathrm{AB}}\right)\rho_{\mathrm{AB}}\right]\\
&+ \nabla_{\boldsymbol{X}}\cdot\boldsymbol{\mu}^{\mathrm{BA}}\cdot\left[k_{\mathrm{B}}T\nabla_{\boldsymbol{R}}\rho_{\mathrm{AB}}+\left(\nabla_{\boldsymbol{R}}W^{\mathrm{AB}}\right)\rho_{\mathrm{AB}}\right]\\
&+ \nabla_{\boldsymbol{X}}\cdot\boldsymbol{\mu}^{\mathrm{BB}}\cdot\left[k_{\mathrm{B}}T\nabla_{\boldsymbol{X}}\rho_{\mathrm{AB}}+\left(\nabla_{\boldsymbol{X}}W^{\mathrm{AB}}\right)\rho_{\mathrm{AB}}\right], \qquad \text{(II.1)}
\end{aligned}$$

where the $\boldsymbol{\mu}$'s are the relevant mobility coefficients that account for the hydrodynamic interactions. Integrating over $\boldsymbol{X}$ yields

$$\begin{aligned}
\partial_t \rho_{\mathrm{A}}(\boldsymbol{R},t) &= \nabla_{\boldsymbol{R}}\cdot\int_{\boldsymbol{X}}\boldsymbol{\mu}^{\mathrm{AA}}(\boldsymbol{R},\boldsymbol{X})\cdot\left[k_{\mathrm{B}}T\nabla_{\boldsymbol{R}}\rho_{\mathrm{AB}}+\left(\nabla_{\boldsymbol{R}}W^{\mathrm{AB}}\right)\rho_{\mathrm{AB}}\right]\\
&+ \nabla_{\boldsymbol{R}}\cdot\int_{\boldsymbol{X}}\boldsymbol{\mu}^{\mathrm{AB}}(\boldsymbol{R},\boldsymbol{X})\cdot\left[k_{\mathrm{B}}T\nabla_{\boldsymbol{X}}\rho_{\mathrm{AB}}+\left(\nabla_{\boldsymbol{X}}W^{\mathrm{AB}}\right)\rho_{\mathrm{AB}}\right], \qquad \text{(II.2)}
\end{aligned}$$

which is not a closed equation because it involves both single-body and two-body distributions. Assuming the solution is dilute and B particles are point-like, $\boldsymbol{\mu}^{\mathrm{AA}}$ will not depend on $\boldsymbol{X}$; in fact, it will be a constant provided there are no boundaries in the system. Then we find

$$\partial_t \rho_{\mathrm{A}}(\boldsymbol{R},t) = k_{\mathrm{B}}T\nabla_{\boldsymbol{R}}\cdot\boldsymbol{\mu}^{\mathrm{AA}}\cdot\nabla_{\boldsymbol{R}}\rho_{\mathrm{A}}+\nabla_{\boldsymbol{R}}\cdot\int_{\boldsymbol{X}}\left(\boldsymbol{\mu}^{\mathrm{AB}}-\boldsymbol{\mu}^{\mathrm{AA}}\right)\cdot\left[k_{\mathrm{B}}T\nabla_{\boldsymbol{X}}\rho_{\mathrm{AB}}+\left(\nabla_{\boldsymbol{X}}W^{\mathrm{AB}}\right)\rho_{\mathrm{AB}}\right], \qquad \text{(II.3)}$$

where we have made use of $\nabla_{\boldsymbol{R}}W^{\mathrm{AB}}=-\nabla_{\boldsymbol{X}}W^{\mathrm{AB}}$ and ignored a boundary term. To close the hierarchy, we use a product approximation

$$\rho_{\mathrm{AB}}(\boldsymbol{R},\boldsymbol{X};t)=\rho_{\mathrm{A}}(\boldsymbol{R},t)\rho_{\mathrm{B}}(\boldsymbol{X},t)\ \mathrm{e}^{-W^{\mathrm{AB}}(\boldsymbol{R}-\boldsymbol{X})/k_{\mathrm{B}}T}, \qquad \text{(II.4)}$$

which gives us the following simplified result:

$$\partial_t \rho_{\mathrm{A}}(\boldsymbol{R},t) = k_{\mathrm{B}}T\nabla_{\boldsymbol{R}}\cdot\boldsymbol{\mu}^{\mathrm{AA}}\cdot\nabla_{\boldsymbol{R}}\rho_{\mathrm{A}}+k_{\mathrm{B}}T\nabla_{\boldsymbol{R}}\cdot\left[\int_{\boldsymbol{X}}\left(\boldsymbol{\mu}^{\mathrm{AB}}-\boldsymbol{\mu}^{\mathrm{AA}}\right)\cdot\left(\nabla_{\boldsymbol{X}}\rho_{\mathrm{B}}\right)\ \mathrm{e}^{-W^{\mathrm{AB}}(\boldsymbol{R}-\boldsymbol{X})/k_{\mathrm{B}}T}\rho_{\mathrm{A}}(\boldsymbol{R},t)\right], \qquad \text{(II.5)}$$

Noting that the mobilities are divergence-free due to the incompressibility constraint and using $\boldsymbol{\mu}^{\mathrm{AB}}=\left(\boldsymbol{\mu}^{\mathrm{BA}}\right)^T$, we find

$$\begin{aligned}
\partial_t \rho_{\mathrm{A}}(\boldsymbol{R},t) &= k_{\mathrm{B}}T\nabla_{\boldsymbol{R}}\cdot\boldsymbol{\mu}^{\mathrm{AA}}\cdot\nabla_{\boldsymbol{R}}\rho_{\mathrm{A}}\\
&+ k_{\mathrm{B}}T\nabla_{\boldsymbol{R}}\cdot\left\{\left[\int_{\boldsymbol{X}}\left(\mathrm{e}^{-W^{\mathrm{AB}}(\boldsymbol{R}-\boldsymbol{X})/k_{\mathrm{B}}T}-1\right)\left(\boldsymbol{\mu}^{\mathrm{AB}}-\boldsymbol{\mu}^{\mathrm{AA}}\right)\cdot\nabla_{\boldsymbol{X}}\rho_{\mathrm{B}}\right]\rho_{\mathrm{A}}\right\}, \qquad \text{(II.6)}
\end{aligned}$$

which is in the form of a drift-diffusion equation

$$\partial_t \rho_{\mathrm{A}}(\boldsymbol{R},t) = \nabla_{\boldsymbol{R}}\cdot\boldsymbol{D}_{\mathbf{c}}\cdot\nabla_{\boldsymbol{R}}\rho_{\mathrm{A}}-\nabla_{\boldsymbol{R}}\cdot\left[\boldsymbol{V}(\boldsymbol{R})\rho_{\mathrm{A}}\right], \qquad \text{(II.7)}$$

with the diffusivity tensor

$$\boldsymbol{D}_{\mathbf{c}} = k_{\mathrm{B}}T\boldsymbol{\mu}^{\mathrm{AA}}, \qquad \text{(II.8)}$$

and the phoretic drift velocity

$$\boldsymbol{V}(\boldsymbol{R}) = -k_{\rm B}T \int_{\boldsymbol{X}} \left[{\rm e}^{-W^{\rm AB}(\boldsymbol{R}-\boldsymbol{X})/k_{\rm B}T} - 1 \right] \left[\boldsymbol{\mu}^{\rm AB}(\boldsymbol{R},\boldsymbol{X}) - \boldsymbol{\mu}^{\rm AA} \right] \cdot \nabla_{\boldsymbol{X}} \rho_{\rm B}. \tag{II.9}$$

For a spherical colloid of radius R, we have the hydrodynamic mobility tensors as

$$\boldsymbol{\mu}^{\rm AA} = \frac{1}{6\pi\eta R}\boldsymbol{I}, \tag{II.10}$$

and

$$\boldsymbol{\mu}^{\rm AB} = \frac{1}{6\pi\eta R}\left[\frac{1}{4}\left(3\frac{R}{r} + \frac{R^3}{r^3}\right)(\boldsymbol{I} - \hat{\boldsymbol{e}}_r\hat{\boldsymbol{e}}_r) + \frac{1}{2}\left(3\frac{R}{r} - \frac{R^3}{r^3}\right)\hat{\boldsymbol{e}}_r\hat{\boldsymbol{e}}_r \right], \tag{II.11}$$

where $r = |\boldsymbol{R} - \boldsymbol{X}|$ is the distance between A and B, and $\hat{\boldsymbol{e}}_r$ is the radial unit vector pointing from A to B. Combining both, we obtain

$$\boldsymbol{\mu}^{\rm AB} - \boldsymbol{\mu}^{\rm AA} = \frac{1}{6\pi\eta R}\left[\left(-1 + \frac{3}{4}\frac{R}{r} + \frac{1}{4}\frac{R^3}{r^3}\right)(\boldsymbol{I} - \hat{\boldsymbol{e}}_r\hat{\boldsymbol{e}}_r) + \left(-1 + \frac{3}{2}\frac{R}{r} - \frac{1}{2}\frac{R^3}{r^3}\right)\hat{\boldsymbol{e}}_r\hat{\boldsymbol{e}}_r \right], \tag{II.12}$$

which is to be inserted in Eq. (II.9). As the expression of the integrand in Eq. (II.9) involves ${\rm e}^{-W^{\rm AB}(r)/k_{\rm B}T} - 1$, the expression in Eq. (II.12) can be expanded near $r = R$ when we are dealing with relatively short-range interactions. Setting $r = R + \delta$, we find

$$\boldsymbol{\mu}^{\rm AB} - \boldsymbol{\mu}^{\rm AA} = -\frac{\delta}{4\pi\eta R^2}(\boldsymbol{I} - \hat{\boldsymbol{e}}_r\hat{\boldsymbol{e}}_r) + O(\delta^2), \tag{II.13}$$

which yields

$$\boldsymbol{V}(\boldsymbol{R}) = -\frac{k_{\rm B}T}{\eta} \times \overbrace{\int_0^\infty {\rm d}\delta\, \delta \left[1 - {\rm e}^{-W^{\rm AB}(\delta)/k_{\rm B}T}\right]}^{\lambda_{\rm D}^2} \times \overbrace{\int \frac{{\rm d}\Omega}{4\pi}\, (\boldsymbol{I} - \hat{\boldsymbol{e}}_r\hat{\boldsymbol{e}}_r)\cdot\nabla_{\boldsymbol{X}}\rho_{\rm B}(R,\Omega)}^{\nabla\rho_{\rm B}^\infty} = -\mu\nabla\rho_{\rm B}^\infty, \tag{II.14}$$

where Ω represents the solid angle and the phoretic mobility μ is given as defined in Eq. (I.17) above. When there is no separation of length scales between the range of the interaction and the radius of the sphere, the full form of the expression in Eqs. (II.9) and (II.12) should be used.

III. SELF-DIFFUSIOPHORESIS

Since diffusiophoresis is force-free—as are all other interfacial phoretic transport mechanisms—it can be used to make self-propelled particles or microswimmers, if the system generates the gradient internally (Golestanian *et al.*, 2005). If we consider the case with a small Peclet number, namely, ${\rm Pe} = \frac{VR}{D} \ll 1$, we can decouple the reaction-diffusion equation that governs the dynamics of the solute molecules from the Stokes equation that governs the dynamics of the (viscous) solvent. The case with finite Peclet number poses additional technical complexities (Michelin and Lauga, 2014).

Since the time scale for solute diffusion around the colloid is considerably shorter than the typical time scale for colloid movement, we can consider the concentration of the solute, $C(\boldsymbol{r})$, to be a quasi-stationary solution of the following reaction-diffusion equation

$$-D\nabla^2 C = 0, \tag{III.1}$$

subject to the boundary condition on the surface of the colloid at position $\boldsymbol{r}_s$

$$-D\boldsymbol{n}\cdot\nabla C|_{\boldsymbol{r}=\boldsymbol{r}_s}=\alpha(\boldsymbol{r}_s), \tag{III.2}$$

where the **activity** α gives the normal flux of solute particles on the surface, and is a measure of the *nonequilibrium activity* of the system. The resulting solution for the concentration can be used together with the slip boundary condition

$$\boldsymbol{v}_s=\mu(\boldsymbol{r}_s)\,(\boldsymbol{I}-\boldsymbol{n}\boldsymbol{n})\cdot\nabla C(\boldsymbol{r}_s), \tag{III.3}$$

to obtain the propulsion (or swimming) velocity of spherical colloids as

$$\boldsymbol{V}=-\frac{1}{4\pi R^2}\int_S \mathrm{d}S\;\mu(\boldsymbol{r}_s)\,(\boldsymbol{I}-\boldsymbol{n}\boldsymbol{n})\cdot\nabla C(\boldsymbol{r}_s), \tag{III.4}$$

using Eq. (I.22). Here, we have allowed for position dependent **mobility** $\mu(\boldsymbol{r}_s)$, which is a local measure of the fluid *response* to the concentration gradient. Note that α and μ can each be positive or negative.

For an axially symmetric distribution of activity, which can be achieved via coatings of catalytic patches with specific patterns, we can describe the activity as an expansion in appropriate harmonic modes (Legendre polynomials), namely, $\alpha(\theta)=\sum_\ell \alpha_\ell P_\ell(\cos\theta)$. The solution for the concentration profile will read $C(r,\theta)=C^\infty+\frac{R}{D}\sum_\ell \frac{\alpha_\ell}{\ell+1}\left(\frac{R}{r}\right)^{\ell+1}P_\ell(\cos\theta)$. Using the mobility profile $\mu(\theta)=\sum_\ell \mu_\ell P_\ell(\cos\theta)$, we can find the following expression for the propulsion velocity

$$\boldsymbol{V}=-\frac{\hat{\boldsymbol{e}}_z}{D}\sum_\ell\left(\frac{\ell+1}{2\ell+3}\right)\alpha_{\ell+1}\left(\frac{\mu_\ell}{2\ell+1}-\frac{\mu_{\ell+2}}{2\ell+5}\right). \tag{III.5}$$

This expression demonstrates the level of symmetry breaking in activity and mobility that is necessary to achieve self-propulsion, as a manifestation of the celebrated Curie principle (Curie, 1894). Calculations performed for other shapes have revealed that geometry can also play a key role in providing the necessary symmetry breaking in combination with activity and mobility (Golestanian *et al.*, 2007; Ibrahim *et al.*, 2018). The symmetry breaking can also be achieved via shape asymmetry (Michelin and Lauga, 2015; Popescu *et al.*, 2011; Reigh *et al.*, 2018) or even spontaneously (Michelin *et al.*, 2013). These effects have also been investigated and verified using Stochastic Rotation Dynamics (SRD) simulations (Reigh *et al.*, 2018; Rückner and Kapral, 2007).

If we have more than one species of chemicals, as it is common with the case of catalytic chemical reactions with several reactants and products, the above calculation should be done for all species k, and the resulting expression for the slip velocity will be a superposition of all the contributions in the form of

$$\boldsymbol{v}_s=\sum_k \mu^{(k)}(\boldsymbol{r}_s)\,(\boldsymbol{I}-\boldsymbol{n}\boldsymbol{n})\cdot\nabla C^{(k)}(\boldsymbol{r}_s). \tag{III.6}$$

For independent species, $C^{(k)}\sim\alpha^{(k)}/D_k$, and the resulting propulsion velocity will be a superposition of the different contributions. For species that are interlinked through catalytic reactions, the resulting correlations will be reflected in the result. This suggests that the direction of propulsion is in general quite sensitive to the details and can even change for the same system under different conditions.

The surface slip velocity profile will lead to a hydrodynamic flow field generated in the vicinity of these swimmers. In free space in 3D, the flow profile decays as $1/r^3$ for Janus particles that are fore-aft symmetric (Golestanian *et al.*, 2005, 2007) whereas a profile that is not fore-aft symmetric will lead to a stronger flow that decays as $1/r^2$ (Jülicher and Prost, 2009). When the swimmers are in contact with a surface, the force monopole that they experience as a result of this contact will change the velocity profile so that it only decays as $1/r$. In Sec. V below we discuss the measured flow field around the

Pt-PS catalytic Janus swimmer.

IV. STOCHASTIC DYNAMICS OF PHORETICALLY ACTIVE PARTICLES

We now investigate how phoretic activity modifies the stochastic dynamics of a colloid (Golestanian, 2009). In essence we would like to know how we can quantify the behaviour of such active colloids and identify deviations from Brownian motion. We can describe the dynamics of a spherical colloid and the chemical field around it in the comoving frame of reference by solving the relevant diffusion equation for the concentration profile of the solute particles

$$\partial_t C(\boldsymbol{r},t) - D\nabla^2 C(\boldsymbol{r},t) = \alpha(\theta,\phi,t)\delta(r-R), \tag{IV.1}$$

where $\alpha(\theta,\phi,t)$ is the surface activity function of the sphere. The time dependent solution to this equation can be inserted into $\boldsymbol{V}(t) = -\int \frac{d\Omega}{4\pi}\ \mu\nabla_{\|}C(R,\theta,\phi,t)$ to obtain the instantaneous velocity of the colloid. The axis of symmetry of the colloid, which points to the direction of propulsion is defined by the unit vector $\mathbf{n}(t) = (\sin\theta_n(t)\cos\phi_n(t), \sin\theta_n(t)\sin\phi_n(t), \cos\theta_n(t))$. The stochastic dynamics of $\mathbf{n}(t)$ due to rotational diffusion causes the cloud of solute particles to constantly redistribute, which will in turn make the velocity of the active colloid fluctuate. We can represent the axially symmetric activity function in terms of the spherical harmonics as

$$\alpha(\theta,\phi,t) = \sum_{\ell,m}\left(\frac{4\pi}{2\ell+1}\right)\alpha_\ell\ Y^*_{\ell m}(\theta_n(t),\phi_n(t))Y_{\ell m}(\theta,\phi). \tag{IV.2}$$

Once we determine the instantaneous velocity, we can calculate the mean-squared displacement via $\Delta L^2(t) = \int_0^t dt_1 \int_0^t dt_2 \left\langle \boldsymbol{V}(t_1)\cdot\boldsymbol{V}(t_2)\right\rangle$.

Equation (IV.1) only gives the average density, and the linear relation between the velocity and the concentration profile suggests that in order to calculate velocity correlations we need to incorporate the density fluctuations as well. To this end, we start from the Langevin equation $\dot{\boldsymbol{r}}_i(t) = \boldsymbol{u}_i(t)$ for the i-th particle whose position is described by $\boldsymbol{r}_i(t)$ and is subject to a random noise $\boldsymbol{u}_i(t)$, which has a Gaussian distribution $P[\boldsymbol{u}] = \exp\left[-\frac{1}{4D}\sum_i\int \mathrm{d}t\ \boldsymbol{u}_i(t)^2\right]$ controlled by the diffusion coefficient. We define a stochastic density $\hat{C}(\boldsymbol{r},t) = \sum_i \delta^3(\boldsymbol{r}-\boldsymbol{r}_i(t))$, which can be seen to satisfy Eq. (IV.1) with a noise term $\hat{Q}(\boldsymbol{r},t)$ added to the right hand side. Using the distribution $P[\boldsymbol{u}]$, we can calculate the moments of the noise term, and show that $\langle\hat{Q}(\boldsymbol{r},t)\rangle = 0$ and $\langle\hat{Q}(\boldsymbol{r},t)\hat{Q}(\boldsymbol{r}',t')\rangle = 2D(-\nabla^2)\delta^3(\boldsymbol{r}-\boldsymbol{r}')\delta(t-t')C(\boldsymbol{r},t)$, where $C(\boldsymbol{r},t) = \langle\hat{C}(\boldsymbol{r},t)\rangle$.

The dynamics of the system involves a number of different regimes due to the existence of a number of intrinsic time scales. The rotational diffusion time, $\tau_r = 4\pi\eta R^3/k_{\mathrm{B}}T$, controls the changes in the orientation of the sphere. The characteristic diffusion time of the chemicals around the sphere $\tau_d = R^2/D$, where $D = k_{\mathrm{B}}T/(6\pi\eta a)$ depends on the radius of the solute particles a. This time scale sets the relaxation time of the redistribution of the particles around the sphere when it changes orientation. Finally, the hydrodynamic time that controls the crossover between the inertial and the viscous regimes is given as $\tau_h = R^2/\nu$, where $\nu = \eta/\rho$ is the kinematic viscosity of water that involves the mass density ρ. We can write the time scales (for water at room temperature and using a typical value of $a = 1$ $\mathring{A}$) in the following convenient forms: $\tau_h = 10^{-6}$ $(R/1\mu\mathrm{m})^2$ s, $\tau_d = 10^{-3}$ $(R/1\mu\mathrm{m})^2$ s, and $\tau_r = 3$ $(R/1\mu\mathrm{m})^3$ s. This shows that we have a clear separation of time scale with $\tau_h \ll \tau_d \ll \tau_r$, and thus a number of different dynamical regimes in between these scales.

There are two independent mechanisms driving stochasticity: (i) density fluctuations, which are relevant for $t \sim \tau_d$ and for both symmetric (apolar) and asymmetric (polar) coatings of the colloid, and (ii) rotational diffusion, which is relevant for asymmetric particles when $t \sim \tau_r$. The interplay between these mechanisms will lead to a number of different dynamical regimes with distinct features, including anomalous diffusion and memory effects, which we highlight below.

A. Anomalous Diffusion

In the the intermediate regime where $\tau_h < t < \tau_d$ a symmetric particle can instantaneously propel itself because of polarization of the cloud of solutes due to density fluctuations. This motion, however, will be decorrelated via density fluctuations themselves, leading to fluctuations without symmetry breaking. We can use a scaling argument to characterize the nature of this anomalous dynamics. To build a scaling relation, we consider $\Delta L^2 \sim v(t)^2 t^2$, and insert $v \sim \mu \nabla C \sim \mu \delta C / R$, to find $\Delta L^2 \sim \mu^2 \langle \delta C(t) \delta C(0) \rangle t^2 / R^2$. The density auto-correlation function can be written as $\langle \delta C(t) \delta C(0) \rangle = \langle \delta C^2 \rangle k(t)$, involving the density fluctuations $\langle \delta C^2 \rangle$ and the kernel $k(t)$ that controls the relevant relaxation mode. Here, relaxation is controlled by diffusion, hence $k(t) \sim 1/(Dt)^{d/2}$ in d-dimensions, and the number fluctuations are controlled by the average number of particles ($\langle \delta N^2 \rangle \sim N_{\text{ave}}$, as inherent to any Poisson process), which yields $\langle \delta C^2 \rangle \sim C_{\text{ave}}$. The average density is controlled by the average particle production rate (per unit area) α_0 as $C_{\text{ave}} \sim (\alpha_0 R^{d-1} t)/R^d$. Putting these all together, we find that the fluctuations exhibit anomalous diffusion

$$\Delta L^2 \sim \frac{\alpha_0 \mu^2}{R^3 D^{d/2}}\, t^{3-d/2} \qquad (t \ll \tau_d). \tag{IV.3}$$

This expression shows that the active velocity fluctuations are controlled by two mechanisms: particle production (that controls the density fluctuations) and diffusion of the produced particles. The exponent $3 - d/2$ indicates superdiffusive behaviour for $d < 4$. At time scales longer than τ_d, there is a crossover to diffusive behaviour

$$\Delta L \sim \frac{\alpha_0 \mu^2}{R^{d-1} D^2}\, t \qquad (t \gg \tau_d). \tag{IV.4}$$

B. Memory Effect

For a given time dependent orientation trajectory, we can calculate the propulsion velocity of the colloid as a function of time by solving Eq. (IV.1) without the noise term. This calculation reveals that the propulsion velocity of the colloid at any instant of time depends on the recent history of the orientation as

$$\mathbf{v}(t) = \frac{v_0}{\tau_d} \int_{-\infty}^{t} \mathrm{d}t' \mathcal{M}(t - t')\, \mathbf{n}(t'), \tag{IV.5}$$

where $v_0 = -\alpha_1 \mu/(3D)$ is the mean propulsion velocity, and the *memory* kernel is given as $\mathcal{M}(t) = \frac{2}{\pi} \int_0^\infty du \frac{u^{3/2}}{(u^2+4)} e^{-u(t/\tau_d)}$, with asymptotic behaviors $\mathcal{M}(t) \sim t^{-1/2}$ for $t \ll \tau_d$ and $\mathcal{M}(t) \sim t^{-5/2}$ for $t \gg \tau_d$. Note that the propulsion velocity is controlled by the $\ell = 1$ term (α_1) in the surface activity profile.

Since the rotational diffusion of the colloid randomizes its orientation over the time scale τ_r, the velocity autocorrelation function takes on the form of a convolution between two memory kernels and the orientation autocorrelation function. Consequently, the mean-squared displacement will have three different regimes. We find the asymptotic form of

$$\Delta L^2 \sim v_0^2 t^2 (t \ll \tau_d \ll \tau_r), \tag{IV.6}$$

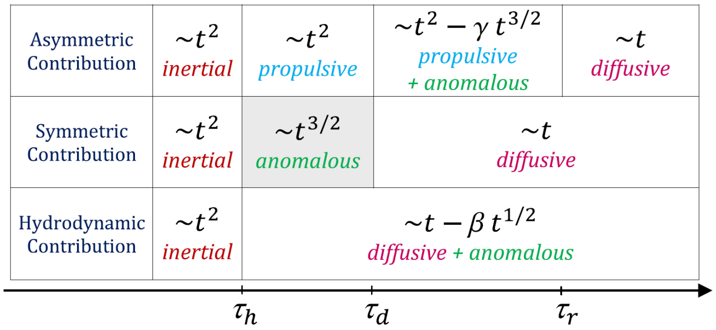


FIG. IV.1 Different contributions to the mean-squared displacement of active colloids. The total mean-squared displacement can be obtained by summing all of the contributions for asymmetric colloids, and the bottom two rows (only) for symmetric colloids (Golestanian, 2009).

at short times,

$$\Delta L^2 \simeq v_0^2 t^2 - \left(\frac{8}{3\sqrt{\pi}}\right) \frac{v_0^2 \tau_d^{3/2}}{\tau_r}\, t^{3/2} \qquad (\tau_d \ll t \ll \tau_r), \tag{IV.7}$$

at intermediate times, and

$$\Delta L^2 \simeq 2 v_0^2 \tau_r\, t \qquad (\tau_d \ll \tau_r \ll t), \tag{IV.8}$$

at long times, with a smooth crossover between them. For $\tau_d < t < \tau_r$ the memory effect that exists for self-propelled asymmetric colloids introduces an anomalous *anti-correlation* (i.e. contribution with negative sign) in the velocity autocorrelation function and the mean-squared displacement [Eq. (IV.7)]. Such anomalous corrections are reminiscent of the effect of the hydrodynamic long-time tail (Alder and Wainwright, 1967; Zwanzig and Bixon, 1970). Note, however, that the anomalous $-\gamma t^{3/2}$ correction in Eq. (IV.7) corresponds to much longer time scales and should be more easily observable than the hydrodynamic long-time tail.

C. Effective Diffusivity

At the longest time scales ($t > \tau_r$), all of the contributions are diffusive, leading to a total effective diffusion coefficient

$$D_{\text{eff}} = \frac{k_{\text{B}}T}{6\pi\eta R} + \frac{4\pi\alpha_1^2\mu^2\eta R^3}{27 D^2 k_{\text{B}} T} + \frac{c_1 \alpha_0 \mu^2}{3\pi^2 D^2 R^2}, \tag{IV.9}$$

where $c_1 = 1.17180$. The different terms in the above expression exhibit different R-dependencies, which causes the asymmetric contribution to be dominant for $R \gtrsim \left[D k_{\text{B}} T/(\alpha_1 \mu \eta)\right]^{1/2}$, while the symmetric contribution takes over when $R \lesssim \alpha_0 \mu^2 \eta^2/(D^2 k_{\text{B}} T)$. At the shortest time scales, on the other hand, the contribution due to phoretic effects will also be dominated by inertial effects that should lead to ballistic contributions (see Fig. IV.1). Moreover, the different terms depend differently on temperature as well, with the active contributions typically decreasing as temperature is increased contrary to the trend observed in the equilibrium Stokes-Einstein relation.

V. EXPERIMENTS ON SELF-PHORESIS

Spherical Janus particles made from polystyrene (PS) beads that are half-coated with platinum have been shown to self-propel because platinum (Pt) catalyzes the breakdown of hydrogen peroxide into

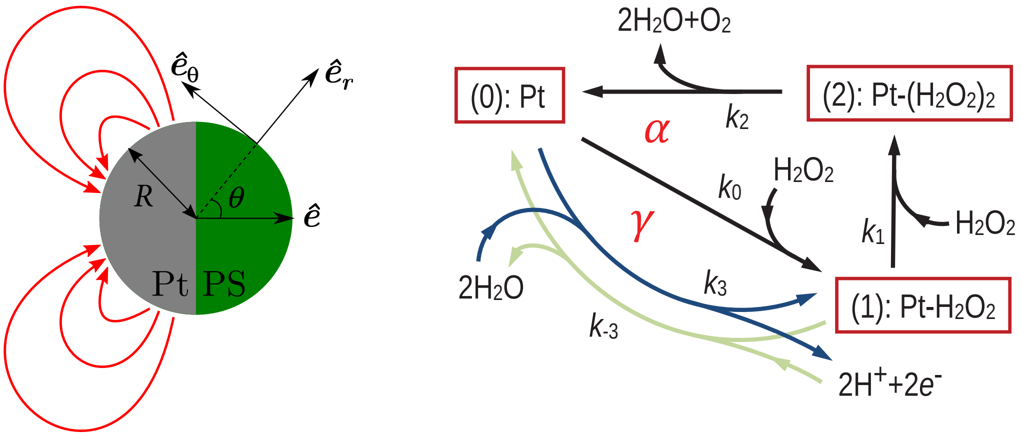


FIG. V.1 Sketch of the Pt-PS Janus particle (left). $\hat{e}$ is the direction of swimming. The red arrows represent the current loops that start at the equator of the particle and end at its pole. The reaction cycle (right), including two sub-cycles with fluxes α and γ, one of which involves ionic intermediates (Ebbens *et al.*, 2014; Ibrahim *et al.*, 2017).

water and oxygen

$$H_2O_2 \xrightarrow{\text{Pt}} H_2O + \frac{1}{2}O_2, \tag{V.1}$$

and the continuous flux of the reaction products establishes a steady gradient across the body of the Janus particle (Howse *et al.*, 2007). Using the Active Brownian Particle model, which was developed for the purpose of analyzing the stochastic trajectories observed from this experiment, it was possible to extract the average propulsion speed of this swimmer as a function of the fuel concentration C. It was observed that the speed depends on the fuel concentration according the Michaelis-Menten rule

$$V(C) = V_\infty \cdot \frac{C}{C + K_M}, \tag{V.2}$$

where K_M is the relevant Michaelis constant. This behaviour is consistent with the catalytic activity that drives the propulsion by setting up the stationary-state gradient, and suggests that the reaction must include a diffusion-limited binding step followed by a reaction-limited step

$$H_2O_2 + \text{Pt} \xrightarrow{k_1} \text{Pt}(H_2O_2) \xrightarrow{k_2} H_2O + \frac{1}{2}O_2 + \text{Pt}. \tag{V.3}$$

The swimming speed was also found to depend on the size of the colloid as $V \sim 1/R$ most of the time, which could be for different reasons, for example a competition between the diffusion- and reaction-limited steps of the catalysis and their interplay with the finite size of the colloid (Ebbens *et al.*, 2012). It is also possible to make self-phoretic active colloids that have spontaneous angular and translational velocities at the same time (Ebbens *et al.*, 2010).

Since the reaction was known to involve only (electrostatically) neutral components, it was a surprise when it was experimentally discovered that adding salt to the solution affects the swimming speed strongly (Brown and Poon, 2014; Ebbens *et al.*, 2014). A possible explanation for this behaviour has been proposed by postulating the existence of ionic intermediates in the catalytic reaction cycle that will give rise to closed loops of current on the platinum coat, as shown in Fig. V.1 (Ebbens *et al.*, 2014; Ibrahim *et al.*, 2017). The specific bi-cyclic topology of the reaction has been constructed based on the experimental observations. For example, it has been observed that addition of salt does not significantly affect the rate of consumption of hydrogen peroxide or the rate of production of oxygen, while it strongly affects the swimming speed (Ebbens *et al.*, 2014). The observations suggest that

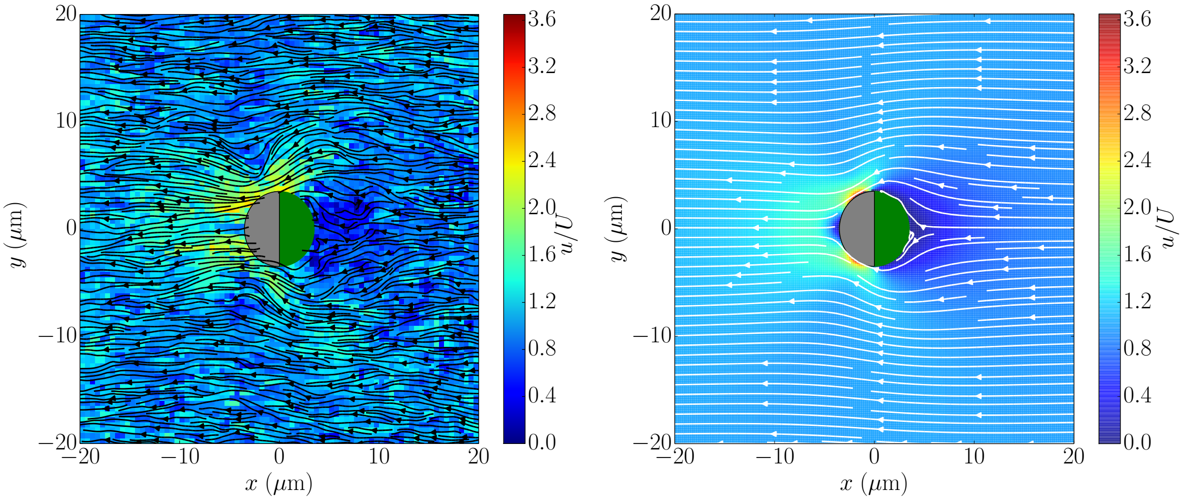


FIG. V.2 Streamlines around freely moving swimming Janus particles obtained experimentally (left) and theoretically (right). The background colors represent the magnitude of the velocity rescaled by the swimming velocity of the Janus particle (Campbell *et al.*, 2019).

the Janus particle employs a self-electrophoretic channel in addition to the neutral self-diffusiophoretic channel for its motility.

The conjectured current loops on the Pt hemisphere and the dominance of the resulting self-electrophoretic contribution have a major implication on the distribution of the effective surface slip velocity: the slip velocity will be concentrated on the Pt side, with a profile that can be represented with the following simplified form (Das *et al.*, 2015)

$$\boldsymbol{v}_s|_{r=R} = \begin{cases} v_0(1+\cos\theta)(-\cos\theta)\hat{e}_\theta & \text{for } \pi/2 < \theta < \pi, \\ 0 & \text{otherwise.} \end{cases} \tag{V.4}$$

This profile has two significant properties: (i) it leads to swimming away from the Pt patch, and (ii) it is not fore-aft symmetric. Interestingly, any surface slip velocity distribution with these two properties should lead to a quenching of the orientation of the swimmer in a direction parallel to the surface due to hydrodynamic interaction. This is indeed observed experimentally (Das *et al.*, 2015). Interactions between phoretic swimmers and surfaces can lead to a wide variety of different behaviours (Bayati *et al.*, 2019; Uspal *et al.*, 2015). In this regards, it has been revealed that the Pt-PS swimmer is a special case. The surface alignment property of the Pt-PS swimmer is not shared by other types, as it arises from its specific form of the slip velocity; usually, other swimmer prototypes do not possess one or both of the above-mentioned necessary properties.

While the alignment property corroborated the proposed structure of the slip velocity due to the current loops, recent measurements of the complete flow field profile around (swimming and stationary) Pt-PS Janus particles provided a direct visualization of the flow and measurement of the slip velocity profile (Campbell *et al.*, 2019). As can be seen in Fig. V.2, the slip velocity is maximum in the middle part of the Pt region, which was in full agreement with the above picture. The measured flow field gave access to the squirmer B_n coefficients, and revealed that for the Pt-PS Janus swimmer $B_2/|B_1| \simeq -2.45 < 0$, which is consistent with a *pusher*, in terms of the classification for hydrodynamic interactions. However, this experiment has provided a more complete picture with regards to the near-field properties of the hydrodynamic interactions than a simplistic squirmer of pusher type, which can be used to build a more faithful representation of the hydrodynamic interactions.

No other prototype phoretic microswimmer has been experimentally characterized as thoroughly and systematically as the Pt-PS Janus particle.

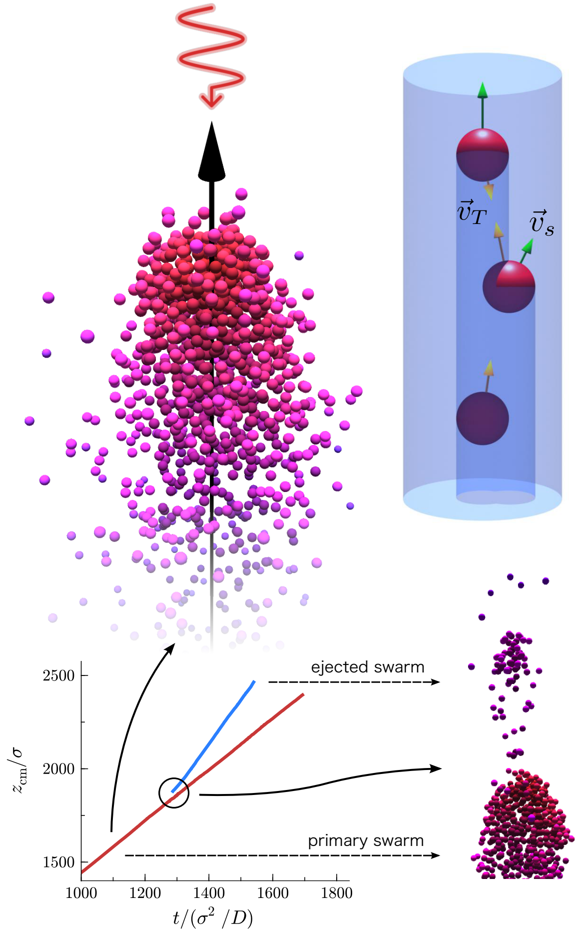


FIG. VI.1 Optically driven thermally active colloids self-organize into a moving comet-like swarm under illumination of a uniform intensity plane wave light source (top left). Colloids can cast partial or complete shadows on those below them and propel through production of a collective thermal field, $\boldsymbol{v}_T$ (yellow arrow), and by local self propulsion, $\boldsymbol{v}_s$ (green arrow), shown here with the magnitude of v_T magnified by a factor of five for presentation. Fluctuations in the swarm shape facilitates ejection of hot colloids at the swarm tip forming faster moving sub swarms (bottom) (Cohen and Golestanian, 2014).

VI. APOLAR ACTIVE COLLOIDS: SWARMING DUE TO EXTERNAL STEERING

Using light as the external source of energy to induce self-thermophoresis is extremely versatile as it can be used to engineer collective swarming behaviour (Cohen and Golestanian, 2014). The colloids then take advantage of the natural asymmetries in the system to create non-equilibrium conditions that drive them into a range of collective behaviour. Here a simple system is discussed where colloids that convert light into heat and move in response to self- and collectively generated thermal gradients. The system exhibits self-organization into a moving comet-like swarm with novel non-equilibrium dynamics. Although these active colloids are controlled by viscous hydrodynamics, their collective behaviour shows very dynamic structures with inertial traits. In particular, it exhibits propagation of transverse waves from back to front of the swarm with no dispersion, ejection of hot colloids from the head of the swarm, and persistent circulation flow within the swarm. The rich behaviour of the dynamic comet-like swarm can be controlled by a single external parameter, the intensity of light.

Consider a collection of particles illuminated with a directed light source with uniform intensity I. The light intensity received at the surface of each colloid is determined by the distribution of the shadows of the colloids above it, as shown in Fig. VI.1. The light received by each colloid is converted into a heat flux that increases the temperature of the colloid and the surrounding fluid in an anisotropic way. A particle with a clear view of the light source will have an illuminated hot top hemisphere and a dark cold bottom

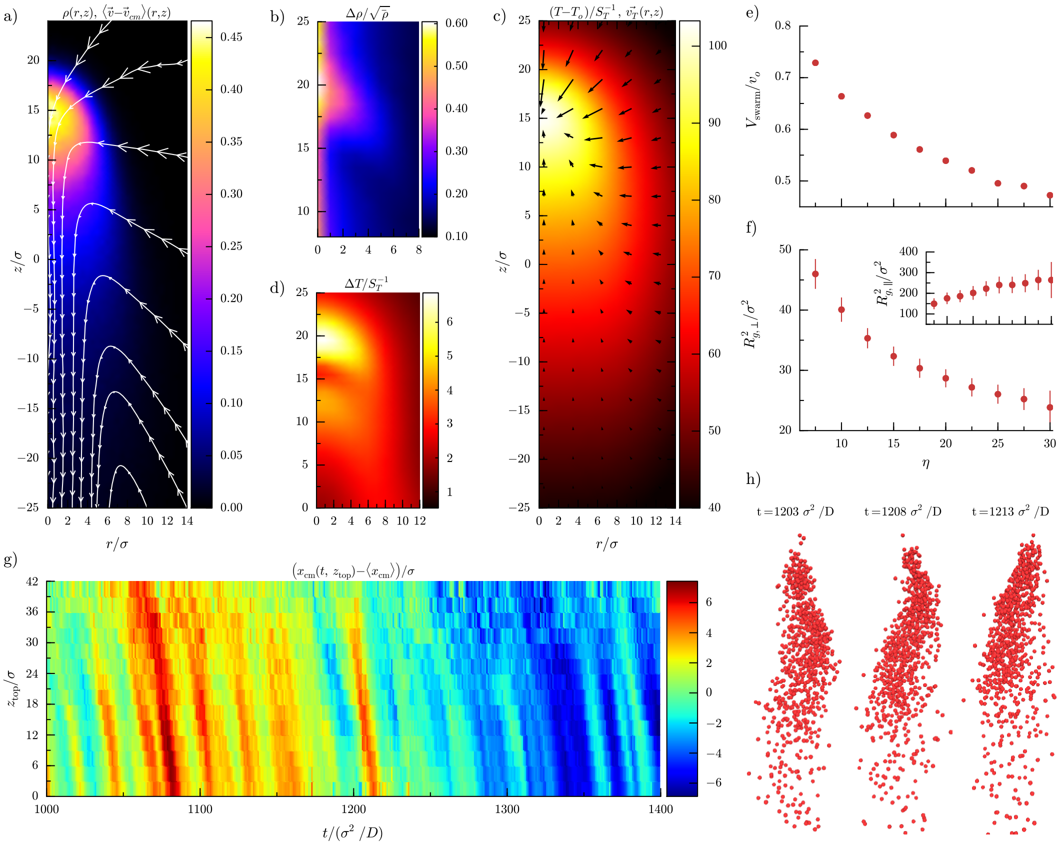


FIG. VI.2 Using the moving centre of mass swarm frame of reference fields can be reconstructed using time and ensemble averaging. The swarm fields are axially averaged and presented in cylindrical coordinates (r, z) with z axis pointing towards the light source and r the radial distance. (a) Density field and average colloid velocity relative to centre of mass velocity. The relative colloid velocity is represented by the size of the arrows placed along the streamlines. (b) Density fluctuations normalized by equilibrium fluctuations. (c) Temperature field with phoretic velocity due to temperature gradients. (d) Temperature fluctuations. The velocity and shape of the swarm is affected by the dimensionless coupling strength η. (e) Swarm velocity normalized to individual colloid drift velocity. (f) Radius of gyration of the swarm in the direction perpendicular to the axis of illumination with error bars show standard deviation over ensemble averaging of 10 simulations (inset showing radius of gyration in the parallel direction). (g) Kymograph of the centre of mass x location for the section of thickness $\Delta z = 3\sigma$ a distance $z_{\rm top}$ from the swarm top at time t for $\eta = 25$. Colour represents the x value of the centre of mass around the time average within the displayed window. (h) Snapshots of the swarm conformation at three successive times as shown by the kymograph in (g) displaying wave propagation from bottom to top. (Cohen and Golestanian, 2014)

hemisphere. This asymmetric temperature distribution results in the self-propulsion of the colloid via a process known as self-thermophoresis, with a maximum velocity of magnitude $v_o = I|D_T|/(9\kappa)$, where D_T is the thermophoretic mobility (also known as the thermodiffusion coefficient) and κ is the thermal conductivity, which is set to be equal for the colloid and the solvent for simplicity. When D_T is negative, which is allowed as it is an off-diagonal Onsager coefficient and possible via appropriate surface treatment of colloids (Piazza and Parola, 2008), the self-propulsion will be predominantly towards the light source with a velocity $\boldsymbol{v}_s$ (see Fig. VI.1), leading to an effective attractive artificial phototaxis. Moreover, all colloids (whether illuminated or not) will experience a thermodiffusion drift velocity due the temperature gradient generated by the illuminated colloids, $\boldsymbol{v}_T = -D_T\nabla T$ (see Fig. VI.1). The choice of negative D_T causes the colloids to act as both heat sources and heat seekers; a combination that could lead to self-organization and instability, as seen in a diverse range of non-equilibrium phenomena.

The behaviour of the system depends on the intensity of the light source, which we can represent using the dimensionless coupling strength $\eta = \sigma I|S_T|/\kappa$, where σ is the diameter of the colloid, $S_T = D_T/D_{\rm c}$ is the Soret coefficient, and $D_{\rm c}$ is the colloid diffusion coefficient. The system self-organizes into a moving swarm of apparent constant centre-of-mass velocity V_{swarm} with a comet-like structure: a high density head region with the outer most illuminated colloids generating a central hot core, and a relatively more dilute trailing aggregate in the form of a tail as shown in Fig. VI.1. Axially averaged fields are presented in Fig. VI.2 along with their fluctuations. The high density head region forms a hot core, as can be clearly seen in Fig. VI.2a and VI.2c, which pulls the tail of the comet along, and also drives the fluctuations. Density fluctuations are normalized by the local equilibrium expectation values in Fig. VI.2b such that any deviation from uniformity indicates non-equilibrium density fluctuations. A particularly interesting mode of density fluctuations occurs at the very tip of the head region as a result of the illuminated self-propelled particles (with the strongest v_s component) attempting to escape the influence of the thermal attraction (also at its strongest, as shown by the vector field in Fig. VI.2c), leading to fluctuations in the swarm shape. These particles usually return to the swarm, although spectacular ejection events are also observed at the tip with likelihood increasing with η; see Fig. VI.1 (bottom). Density fluctuations at the swarm tip lead to novel temperature fluctuations due to the transient appearance of heat sources as seen in Fig. VI.2d.

A circulation can be observed in the average colloid velocity streamlines in the swarm centre-of-mass frame, as shown in Fig. VI.2a; the colloids that are attracted to the hot core reverse their direction on crossing the shadow boundary. This phenomenon also results from the competition between the strong thermally induced drift velocity towards the core shown in the vector field of Fig. VI.2c and the propulsion of individual colloids towards the light source. The partially illuminated colloids that are near boundary of the swarm introduce a "thermal drag" that slows down the swarm as compared to the external fully illuminated isolated colloids. Figure VI.2e shows how this slowing down becomes more prominent as the coupling strength is increased, leading to an effectively sub-linear increase of V_{swarm} with respect to η. This suggests the following explanation for the observed circulation. A particle in the shadowed tail of the swarm, where the thermal attraction of the core is not strong enough to keep particles within the bulk of the swarm, may be left behind but remain in the shadow. At some point this inactive colloid will diffuse out of the shadow to become active again propelling towards the source. As it moves faster individually than in the swarm it may catch up and find itself attracted back to the hot core creating a circulation. Alternatively, the colloid may escape the influence of thermal attraction and propel past the swarm. The average shape of the swarm is also affected by the value of η in line with the above picture, as shown in Fig. VI.2f. The radius of gyration perpendicular (parallel) to the axis of illumination becomes smaller (larger) as η is increased, resulting in an increased aspect ratio.

Increasing the coupling strength will also make the swarm more dynamic. Transverse waves of colloids can be observed (Fig. VI.2g and Fig. VI.2h) propagating from tail to head in randomly selected azimuthal directions, with pronunciation increasing with higher η. The waves appear to be randomly initiated at the back of the swarm, and propagate with a constant speed (that increases with η) towards the front without any dispersion, as can be seen from the kymograph displayed in Fig. VI.2g. The existence of these waves is a result of the competition between the transverse (xy) components of the self-propulsion and thermal drift velocities, as shown in Fig. VI.1. The undulations arise from the colloids in the tail region diffusing out of the shadow, aided by the effect of partial illumination upon crossing the shadow boundary that further drives their motion away from the swarm. The colloids then become thermally active and attract higher up colloids out of the shadow, thereby initiating a propagating wave along the swarm length.

VII. MIXTURES OF APOLAR ACTIVE COLLOIDS: ACTIVE MOLECULES

After understanding how nonequilibrium phoretic activity affects the dynamics of single colloidal particles, we can now study the nonequilibrium interactions between these particles. The simplest case to consider is the interaction between two different types of apolar (symmetric) active colloids (Soto and Golestanian, 2014).

Consider a concentration field $C(\boldsymbol{r})$ that satisfies the stationary diffusion equation $\nabla^2 C = 0$, subject to the boundary condition of production or consumption on the surfaces of colloidal particles. For activity α, the boundary condition is $-D\partial_r C|_{r=R} = \alpha$, which leads to a contribution to the concentration field of the form $C(\boldsymbol{r}) = \frac{\alpha R^2}{Dr}$. At a distance $\boldsymbol{r}$ away from the colloid that acts as a source or a sink for the chemical, a colloidal particle will experience a drift velocity of $\boldsymbol{v} = -\mu \nabla C = -\frac{\mu\alpha}{D}\nabla\left(\frac{1}{r}\right)$. Therefore, our system is governed by a dissipative equivalent of gravitational or Coulomb interactions with a $1/r$ potential. There is, however, a peculiar feature when we consider a mixture of, say, type-A and type-B active colloids. In this case, the symmetry between action and reaction will be, in general, broken. This can be observed from the drift velocities of A and B particles:

$$\begin{aligned} \boldsymbol{v}_{\mathrm{A}} &= -\mu_{\mathrm{A}}\nabla C|_{\mathrm{A}} = -\mu_{\mathrm{A}}\alpha_{\mathrm{B}}\nabla_{\mathrm{A}}\left(\frac{R^2}{D|\boldsymbol{r}_{\mathrm{B}}-\boldsymbol{r}_{\mathrm{A}}|}\right) = -\mu_{\mathrm{A}}\alpha_{\mathrm{B}}\cdot\frac{R^2}{D}\cdot\frac{\boldsymbol{r}_{\mathrm{B}}-\boldsymbol{r}_{\mathrm{A}}}{|\boldsymbol{r}_{\mathrm{B}}-\boldsymbol{r}_{\mathrm{A}}|^3}, \qquad \text{(VII.1)} \\ \boldsymbol{v}_{\mathrm{B}} &= -\mu_{\mathrm{B}}\nabla C|_{\mathrm{B}} = -\mu_{\mathrm{B}}\alpha_{\mathrm{A}}\nabla_{\mathrm{B}}\left(\frac{R^2}{D|\boldsymbol{r}_{\mathrm{B}}-\boldsymbol{r}_{\mathrm{A}}|}\right) = +\mu_{\mathrm{B}}\alpha_{\mathrm{A}}\cdot\frac{R^2}{D}\cdot\frac{\boldsymbol{r}_{\mathrm{B}}-\boldsymbol{r}_{\mathrm{A}}}{|\boldsymbol{r}_{\mathrm{B}}-\boldsymbol{r}_{\mathrm{A}}|^3}. \qquad \text{(VII.2)} \end{aligned}$$

Unless we have a specially fine-tuned system, we have $\mu_{\mathrm{A}}\alpha_{\mathrm{B}} \neq \mu_{\mathrm{B}}\alpha_{\mathrm{A}}$, and therefore, the action-reaction symmetry is broken. We can understand this property as a generalization of electrostatics or gravity in which for every particle the charge or mass that creates the field is different from the charge or mass that responds to the field (created by others).

To examine the behaviour of a mixture in the dilute limit, we start by considering two colloids. Assuming an additional equilibrium interaction potential $U(|\boldsymbol{r}_{\mathrm{B}}-\boldsymbol{r}_{\mathrm{A}}|)$, which can arise from excluded volume interaction for example, the particles will experience an additional contribution to the drift velocity

$$\boldsymbol{v}_{\mathrm{A}}^{\mathrm{eq}} = -D_{\mathrm{c}}\beta\nabla_{\mathrm{A}}U = -\boldsymbol{v}_{\mathrm{B}}^{\mathrm{eq}}, \qquad \text{(VII.3)}$$

where D_{c} is the diffusion coefficient of the colloids. We can change coordinates from $\boldsymbol{r}_{\mathrm{A}}$ and $\boldsymbol{r}_{\mathrm{B}}$ to the relative and centre of mass coordinates, $\boldsymbol{r} = \boldsymbol{r}_{\mathrm{B}} - \boldsymbol{r}_{\mathrm{A}}$ and $\boldsymbol{r}_{\mathrm{CM}} = \frac{1}{2}(\boldsymbol{r}_{\mathrm{A}} + \boldsymbol{r}_{\mathrm{B}})$. For the velocities, we obtain

$$\begin{aligned} \boldsymbol{v} &= \boldsymbol{v}_{\mathrm{B}} - \boldsymbol{v}_{\mathrm{A}} = (\mu_{\mathrm{B}}\alpha_{\mathrm{A}} + \mu_{\mathrm{A}}\alpha_{\mathrm{B}})\cdot\frac{R^2}{D}\cdot\frac{\boldsymbol{r}}{r^3} - 2D_{\mathrm{c}}\beta\nabla U, \qquad \text{(VII.4)} \\ \boldsymbol{V} &= \frac{1}{2}(\boldsymbol{v}_{\mathrm{A}} + \boldsymbol{v}_{\mathrm{B}}) = \frac{1}{2}(\mu_{\mathrm{B}}\alpha_{\mathrm{A}} - \mu_{\mathrm{A}}\alpha_{\mathrm{B}})\cdot\frac{R^2}{D}\cdot\frac{\boldsymbol{r}}{r^3}. \qquad \text{(VII.5)} \end{aligned}$$

Equilibration in the relative distance yields

$$\mathcal{P}(r) = \mathcal{A}r^{d-1}\exp\left\{-(\mu_{\mathrm{B}}\alpha_{\mathrm{A}} + \mu_{\mathrm{A}}\alpha_{\mathrm{B}})\cdot\frac{R^2}{DD_{\mathrm{c}}r} - 2\beta U\right\}, \qquad \text{(VII.6)}$$

where $\mathcal{A}$ is a normalization constant. Invoking the analogy to electrolytes, we can define a generalized Bjerrum length as $\ell_{\mathrm{B}} = |\mu_{\mathrm{B}}\alpha_{\mathrm{A}} + \mu_{\mathrm{A}}\alpha_{\mathrm{B}}|R^2/(DD_{\mathrm{c}})$, which represents the distance at which "energy" and "entropy" are comparable. Using the Bjerrum length, the stationary distribution for the relative distance is $\mathcal{P}(r) \sim r^{d-1}\exp\left(\frac{\ell_{\mathrm{B}}}{r} - 2\beta U\right)$. Since the centre of mass speed is determined by r as $V = \frac{1}{2}|\mu_{\mathrm{B}}\alpha_{\mathrm{A}} - \mu_{\mathrm{A}}\alpha_{\mathrm{B}}|\cdot\frac{R^2}{Dr^2}$, we can deduce the distribution of swimming speed as follows

$$\mathcal{P}(V) = \mathcal{B}\cdot\frac{e^{\sqrt{V/V_0}}}{V^{1+d/2}} \qquad \text{(VII.7)}$$

where $V_0 \equiv \frac{|\mu_{\mathrm{B}}\alpha_{\mathrm{A}} - \mu_{\mathrm{A}}\alpha_{\mathrm{B}}|DD_{\mathrm{c}}^2}{2|\mu_{\mathrm{B}}\alpha_{\mathrm{A}} + \mu_{\mathrm{A}}\alpha_{\mathrm{B}}|^2R^2}$ and $\mathcal{B}$ is a normalization constant. The distribution is cut off at $V_{\max} = \frac{1}{8D}|\mu_{\mathrm{B}}\alpha_{\mathrm{A}} - \mu_{\mathrm{A}}\alpha_{\mathrm{B}}|$.

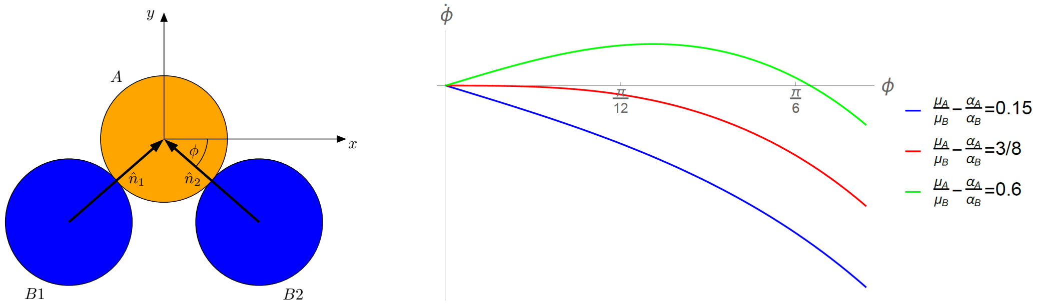


FIG. VII.1 Stability of AB_2 molecule. The angle ϕ parametrizes the deviation from a linear chain of the AB_2 molecule (left). Plots of Eq. (VII.12) for different values of the parameter $\frac{\mu_\mathrm{A}}{\mu_\mathrm{B}} - \frac{\alpha_\mathrm{A}}{\alpha_\mathrm{B}}$.

A. Designing the Configurations of Active Molecules

When the phoretic interaction strengths are sufficiently strong, the active colloids will form clusters. We can examine the configurations of different clusters and determine which ones have the potential to form stable configurations, which can be regarded as stable active molecules. Let us consider a cluster formed with one A particle and two B particles. We can parametrize reflection-symmetric configurations using the angle ϕ defined in Fig. VII.1. We can write down the following expressions for various contributions to the velocities of the three particles

$$\begin{aligned}
\boldsymbol{v}_\mathrm{A} &= \frac{\mu_\mathrm{A}\alpha_\mathrm{B}}{D}\boldsymbol{n}_1 + \frac{\mu_\mathrm{A}\alpha_\mathrm{B}}{D}\boldsymbol{n}_2 + v^\mathrm{eq}\boldsymbol{n}_1 + v^\mathrm{eq}\boldsymbol{n}_2, && \text{(VII.8)}\\
\boldsymbol{v}_{\mathrm{B}_1} &= -\frac{\mu_\mathrm{B}\alpha_\mathrm{A}}{D}\boldsymbol{n}_1 - v^\mathrm{eq}\boldsymbol{n}_1 - \frac{\mu_\mathrm{B}\alpha_\mathrm{B}}{D}\cdot\frac{1}{(2\cos\phi)^2}\boldsymbol{e}_x, && \text{(VII.9)}\\
\boldsymbol{v}_{\mathrm{B}_2} &= -\frac{\mu_\mathrm{B}\alpha_\mathrm{A}}{D}\boldsymbol{n}_2 - v^\mathrm{eq}\boldsymbol{n}_2 + \frac{\mu_\mathrm{B}\alpha_\mathrm{B}}{D}\cdot\frac{1}{(2\cos\phi)^2}\boldsymbol{e}_x. && \text{(VII.10)}
\end{aligned}$$

We can invoke the constraint of no penetration between the particles via $\boldsymbol{n}_i\cdot(\boldsymbol{v}_{\mathrm{B}_i} - \boldsymbol{v}_\mathrm{A}) = 0$ and obtain the equilibrium contribution of the velocities as

$$v^\mathrm{eq} = \frac{\left(\frac{\mu_\mathrm{B}\alpha_\mathrm{B}}{D}\right)}{2-\cos 2\phi}\left[-\left(\frac{\alpha_\mathrm{A}}{\alpha_\mathrm{B}} + \frac{\mu_\mathrm{A}}{\mu_\mathrm{B}}\right) + \left(\frac{\mu_\mathrm{A}}{\mu_\mathrm{B}}\right)\cos 2\phi - \frac{1}{4\cos\phi}\right]. \tag{VII.11}$$

Using a kinematic definition $R\dot{\phi} = -\boldsymbol{n}_i \times (\boldsymbol{v}_{\mathrm{B}_i} - \boldsymbol{v}_\mathrm{A})\cdot\boldsymbol{e}_z$, we find the following dynamical equation for the configuration of the molecule

$$\dot{\phi} = \frac{2}{R}\left(\frac{\mu_\mathrm{B}\alpha_\mathrm{B}}{D}\right)\frac{\sin\phi}{(2-\cos 2\phi)}\left[\left(\frac{\mu_\mathrm{A}}{\mu_\mathrm{B}} - \frac{\alpha_\mathrm{A}}{\alpha_\mathrm{B}}\right)\cos\phi - \frac{3}{8\cos^2\phi}\right]. \tag{VII.12}$$

Figure VII.1 shows the behaviour of this dynamical system for different values of the tuning parameter. When $\frac{\mu_\mathrm{A}}{\mu_\mathrm{B}} - \frac{\alpha_\mathrm{A}}{\alpha_\mathrm{B}} < \frac{3}{8}$ the dynamical system in Eq. (VII.12) has only one stable fixed point at $\phi = 0$, which corresponds to a linear B–A–B conformation for the AB_2 molecule; since the conformation is symmetric, the molecule is not self-propelled. At $\frac{\mu_\mathrm{A}}{\mu_\mathrm{B}} - \frac{\alpha_\mathrm{A}}{\alpha_\mathrm{B}} = \frac{3}{8}$ the dynamical system exhibits a supercritical pitchfork bifurcation, and for $\frac{\mu_\mathrm{A}}{\mu_\mathrm{B}} - \frac{\alpha_\mathrm{A}}{\alpha_\mathrm{B}} > \frac{3}{8}$ two stable fixed points appear at $\phi = \pm\phi_s$, defined via $\cos\phi_s = \left(\frac{3/8}{\mu_\mathrm{A}/\mu_\mathrm{B} - \alpha_\mathrm{A}/\alpha_\mathrm{B}}\right)^{1/3}$, while the $\phi = 0$ fixed point becomes unstable. Due to the symmetry breaking in the conformation, the AB_2 molecule will now be self-propelled, with a speed that can be determined from the above equations in terms of the parameters.

This calculation demonstrates how it is possible to design stable shapes for various active colloidal molecules using the parameters of the system, namely the values of the activities and the mobilities.

B. Dynamic Function

For sufficiently large molecules, it is possible to have cases where the conformations of the molecule change dynamically, and consequently this will be translated to the manifested non-equilibrium function exhibited by those molecules. To have such dynamical changes, one possibility is to have multiple stable fixed points and noise-induced transitions between them, presumably across barrier. An interesting case with such behaviour is the AB_3 molecule, which has a stable Y-isomer with no self-propulsion, and a stable T-isomer with self-propulsion; stochastic switching between them leads to an emergent run-and-tumble behaviour in a system with a continuous configuration space. Another possibility is the existence of an oscillatory conformation. When such conformations are symmetric, such as the case for A_4B_8, the molecule will exhibit spontaneous oscillations without self-propulsion. In asymmetric cases, such A_5B_8, the oscillations can lead to self-propulsion, in a way that is reminiscent of the swimming of sperm (Soto and Golestanian, 2015).

C. From Structure to Function: A New Non-equilibrium Paradigm

The framework described above can be generalized to cases where different parameters such as size, surface chemistry, and surface activity are tuned in order to achieve desired clusters and molecules. With such capabilities, the framework provides a paradigm in which we can design certain *structures*—i.e. 3D geometry and conformation—that will exhibit certain non-equilibrium *functions* entirely due to their shape. The function can be derived from symmetry properties of the conformations. For example, axially symmetric molecules will exhibit an intrinsic (self-propulsion) translational velocity, whereas non-axially symmetric molecules will have intrinsic angular velocity or spin. If the molecules are "too symmetric" they might not exhibit any mechanical function and can be categorized as inert. Sufficiently large complexes can spontaneously break time-translation invariance and exhibit oscillations. The paradigm has similarities to the way proteins are designed from sequences to shapes to biological function.

VIII. MIXTURES OF APOLAR ACTIVE COLLOIDS: STABILITY OF SUSPENSIONS

As described in the previous section, two different apolar chemically active colloids interact with each other through the chemical fields that they themselves produce, and a key feature of these interactions is that they are in general non-reciprocal, see Eqs. (VII.1) and (VII.2). It is therefore pertinent to investigate the phase behaviour of mixtures of several species of active colloids (Agudo-Canalejo and Golestanian, 2019). Brownian dynamics simulations of a dilute suspention of many colloids belonging to different species and interacting through Eqs. (VII.1) and (VII.2) show a variety of phase separation phenomena. For binary mixtures, the simulations reveal that, while in a large region of the parameter space the mixtures remain homogeneous, the homogeneous state can also become unstable leading to a great variety of phase separation phenomena; see Figs. VIII.1(c–e). Here, phase separation is used in the sense of macroscopic (system-spanning) separation typically into a single large cluster [occasionally into two; see Fig. VIII.1(c)] that coexists with a dilute (or empty) phase. The phase separation process may lead to aggregation of the two species into a single mixed cluster, or to separation of the two into either two distinct clusters or into a cluster of a given stoichiometry and a dilute phase. The resulting configurations are qualitatively distinct for mixtures of one chemical-producer and one chemical-consumer species, as opposed to mixtures of two producer (or consumer) species; compare Fig. VIII.1(c) and Fig. VIII.1(e). While the typical steady-state configurations are static, for mixtures of producer and consumer species it is observed that static clusters can undergo a shape-instability that breaks their symmetry, leading to a self-propelling cluster. Randomly-generated highly-polydisperse mixtures of up

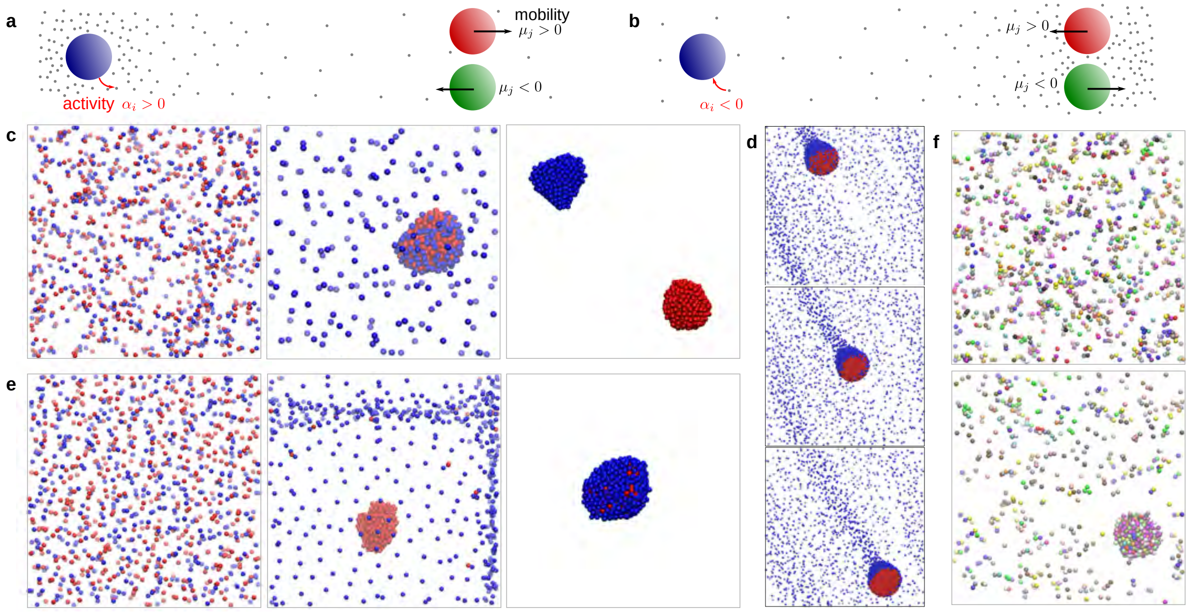


FIG. VIII.1 Active phase separation phenomena in mixtures of chemically-interacting particles. (a) Producer particles repel (attract) particles with positive (negative) mobility, while (b) the opposite is true for consumer particles. (c) Binary mixtures of producer (blue) and consumer (red) species show, from left to right, homogeneous states with association of particles into small molecules, see Section VII, aggregation into a static dense phase that coexists with a dilute phase, and separation into two static collapsed clusters. (d) The static aggregate [(c), centre] can undergo symmetry breaking to form a self propelled macroscopic cluster. (e) Binary mixtures of producer species (blue and red) show homogeneous states without molecule formation, separation into a static dense phase and a dilute phase that are pushed away from each other, and aggregation into a static collapsed cluster. (f) Randomly-generated highly polydisperse mixtures (20 different species) can remain homogeneous or undergo macroscopic phase separation.

to 20 species also show homogeneous as well as phase-separated states [Fig. VIII.1(f)].

This variety of phase separation phenomena can be understood within a continuum theory of the mixture. Let us consider a system consisting of M different species of chemically-interacting particles, with concentrations $\rho_i(\boldsymbol{r}, t)$ for $i = 1, ..., M$; and a messenger chemical with concentration $C(\boldsymbol{r}, t)$. The concentration of species i is described by

$$\partial_t \rho_i(\boldsymbol{r}, t) - \nabla \cdot [D_{\mathrm{c}} \nabla \rho_i + (\mu_i \nabla C)\rho_i] = 0, \tag{VIII.1}$$

which includes a diffusive term with diffusion coefficient D_{c}, which for simplicity is taken to be equal for all types, as well as the phoretic drift term with mobility μ_i which is positive or negative if the particle is repelled or attracted to the chemical, respectively; see Figs. VIII.1(a) and VIII.1(b). The concentration of the chemical is described by

$$\partial_t C(\boldsymbol{r}, t) - D\nabla^2 C = \sum_i \alpha_i \rho_i, \tag{VIII.2}$$

where the first term is diffusive with coefficient D, and the second represents production or consumption of the chemical by all particle species. Within this continuum theory for the mixture, we can study the stability of the homogeneous state, and show that under certain conditions the system undergoes macroscopic phase separation.

A. Linear Stability Analysis of the Homogeneous Mixture

We consider small deviations from the homogeneous state, so that the colloid density is described by $\rho_i(\boldsymbol{r},t) = \rho_{0i} + \delta\rho_i(\boldsymbol{r},t)$. The net catalytic activity of a mixture is defined as $A \equiv \sum_i \alpha_i \rho_{0i}$, where we note that A represents activity in the homogeneous state, while locally we have $\sum_i \alpha_i \rho_i = A + \sum_i \alpha_i \delta\rho_i$. The chemical concentration can be separated into a (time-dependent) uniform value and the deviations from this uniform value in response to nonuniformities of the colloid distribution, so that $C(\boldsymbol{r},t) = C_0 + At + \delta C(\boldsymbol{r},t)$. Introducing this into the evolution equation for $C(\boldsymbol{r},t)$ we obtain an equation for the deviations $\delta C(\boldsymbol{r},t)$ given by

$$\partial_t \delta C(\boldsymbol{r},t) - D\nabla^2 \delta C = \sum_i \alpha_i \delta\rho_i. \tag{VIII.3}$$

Because the small chemical diffuses much faster than the large colloids, the deviations $\delta C(\boldsymbol{r},t)$ of the chemical concentration from the uniform value $C_0 + At$ can be assumed to reach a steady state instantaneously for each configuration of the colloids, so that from Eq. (VIII.3) we obtain

$$-D\nabla^2 \delta C = \sum_i \alpha_i \delta\rho_i. \tag{VIII.4}$$

Introducing this into the evolution equation for $\rho_i(\boldsymbol{r},t)$, and staying only to linear order in $\delta\rho_i(\boldsymbol{r},t)$, we obtain

$$\partial_t \delta\rho_i(\boldsymbol{r},t) = D_\mathrm{c}\nabla^2 \delta\rho_i - \frac{\mu_i \rho_{0i}}{D} \sum_j \alpha_j \delta\rho_j. \tag{VIII.5}$$

The linearized system of equations [Eqs. (VIII.5)] with $i = 1, ..., M$ describes the evolution of the deviations of the colloid density around the homogeneous state. This result is valid at all times for mixtures with net positive or zero production $A \geq 0$; and for mixtures with net consumption $A < 0$ as long as the chemical concentration is still large enough that the consumption of chemical by the colloids can be considered to be taking place in the saturated regime, i.e. at a rate $\alpha_i < 0$ independent of the local chemical concentration. If K is the (largest) equilibrium constant of the consumption reaction at the surface of the colloids, then this approach is valid as long as $C_0 - |A|t \gg K$, i.e. for sufficiently short experiments with $t \ll (C_0 - K)/|A|$.

The stability analysis of Eqs. (VIII.5) is done most conveniently by defining the new variables $U_i \equiv \alpha_i \delta\rho_i$ and the parameters $\gamma_i \equiv \mu_i \alpha_i \rho_{0i}/D$. The system of equations (VIII.5) can be rewritten as

$$[\partial_t - D_\mathrm{c}\nabla^2 + \gamma_i]U_i + \gamma_i \sum_{j \neq i} U_j = 0 \tag{VIII.6}$$

the solution of which is given by a sum of Fourier modes of the form $U_i(\boldsymbol{r},t) = U_{\boldsymbol{q}i} \mathrm{e}^{i\boldsymbol{q}\cdot\boldsymbol{r}} \mathrm{e}^{\lambda t}$. Introducing this into (VIII.6) finally results in the eigenvalue problem

$$[\lambda + D_\mathrm{c}\boldsymbol{q}^2 + \gamma_i]U_{\boldsymbol{q}i} + \gamma_i \sum_{j \neq i} U_{\boldsymbol{q}j} = 0 \tag{VIII.7}$$

for the growth rate λ of the perturbation modes with wavenumber $\boldsymbol{q}$.

By defining $\tilde{\lambda} \equiv -(\lambda + D_\mathrm{c}\boldsymbol{q}^2)$, the eigenvalue problem (VIII.7) is equivalent to finding the eigenvalues of a $M \times M$ matrix with M identical rows each given by $[\gamma_1\ \gamma_2\ ...\ \gamma_M]$. Such a matrix has rank 1 and therefore at least $M - 1$ of its eigenvalues are equal to zero, $\tilde{\lambda}_- = 0$. Because the trace of a matrix is equal to the sum of its eigenvalues, the remaining eigenvalue is equal to the trace of the matrix, so that $\tilde{\lambda}_+ = \sum_i \gamma_i$.

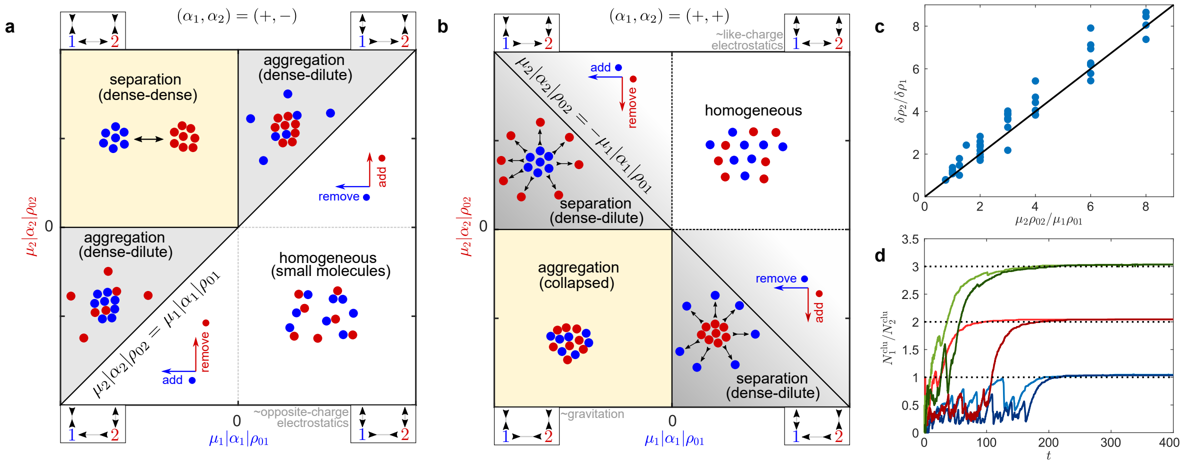


FIG. VIII.2 Stability diagrams and stoichiometry for binary mixtures. (a) Stability diagram for mixtures of one producer and one consumer species [cf. Fig. VIII.1(c)], and (b) for mixtures of two producer species [cf. Fig. VIII.1(e)]. In (a) and (b) the boxed legends attached to each quadrant symbolize the "interaction network" representing the sign of interactions between each species in the system. Phase separation [aggregation in (a), separation in (b)] can be triggered by addition or removal of particles (density changes) only when interactions between the two species are intrinsically non-reciprocal. (c) Stoichiometry at the onset of the instability, obtained from 44 simulations (blue circles) compared to the stability analysis prediction [Eq. (VIII.10)]. (d) Time evolution of the stoichiometry of the biggest cluster arising from aggregation of $(\alpha_1, \alpha_2) = (+, -)$ mixtures, demonstrating that the long time stoichiometry is predicted by the neutrality rule [Eq. (VIII.11)] and is independent of the species' mobility (blue: $\tilde{\alpha_2} = -1$, $\tilde{\mu_2} = 8$ and 12; red: $\tilde{\alpha_2} = -2$, $\tilde{\mu_2} = 4$ and 8; green: $\tilde{\alpha_2} = -3$, $\tilde{\mu_2} = 3$ and 5; in all cases $N_1 = 800$, $N_2 = 200$, $\tilde{\alpha}_1 = \tilde{\mu}_1 = 1$).

Transforming from $\tilde{\lambda}$ back to λ, we finally find $M - 1$ identical eigenvalues $\lambda_- = -D_c\boldsymbol{q}^2$, and one eigenvalue $\lambda_+ = -D_c\boldsymbol{q}^2 - \sum_i \gamma_i$. The latter eigenvalue can become positive, indicating an instability. When rewritten in the original variables, we find that the homogeneous state becomes unstable towards a spatially-inhomogeneous state when the following condition holds

$$\sum_i \mu_i \alpha_i \rho_{0i} < 0. \tag{VIII.8}$$

The instability corresponds to macroscopic phase separation, in the sense that it occurs for perturbations of infinite wave length, specifically for perturbations with wave number

$$\boldsymbol{q}^2 < -(DD_c)^{-1} \sum_i \mu_i \alpha_i \rho_{0i}, \tag{VIII.9}$$

with those having infinite wave length $\boldsymbol{q} \to 0$ being the first and most unstable. Importantly, the stability analysis also tells us about the stoichiometry of the different particle species at the onset of growth of the perturbation, which follows

$$(\delta\rho_1, \delta\rho_2, ..., \delta\rho_M) = \left(1, \frac{\mu_2\rho_{02}}{\mu_1\rho_{01}}, ..., \frac{\mu_M\rho_{0M}}{\mu_1\rho_{01}}\right)\delta\rho_1. \tag{VIII.10}$$

If only a single particle species is present ($M = 1$), the instability criterion (VIII.8) describes the well-known Keller-Segel instability (Keller and Segel, 1970), which simply says that the homogeneous state is stable for particles that repel each other ($\mu_1\alpha_1 > 0$), whereas particles that attract each other ($\mu_1\alpha_1 < 0$) tend to aggregate, with the end state being a featureless macroscopic cluster containing all particles. In contrast, as soon as we have mixtures of more than one species, the combination of the instability criterion (VIII.8) and the stoichiometric relation (VIII.10) predicts a wealth of new phase separation phenomena.

B. Phase Separation in Binary Mixtures

For binary mixtures ($M = 2$), the instability condition (VIII.8) becomes $\mu_1\alpha_1\rho_{01} + \mu_2\alpha_2\rho_{02} < 0$, and the stoichiometric constraint (VIII.10) implies that when μ_1 and μ_2 have equal or opposite sign, the instability will lead respectively to aggregation or separation of the two species. Combining these criteria we can construct a stability diagram for the binary mixture, although we must distinguish between two qualitatively-different kinds of mixtures: those of one producer and one consumer species, see Fig VIII.2(a) where we can choose $(\alpha_1, \alpha_2) = (+, -)$ without loss of generality; and those of two producer species, see Fig VIII.2(b). The case of two consumer species is related to the latter by the symmetry $(\mu_1, \mu_2) \to -(\mu_1, \mu_2)$. In this way, the parameter space for each type of mixture can be divided into regions leading to homogeneous, aggregated, or separated states, which correspond directly to those observed in simulations; compare Figs. VIII.2(a) and VIII.2(b) to Figs. VIII.1(c) and VIII.1(e). We note, however, that while for $(\alpha_1, \alpha_2) = (+, -)$ mixtures the simulations are always seen to match the predicted phase behaviour, for $(\alpha_1, \alpha_2) = (+, +)$ mixtures once can observe separation in the simulations even when the continuum theory predicts the homogeneous state to be linearly stable, although proceeding much more slowly, indicating that in this region separation may be occurring through a nucleation-and-growth process controlled by fluctuations. This is denoted as the shaded gray region extending past the instability line in Fig VIII.2(b).

The wide variety of phase separation phenomena arising in these mixtures is intimately related to the active, non-reciprocal character of the chemical interactions. In particular, it is useful to consider the sign of both inter-species as well as intra-species interactions. In the stability diagrams in Figs. VIII.2(a) andVIII.2(b), one finds that each quadrant corresponds to a distinct "interaction network" between species, as depicted in the boxed legends attached to every quadrant (as an example, the top-right interaction network in Fig. VIII.2(a) can be read as "1 is attracted to 2, 2 is repelled from 1, 1 is repelled from 1, and 2 is attracted to 2"). Note that only three regions in the parameter space have passive analogs: (i) The bottom-right of VIII.2(a) corresponds to electrostatics with opposite charges, where equals repel and opposites attract, allowing for the formation of small active molecules as studied in Section VII. (ii) The top-right of Fig. VIII.2(b) corresponds to electrostatics with like charges, where all interactions are repulsive leading to a homogeneous state. (iii) The bottom-left of Fig. VIII.2(b) corresponds to gravitation, where all interactions are attractive. The top-left of Fig. VIII.2(a) can be thought of as the opposite of electrostatics (or as gravitation including a negative mass species), where equals attract and opposites repel. The remaining four quadrants involve intrinsically non-reciprocal interactions where one species chases after the other: in Fig. VIII.2(a), a self-repelling species chases after a self-attracting species; whereas in Fig. VIII.2(b), a self-attracting species chases after a self-repelling species. Importantly, the most non-trivial instances of phase separation, which are also those that can be triggered simply by density changes (e.g. by addition or removal of particles), occur in regions with such chasing interactions, which are in turn a direct signature of non-equilibrium activity.

Fourier analysis of the Brownian dynamics simulations (44 simulations with varying N_i, α_i, and μ_i) agrees quantitatively with the theoretical prediction (VIII.10) for the stoichiometry at the onset of the instability; see Fig. VIII.2(c). However, this initial value is not representative of the long-time stoichiometry of the phases. For $(\alpha_1, \alpha_2) = (+, +)$ mixtures, shown in Figs. VIII.1(e) and VIII.2(b), we always observe final configurations with either complete aggregation or separation of the two species. For $(\alpha_1, \alpha_2) = (+, -)$ mixtures, shown in Figs. VIII.1(c) and VIII.2(a), we always observe complete separation, but in this case aggregation leads to a cluster with non-trivial stoichiometry [Fig. VIII.1(c), centre]. Phenomenologically, we observe that cluster formation in this case proceeds by fast initial aggregation of the particles of the self-attractive species ($\alpha_i\mu_i < 0$) followed by slower recruitment of particles of the self-repelling species ($\alpha_i\mu_i > 0$) until the cluster is chemically "neutral", in the sense that its net consumption or production of chemicals vanishes, namely

$$\alpha_1 N_1^{\mathrm{clu}} + \alpha_2 N_2^{\mathrm{clu}} = 0, \tag{VIII.11}$$

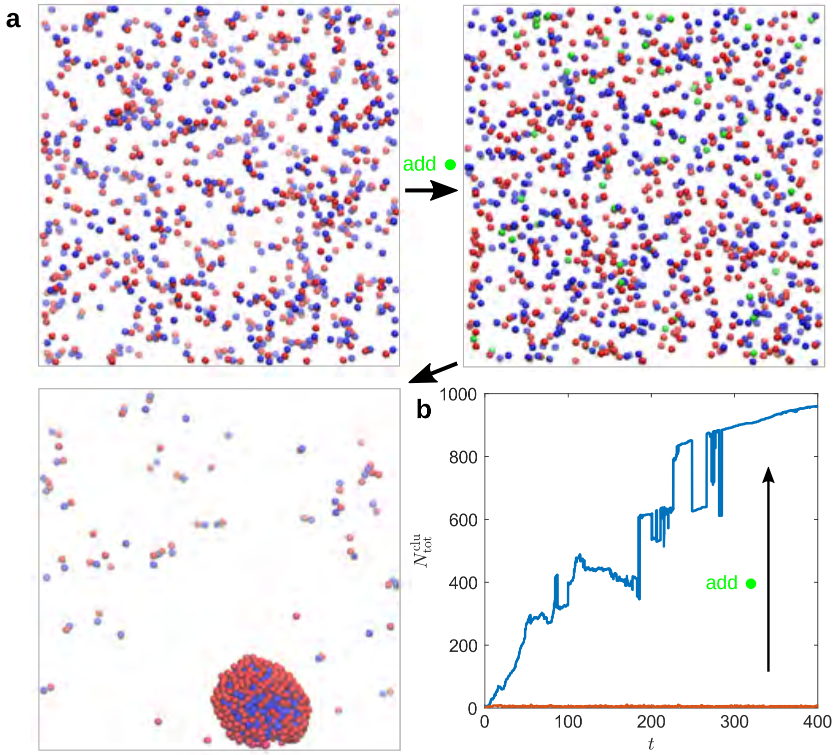


FIG. VIII.3 Phase separation induced by a small amount of an active "doping agent". (a) Simulation snapshots showing macroscopic aggregation of a previously homogeneous mixture ($N_1 = N_2 = 500$, $\tilde{\alpha}_1 = \tilde{\mu}_1 = 1$, $\tilde{\alpha}_2 = -1$, $\tilde{\mu}_2 = 1/2$) after addition of 5 % of a third species ($N_3 = 50$, $\tilde{\alpha}_3 = -5$, $\tilde{\mu}_3 = 2$). (b) Time evolution of the size of the largest cluster (total number of particles), in the absence and presence of the third species.

where $N_i^{\rm clu}$ is the number of particles of species i in the cluster. The long-time stoichiometry of the clusters thus depends on the activity of the species, but it is independent of their mobility; see Fig. VIII.2(d). An intuitive explanation for this observation can be provided as follows: once the cluster becomes neutral, the remaining self-repelling particles will no longer "sense" its presence and stay in a dilute phase. At high values of activity and mobility for the self-attractive species, however, these static neutral clusters can become unstable *via* shape-symmetry breaking towards a self-propelled asymmetric cluster [Fig. VIII.1(d)], which also involves the "shedding" of some of the self-repelling particles. Such self-propelled clusters are possible only thanks to the existence of non-reciprocal interactions.

C. Beyond Binary Mixtures

Going beyond binary mixtures ($M > 2$), the phase separation phenomenology becomes even more complex due to the increasing number of parameter combinations, leading to a large variety of possible interaction networks between the different species. The instability condition (VIII.8), however, remains extremely useful. Figure VIII.3 demonstrates as an example how a small amount of a highly active "dopant" third species can be added to an otherwise homogeneous binary mixture in order to trigger macroscopic phase separation of the whole mixture on demand. Moreover, the instability condition (VIII.8) can also be used to predict whether highly polydisperse mixtures will phase separate or remain homogeneous, see Fig. VIII.1(f).

IX. POLAR ACTIVE COLLOIDS: MOMENT EXPANSION

The description of the collective behaviour of polar active colloids is considerably more complicated than apolar particles due to the additional complexity that arises from the coupling between polarity and motion. Here we develop a systematic framework that can accommodate this complexity in terms of a hierarchical expansion in terms of the moments of the distribution in the Fokker-Planck equation (Golestanian, 2012; Saha *et al.*, 2014).

A. From Trajectories to Hydrodynamic Equations

We consider a collections of N polar spherical particles of radius R and describe the configuration of a particle labeled i with position $\boldsymbol{r}_i(t)$ and orientation $\boldsymbol{n}_i(t)$. The particle experiences deterministic translational velocity $\boldsymbol{v}_i$ and angular velocity $\boldsymbol{\omega}_i$, as well as noise characterized by $D_{\rm c}$ and $D_{\rm r}$, which are the translational and rotational diffusion coefficients. In a medium with uniform temperature T, we have $D_{\rm c} = \frac{k_{\rm B}T}{6\pi\eta R}$ and $D_{\rm r} = \frac{k_{\rm B}T}{8\pi\eta R^3}$, where η is the viscosity of water. The resulting general Langevin equations for the translational and rotational degrees of freedom are as follows

$$\frac{\rm d}{{\rm d}t}\boldsymbol{r}_i(t) = \boldsymbol{v}_i + \boldsymbol{\xi}_i, \tag{IX.1}$$

$$\frac{\rm d}{{\rm d}t}\boldsymbol{n}_i(t) = \boldsymbol{\omega}_i + \boldsymbol{\eta}_i \times \mathbf{n}_i, \tag{IX.2}$$

where $\boldsymbol{\xi}_i$ and $\boldsymbol{\eta}_i$ are Gaussian-distributed white noise terms of unit strength. From the stochastic trajectories, we can define the probability distribution

$$\mathcal{P}(\boldsymbol{r}, \boldsymbol{n}, t) \equiv \left\langle \sum_{i=1}^{N} \delta\left(\boldsymbol{r} - \boldsymbol{r}_i(t)\right) \delta\left(\boldsymbol{n} - \boldsymbol{n}_i(t)\right) \right\rangle,$$

and recast the Langevin equations, which are governing equations for the trajectories, into an evolution equation for the probability distribution. Defining the rotational gradient operator $\boldsymbol{\mathfrak{R}} \equiv \boldsymbol{n} \times \partial_{\boldsymbol{n}}$, which has the properties $\mathcal{R}_\alpha n_\beta = -\epsilon_{\alpha\beta\gamma} n_\gamma$ and $\mathcal{R}^2 n_\beta = -2n_\beta$, allows us to construct the translational and the rotational fluxes as follows

$$\boldsymbol{J} = \boldsymbol{v}(\boldsymbol{r}, \boldsymbol{n})\ \mathcal{P}(\boldsymbol{r}, \boldsymbol{n}) - D_{\rm c} \boldsymbol{\nabla} \mathcal{P}(\boldsymbol{r}, \boldsymbol{n}), \tag{IX.3}$$

$$\boldsymbol{J}_r = \boldsymbol{\omega}(\boldsymbol{r}, \boldsymbol{n})\ \mathcal{P}(\boldsymbol{r}, \boldsymbol{n}) - D_{\rm r} \boldsymbol{\mathfrak{R}} \mathcal{P}(\boldsymbol{r}, \boldsymbol{n}). \tag{IX.4}$$

Then, we can write the Fokker-Planck equation as a conservation law

$$\partial_t \mathcal{P} + \nabla \cdot \boldsymbol{J} + \boldsymbol{\mathfrak{R}} \cdot \boldsymbol{J}_r = 0. \tag{IX.5}$$

To describe the collective behaviour of active particles with phoretic interactions plus translational self-propulsion, we choose the following forms for the velocities

$$\boldsymbol{v}(\boldsymbol{r}, \boldsymbol{n}) = v_0 \boldsymbol{n} - \mu \boldsymbol{\nabla} \psi, \tag{IX.6}$$

$$\boldsymbol{\omega}(\boldsymbol{r}, \boldsymbol{n}) = \chi \boldsymbol{n} \times \boldsymbol{\nabla} \psi, \tag{IX.7}$$

where v_0 is the self-propulsion speed and represents a thermodynamic potential such as solute concentration (diffusiophoresis), electrostatic potential (electrophoresis), or temperature (thermophoresis). The resulting Fokker-Planck equation reads

$$\partial_t \mathcal{P} + \nabla \cdot \left[\left(v_0 \boldsymbol{n} - \mu \boldsymbol{\nabla} \psi \right) \mathcal{P} - D_{\rm c} \boldsymbol{\nabla} \mathcal{P} \right] + \boldsymbol{\mathfrak{R}} \cdot \left[\left(\chi \boldsymbol{n} \times \boldsymbol{\nabla} \psi \right) \mathcal{P} - D_{\rm r} \boldsymbol{\mathfrak{R}} \mathcal{P} \right] = 0. \tag{IX.8}$$

Equation (IX.8) is rather complex as it mixes orientation and position. A systematic approximation framework called moment expansion helps us to tackle this complication. The method builds on the orientation moments of the distribution, namely, the density $\rho(\boldsymbol{r}) = \int_{\boldsymbol{n}} \mathcal{P}(\boldsymbol{r}, \boldsymbol{n})$, the polarization field $\boldsymbol{p}(\boldsymbol{r}) = \int_{\boldsymbol{n}} \boldsymbol{n}\, \mathcal{P}(\boldsymbol{r}, \boldsymbol{n})$, the nematic order parameter $\boldsymbol{Q}(\boldsymbol{r}) = \int_{\boldsymbol{n}} \left[\boldsymbol{n}\boldsymbol{n} - \frac{1}{3}\boldsymbol{I}\right] \mathcal{P}(\boldsymbol{r}, \boldsymbol{n})$ etc, and a hierarchy of equations derived from Eq. (IX.8), which connect them.

The governing equation for the zeroth moment of orientation is obtain by integrating Eq. (IX.8) over $\boldsymbol{n}$. This gives

$$\partial_t \rho + v_0 \boldsymbol{\nabla} \cdot \boldsymbol{p} - \mu \boldsymbol{\nabla} \cdot \left[(\boldsymbol{\nabla}\psi)\, \rho\right] - D_{\mathrm{c}} \boldsymbol{\nabla}^2 \rho = 0, \tag{IX.9}$$

which has a source term in the form of $-v_0 \nabla \cdot \mathbf{p}$ due to the self-propulsion of the colloids. Performing $\int_{\mathbf{n}} \mathbf{n}\times$ Eq. (IX.8), we can obtain an equation for the polarization field as

$$\partial_t \boldsymbol{p} + \frac{v_0}{3} \boldsymbol{\nabla} \rho + v_0 \boldsymbol{\nabla} \cdot \boldsymbol{Q} - \mu \partial_\alpha \left[(\partial_\alpha \psi)\, \boldsymbol{p}\right] - D_{\mathrm{c}} \boldsymbol{\nabla}^2 \boldsymbol{p} + \chi \boldsymbol{Q} \cdot \boldsymbol{\nabla}\psi - \frac{2}{3}\chi \rho \boldsymbol{\nabla}\psi + 2D_{\mathrm{r}} \boldsymbol{p} = 0. \tag{IX.10}$$

Continuing this process will produce the interconnected hierarchy of equations for the moments. To make further progress, we truncate the hierarchy so that we can deal with a finite number of equations. For sufficiently dilute solutions (i.e. when $\rho R^3 \ll 1$) and in the absence of any external mechanisms that can lead to alignment, such as external fields or boundaries (Enculescu and Stark, 2011; Palacci *et al.*, 2010), we can ignore the nematic order and set $\boldsymbol{Q} \simeq 0$. This yields

$$\partial_t \boldsymbol{p} + \frac{v_0}{3} \boldsymbol{\nabla} \rho - \mu \partial_\alpha \left[(\partial_\alpha \psi)\, \boldsymbol{p}\right] - D_{\mathrm{c}} \boldsymbol{\nabla}^2 \boldsymbol{p} - \frac{2}{3}\chi \rho \boldsymbol{\nabla}\psi + 2D_{\mathrm{r}} \boldsymbol{p} = 0. \tag{IX.11}$$

B. Self-consistent Field Equations

To complete the description of the system, we need to specify how the field is generated by the phoretically active particles. A generic governing equation for the field can be written as

$$\cancel{\partial_t \psi} - K \boldsymbol{\nabla}^2 \psi = \text{sources and sinks}, \tag{IX.12}$$

where K represents the solute diffusion coefficient (diffusiophoresis) or the heat conductivity (thermophoresis) etc. The time derivative term is ignored because we are interested in the long time limit and assume that solute, heat, etc diffuse much faster than the colloids. Assuming a surface activity coverage $\alpha_i(\Omega_i)$ for the ith colloid, we can describe the right hand side of Eq. (IX.12) as follows

$$\begin{aligned} -K \boldsymbol{\nabla}^2 \psi &= \sum_i \int \mathrm{d}S_i\ \alpha_i(\Omega_i)\ \delta(\boldsymbol{r} - \boldsymbol{r}_i - R\hat{\boldsymbol{R}}_i) \\ &= \sum_i \int \mathrm{d}S_i\ \alpha_i(\Omega_i) \left[\delta(\boldsymbol{r} - \boldsymbol{r}_i) - R\hat{\boldsymbol{R}}_i \cdot \boldsymbol{\nabla}\delta(\boldsymbol{r} - \boldsymbol{r}_i) + \cdots\right] \\ &= R^2 \sum_i \left[\int \mathrm{d}\Omega_i\ \alpha_i(\Omega_i)\right] - R^3 \sum_i \left[\int \mathrm{d}\Omega_i\ \alpha_i(\Omega_i)\hat{\boldsymbol{R}}_i\right] \cdot \delta(\boldsymbol{r} - \boldsymbol{r}_i) + \cdots . \end{aligned} \tag{IX.13}$$

Assuming all colloids are the same and using the expansion of Eq. (IV.2), we find $\int \mathrm{d}\Omega \alpha(\Omega) = 4\pi\alpha_0$ and $\int \mathrm{d}\Omega \alpha(\Omega)\hat{\boldsymbol{R}} = \frac{4\pi}{3}\alpha_1 \boldsymbol{n}$. Therefore, Eq. (IX.13) reads

$$-K \boldsymbol{\nabla}^2 \psi = 4\pi R^2 \alpha_0 \sum_i \delta(\boldsymbol{r} - \boldsymbol{r}_i) - \frac{4\pi}{3} R^3 \alpha_1 \boldsymbol{\nabla} \cdot \left[\sum_i \boldsymbol{n}_i \delta(\boldsymbol{r} - \boldsymbol{r}_i)\right] + \cdots . \tag{IX.14}$$

The above form of the equation for highlights its stochastic nature. To proceed, we implement a mean-field approximation and replace ψ by its average over the trajectories, which yields

$$-K\boldsymbol{\nabla}^2\psi = 4\pi R^2\alpha_0\rho - \frac{4\pi}{3}R^3\alpha_1\boldsymbol{\nabla}\cdot\boldsymbol{p} + \cdots\,. \tag{IX.15}$$

This equation now complements Eqs (IX.9) and (IX.11) for a complete approximate description of the system.

C. Behaviour at Long Times and Large Length Scales

We are interested in the behaviour of the system at time scales sufficiently longer than the rotational diffusion time of the colloids D_r^{-1} and lengths much larger than the size of the colloid R. A description of this regime can achieved by ignoring a number of terms in Eq. (IX.11) as follows

$$\xcancel{\partial_t\boldsymbol{p}} + \frac{v_0}{3}\boldsymbol{\nabla}\rho - \mu\left(\boldsymbol{\nabla}^2\psi\right)\boldsymbol{p} - \xcancel{\mu\left(\partial_\alpha\psi\right)\left(\partial_\alpha\boldsymbol{p}\right)} - \xcancel{D_c\boldsymbol{\nabla}^2\boldsymbol{p}} - \frac{2}{3}\chi\rho\boldsymbol{\nabla}\psi + 2D_\mathrm{r}\boldsymbol{p} = 0. \tag{IX.16}$$

Note that the 4th term is similar to the 3rd term, but it adds a tensorial structure to the equation. We have ignored it here for simplicity. Inserting an approximate form of Eq. (IX.15), namely $\boldsymbol{\nabla}^2\psi \simeq -4\pi R^2\alpha_0\rho/K$, in Eq. (IX.16), we obtain

$$\left[(d-1)D_\mathrm{r} + S_d R^{d-1}\,\frac{\alpha_0\mu}{K}\,\rho\right]\boldsymbol{p} = -\frac{v_0}{d}\boldsymbol{\nabla}\rho + \frac{d-1}{d}\,\chi\rho\boldsymbol{\nabla}\psi, \tag{IX.17}$$

where we have written the coefficients explicitly in terms of the dimensionality of space d. Here $S_d = 2\pi^{d/2}/\Gamma(d/2)$ is the surface area of the unit sphere embedded in d dimensions. From Eq. (IX.17), we can find an explicit expression for the polarization in terms of the density and the field, which reads

$$\boldsymbol{p} = \frac{\left(-\frac{v_0}{3}\boldsymbol{\nabla}\rho + \frac{2}{3}\chi\rho\boldsymbol{\nabla}\psi\right)}{\left(2D_\mathrm{r} + \frac{4\pi R^2\alpha_0\mu}{K}\,\rho\right)}, \tag{IX.18}$$

in $d = 3$. Note that phoretic interaction renormalizes the rotational diffusion of the colloids. Setting $\rho \simeq \rho_0 = \mathrm{const}$ in the denominator of Eq. (IX.18) and inserting the resulting form for $\boldsymbol{p}$ back into Eq. (IX.9), we obtain the following equation for the density field

$$\partial_t\rho + \boldsymbol{\nabla}\cdot\boldsymbol{J}_\mathrm{eff} = 0, \tag{IX.19}$$

where the effective flux is defined as

$$\boldsymbol{J}_\mathrm{eff} = -D_\mathrm{eff}\boldsymbol{\nabla}\rho - \mu_\mathrm{eff}(\boldsymbol{\nabla}\psi)\rho, \tag{IX.20}$$

in terms of the effective diffusion coefficient

$$D_\mathrm{eff} = D_\mathrm{c} + \frac{v_0^2}{6(D_\mathrm{r} + 2\pi R^2\alpha_0\mu\rho_0/K)}, \tag{IX.21}$$

and the effective phoretic mobility

$$\mu_\mathrm{eff} = \mu - \frac{v_0\chi}{3(D_\mathrm{r} + 2\pi R^2\alpha_0\mu\rho_0/K)}. \tag{IX.22}$$

We thus find that self-propulsion leads to an enhancement of the translational diffusion of the colloid on time scales longer than the rotational diffusion, while the combination of phoretic alignment and self-propulsion leads to a renormalization of the translational phoretic mobility at long times.

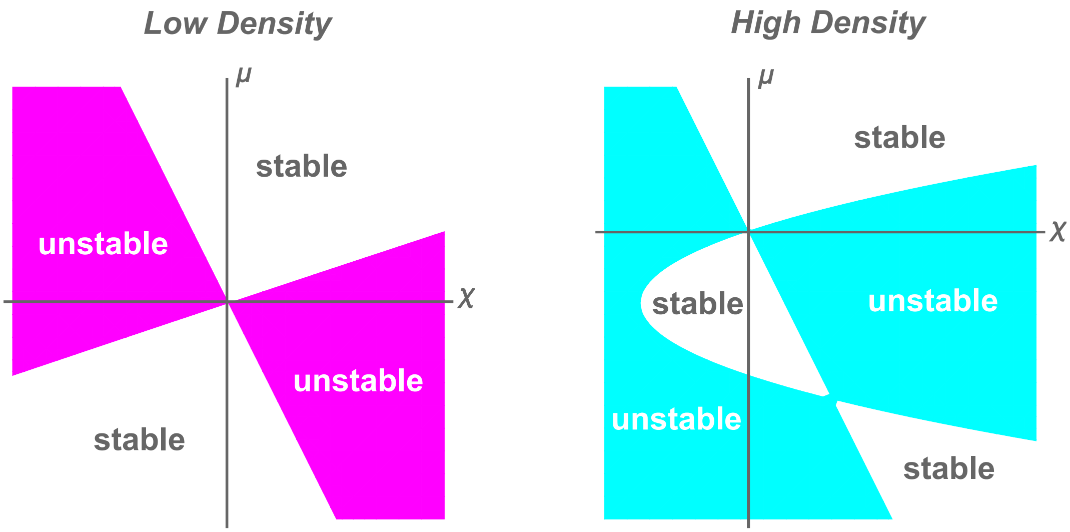


FIG. IX.1 Stability of the suspension of polar phoretic active colloids as determined by the condition $\Gamma < 0$, which ensures that the colloids will point in the direction of higher concentrations and actively swim towards the denser region.

In stationary state, Eq. (IX.19) is satisfied if $\boldsymbol{J}_{\text{eff}} = 0$, which yields

$$\rho(\boldsymbol{r}) = \rho_0 \, \exp\left[-\frac{\mu_{\text{eff}}}{D_{\text{eff}}}\psi(\boldsymbol{r})\right]. \tag{IX.23}$$

This is a generalized Boltzmann distribution, which will allow us to use analogies to equilibrium theories of electrolytes.

1. Stationary State Polarization

We can insert the stationary distribution into Eq. (IX.18) to find a direct relationship between the polarization and the density gradient as follows

$$\boldsymbol{p} = -\Gamma \boldsymbol{\nabla}\rho, \tag{IX.24}$$

where the response coefficient is given as

$$\Gamma = \frac{\frac{D_{\text{c}}}{v_0}\chi + \mu/2}{3\frac{D_{\text{r}}}{v_0}\left(1 + \frac{2\pi R^2 \alpha_0 \mu \rho_0}{K D_{\text{r}}}\right)\mu - \chi}, \tag{IX.25}$$

in terms of μ and χ, which can both be either positive or negative. The parameters can thus be tuned such that $\Gamma > 0$, in which case polarization tends to stabilize accumulation of particles via a tendency for the particles to swim away from high density regions. When $\Gamma < 0$, on the other hand, the particles tend to be aligned with the concentration gradient and the particles tend to swim towards already crowded regions, hence instigating an instability. Therefore, the alignment or polarization tendencies of the system as controlled by χ will determine the phase behaviour of the system in competition with the translational or positional tendencies that are controlled by μ (see Fig. IX.1).

Going back to Eq. (IX.15), we can now eliminate the polarization by inserting its explicit form from Eq. (IX.18). This yields

$$-\boldsymbol{\nabla}\cdot\big(K_{\text{eff}}\boldsymbol{\nabla}\psi\big) = 4\pi R^2\alpha_0\rho + \xcancel{\frac{2\pi}{9}\frac{R^2\alpha_1 v_0}{(D_{\text{r}}+2\pi R^2\alpha_0\mu\rho_0/K)}}\boldsymbol{\nabla}^2\rho + \cdots\,, \tag{IX.26}$$

where the K coefficient is renormalized due to the polarization of the Janus particles as follows

$$K_{\text{eff}} = K - \frac{R^3\alpha_1\chi\rho}{(D_{\text{r}}+2\pi R^2\alpha_0\mu\rho_0/K)}. \tag{IX.27}$$

This phenomenon is analogous to the emergence of the polarization field inside dielectric material, which is accounted for by an effective dielectric constant that reduces or screens the field. We can simplify further and assume a constant density profile in Eq. (IX.27), and therefore treat K_{eff} as a constant. Putting Eq. (IX.23) into this simplified form of Eq. (IX.26), we find

$$-K_{\text{eff}}\boldsymbol{\nabla}^2\psi = 4\pi R^2\alpha_0\rho_0\ \exp\left[-\frac{\mu_{\text{eff}}}{D_{\text{eff}}}\psi(\boldsymbol{r})\right]. \tag{IX.28}$$

This equation is reminiscent of the Poisson-Boltzmann equation for electrolytes, which should be solved subject to the normalization constraint

$$N = \int \mathrm{d}^d\boldsymbol{r}\ \rho(\boldsymbol{r}) = \rho_0\int \mathrm{d}^d\boldsymbol{r}\ \exp\left[-\frac{\mu_{\text{eff}}}{D_{\text{eff}}}\psi(\boldsymbol{r})\right]. \tag{IX.29}$$

We can identify two distinct classes described by the above equations:

- **Electrostatic**, in which like charges predominantly repel. This corresponds to $\alpha_0\mu_{\text{eff}} > 0$.
- **Gravitational**, in which like charges predominantly attract. This corresponds to $\alpha_0\mu_{\text{eff}} < 0$.

We can define a dimensionless field as

$$\Psi \equiv \frac{\mu_{\text{eff}}\psi}{D_{\text{eff}}}\cdot\text{sgn}(\alpha_0\mu_{\text{eff}}), \tag{IX.30}$$

and a characteristic Bjerrum length scale

$$\ell \equiv \frac{R^2|\alpha_0\mu_{\text{eff}}|}{K_{\text{eff}}D_{\text{eff}}}, \tag{IX.31}$$

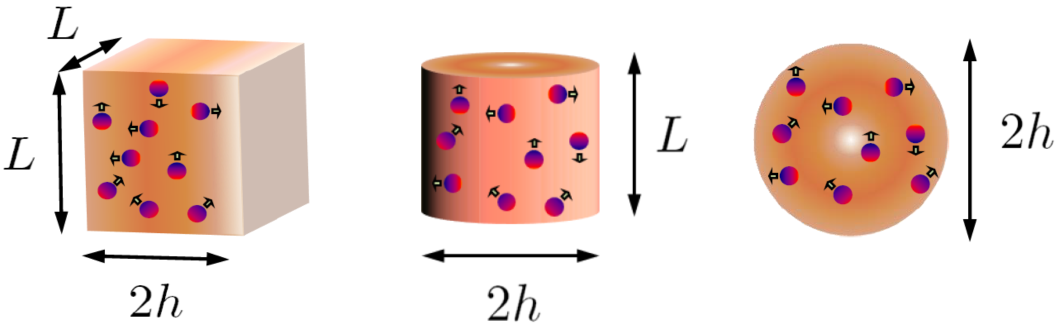


FIG. IX.2 Different geometries for the electrostatic and gravitational cases in analogy with cases where Poisson-Boltzmann equation has been studied.

as well as a corresponding Debye length κ^{-1} defined via

$$\kappa^2 = 4\pi\ell\rho_0. \tag{IX.32}$$

Then our Poisson-Boltzmann equation reads

$$-\nabla^2\Psi = \kappa^2 e^{\mp\Psi}, \tag{IX.33}$$

where the sign choice is $\mp = -\mathrm{sgn}(\alpha_0\mu_{\text{eff}})$. Equation (IX.33) is subject to the constraint $N = \rho_0 \int \mathrm{d}^d\boldsymbol{r}\ e^{\mp\Psi}$.

In analogy with studies of Poisson-Boltzmann equation, we can look for exact solutions of the above equation under confinement, by applying the following boundary condition

$$-\partial_\perp\Psi|_S = 4\pi\ell N/A, \tag{IX.34}$$

which we obtain by invoking Gauss theorem, as well as symmetry considerations. We will now consider a number of different geometries as shown in Fig. IX.2.

In the electrostatic case, we can obtain exact solutions in cases with 1D and 2D confinement (Levin, 2002). When the colloids are confined between two plates of lateral size L and distance $2h$, the exact density profile is found as

$$\rho(x) = \frac{\rho_0}{\left[1 + \frac{2\pi^2\ell^2}{k^2}\left(\frac{N}{L^2}\right)^2\right]\cos^2\left(\frac{kx}{\sqrt{2}}\right)}, \tag{IX.35}$$

where ρ_0 is the concentration at the edge of the confining wall, and k satisfies the following transcendental equation

$$\left(\frac{kh}{\sqrt{2}}\right)\tan\left(\frac{kh}{\sqrt{2}}\right) = \pi\ell h\left(\frac{N}{L^2}\right), \tag{IX.36}$$

The profile of Eq. (IX.35) describes an accumulation of the colloids near the confining boundary that is analogous to the phenomenon of *counterion condensation* (Levin, 2002), and a resulting depletion zone in the central region of the system. In the strong coupling limit when $N\ell h/L^2 \gg 1$, we can obtain an approximate solution to Eq. (IX.36) as $kh \simeq \frac{\pi}{\sqrt{2}}\left[1 - \frac{1}{\pi(N\ell h/L^2)}\right]$. In this limit, the ratio between the density of the colloids in the middle and at the edge can be found as $\frac{\rho_{\mathrm{m}}}{\rho_0} \simeq \frac{1}{4}\left(N\ell h/L^2\right)^{-2}$, which shows

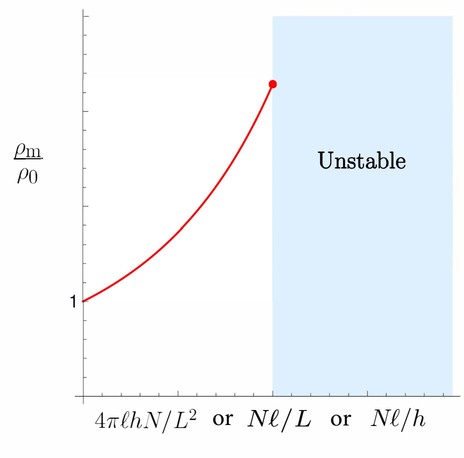


FIG. IX.3 The density of colloids in the middle of the confined space relative to the density at the edge as a function of the dimensionless phoretic coupling constant for the gravitational case.

a significant depletion effect. Note that the depletion becomes stronger as h is increased, when other parameters are kept fixed. The length scale $1/[2\pi\ell(N/L^2)]$ is equivalent to the Gouy-Chapman length in the electrostatic analogy (Levin, 2002).

For an active colloidal solution confined in a cylindrical cage of length L and width $2h$, the density profile is obtained as

$$\rho(r) = \frac{\rho_0}{\left[1+\frac{1}{2}\left(\frac{N\ell}{L}\right)\right]^2\left[1-\frac{1}{8}k^2r^2\right]^2}, \tag{IX.37}$$

where k is given by following closed-form expression

$$kh = \sqrt{\frac{8(N\ell/L)}{2+(N\ell/L)}}. \tag{IX.38}$$

In this geometry, the strong coupling limit corresponds to $N\ell/L \gg 1$, in which case we obtain a measure of depletion as follows $\frac{\rho_{\mathrm{m}}}{\rho_0} \simeq 4\,(N\ell/L)^{-2}$, which is independent of the confinement size in this geometry. The ratio $\ell N/L$ is analogous to the so-called Manning-Oosawa parameter for highly charged rodlike polyelectrolytes (Levin, 2002). The same type of profile is obtained when the colloidal solution is confined in 3D to a spherical cage of diameter $2h$, where in the strong coupling limit defined via $N\ell/h \gg 1$, we have $\frac{\rho_{\mathrm{m}}}{\rho_0} \simeq 21.4\,(N\ell/h)^{-2}$. Here, the depletion is inversely related to the size of the cage, i.e. it decreases for larger cages.

In the gravitational case, we observe accumulation of the colloidal particles at the centre of the confined region, which contrasts from the electrostatic case, while the potential profile $\Psi(\boldsymbol{r})$ is still peaked at the centre as in the electrostatic case. The relevant coupling constants denoted as $g(d)$ are, $g(1) = N\ell h/L^2$ (1D), $g(2) = N\ell/L$ (2D), $g(3) = N\ell/h$ (3D), as discussed above. As $g(d)$ increases, the ratio ρ_{m}/ρ_0 increases as well, signalling accumulation at the centre. This structure, which is still a relatively dilute as of particles that are free to diffuse within the confined region, is analogous to a "colloidal star", in the gravitational analogy. This dilute structure is stable up to a critical point $g_c(d)$ beyond which a stable (stationary-state) solution no longer exists; see Fig. IX.3. For example, in 1D the onset of instability occurs at $g_c(1) = 1/2\pi$, at which $(\rho_{\mathrm{m}}/\rho_0)_c = 3.29$, while similar thresholds hold for the 2D and 3D confinement cases (Landau and Lifshitz, 2013). The instability occurs because the particles that act as sources for the potential attract each other and result in a suspension that becomes increasingly denser and more attractive. In this case, the flux at the outer boundary of the system cannot balance the field generated inside the confined region, which leads to an uncontrolled buildup of thermodynamic energy associated with the potential Ψ. This state of the system can be called a "colloidal supernova" in our gravitational analogy. However, we should bear in mind that the analogy is not exact as the colloidal system operates in the dissipative regime, in contrast with the inertial and conserved dynamics of the gravitational system.

3. Additional Generalizations

In our simplified description, we have so far ignored a number of important features that will affect the dynamics of catalytically active Janus particles. Typically, the catalytic activity on Janus particles involves reactions that convert *substrates* (as reactants are called in the biochemistry literature) into *products*, i.e. S→P. This means that there a number of chemical species in the solution, each producing their own gradients and contributing to phoretic transport with their corresponding mobilities, as shown in Eq. (III.6).

Moreover, the mobilities will in general be tensors for Janus particles, allowing in general different drift velocities along the polar axis of the particle and perpendicular to it. Additionally, the Janus structure

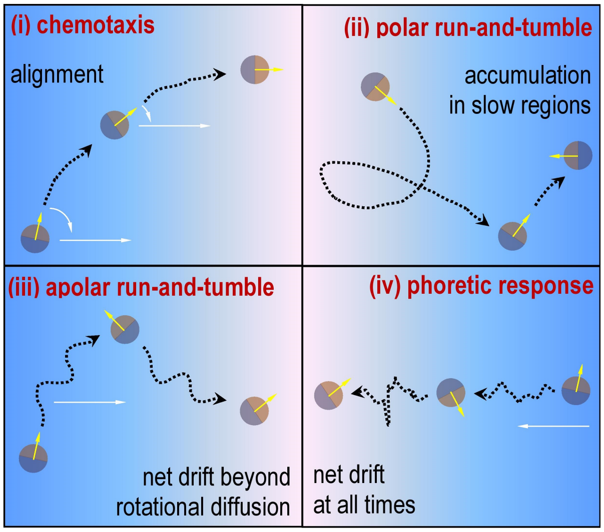


FIG. IX.4 Four different mechanisms or channels through which a Janus particle responds to gradients in substrate concentration (Saha *et al.*, 2014).

will introduce an alignment response to a concentration gradient due to the angular velocity given in Eq. (I.25). Finally, the variations in the local concentration of the substrate molecule that fuels the propulsion will also modulate the swimming velocity v_0. Putting together all these contributions, we find that a single Janus particle can respond to variations in substrate concentrate via four different channels or mechanisms as summarized in Fig. IX.4.

Taking these effects into consideration will give us a complex phase diagram that includes a range of collective dynamical regimes including clustering, pattern formation, aster condensation, plasma oscillations, and spontaneous oscillations (Saha *et al.*, 2014). The existence of such a range of different regimes can be traced back to different possibilities provided by the positional and orientational interactions, as discussed in Sec. IX.A above. When $\mu < 0$ and $\chi > 0$, the particles are translationally attracted to one another while they would also tend to orient towards each other and swim towards one another; this is a clear cut case of collapse instability. If, on the other hand, $\mu > 0$ and $\chi > 0$, they repel each while they tend to orient towards each other and swim to one another; this is a frustrated case, which can lead to oscillations and pattern formation. Similar observations have been reported from studies using Brownian dynamics simulations (Pohl and Stark, 2014; Stark, 2018). Enhanced density fluctuations and clustering that can arise from phoretic instabilities as discussed above have been observed experimentally in suspensions of catalytic Janus swimmers (Palacci *et al.*, 2010, 2013).

X. POLAR ACTIVE COLLOIDS: SCATTERING AND ORBITING

The existence of different modes of chemotactic coupling to position and orientation has interesting implications on how two active colloids interact with one another (Saha *et al.*, 2019).

When a polar active colloids interacts with an apolar source of chemical, two different types of behaviour can emerge (see Fig. X.1). An active colloid that tends to align with the local gradient of an externally imposed chemical field can be trapped by a source of fuel. A trapped swimmer either comes to rest at a fixed distance from the source or executes periodic orbits. By tuning initial condition, it is possible to transition to a state where the swimmer interacts with the source for a short period before running away, thereby undergoing scattering.

Two interacting chemotactic active colloids, which can rotate their polar axis to align with an external chemical gradient, form new bound states by cancellation of velocities rather than by minimization of a

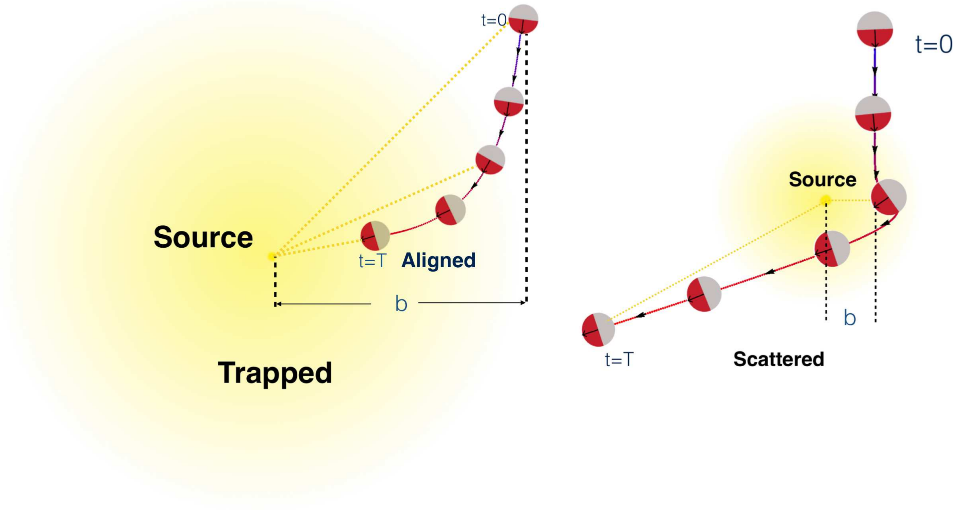


FIG. X.1 Typical trajectories of an active colloid in an isotropic source of fuel, illustrating the role of initial conditions and active rotation of the polar axis. The swimmer can be captured by a source of chemicals or interact with it briefly before scattering, depending on initial conditions. If it approaches the source with an impact parameter b lower than a threshold, a bound state is formed when self-propulsion and phoretic interactions balance each other. The swimmer on the left is trapped when the polarity of the aligns with the local chemical gradient marked with dotted orange lines. The swimmer on the right is scattered as it cannot be trapped. (Figure by Suropriya Saha).

free energy. The interactions are dynamical in origin, resulting from an interplay of self propulsion and gradient-seeking mechanisms, and are thus non-central and non-reciprocal. Bound states are formed where the distance between their centres and relative orientation of their polarity remains fixed or traces a periodic cycle. These states fall in two broad classes: (i) active dimers, where the centre of mass translates linearly and (ii) orbits, where the centre of mass moves in a closed orbit. A necessary condition is that the chemotactic alignment response of at least one colloid in the pair must be positive. Similarly to the case of a single swimmer near a source, they can unbind and scatter when the surface activity is changed. The fixed points underlying the bound states correspond to the case when exactly one of the two colloids is stationary, and show that the transition happens through bifurcations. These findings are robust upon the introduction of hydrodynamic interactions and (relevant) thermal fluctuations (Saha *et al.*, 2019). Similar effects have been studied for a system of active colloids under confinement (Kanso and Michelin, 2019).

XI. NON-EQUILIBRIUM DYNAMICS OF ACTIVE ENZYMES

Enzymes are molecular machines that catalyze chemical reactions. The appropriate description of a chemical reaction is a Kramers (escape) process in the reaction space in which the system aims to go from an initial higher energy state, which corresponds to the substrate to a final lower energy state, corresponding to the product, by overcoming an energy barrier. A schematic description of how enzymes work can be constructed as follows (see Fig. XI.1). Consider a chemical reaction S→P as an activated process along a specific reaction coordinate, with a barrier that is considerably larger than $k_{\rm B}T$; this transition will happen very slowly on its own. An enzyme can speed up this reaction if upon binding to the substrate it can effectively lower the barrier, or perhaps more accurately, open up a new trajectory with a lower barrier that was not accessible before. Therefore, enzymes are drivers of non-equilibrium activity at the right time at the right place.

For the reaction path described in Fig. XI.1, the overall rate of product formation follows the so-called

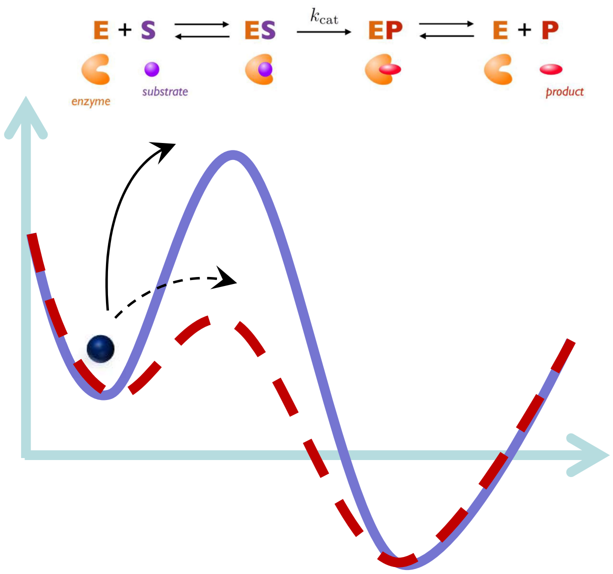


FIG. XI.1 The interaction of an enzyme and its substrate effectively lowers the energy barrier along the relevant reaction coordinate and speed up a chemical reaction, which would normally happen, but at a much slower rate. Solid line represents a free reaction and dashed lined represents a reaction that is facilitated by an enzyme.

Michaelis-Menten rule

$$\frac{\mathrm{d}}{\mathrm{d}t}P = kE_0, \tag{XI.1}$$

where E_0 is the bulk concentration of enzymes, and k is the effective catalytic reaction for a single enzyme, given as

$$k = k_{\text{cat}}\,\frac{S}{K+S}, \tag{XI.2}$$

with K being the Michaelis constant.

A. Enhanced Diffusion of Enzymes

There have been a number of experimental reports on the effect of catalytic activity on the diffusion of enzymes (Muddana *et al.*, 2010; Riedel *et al.*, 2015; Sengupta *et al.*, 2013, 2014). Typically, the enzymes have been found to undergo diffusion with an effective diffusion coefficient that depends on the substrate concentration, which can be approximately described via

$$D_{\text{eff}}(S) = D_0 + \Delta D\,\frac{S}{K+S}, \tag{XI.3}$$

where $D_0 \equiv D(S=0)$, and $\Delta D/D_0$ is often of the order of a fraction of one (ten percent or so) (Muddana *et al.*, 2010; Riedel *et al.*, 2015; Sengupta *et al.*, 2013, 2014).

There are several mechanisms that can contribute to enhanced diffusion of enzymes with varying degrees of significance (Agudo-Canalejo *et al.*, 2018a; Golestanian, 2015):

(i) self-phoresis, due to self-generated chemical gradients or temperature gradients if the reactions are exothermic. This contribution can typically yield $\Delta D/D_0 \sim 10^{-16}$ for fast enzymes.

(ii) boost in kinetic energy, as caused by the release of the energy of the reaction to the centre of mass translational degrees of freedom by equipartition. This mechanism leads to an effective diffusion coefficient given by

$$D_{\text{eff}} = D_0 \left[1 + \frac{2}{3}\frac{\gamma Q}{k_{\text{B}}T}\, k\tau_b\right], \tag{XI.4}$$

where γ represents the fraction of the released energy of reaction Q that is transferred to the centre of mass, and τ_b is the time scale characterizing the decay of the inertial boost. Using an estimate of $\gamma = 10^{-4}$ based on the number of degrees of freedom in a typical enzyme, as well as $k_{\text{cat}} = 10^5\ \text{s}^{-1}$ and $Q = 40k_{\text{B}}T$ for a fast exothermic enzyme like catalase, we find $\Delta D/D_0 \sim 10^{-9}$.

(iii) stochastic swimming due to cyclic stochastic conformational changes associated with the catalytic activity of the enzyme (Bai and Wolynes, 2015; Golestanian and Ajdari, 2008; Najafi and Golestanian, 2010). We can use a simple bead-spring model to estimate this effect. Let us consider two spherical beads of radius R that are attached to each by a linker, that can undergo stochastic elongations with amplitude b. The amplitude of these deformations is typically much smaller than the size of the enzyme, e.g. when they arise from mechanochemical coupling of electrostatic nature (Golestanian, 2010) (analogous to phosphorylation) or structural changes due to ligand binding (Sakaue *et al.*, 2010). However, it is possible that the local heat release could disturb the relatively more fragile tertiary structure of the folded protein for a short while, leading to large amplitudes; thereby suggesting $b \lesssim R$.

To calculate the contribution of such conformational changes to effective diffusion coefficient, we use a simple model in which the conformational change is described by one degree of freedom $L(t)$ representing elongation of the structure along an axis defined by a unit vector $\boldsymbol{n}(t)$. To achieve directed swimming, we need at least two degrees of freedom to incorporate the coherence needed for breaking the time-reversal symmetry at a stochastic level (Najafi and Golestanian, 2004), and we know that realistic conformational changes must involve many degrees of freedom. The randomization of the orientation, described via $\langle \boldsymbol{n}(t) \cdot \boldsymbol{n}(t')\rangle = e^{-2D_r|t-t'|}$, will turn the directed motion into enhanced diffusion over the time scales longer than $1/D_r$. Since the same can be achieved through reciprocal conformational changes described by one compact degree of freedom, we will adopt this simpler form. The stochastic motion of the enzyme can be described by the Langevin equation

$$\boldsymbol{v}(t) \simeq g\left(\frac{\mathrm{d}}{\mathrm{d}t}L(t)\right)\boldsymbol{n}(t) + \boldsymbol{\xi}(t), \tag{XI.5}$$

where α is a numerical pre-factor that depends on the geometry of the enzyme and $\boldsymbol{\xi}(t)$ is the Gaussian white noise that will give us the intrinsic translational diffusion coefficient D_0.

We can describe the combined mechanochemical cycle using a two-step process, which takes the enzyme from its free state through the reaction that is accompanied by the deformation with rate k, and a relaxation back to it native state with rate k_r. This is a simplification of a more realistic model with three states (free, substrate-bound, and reacted-deformed) and k is to be understood as the combined catalytic rate that has the Michaelis-Menten form as defined above. Therefore, we can describe $L(t)$ as a telegraph process and calculate the elongation speed auto-correlation function as

$$\left\langle \frac{\mathrm{d}}{\mathrm{d}t}L(t) \cdot \frac{\mathrm{d}}{\mathrm{d}t'}L(t')\right\rangle = 2b^2\left(\frac{kk_r}{k+k_r}\right)\left[\delta(t-t') - \frac{1}{2}(k+k_r)e^{-(k+k_r)|t-t'|}\right]. \tag{XI.6}$$

By combining this with the orientation auto-correlation, we can calculate the effective diffusion coefficient of the enzyme, which gives the following correction

$$D_{\text{eff}} = D_0 + \frac{1}{3}g^2b^2\left(\frac{kk_r}{k+k_r}\right)\frac{2D_r}{2D_r+k+k_r}, \tag{XI.7}$$

Even for the fastest enzymes, we typically have $k_r \approx D_r \gg k$. Using an upper bound of $b \sim R$, we can approximate Eq. (XI.7) as $\Delta D \approx kR^2$. For catalase, we obtain $\Delta D \approx 1\,\mu\text{m}^2\text{s}^{-1}$, which gives an upper

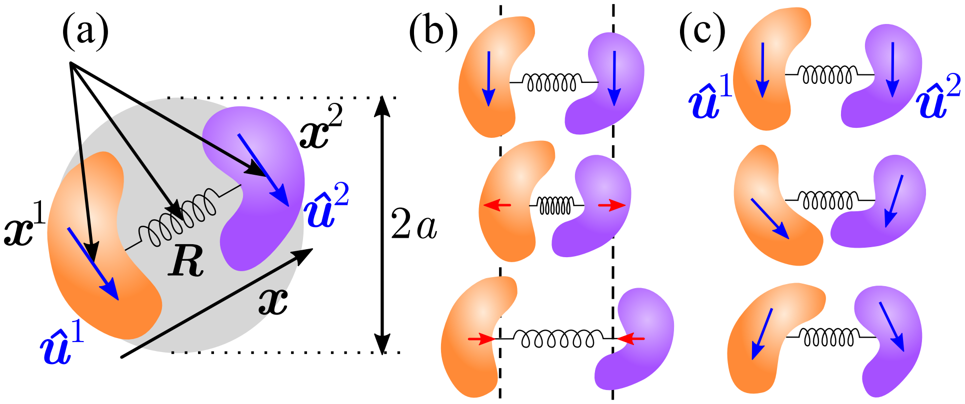


FIG. XI.2 (a) The asymmetric dumbbell model: a is the typical size of the protein, which is made of two subunits with orientations $\boldsymbol{u}^1$ and $\boldsymbol{u}^2$, and located at positions $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$. $\boldsymbol{R}$ denotes the center of mass of the protein and $\boldsymbol{x}$ its elongation. (b) The dumbbell fluctuates around the equilibrium position of the interaction potential. The red arrows represent the forces experienced by the subunits when the dumbbell is contracted or extended. (c) The orientations of the subunits fluctuate around an equilibrium configuration.

bound of $\Delta D/D_0 \approx 10^{-2}$. This is one order of magnitude smaller than the observed values.

(iv) collective heating caused by treating enzymes as mobile sources of heat for exothermic reactions (Golestanian, 2015). This effect will cause the global temperature of the system to raise, leading to an enhancement of the fluctuations as well as simultaneously decreasing the viscosity of the solution. The combined effect can lead to $\Delta D/D_0 \sim 10^{-2} - 10^{-1}$ depending on the size of the container. The nonlinearity in the heat conduction equation arising from the dependence of the heat source on temperature can lead to the possibility of the formation of shock waves and fronts, which could contribute to observation of enhanced diffusion after time-averaging.

(v) modified equilibrium due to the changes in the hydrodynamic couplings between the different modules of an enzyme while undergoing thermal fluctuations, caused by binding and unbinding of chemicals (Adeleke-Larodo *et al.*, 2019b; Illien *et al.*, 2017a,b). This effect originates from the observation that the centre of mass diffusion of compound asymmetric objects depends on their configuration because their internal degrees of freedom are coupled to the centre of mass translation. This phenomenon can be studied using a simple asymmetric dumbbell model, which represent the modular structure of a generic enzyme; see Fig. XI.2(a).

We begin our analysis with the Smoluchowski equation for a pair of Brownian particles interacting with the potential U. The probability $P(\boldsymbol{x}^1, \boldsymbol{x}^2, \boldsymbol{u}^1, \boldsymbol{u}^2; t)$ of finding subunit α at position $\boldsymbol{x}^\alpha$ and with orientation $\boldsymbol{u}^\alpha$ at time t has the following evolution equation

$$\begin{aligned}\partial_t P \;=\; & \sum_{\alpha,\beta=1,2} \Big\{ \nabla_\alpha \cdot \mathbf{M}^{\alpha\beta}_{\mathrm{TT}} \cdot [(\nabla_\beta U)P + k_{\mathrm{B}}T\nabla_\beta P] + \nabla_\alpha \cdot \mathbf{M}^{\alpha\beta}_{\mathrm{TR}} \cdot \Big[(\boldsymbol{\mathfrak{R}}^\beta U)P + k_{\mathrm{B}}T\boldsymbol{\mathfrak{R}}^\beta P\Big] \\ & + \boldsymbol{\mathfrak{R}}^\alpha \cdot \mathbf{M}^{\alpha\beta}_{\mathrm{RT}} \cdot [(\nabla_\beta U)P + k_{\mathrm{B}}T\nabla_\beta P] + \boldsymbol{\mathfrak{R}}^\alpha \cdot \mathbf{M}^{\alpha\beta}_{\mathrm{RR}} \cdot \Big[(\boldsymbol{\mathfrak{R}}^\beta U)P + k_{\mathrm{B}}T\boldsymbol{\mathfrak{R}}^\beta P\Big] \Big\}, \end{aligned} \qquad \text{(XI.8)}$$

where $\mathbf{M}^{\alpha\beta}_{\mathrm{AB}}$ are elements of a mobility matrix which couples the interactions between the translational (T) and rotational (R) modes of the subunits. We now use the centre of mass and elongation coordinates, $\boldsymbol{R} = (\boldsymbol{x}^1 + \boldsymbol{x}^2)/2$ and $\boldsymbol{x} = \boldsymbol{x}^2 - \boldsymbol{x}^1$, to find the Smoluchowski equation for P

$$\begin{aligned}\partial_t P \;=\; & \frac{k_{\mathrm{B}}T}{4}\nabla_{\boldsymbol{R}} \cdot \mathbf{M} \cdot \nabla_{\boldsymbol{R}} P + \frac{1}{2}\nabla_{\boldsymbol{R}} \cdot \boldsymbol{\Gamma} \cdot (\nabla_{\boldsymbol{x}} U)P + \frac{k_{\mathrm{B}}T}{2}(\nabla_{\boldsymbol{R}} \cdot \boldsymbol{\Gamma} \cdot \nabla_{\boldsymbol{x}} P + \nabla_{\boldsymbol{x}} \cdot \boldsymbol{\Gamma} \cdot \nabla_{\boldsymbol{R}} P) \\ & + \nabla_{\boldsymbol{x}} \cdot \mathbf{W} \cdot [k_{\mathrm{B}}T\nabla_{\boldsymbol{x}} P + (\nabla_{\boldsymbol{x}} U)P] + \sum_{\alpha,\beta=1,2} \boldsymbol{\mathfrak{R}}^\alpha \cdot \mathbf{M}^{\alpha\beta}_{\mathrm{RR}} \cdot [k_{\mathrm{B}}T\boldsymbol{\mathfrak{R}}^\beta P + (\boldsymbol{\mathfrak{R}}^\beta U)P] \\ & + \sum_{\alpha=1,2} \Big\{ \nabla_\alpha \cdot \mathbf{M}^{\alpha\beta}_{\mathrm{TR}} \cdot [(\boldsymbol{\mathfrak{R}}^\beta U)P + k_{\mathrm{B}}T\boldsymbol{\mathfrak{R}}^\beta P] + \boldsymbol{\mathfrak{R}}^\alpha \cdot \mathbf{M}^{\alpha\beta}_{\mathrm{RT}} \cdot [(\nabla_\beta U)P + k_{\mathrm{B}}T\nabla_\beta P] \Big\}, \end{aligned} \qquad \text{(XI.9)}$$

where the new mobility coefficients are mixtures of the previous ones (Adeleke-Larodo *et al.*, 2019b; Illien *et al.*, 2017a). Using a generic form for the interaction potential $U = V_0(x) + V_{01}(x)\boldsymbol{n}\cdot\boldsymbol{u}^1 + V_{02}(x)\boldsymbol{n}\cdot\boldsymbol{u}^2 + V_{12}(x)\boldsymbol{u}^1\cdot\boldsymbol{u}^2 + \cdots$, we can average Eq. (XI.9) over x and obtain a reduced equation for the orientational degrees of freedom, involving the definitions $\mathcal{P} = \int \mathrm{d}x\, x^2 P$ and $\langle\cdot\rangle = \frac{1}{\mathcal{P}}\int \mathrm{d}x\, x^2 \cdot P$. The averaging procedure is motivated by the observation of a clear separation of the time-scales in the problem. The dumbbell possesses three time-scales, each describing the relaxation time of one of its degrees of freedom. The slowest of the three is the relaxation time of the centre-of-mass coordinate $\boldsymbol{R}$. Given a potential that can be Taylor expanded around the minimum, the quadratic term gives the effective spring constant of the potential (k) and the time-scale for x to return to its equilibrium value, $\tau_s = \xi/k$, where ξ is the friction coefficient of the enzyme. The rotational diffusion time of the enzyme τ_r, which determines the rate of the loss of memory of the orientation, is of the order $\xi a^2/k_{\mathrm{B}}T$. The ratio of the two times τ_s/τ_r goes as $k_{\mathrm{B}}T/ka^2 \sim \delta x/a$, which is a the relative deformation of the enzyme due to thermal fluctuations and is therefore bounded by unity. With this consideration, we can average over the separation of the subunits assuming $\boldsymbol{n}$, $\boldsymbol{u}^1$ and $\boldsymbol{u}^2$ to be fixed.

We define the lowest order moments of the average distribution with respect to the three unit vectors $\boldsymbol{n}$, $\boldsymbol{u}^1$ and $\boldsymbol{u}^2$ as $\rho \equiv \int_{\boldsymbol{n},\boldsymbol{u}^1,\boldsymbol{u}^2} \mathcal{P}$, $\boldsymbol{p} \equiv \int_{\boldsymbol{n},\boldsymbol{u}^1,\boldsymbol{u}^2} \boldsymbol{n}\mathcal{P}$ and $\boldsymbol{p}^\alpha \equiv \int_{\boldsymbol{n},\boldsymbol{u}^1,\boldsymbol{u}^2} \boldsymbol{u}^\alpha\mathcal{P}$ and obtain the respective evolution equations in the moment expansion of the resulting equation (Golestanian, 2012). To the lowest order, this yields

$$\partial_t\rho = \frac{k_{\mathrm{B}}T}{4}\langle m_0\rangle \nabla^2_{\boldsymbol{R}}\rho + k_{\mathrm{B}}T\left\langle\frac{\gamma_0}{x}\right\rangle \nabla_{\boldsymbol{R}}\cdot\boldsymbol{p} + \frac{1}{3}\left[\left\langle\frac{\gamma_0 V_{01}}{x}\right\rangle\nabla_{\boldsymbol{R}}\cdot\boldsymbol{p}^1 + \left\langle\frac{\gamma_0 V_{02}}{x}\right\rangle\nabla_{\boldsymbol{R}}\cdot\boldsymbol{p}^2\right], \tag{XI.10}$$

$$\partial_t p_i = -\frac{k_{\mathrm{B}}T}{3}\left\langle\frac{\gamma_0}{x}\right\rangle \partial_{R_i}\rho - 2k_{\mathrm{B}}T\left\langle\frac{w_0}{x^2}\right\rangle p_i - \frac{2}{3}\sum_{\alpha=1,2}\left\langle\frac{w_0 V_{0\alpha}}{x^2}\right\rangle p_i^\alpha, \tag{XI.11}$$

$$\partial_t p_i^\alpha = -2k_{\mathrm{B}}T\langle\psi_0^{(\alpha)}\rangle p_i^\alpha + \frac{1}{9}\left\langle\frac{\gamma_0 V_{0\alpha}}{x}\right\rangle\partial_{R_i}\rho - \frac{2}{3}[\langle\psi_0^{(\alpha)}V_{12}\rangle p_i^\beta + \langle\psi_0^{(\alpha)}V_{0\alpha}\rangle p_i - \langle M^{\alpha\beta}_{\mathrm{RR}\ 0}V_{12}\rangle p_i^\beta], \tag{XI.12}$$

where the coefficients are the the corresponding amplitudes of the mobility tensors in a harmonic expansion (Adeleke-Larodo *et al.*, 2019b; Illien *et al.*, 2017a). A closed equation for the density ρ can be now obtained by taking the stationary limits of these equations, which yields

$$\partial_t\rho(\boldsymbol{R};t) = D_{\mathrm{eff}}\nabla^2_{\boldsymbol{R}}\rho, \tag{XI.13}$$

where the effective diffusion has the form

$$D_{\mathrm{eff}} = D_{\mathrm{ave}} - \delta D_{\mathrm{fluc}} = \frac{k_{\mathrm{B}}T}{4}\langle m_0\rangle - \frac{k_{\mathrm{B}}T}{6}\frac{\langle\gamma_0/x\rangle^2}{\langle w_0/x^2\rangle}\left[1 + \text{corrections}\right]. \tag{XI.14}$$

This result highlights the fact that due to the coupling between the internal degrees of freedom and the centre of mass translation, the effective diffusion coefficient is composed of an average contribution that relates to the average mobility of the parts that make the enzyme, and a fluctuation–induced correction, which is universally negative. To obtain a physical intuition for the result, consider the force dipole created by the dumbbell on the fluid; see Fig. XI.2(b). If this fluctuating dipole has a nonvanishing average, the asymmetry of the dumbbell couples the dipole to a net drift velocity. Considering, for instance, the force experienced by the second subunit over the timescale when the radial coordinate has equilibrated but the orientation of the dumbbell is frozen, $\langle\boldsymbol{F}_2\rangle = -\langle U'\rangle\,\boldsymbol{n}$, we notice that a nonzero average dipole will result provided $\langle U'\rangle \neq 0$. Assuming that the average at the relevant timescale is taken with the weight $\mathrm{e}^{-U/k_{\mathrm{B}}T}$, it is easy to see that $\langle U'\rangle$ does not vanish in any dimension other than one (even though $U'|_{x=x_{\mathrm{eq}}} = 0$) due to the entropic contributions to sampling of the configuration space when using radial coordinates. The resulting contribution to the diffusion coefficient is proportional to $-\langle U'\rangle^2$, which highlights a similarity to dispersion forces (London, 1937). The fluctuations can also explore the orientational degrees of freedom of the sub-units, as shown on Fig. XI.2(c), leading to the

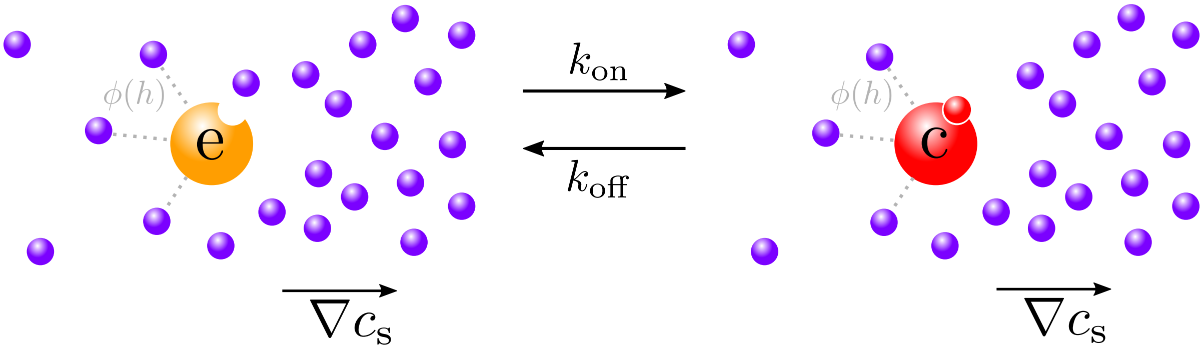


FIG. XI.3 Microscopic model for chemotaxis. The free enzyme (yellow) is in a gradient of substrate molecules (purple), with concentration $c_{\text{s}}(\boldsymbol{R})$. The enzyme can bind one-on-one to a substrate molecule to form a complex (red), with binding rate k_{on} and unbinding rate k_{off}, and also interacts with all other substrate molecules around it through the non-specific potential $\phi(h)$.

corrections highlighted in Eq. (XI.14).

The above result is obtained in the longest time limit, as set out by the moment expansion technique. Using a path integral description, it is possible to show that the diffusion coefficient of an enzyme at short time scales is given by the average contribution only, and that the crossover to the long time behaviour—that includes the fluctuation–induced reduction of the diffusion coefficient—happens at the expected time scales (Illien *et al.*, 2017a).

The above formulation lends itself naturally to a new mechanism contributing to enhanced diffusion of catalytically active enzymes. Since the diffusion coefficient has a component that depends on fluctuations, any process that modifies the fluctuations will change the diffusion coefficient. In particular, substrate binding can generically supress elongation and orientation fluctuations of an enzyme, and thereby reduce the negative fluctuation–induced correction in Eq. (XI.14), leading to an enhanced diffusion as given by Eq. (XI.3). The Michaelis-Menten form of the substrate concentration dependence originates from the probability of substrate binding in stationary state, and not from the catalytic rate of reaction, which happens to have the same form [see Eq. (XI.2)], because it is also proportional to the probability of substrate binding. Note that while this calculation was performed only using the equilibrium components of the forces between the sub-units, the nonequilibrium forces that have been dealt with for the stochastic swimming scenario can also be taken into consideration within this more elaborate formalism.

The existence of multiple mechanisms highlights that it is perhaps misleading to think about enhanced diffusion of catalytically active enzymes as one phenomenon with universal features.

B. Chemotaxis of Enzymes

In recent years, it has been observed experimentally that the non-equilibrium chemical activity of enzymes leads to intrinsically non-equilibrium behaviour and interactions already at the microscopic scale, and in particular chemotaxis towards (Dey *et al.*, 2014; Guha *et al.*, 2017) and away (Jee *et al.*, 2018) from the substrate source. Similarly to the case of enhanced diffusion, it can be argued that several mechanisms contribute to enzyme chemotaxis and that the different tendencies observed in the experiments can be explained as a competition between these contributions (Agudo-Canalejo *et al.*, 2018a,b). In particular, phoresis tends to direct enzymes towards their substrates, due to the predominantly attractive interaction between them, whereas enhanced diffusion tends to have the opposite effect as the enzymes will evacuate the region with higher substrate concentration more quickly due to enhanced diffusive activity.

We can use a minimal model that takes into account the following ingredients: specific binding between the substrate and the enzyme, nonspecific interactions such as electrostatic, steric and van den Waals interactions, as well as hydrodynamics interactions; see Fig. XI.3. The setup the model, we start from the full N-particle Fokker-Planck equation of the system of enzyme and substrate molecules and integrate out the substrate degrees of freedom to arrive at the following effective one-particle description for the free enzyme and complex probability distributions

$$\begin{aligned}\partial_t C_{\rm e}(\boldsymbol{R};t) &= \nabla_{\boldsymbol{R}}\cdot\left[D_{\rm e}\nabla_{\boldsymbol{R}}C_{\rm e}-\boldsymbol{v}_{\mathbf{e}}(\boldsymbol{R})C_{\rm e}\right]-k_{\rm on}C_{\rm e}C_{\rm s}+k_{\rm off}C_{\rm c} &\text{(XI.15)}\\ \partial_t C_{\rm c}(\boldsymbol{R};t) &= \nabla_{\boldsymbol{R}}\cdot\left[D_{\rm c}\nabla_{\boldsymbol{R}}C_{\rm c}-\boldsymbol{v}_{\mathbf{c}}(\boldsymbol{R})C_{\rm c}\right]+k_{\rm on}C_{\rm e}C_{\rm s}-k_{\rm off}C_{\rm c} &\text{(XI.16)}\end{aligned}$$

where $C_{\rm s}(\boldsymbol{R})$ is the concentration of substrate molecules. Equations (XI.15–XI.16) capture the effects of three important physical mechanisms. First, the free enzyme can turn into a complex and vice versa through the *binding and unbinding* of a substrate molecule, with rates $k_{\rm on}$ and $k_{\rm off}$. Second, the free enzyme and complex *diffuse* with diffusion coefficients respectively given by $D_{\rm e} = k_{\rm B}T\boldsymbol{\mu}^{\rm ee}$ and $D_{\rm c} = k_{\rm B}T\boldsymbol{\mu}^{\rm cc}$. Finally, the combination of non-specific and hydrodynamic interactions between the free enzyme or complex and the substrate molecules leads to a *diffusiophoretic drift* of the free enzyme and complex with velocities respectively given by

$$\begin{aligned}\boldsymbol{v}_{\mathbf{e}}(\boldsymbol{R}) &\approx -\frac{k_{\rm B}T}{\eta}\left[\int_0^\infty {\rm d}h h\left(1-{\rm e}^{-\phi^{\rm es}(h)/k_{\rm B}T}\right)\right]\nabla_{\boldsymbol{R}}C_{\rm s}\equiv-\frac{k_{\rm B}T}{\eta}\lambda_{\rm e}^2\nabla_{\boldsymbol{R}}C_{\rm s} &\text{(XI.17)}\\ \boldsymbol{v}_{\mathbf{c}}(\boldsymbol{R}) &\approx -\frac{k_{\rm B}T}{\eta}\left[\int_0^\infty {\rm d}h h\left(1-{\rm e}^{-\phi^{\rm cs}(h)/k_{\rm B}T}\right)\right]\nabla_{\boldsymbol{R}}C_{\rm s}\equiv-\frac{k_{\rm B}T}{\eta}\lambda_{\rm c}^2\nabla_{\boldsymbol{R}}C_{\rm s} &\text{(XI.18)}\end{aligned}$$

where η is the viscosity of the fluid. Equations (XI.17–XI.18) are approximate forms of the diffusiophoretic velocity valid for the typical case in which the range of the non-specific interactions is shorter than the size of the enzyme or the complex. We have defined here the Derjaguin length λ_α as before. Remember that, in our convention, λ_α^2 may be positive or negative, with positive (negative) values corresponding to overall repulsive (attractive) interactions that lead to a depletion (an accumulation) of substrate molecules in the proximity of the enzyme. Because enzyme–substrate interactions are generally attractive, we expect $\lambda_\alpha^2 < 0$, implying that for typical enzymes the phoretic velocity is directed towards higher concentrations of the substrate.

Equations (XI.15–XI.16) already contain all the ingredients necessary to describe enzyme–substrate interactions. Nevertheless, in order to describe chemotaxis we are actually interested in the *total* concentration of enzyme, both free and bound, given by

$$C_{\rm e}^{\rm tot}(\boldsymbol{R};t) = C_{\rm e}(\boldsymbol{R};t)+C_{\rm c}(\boldsymbol{R};t) \tag{XI.19}$$

which corresponds to what is measured in the experiments with fluorescently tagged enzymes (free enzyme and complex cannot be distinguished). Furthermore, the typical timescale of diffusion and phoretic drift is much longer than the typical timescale of binding and unbinding. We can therefore assume that the enzyme is locally and instantaneously at binding equilibrium with the substrate, so that at any position $\boldsymbol{R}$ we will have

$$k_{\rm on}C_{\rm e}(\boldsymbol{R};t)C_{\rm s}(\boldsymbol{R};t)\approx k_{\rm off}C_{\rm c}(\boldsymbol{R};t) \tag{XI.20}$$

at time t. Combining Eq. (XI.19) and Eq. (XI.20), we find the typical Michaelis-Menten kinetics for the free enzyme and the complex

$$C_{\rm e} = \frac{K}{K+C_{\rm s}}C_{\rm e}^{\rm tot} \quad\text{and}\quad C_{\rm c} = \frac{C_{\rm s}}{K+C_{\rm s}}C_{\rm e}^{\rm tot} \tag{XI.21}$$

where we have defined the Michaelis constant $K \equiv k_{\rm off}/k_{\rm on}$.

Adding together Equations (XI.15) and (XI.16), and using (XI.21), we finally obtain an expression for the time evolution of the total enzyme concentration

$$\partial_t C_{\mathrm{e}}^{\mathrm{tot}}(\boldsymbol{R};t) = \nabla_{\boldsymbol{R}} \cdot \left\{ D(\boldsymbol{R}) \cdot \nabla_{\boldsymbol{R}} C_{\mathrm{e}}^{\mathrm{tot}} - [\boldsymbol{V}_{\mathrm{ph}}(\boldsymbol{R}) + \boldsymbol{V}_{\mathrm{bi}}(\boldsymbol{R})] C_{\mathrm{e}}^{\mathrm{tot}} \right\} \tag{XI.22}$$

with the space-dependent diffusion coefficient

$$D(\boldsymbol{R}) \;=\; D_{\mathrm{e}} + (D_{\mathrm{c}} - D_{\mathrm{e}}) \frac{C_{\mathrm{s}}(\boldsymbol{R})}{K + C_{\mathrm{s}}(\boldsymbol{R})}, \tag{XI.23}$$

a drift velocity arising from phoretic effects

$$\boldsymbol{V}_{\mathrm{ph}}(\boldsymbol{R}) \;=\; \boldsymbol{v}_{\mathbf{e}}(\boldsymbol{R}) + [\boldsymbol{v}_{\mathbf{c}}(\boldsymbol{R}) - \boldsymbol{v}_{\mathbf{e}}(\boldsymbol{R})] \frac{C_{\mathrm{s}}(\boldsymbol{R})}{K + C_{\mathrm{s}}(\boldsymbol{R})}, \tag{XI.24}$$

as well as a drift velocity arising from the changes in diffusion coefficient due to substrate binding and unbinding

$$\boldsymbol{V}_{\mathrm{bi}}(\boldsymbol{R}) \;=\; -(D_{\mathrm{c}} - D_{\mathrm{e}}) \nabla_{\boldsymbol{R}} \left(\frac{C_{\mathrm{s}}(\boldsymbol{R})}{K + C_{\mathrm{s}}(\boldsymbol{R})} \right) = -\nabla_{\boldsymbol{R}} D(\boldsymbol{R}). \tag{XI.25}$$

The drift velocity arising from diffusiophoresis $\boldsymbol{V}_{\mathrm{ph}}(\boldsymbol{R})$ corresponds to an 'average' of the phoretic velocities of the free enzyme and the complex. With increasing substrate concentration, a smooth Michaelis-Menten-like crossover is observed between the velocity of the free enzyme $\boldsymbol{v}_{\mathbf{e}}$ and that of the complex $\boldsymbol{v}_{\mathbf{c}}$. In principle, this velocity may be directed towards or away from higher concentrations of substrate, depending on the details of the non-specific interactions. Because enzyme–substrate interactions are generally attractive, however, we expect that the typical phoretic velocity for enzymes will be directed towards higher substrate concentrations. The drift velocity $\boldsymbol{V}_{\mathrm{bi}}(\boldsymbol{R})$ is a direct consequence of the changes in the diffusion coefficient of the enzyme due to binding and unbinding of the substrate. This drift velocity is directed towards higher concentrations of substrate in the case of inhibited diffusion with $D_{\mathrm{c}} < D_{\mathrm{e}}$, and towards lower concentrations of substrate in the case of enhanced diffusion with $D_{\mathrm{c}} > D_{\mathrm{e}}$. In the absence of phoresis with $\boldsymbol{V}_{\mathrm{ph}}(\boldsymbol{R}) = 0$, Equation (XI.22) can then be written as $\partial_t C_{\mathrm{e}}^{\mathrm{tot}}(\boldsymbol{R};t) = \nabla_{\boldsymbol{R}}^2 [D(\boldsymbol{R}) C_{\mathrm{e}}^{\mathrm{tot}}]$, and we would thus expect $C_{\mathrm{e}}^{\mathrm{tot}}(\boldsymbol{R}) \propto 1/D(\boldsymbol{R})$ in the steady state, i.e. the enzyme tends to concentrate in regions where its diffusion is slowest. This type of behaviour was recently reported experimentally in Ref. (Jee *et al.*, 2018) for urease, in apparent conflict with older results in the literature, Ref. (Sengupta *et al.*, 2013), in which urease was observed to chemotax towards higher concentrations of urea. The existence of two distinct mechanisms for chemotaxis, namely phoresis and binding-induced changes in diffusion as just described, may explain the seemingly contradictory observations. It is interesting to note that chemotactic alignment can also contribute to the process of net chemotactic drift as well as enhanced diffusion on time scales longer than the rotational diffusion time (Adeleke-Larodo *et al.*, 2019a).

XII. PHORESIS ON THE SLOW LANE: TRAIL-FOLLOWING BACTERIA

When bacteria live in the bulk, they use swimming motility via flagella, and search strategies like run-and-tumble to provide for themselves, as single entities. The planktonic life of bacteria is lonely, because random walks in 3D are terse. When bacteria find a surface and interact with it mechanically, at some point they "decide" to settle on it. The decision making process is stochastic, and involves a cross-correlation between cAMP signal and type-IV pili (TFP) mechanical activity, which lasts through several generations of cell division (Lee *et al.*, 2018). The settled bacteria lose their flagella and grow pili, and start exploring the 2D surface using surface attachment motility: pili elongate and retract through polymerization and depolymerization, and they stochastically attach to the surface and detach from it. On surfaces, bacteria live communal lives, because random walks in 2D are compact, which implies that bacteria meet very often even if they do not signal each other.

In practice, however, bacteria do signal each other and use it to self-organize much more efficiently than random walks can provide for them. They achieve this by leaving behind a metabolically expensive polysaccharide trail. While this might appear to be an inefficient use of individual resources, the communal lifestyle that follows from it benefits them immensely. This process belongs to the general class of motility with chemical sensing, in the extreme limit where the chemicals do not diffuse. Here we describe the detailed derivation of a set of stochastic equations of motion for bacteria that phoretically interact via non-diffusing trails they leave behind themselves. The theoretical formulation is based on a microscopic motility model that takes into account stochastic attachment and detachment of pili, as well as friction coefficients and pili contraction forces that depend on the amount of polysaccharide (Gelimson *et al.*, 2016; Kranz *et al.*, 2016).

A. Systematic Derivation of Stochastic Dynamical Equations

We consider a bacterium of length ℓ (Fig. XII.1). We will assume that each of the N pili located at one tip (or at both tips) of a bacterium will attach to, and detach from, the surface and pull in a polysaccharide-dependent way. The direction unit vectors of the pili $\boldsymbol{e}_i,\ i = 1, \dots, N$ are assumed to originate from the bacterial tip $\boldsymbol{r}_0$. One could consider variants of this model where the pili distribution is different. A single pilus tip, located at $\boldsymbol{r}_0 + \ell_p \hat{\boldsymbol{e}}_i$, will randomly attach to the surface with a polysaccharide-dependent rate $\lambda(\psi_i)$, detach with a rate $\mu(\psi_i)$ and pull in its direction with a force $f(\psi_i)$ if it is attached to the surface. Here, ψ_i is the polysaccharide concentration at the tip of pilus i, $\psi_i = \psi(\boldsymbol{r}_0 + \ell_p \hat{\boldsymbol{e}}_i)$. For each pilus i we define $\Theta_i = 1$ if the pilus is attached and $\Theta_i = 0$ if it is detached. Θ_i will have a mean

$$\bar{\Theta} = \langle \Theta_i \rangle = \frac{\lambda(\psi)}{\lambda(\psi) + \mu(\psi)}, \tag{XII.1}$$

and a variance

$$\sigma^2 = \langle (\Theta_i - \bar{\Theta})^2 \rangle = \frac{\lambda(\psi)\mu(\psi)}{[\lambda(\psi) + \mu(\psi)]^2}. \tag{XII.2}$$

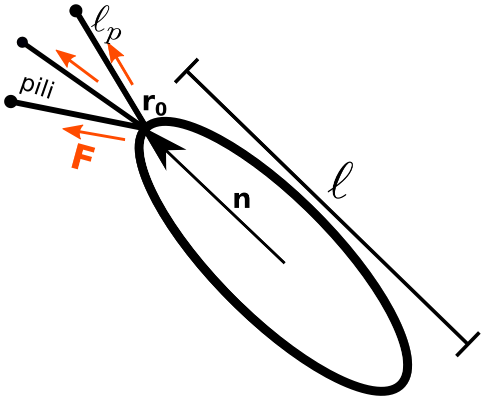


FIG. XII.1 Microscopic model: the bacterium has type-IV pili of length ℓ_p, which can attach to a surface and pull the bacterium forward with a generically polysaccharide-dependent force. Originating at a bacterial pole $\boldsymbol{r}_0$, the pili point in different directions $\hat{\boldsymbol{e}}_i,\ i = 1, \dots, N$, such that the pili tips are located at the positions $\boldsymbol{r}_0 + \ell_p \hat{\boldsymbol{e}}_i$. A pilus pulling force $f(\psi_i)\hat{\boldsymbol{e}}_i$ on the bacterial tip $\boldsymbol{r}_0$ is only generated if the pilus is attached to the surface and in general the attachment and detachment rates $\lambda(\psi_i)$ and $\mu(\psi_i)$ will be dependent on the polysaccharide concentration $\psi_i = \psi(\boldsymbol{r}_0 + \ell_p \hat{\boldsymbol{e}}_i)$ at the end of pilus i

The total force exerted at the bacterial tip will then be

$$\boldsymbol{F} = \sum_i \hat{\boldsymbol{e}}_i f(\psi_i)\Theta_i. \tag{XII.3}$$

Using the form $\Theta_i = \bar{\Theta} + \delta\Theta_i$, and the Taylor expansion $\psi_i = \psi(\boldsymbol{r}_0 + \ell_p \hat{\boldsymbol{e}}_i) \approx \psi(\boldsymbol{r}_0) + \ell_p(\boldsymbol{\nabla}\, \psi|_{\boldsymbol{r}_0} \cdot \hat{\boldsymbol{e}}_i)$, we obtain

$$\boldsymbol{F} = \sum_i \hat{\boldsymbol{e}}_i f(\psi)\bar{\Theta}(\psi) + \sum_i \hat{\boldsymbol{e}}_i \ell_p (\boldsymbol{\nabla}\psi \cdot \hat{\boldsymbol{e}}_i)\partial_\psi (f\bar{\Theta}) + \sum_i \hat{\boldsymbol{e}}_i f(\psi_i)\delta\Theta_i. \tag{XII.4}$$

We can express the pili orientation vectors in terms of the bacterial body orientation $\hat{\boldsymbol{n}}$ and the orientation $\hat{\boldsymbol{n}}_\perp = \hat{\boldsymbol{e}}_z \times \hat{\boldsymbol{n}}$ orthogonal to it as $\hat{\boldsymbol{e}}_i = \cos\vartheta_i \hat{\boldsymbol{n}} + \sin\vartheta_i \boldsymbol{n}_\perp$. It is reasonable to assume that the pseudomonas bacteria mainly contributing to surface-mediated chemotaxis will have a pili distribution that is approximately symmetrical with respect to the body orientation (if this was not the case the bacterium would generate a torque in one preferred direction, rotate around itself and effectively stay in one point). Therefore, we can neglect terms in the sum over the pili that are odd in $\sin\vartheta_i$, like $\langle \sin\vartheta_i \rangle = \frac{1}{N}\sum_i \sin\vartheta_i$ or $\langle \cos\vartheta_i \sin\vartheta_i \rangle = \frac{1}{N}\sum_i \cos\vartheta_i \sin\vartheta_i$. The force can then be written as

$$\begin{aligned}\boldsymbol{F} \;&=\; N\langle\cos\vartheta_i\rangle f(\psi)\bar{\Theta}(\psi)\hat{\boldsymbol{n}} + N\langle\cos^2\vartheta_i\rangle \ell_p \partial_\psi \left[f(\psi)\bar{\Theta}(\psi)\right](\boldsymbol{\nabla}\psi\cdot\hat{\boldsymbol{n}})\,\hat{\boldsymbol{n}} + \hat{\boldsymbol{n}}\overbrace{\sum_i \cos\vartheta_i f(\psi_i)\delta\Theta_i}^{\delta F_\parallel} \\ &\quad + N\langle\sin^2\vartheta_i\rangle \ell_p \partial_\psi \left[f(\psi)\bar{\Theta}(\psi)\right](\boldsymbol{\nabla}\psi\cdot\hat{\boldsymbol{n}}_\perp)\,\hat{\boldsymbol{n}}_\perp + \hat{\boldsymbol{n}}_\perp \underbrace{\sum_i \sin\vartheta_i f(\psi_i)\delta\Theta_i}_{\delta F_\perp}. \end{aligned} \tag{XII.5}$$

If the pili attachment events of two different pili are independent of each other we have $\langle \delta\Theta_i \delta\Theta_j \rangle = \sigma^2(\psi)\delta_{ij}$. Moreover, we can calculate the auto-correlation of the attachment as $\langle \delta\Theta_i(t)\delta\Theta_i(t')\rangle = \sigma^2\, e^{-(\mu+\lambda)|t-t'|}\,\delta_{ij}$. If we focus on time scales that are longer than the average attachment/detachment time, we can represent this auto-correlation as a delta function

$$\langle \delta\Theta_i(t)\delta\Theta_i(t')\rangle = \frac{\sigma^2(\psi)}{[\lambda(\psi)+\mu(\psi)]}\,\delta_{ij}\delta(t-t') = \frac{\lambda(\psi)\mu(\psi)}{[\lambda(\psi)+\mu(\psi)]^3}\,\delta_{ij}\delta(t-t'). \tag{XII.6}$$

With this, we get the mean square fluctuations of the parallel force as

$$\left\langle \delta F_\parallel^2 \right\rangle = N\langle\cos^2\vartheta_i\rangle f^2(\psi)\sigma^2(\psi) + O\left(\boldsymbol{\nabla}\psi\cdot\ell_p\right). \tag{XII.7}$$

Analogously, for the mean square fluctuations of the perpendicular force we obtain

$$\left\langle \delta F_\perp^2 \right\rangle = N\langle\sin^2\vartheta_i\rangle f^2(\psi)\sigma^2(\psi) + O\left(\boldsymbol{\nabla}\psi\cdot\ell_p\right). \tag{XII.8}$$

The attachment/detachment of pili on the surface can be considered a set of N random events and in case of $N \gg 1$ we can expect that the overall fluctuations in the force can be well approximated by a Gaussian with a variance of $N\langle\cos^2\vartheta_i\rangle f^2(\psi)\sigma^2(\psi)$ in the direction $\hat{\boldsymbol{n}}$ and a variance of $N\langle\sin^2\vartheta_i\rangle f^2(\psi)\sigma^2(\psi)$ in the direction $\hat{\boldsymbol{n}}_\perp$.

The pili-generated force will propel the bacterium and also generate a torque. Therefore, we obtain the following translational equation of motion

$$\frac{d\boldsymbol{r}}{dt} = \frac{1}{\gamma_\parallel}\boldsymbol{F}_\parallel + \frac{1}{\gamma_\perp}\boldsymbol{F}_\perp = v(\psi)\hat{\boldsymbol{n}} + A(\psi)(\boldsymbol{\nabla}\psi\cdot\hat{\boldsymbol{n}}_\perp)\hat{\boldsymbol{n}}_\perp + B(\psi)(\boldsymbol{\nabla}\psi\cdot\hat{\boldsymbol{n}})\hat{\boldsymbol{n}} + \sqrt{2D_\parallel}\,\eta^\parallel\hat{\boldsymbol{n}} + \sqrt{2D_\perp}\,\eta^\perp\hat{\boldsymbol{n}}_\perp, \tag{XII.9}$$

and the following rotational equation of motion

$$\frac{d\hat{\boldsymbol{n}}}{dt} = \boldsymbol{\omega} \times \hat{\boldsymbol{n}} = -\frac{1}{\gamma_{\text{rot}}}\,\hat{\boldsymbol{n}} \times \boldsymbol{\tau} = -\chi(\psi)\hat{\boldsymbol{n}} \times [\hat{\boldsymbol{n}} \times \boldsymbol{\nabla}\psi] + \sqrt{2D_r(\psi)}\,\eta^{\perp}\hat{\boldsymbol{n}}_{\perp}. \tag{XII.10}$$

The parameters are known functions of the rates, but there are only two independent parameters

$$v(\psi) = N\,c_1\,\frac{f(\psi)}{\gamma_{\|}}\,\frac{\lambda(\psi)}{\lambda(\psi)+\mu(\psi)}, \tag{XII.11}$$

$$D_r = \frac{1}{8}N\ell^2(1-c_2)\,\frac{f^2(\psi)}{\gamma_{\text{rot}}^2}\,\frac{\lambda(\psi)\mu(\psi)}{[\lambda(\psi)+\mu(\psi)]^3}. \tag{XII.12}$$

$$\tag{XII.13}$$

The rest of the parameters can be expressed in terms of $v(\psi)$ and $D_r(\psi)$ in the following way

$$D_{\|}(\psi) = \left[\frac{4\gamma_{\text{rot}}^2 c_2}{\ell^2\gamma_{\|}^2(1-c_2)}\right] D_r(\psi), \tag{XII.14}$$

$$D_{\perp}(\psi) = \left(\frac{2\gamma_{\text{rot}}}{\gamma_{\perp}\ell}\right)^2 D_r(\psi), \tag{XII.15}$$

$$\chi(\psi) = \left[\frac{\gamma_{\|}\ell\ell_p(1-c_2)}{2\gamma_{\text{rot}}c_1}\right]\partial_\psi v(\psi) + \theta(0)\partial_\psi D_r(\psi), \tag{XII.16}$$

$$A(\psi) = \left[\frac{\ell_p(1-c_2)\gamma_{\|}}{c_1\gamma_{\perp}}\right]\partial_\psi v(\psi) + \theta(0)\left(\frac{2\gamma_{\text{rot}}}{\gamma_{\perp}\ell}\right)^2\partial_\psi D_r(\psi), \tag{XII.17}$$

$$B(\psi) = \left[\frac{\ell_p c_2}{c_1}\right]\partial_\psi v(\psi) + \theta(0)\left[\frac{4\gamma_{\text{rot}}^2 c_2}{\ell^2\gamma_{\|}^2(1-c_2)}\right]\partial_\psi D_r(\psi), \tag{XII.18}$$

where $c_1 = \langle\cos\vartheta_i\rangle$ and $c_2 = \langle\cos^2\vartheta_i\rangle$ are assumed to be polysaccharide independent.

The noise terms $\eta_{\|}$, $\eta_{\perp}$ are Gaussian noise components parallel and orthogonal to $\hat{\boldsymbol{n}}$ with $\langle\eta^{\|/\perp}\rangle = 0$, $\langle\eta^{\|}(t)\eta^{\|}(t')\rangle = \delta(t-t')$, $\langle\eta^{\perp}(t)\eta^{\perp}(t')\rangle = \delta(t-t')$, and $\langle\eta^{\|}(t)\eta^{\perp}(t')\rangle = 0$. Our formulation contains multiplicative noise and therefore needs an additional interpretation rule as discussed in (Van Kampen, 1981): the Itô interpretation corresponds to $\theta(0) = 0$ whereas $\theta(0) = 1/2$ corresponds to Stratonovich.

B. Single-particle Dynamics

The microscopic derivation provides us with a powerful framework that can be used towards a deeper understanding of the behaviour of trail-following bacteria. The following two aspects are particularly interesting for consideration.

1. Dependence of the Parameters on the Chemical Concentration

We are interested to address generic questions about this system, such as the influence of the polysaccharide on the speed of bacteria. From the expression for $v(\psi)$ in Eq. (XII.11), we can see that the answer to this question is not trivial. The expression for speed involves the force $f(\psi)$ that is likely to increase with ψ due to stronger pulling (although it does not necessarily need to change), the friction coefficient $\gamma(\psi)$ that is likely to increase with ψ due to the stickiness of the polysaccharide, the attachment rate $\lambda(\psi)$ that is likely to increase due to preferential attachment, and the detachment rate $\mu(\psi)$ that is likely to decrease to preferential attachment. Given the combination with which these parameters appear in the expression for speed in Eq. (XII.11), it is not straightforward to predict the overall trend, as there are competing contributions.

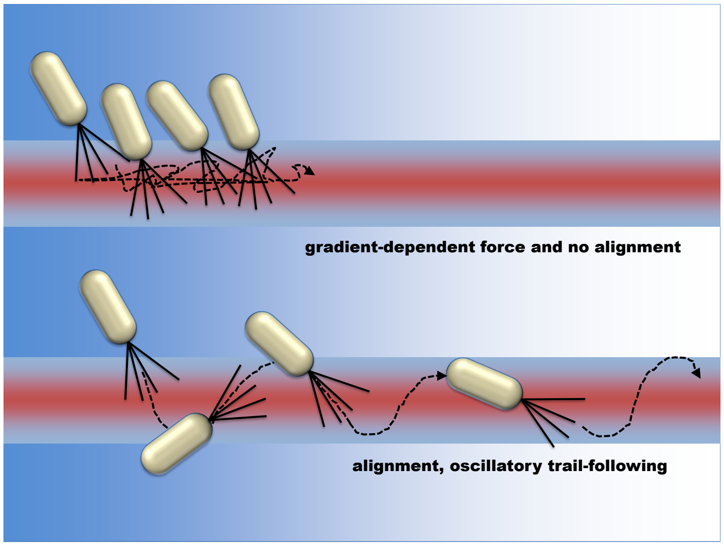


FIG. XII.2 Translational gradient sensing versus orientational gradient sensing: The perpendicular alignment strategy of chemotaxis allows for an efficient trail-following mechanism via oscillatory trajectories.

By fitting single-particle trajectories from the experiment to Eqs. (XII.11)–(XII.18), it was possible to deduce that $v(\psi) \propto \frac{f}{\gamma_{\parallel}} \frac{\lambda}{\lambda+\mu}$ increases significantly with ψ, whereas $D_r(\psi) \propto \frac{f^2}{\gamma_{\text{rot}}^2} \frac{\lambda\mu}{[\lambda+\mu]^3}$ is relatively insensitive to at the lowest order. While this information is not sufficient to help us extract all the functions, it suggests that the observations are consistent with the following pattern: $\lambda \approx \kappa_0 + \kappa_1 \psi$, $\mu \approx \kappa_0 - \kappa_1 \psi$, and $\frac{f}{\gamma} \approx$ constant (Gelimson *et al.*, 2016; Kranz *et al.*, 2016). This provides us with novel insight into the behaviour of trail-following bacteria.

2. Perpendicular Alignment Trail-following Strategy

For trail-following bacteria, it is not clear how a good strategy for following an existing trail can be devised using gradient sensing, in line with common search strategies of microorganisms. The reason is that the main gradient in a trail is setup perpendicularly to the main axis (see Fig. XII.2). Therefore, a conventional or “translational” gradient sensing will only ensure that the bacterium is attracted to the centre of the trail cross section; to move along trail, it will need to rely on random fortuitous alignment of the body that will give it the appropriate velocity along the trail via self-propulsion. Our model reveals an alternative and more efficient strategy, via alignment along the gradient. This “orientational” gradient sensing strategy ensures that the bacterium is always aligned perpendicularly to the trail heading towards its centre. When crossing through the maximum at the centre, the alignment strategy will make the bacterium turn, resulting in an oscillatory or zigzag trajectory along the trail. This provides the microorganism with an efficient trail-following strategy without any need for a complex feedback mechanism between motility and sensing. This remarkable scenario is a natural consequence of the microscopic interactions that are incorporated in the model, even though it might appear to be counter-intuitive at first.

C. Many-particle Dynamics

The above equations need to be solved self-consistently together with the equation that describes trail deposition at rate k by the moving bacteria

$$\partial_t \psi - \cancel{\mathcal{D}_{\mathrm{p}} \nabla^2 \psi} = k \sum_a \delta(\boldsymbol{r} - \boldsymbol{r}_a(t)), \tag{XII.19}$$

where the second term can be ignored as for the polysaccharide trails $\mathcal{D}_{\mathrm{p}} \approx 0$. Solving the above equations for the many-body bacterial system provides a quantitative account of the process of early biofilm nucleus formation that agrees well with experiments (Gelimson *et al.*, 2016). In particular, it unravels a hierarchical self-organization of the multi-cellular system at very low densities, which exhibits a power law visit distribution of space, whose prediction from the theory is quantitatively matched with the experimental observations without any fitting parameters for the many-body dynamics.

D. Chemotactic Localization Transition

An intriguing outcome of the above formulation for trail-following bacteria is the emergence of a localization transition when a microorganism can interact with its own trail (Kranz *et al.*, 2016). Let us consider the trail excreted from a microorganism as characterized by the equation

$$\partial_t \psi(\boldsymbol{r}, t) = k \delta_R^2 \left(\boldsymbol{r} - \boldsymbol{r}(t)\right), \tag{XII.20}$$

where $\delta_R^2 \left(\boldsymbol{r} - bfr(t)\right) \equiv \theta(R^2 - r^2)/\pi R^2$ is a "regularized delta function" that accounts for the finite size R. Integrating this equation, we find for the trail profile at time t and position $\boldsymbol{x}$ as

$$\psi(\boldsymbol{x}, t) = k \int_0^t \mathrm{d}t' \, \delta_R^2 \left(\boldsymbol{x} - \boldsymbol{r}(t')\right) = \frac{k}{\pi R^2} \int_0^t \mathrm{d}t' \, \theta(R^2 - |\boldsymbol{x} - \boldsymbol{r}(t')|^2). \tag{XII.21}$$

The trail width $2R$ defines a microscopic time scale $\tau = R/v_0$, which gives the trail crossing time.

The orientation dynamics written in terms of the angle $\varphi(t)$ defined via $\boldsymbol{n} = (\cos\varphi, \sin\varphi)$ is governed by the following Langevin equation

$$\partial_t \varphi(t) = \chi \partial_\perp \psi(\boldsymbol{r}(t), t) + \xi(t), \tag{XII.22}$$

where $\partial_\perp \psi = \boldsymbol{n}_\perp(t) \cdot \nabla\psi(\boldsymbol{r}(t), t)$, with the lateral unit vector given as $\boldsymbol{n}_\perp = (-\sin\varphi, \cos\varphi)$. Here $\xi(t)$ is a Gaussian random variable obeying $\langle \xi(t)\xi(t')\rangle = 2D_r^0 \delta(t - t')$ and D_r^0 is the microscopic rotational diffusion coefficient controlling the persistence time $1/D_r^0$.

The lateral gradient can be calculated by projecting the gradient of the trail profile onto the lateral unit vector. From the definition of the trail field we have

$$\nabla_{\boldsymbol{x}} \psi(\boldsymbol{x}, t) = -\frac{2k}{\pi R^2} \int_0^t \mathrm{d}t' [\boldsymbol{x} - \boldsymbol{r}(t')] \delta(R^2 - |\boldsymbol{x} - \boldsymbol{r}(t')|^2), \tag{XII.23}$$

and with a change of variable $t' \to t - t'$, we find

$$\partial_\perp \psi(\boldsymbol{r}(t), t) = -\frac{2k}{\pi R^2} \int_0^t \mathrm{d}t' [\boldsymbol{r}(t) - \boldsymbol{r}(t - t')] \cdot \boldsymbol{n}_\perp(t) \delta(R^2 - |\boldsymbol{r}(t) - \boldsymbol{r}(t - t')|^2). \tag{XII.24}$$

The translational dynamical equation described by

$$\frac{\mathrm{d}\boldsymbol{r}(t)}{\mathrm{d}t} = v_0 \boldsymbol{n}(t), \tag{XII.25}$$

can be recursively integrated as

$$\boldsymbol{r}(t - \tau) - \boldsymbol{r}(t) = v_0 \int_t^{t-\tau} \mathrm{d}u \, \boldsymbol{n}(u) = v_0 \int_t^{t-\tau} \mathrm{d}u \left[\boldsymbol{n}(t) + \int_t^u \mathrm{d}w \, \dot{\boldsymbol{n}}(w)\right], \tag{XII.26}$$

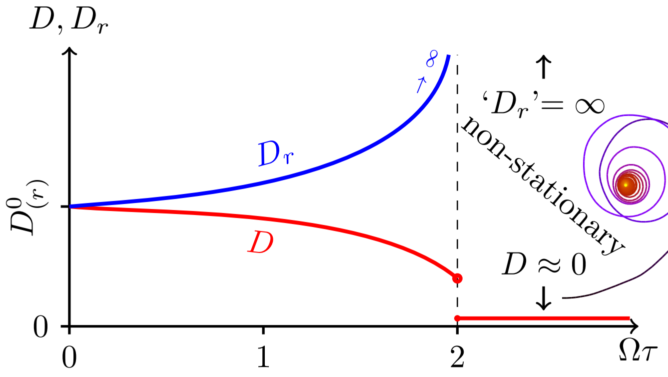


FIG. XII.3 Phase diagram of the dynamics of the microorganism with trail-mediated self-interaction, the effective translational diffusion coefficient D, and the effective rotational diffusion coefficient D_r, as functions of the dimensionless turning frequency $\Omega\tau$ that is a measure of the rate of trail deposition. A location transition is observed at $\Omega\tau = 2$.

one finds $[\boldsymbol{r}(t-\tau) - \boldsymbol{r}(t)]^2 = v_0^2\tau^2 + \mathcal{O}(\tau^3)$ and

$$[\boldsymbol{r}(t-\tau)-\boldsymbol{r}(t)]\cdot\boldsymbol{n}_\perp(t) = -v_0\int_0^\tau \mathrm{d}u\int_0^u \mathrm{d}w\,\{\chi\hat{\boldsymbol{e}}_z\cdot[\boldsymbol{n}(t-w)\times\nabla\psi(t-w)] + \xi(t-w)\}\,\boldsymbol{n}_\perp(t-w)\cdot\boldsymbol{n}_\perp(t). \tag{XII.27}$$

An identical iteration shows $\boldsymbol{n}_\perp(t-w)\cdot\boldsymbol{n}_\perp(t) = 1 + \mathcal{O}(w)$ and thus

$$[\boldsymbol{r}(t-\tau) - \boldsymbol{r}(t)]\cdot\boldsymbol{n}_\perp(t) = -v_0\int_0^\tau \mathrm{d}u\int_0^u \mathrm{d}w\,[\chi\partial_\perp\psi(t-w) + \xi(t-w)] + \mathcal{O}(\tau^3). \tag{XII.28}$$

Using the above two approximations in Eq. (XII.24) and performing one of the integrals yields a stochastic integral equation

$$\partial_\perp\psi(t) = \frac{\Omega}{\tau}\int_0^\tau \mathrm{d}u\,(\tau-u)\,\left[\partial_\perp\psi(t-u) + \xi(t-u)/\chi\right], \tag{XII.29}$$

where $\Omega = k\chi\tau/(\pi R^3)$ is an effective turning rate, which increases for more intense trails (larger k) and for more sensitive organisms (larger χ). The delay τ reflects the memory imparted by the trail. The closed set of equations (XII.22), (XII.25), and (XII.29) constitute our effective dynamical description of the system.

For the average gradient one finds $\langle\partial_\perp\psi\rangle \sim e^{\alpha t}$ where the rate α is given implicitly as the solution of $\lambda(\alpha) = 0$ where

$$\lambda(\alpha) = 1 - \frac{\Omega\tau}{\alpha\tau}\left[1 + \frac{1}{\alpha\tau}\left(e^{-\alpha\tau} - 1\right)\right]. \tag{XII.30}$$

The relevant dimensionless parameter controlling the behaviour of our system is $\Omega\tau = k\chi/(\pi R v_0^2)$. For $\Omega\tau < 2$, $\alpha < 0$ and Eq. (XII.29) defines a random process with zero mean that leads to a stationary dynamics which is time-translation invariant. For $\Omega\tau > 2$ one finds $\alpha > 0$, i.e., the gradient (angular velocity) diverges exponentially in time. This implies that the trajectory develops a spiral form and converges to a localized point. We thus find that there is a maximum value of the product of trail deposition rate and sensitivity, $k\chi$, that allows steady-state motion. The phase diagram of the system is shown in Fig. XII.3.

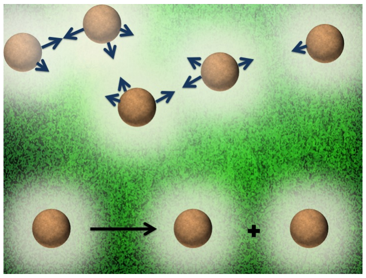


FIG. XIII.1 The stochastic process of birth and death can compete with chemical interactions.

XIII. CHEMOTAXIS AND CELL DIVISION

Chemotaxis plays a crucial role in living systems even at time scales relevant to cell division. Examples of such cases include cancer metastasis, the early stages of bacterial colony formation, wound healing, and development of embryos. However, the underlying mechanisms of these important processes are not fully understood due to the high complexity of these living many-body systems.

It is possible to shed some light on the dynamics of a system of chemically interacting cells, by taking into account cellular growth and death (Fig. XIII.1). Using a stochastic field theoretical framework, one can construct a dynamical equation for the density fluctuations of such a system as follows

$$\partial_t \rho = D_{\mathrm{c}} \nabla^2 \rho - \theta \rho - \nu_1 \nabla \cdot \left[\rho \nabla \left(\frac{1}{\nabla^2} \right) \rho \right] - \frac{\nu_2}{2} \rho^2 + \eta. \tag{XIII.1}$$

In this equation, ν_1 is the coupling constant for the term that represents chemotaxis and ν_2 is the coupling constant for the term that represents growth and death. One observes that the two seemingly unrelated nonlinearities have the same strength in terms of power counting, but possess different symmetries: the chemotactic term is conserved and the growth term is not (Gelimson and Golestanian, 2015). The noise correlator in Fourier space

$$\langle \eta(\mathbf{k}, \omega) \eta(\mathbf{k}', \omega') \rangle = 2 \left[D_0 + D_2 k^2 \right] (2\pi)^{d+1} \delta(\mathbf{k} + \mathbf{k}') \delta(\omega + \omega'), \tag{XIII.2}$$

highlights conserved and non-conserved contributions to the noise strength as well. The large scale behaviour of a population of cells that grow and interact through the concentration field of the chemicals they secrete can be studied using the method of dynamical renormalization group (RG). The combination of the effective long-range chemotactic interaction and lack of number conservation leads to a rich variety of phase behaviour in the system, including anomalous diffusion and number fluctuations at the critical point.

XIV. CONCLUDING REMARKS

A collection of interacting active colloids could serve as a promising model system to study collective non-equilibrium dynamics, as both the single-particle activity and the interactions could be controlled by construction. This approach will be directly relevant to studies of chemically active processes in biological systems, from enzyme cluster formation to multi-cellular organization. Moreover, it helps solve one of the challenges in making synthetic autonomous microscopic systems with non-trivial collective dynamics, which is the injection of energy at the level of individual objects and the ability to engineer the emergent properties of a collection of such objects.

Acknowledgements

I would like to acknowledge my many collaborators over the last few years with whom I have had the pleasure of working on Phoretic Active Matter: Tunrayo Adeleke-Larodo, Jaime Agudo-Canalejo, Armand Ajdari, Andrew Campbell, Jack Cohen, Stephen Ebbens, Anatolij Gelimson, Jon Howse, Yahaya Ibrahim, Pierre Illien, Richard Jones, Till Kranz, Tannie Liverpool, K.R. Prathyusha, Sriram Ramaswamy, Suropriya Saha, Rodrigo Soto, and Gerard Wong.

References

Adeleke-Larodo, T., J. Agudo-Canalejo, and R. Golestanian, 2019a, The Journal of Chemical Physics **150**(11), 115102.

Adeleke-Larodo, T., P. Illien, and R. Golestanian, 2019b, The European Physical Journal E **42**(3).

Agudo-Canalejo, J., T. Adeleke-Larodo, P. Illien, and R. Golestanian, 2018a, Acc. Chem. Res. **51**, 2365.

Agudo-Canalejo, J., and R. Golestanian, 2019, Phys. Rev. Lett. **123**, 018101.

Agudo-Canalejo, J., P. Illien, and R. Golestanian, 2018b, Nano Lett. **18**(4), 2711.

Alder, B. J., and T. E. Wainwright, 1967, Phys. Rev. Lett. **18**, 988.

Anderson, J. L., 1989, Annual Review of Fluid Mechanics **21**(1), 61.

Anderson, J. L., M. E. Lowell, and D. C. Prieve, 1982, J. Fluid Mech. **117**, 107.

Anderson, J. L., and D. C. Prieve, 1984, Sep. Purif. Rev. **13**(1), 67.

Bai, X., and P. G. Wolynes, 2015, J. Chem. Phys. **143**(16), 165101.

Bayati, P., M. N. Popescu, W. E. Uspal, S. Dietrich, and A. Najafi, 2019, Soft Matter **15**(28), 5644.

Brown, A., and W. Poon, 2014, Soft Matter **10**(22), 4016.

Campbell, A. I., S. J. Ebbens, P. Illien, and R. Golestanian, 2019, Nature Communications **10**(1).

Cohen, J. A., and R. Golestanian, 2014, Phys. Rev. Lett. **112**, 068302.

Curie, P., 1894, Journal de Physique Théorique et Appliquée **3**(1), 393.

Das, S., A. Garg, A. I. Campbell, J. Howse, A. Sen, D. Velegol, R. Golestanian, and S. J. Ebbens, 2015, Nature Communications **6**(1).

Derjaguin, B. V., 1987, in *Surface Forces and Surfactant Systems* (Steinkopff, Darmstadt), pp. 17–30.

Derjaguin, B. V., G. P. Sidorenkov, E. A. Zubashchenkov, and E. V. Kiseleva, 1947, Kolloidn. Zh **9**, 335.

Dey, K. K., S. Das, M. F. Poyton, S. Sengupta, P. J. Butler, P. S. Cremer, and A. Sen, 2014, ACS Nano **8**(12), 11941.

Ebbens, S., D. A. Gregory, G. Dunderdale, J. R. Howse, Y. Ibrahim, T. B. Liverpool, and R. Golestanian, 2014, EPL (Europhysics Letters) **106**(5), 58003.

Ebbens, S., R. A. L. Jones, A. J. Ryan, R. Golestanian, and J. R. Howse, 2010, Phys. Rev. E **82**, 015304.

Ebbens, S., M.-H. Tu, J. R. Howse, and R. Golestanian, 2012, Phys. Rev. E **85**, 020401.

Enculescu, M., and H. Stark, 2011, Phys. Rev. Lett. **107**, 058301.

Gelimson, A., and R. Golestanian, 2015, Phys. Rev. Lett. **114**(2), 028101.

Gelimson, A., K. Zhao, C. K. Lee, W. T. Kranz, G. C. L. Wong, and R. Golestanian, 2016, Phys. Rev. Lett. **117**, 178102.

Golestanian, R., 2009, Phys. Rev. Lett. **102**, 188305.

Golestanian, R., 2010, Phys. Rev. Lett. **105**(1), 1.

Golestanian, R., 2012, Phys. Rev. Lett. **108**, 038303.

Golestanian, R., 2015, Phys. Rev. Lett. **115**(10), 1.

Golestanian, R., and A. Ajdari, 2008, Phys. Rev. Lett. **100**, 038101.

Golestanian, R., T. B. Liverpool, and A. Ajdari, 2005, Phys. Rev. Lett. **94**, 220801.

Golestanian, R., T. B. Liverpool, and A. Ajdari, 2007, New Journal of Physics **9**(5), 126.

Guha, R., F. Mohajerani, M. Collins, S. Ghosh, A. Sen, and D. Velegol, 2017, J. Am. Chem. Soc. , jacs.7b08783.

Happel, J., and H. Brenner, 1981, *Low Reynolds number hydrodynamics* (Springer Netherlands).

Howse, J. R., R. A. L. Jones, A. J. Ryan, T. Gough, R. Vafabakhsh, and R. Golestanian, 2007, Phys. Rev. Lett. **99**, 048102.

Ibrahim, Y., R. Golestanian, and T. B. Liverpool, 2017, Journal of Fluid Mechanics **828**, 318.

Ibrahim, Y., R. Golestanian, and T. B. Liverpool, 2018, Phys. Rev. Fluids **3**, 033101.

Illien, P., T. Adeleke-Larodo, and R. Golestanian, 2017a, Europhys. Lett. **119**, 40002.

Illien, P., X. Zhao, K. K. Dey, P. J. Butler, A. Sen, and R. Golestanian, 2017b, Nano Lett. **17**, 4415.

Jee, A.-Y., S. Dutta, Y.-K. Cho, T. Tlusty, and S. Granick, 2018, Proc. Natl. Acad. Sci. U. S. A. **115**(1), 14.

Jülicher, F., and J. Prost, 2009, The European Physical Journal E **29**(1), 27.

Kanso, E., and S. Michelin, 2019, The Journal of Chemical Physics **150**(4), 044902.

Keller, E. F., and L. A. Segel, 1970, Journal of Theoretical Biology **26**(3), 399.

Kranz, W. T., A. Gelimson, K. Zhao, G. C. L. Wong, and R. Golestanian, 2016, Phys. Rev. Lett. **117**, 038101.

Landau, L. D., and E. M. Lifshitz, 2013, *Fluid Mechanics: Landau and Lifshitz: Course of Theoretical Physics, Volume 6* (Pergamon).

Lee, C. K., J. de Anda, A. E. Baker, R. R. Bennett, Y. Luo, E. Y. Lee, J. A. Keefe, J. S. Helali, J. Ma, K. Zhao, R. Golestanian, G. A. O'Toole, *et al.*, 2018, Proceedings of the National Academy of Sciences **115**(17), 4471.

Levin, Y., 2002, Reports on Progress in Physics **65**(11), 1577.

London, F., 1937, Transactions of the Faraday Society **33**, 8b.

Lorentz, H., 1896, Zittingsverslag van deKoninklijke Akademie van Wetenschappen te Amsterdam **5**, 168.

Michelin, S., and E. Lauga, 2014, Journal of Fluid Mechanics **747**, 572.

Michelin, S., and E. Lauga, 2015, The European Physical Journal E **38**(2).

Michelin, S., E. Lauga, and D. Bartolo, 2013, Physics of Fluids **25**(6), 061701.

Muddana, H. S., S. Sengupta, T. E. Mallouk, A. Sen, and P. J. Butler, 2010, J. Am. Chem. Soc. **132**, 2110.

Najafi, A., and R. Golestanian, 2004, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics **69**(February), 3.

Najafi, A., and R. Golestanian, 2010, EPL (Europhysics Letters) **90**(6), 68003.

Palacci, J., B. Abécassis, C. Cottin-Bizonne, C. Ybert, and L. Bocquet, 2010, Phys. Rev. Lett. **104**(13), 138302.

Palacci, J., S. Sacanna, A. P. Steinberg, D. J. Pine, and P. M. Chaikin, 2013, Science **339**(6122), 936.

Piazza, R., and A. Parola, 2008, Journal of Physics: Condensed Matter **20**(15), 153102.

Pohl, O., and H. Stark, 2014, Phys. Rev. Lett. **112**, 238303.

Popescu, M. N., M. Tasinkevych, and S. Dietrich, 2011, EPL (Europhysics Letters) **95**(2), 28004.

Prieve, D. C., J. L. Anderson, J. P. Ebel, and M. E. Lowell, 1982, J. Fluid Mech. **148**, 247.

Reigh, S. Y., P. Chuphal, S. Thakur, and R. Kapral, 2018, Soft Matter **14**(29), 6043.

Riedel, C., R. Gabizon, C. A. M. Wilson, K. Hamadani, K. Tsekouras, S. Marqusee, S. Pressé, and C. Bustamante, 2015, Nature **517**(7533), 227.

Rückner, G., and R. Kapral, 2007, Phys. Rev. Lett. **98**, 150603.

Saha, S., R. Golestanian, and S. Ramaswamy, 2014, Phys. Rev. E **89**, 062316.

Saha, S., S. Ramaswamy, and R. Golestanian, 2019, New Journal of Physics **21**(6), 063006.

Sakaue, T., R. Kapral, and A. S. Mikhailov, 2010, The European Physical Journal B **75**(3), 381.

Sengupta, S., K. K. Dey, H. S. Muddana, T. Tabouillot, M. E. Ibele, P. J. Butler, and A. Sen, 2013, J. Am. Chem. Soc. **135**(4), 1406.

Sengupta, S., M. M. Spiering, K. K. Dey, W. Duan, D. Patra, P. J. Butler, R. D. Astumian, S. J. Benkovic, and A. Sen, 2014, ACS Nano **8**(3), 2410.

Soto, R., and R. Golestanian, 2014, Phys. Rev. Lett. **112**, 068301.

Soto, R., and R. Golestanian, 2015, Phys. Rev. E **91**, 052304.

Stark, H., 2018, Accounts of Chemical Research **51**(11), 2681.

Stone, H. A., and A. D. T. Samuel, 1996, Phys. Rev. Lett. **77**, 4102.

Uspal, W. E., M. N. Popescu, S. Dietrich, and M. Tasinkevych, 2015, Soft Matter **11**(3), 434.

Van Kampen, N., 1981, J. Stat. Phys. **24**(1), 175.

Zwanzig, R., and M. Bixon, 1970, Phys. Rev. A **2**, 2005.

Flows in foams

Elise Lorenceau

- Foam structure.
- Dissipation in foams
- Non aqueous flow in foams

- ○ Wednesday : 09h00-10h30
- ○ Thursday : 14h0-15h30
- ○ Friday : 14h00-15h30

Flow in Foams and Flowing Foams

Sylvie Cohen-Addad,[1] Reinhard Höhler,[1]
and Olivier Pitois[2]

[1]Institut des NanoSciences de Paris, UMR 7588 CNRS-UPMC, Université Paris 6, 75005 Paris, France, and Laboratoire de Physique des Matériaux Divisés et des Interfaces, EA7264, Université Paris-Est, 77454 Marne-la-Vallée, France; email: sylvie.cohen-addad@insp.upmc.fr

[2]Laboratoire Navier, UMR 8205 CNRS – IFSTTAR – École des Ponts-ParisTech, Université Paris-Est, 77420 Champs-sur-Marne, France

Annu. Rev. Fluid Mech. 2013. 45:241–67

The *Annual Review of Fluid Mechanics* is online at fluid.annualreviews.org

This article's doi:
10.1146/annurev-fluid-011212-140634

Keywords

drainage, rheology, interfacial rheology, non-Newtonian flow, permeability, multiscale modeling

Abstract

Aqueous foams are complex fluids composed of gas bubbles tightly packed in a surfactant solution. Even though they generally consist only of Newtonian fluids, foam flow obeys nonlinear laws. This can result from nonaffine deformations of the disordered bubble packing as well as from a coupling between the surface flow in the surfactant monolayers and the bulk liquid flow in the films, channels, and nodes. A similar coupling governs the permeation of liquid through the bubble packing that is observed when foams drain due to gravity. We review the experimental state of the art as well as recent models that describe the interplay of the processes at multiple length scales involved in foam drainage and rheology.

1. INTRODUCTION

Aqueous foams are concentrated dispersions of gas bubbles in a surfactant solution. Their structures are organized over a large range of length scales, from the size of a bubble down to the scale of the surfactant molecules adsorbed on liquid-gas interfaces. Surfactants confer specific rheological properties to the interfaces that are coupled to the flow of the underlying bulk liquid. Thus foam drainage and rheology result from interplay between processes over different length scales and timescales. Many open questions about foam rheology were identified in an article in this series by Kraynik (1988) 25 years ago. In this review, we present recent experimental and theoretical progress in foam drainage and rheology, and we emphasize the link between micro- and macroscales as an essential feature of their understanding.

2. FOAM STRUCTURE AND INTERFACES

2.1. Structure and Aging

Equilibrium foam structures have a minimal interfacial energy density for fixed bubble volumes. This requirement determines the dependency of the foam structure on the liquid volume fraction ϕ_l illustrated in **Figure 1**. For ϕ_l larger than the close-packing fraction $\phi_{l,c}$, the dispersion is a bubbly liquid. At $\phi_l = \phi_{l,c}$, the bubble packing jams, and a wet foam is formed in which the bubbles are almost spheres. With decreasing ϕ_l, they progressively become polyhedra separated by thin films. Close to the dry limit at which $\phi_l \to 0$, the equilibrium foam structure satisfies Plateau's rules: Films join three by three at equal angles of 120°, and their junctions, called Plateau borders, meet four by four at nodes, forming equal angles of 109.5° (**Figure 1*b*,*d*,*e***) (Weaire & Hutzler 1999, Cantat et al. 2010).

Foam structures can be either disordered (**Figure 1*a*,*b***) or ordered (**Figure 1*c*,*d***). Monodisperse dry foams crystallize into the Kelvin [body-centered cubic (bcc)] structure, whereas wet monodisperse foams have a face-centered cubic (fcc) or hexagonal close-packed (hcp) structure (Weaire & Hutzler 1999, Cantat et al. 2010). These two structures have a close-packing fraction $\phi_{l,fcc} = 0.26$, smaller than that of monodisperse disordered foams $\phi_{l,rcp} = 0.36$. Laplace's law $\Delta P = 2\gamma\kappa$ relates the mean curvature κ of the gas-liquid interfaces to the pressure difference ΔP across them and the surface tension γ. In dry foams, the cross section of a Plateau border is a concave triangle with curvature $1/R_{Pb}$ (**Figure 1*e***), and ΔP is called the capillary pressure, $P_c = \gamma/R_{Pb}$. The cross-sectional area $A = (\sqrt{3} - \pi/2)\, R_{Pb}^2$ of such a Plateau border increases with ϕ_l and with the average bubble radius R. For $\phi_l \leq 2\%$–3%, Plateau borders are slender and $R_{Pb} \cong R\sqrt{\phi_l/c}$, where c is a geometrical constant ($c \cong 0.33$ for the Kelvin structure) (Koehler et al. 2000). For larger ϕ_l, borders and nodes are no longer well defined (**Figure 1*f***).

If foam is in contact with a liquid reservoir via a semipermeable membrane, it will suck the liquid because this allows the bubbles to become more spherical and to decrease their interfacial energy. The pressure that must be exerted to prevent such a flow is called the foam osmotic pressure, Π (Princen 1986). At the jamming transition $\phi_{l,c}$, bubbles are spherical, and the surface energy density is minimal so that $\Pi = 0$. The variations of Π found in experiments and simulations over the full range of liquid fractions illustrated in **Figure 1*g*** are described by the semiempirical law

$$\Pi = k\frac{\gamma}{R}\frac{(\phi_{l,c} - \phi_l)^{\beta}}{\sqrt{\phi_l}}, \tag{1}$$

with $k \cong 7.3$, $\beta \cong 2$, and $\phi_{l,c} = \phi_{l,fcc}$ for an fcc structure (Höhler et al. 2008). The data for disordered structures yield the rough estimates $k \cong 3.5$, $\beta \cong 2.5$, and $\phi_{l,c} = \phi_{l,rcp}$ (Princen & Kiss 1987). Note that these parameter values may not be valid in the asymptotic limit $\phi_l \to \phi_{l,c}$.

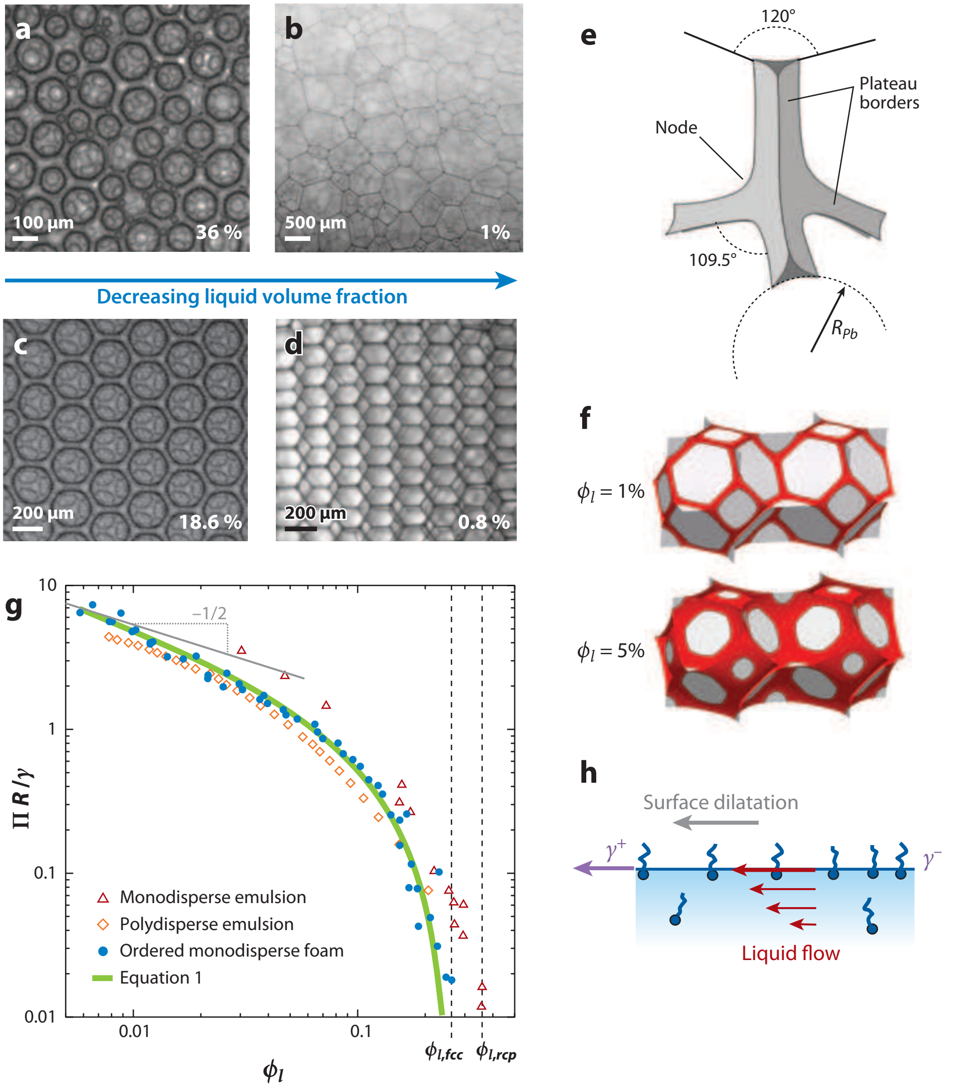


Figure 1

Disordered foams for the indicated liquid volume fractions: (*a*) foam at the jamming transition $\phi_l = \phi_{l,rcp}$ and (*b*) dry foam satisfying Plateau's laws. Panel *a* courtesy of R. Lespiat. Ordered monodisperse foams for the indicated liquid volume fractions: (*c*) fcc structure of a wet foam and (*d*) bcc Kelvin structure of a dry foam. The transition between fcc and Kelvin structures occurs for $\phi_l = 0.06$. Panels *c* and *d* courtesy of Y. Yip Cheung Sang. (*e*) Four Plateau borders with radius of curvature R_{Pb} joining at a node. The angles are those predicted by Plateau's rules for a dry equilibrium foam structure. (*f*) Structure of Kelvin foam simulated with the Surface Evolver software for two liquid fractions. (*g*) Variations of the osmotic pressure Π normalized by γ/R with the liquid volume fraction ϕ_l, for foams and emulsions. In the limit of dry foams, $\Pi \cong 0.6\gamma/(R\phi_l^{1/2})$. The green line corresponds to Equation 1. Data taken from Höhler et al. (2008). Panel *g* adapted with permission from Höhler et al (2008). Copyright 2008 American Chemical Society.
(*h*) Schematic illustration of a Marangoni flow. As a surfactant monolayer is dilated, surface tension gradients are created. Areas of high surface tension γ^+ contract at the expense of areas of low surface tension γ^-. The surface stress also induces flow in the liquid underneath.

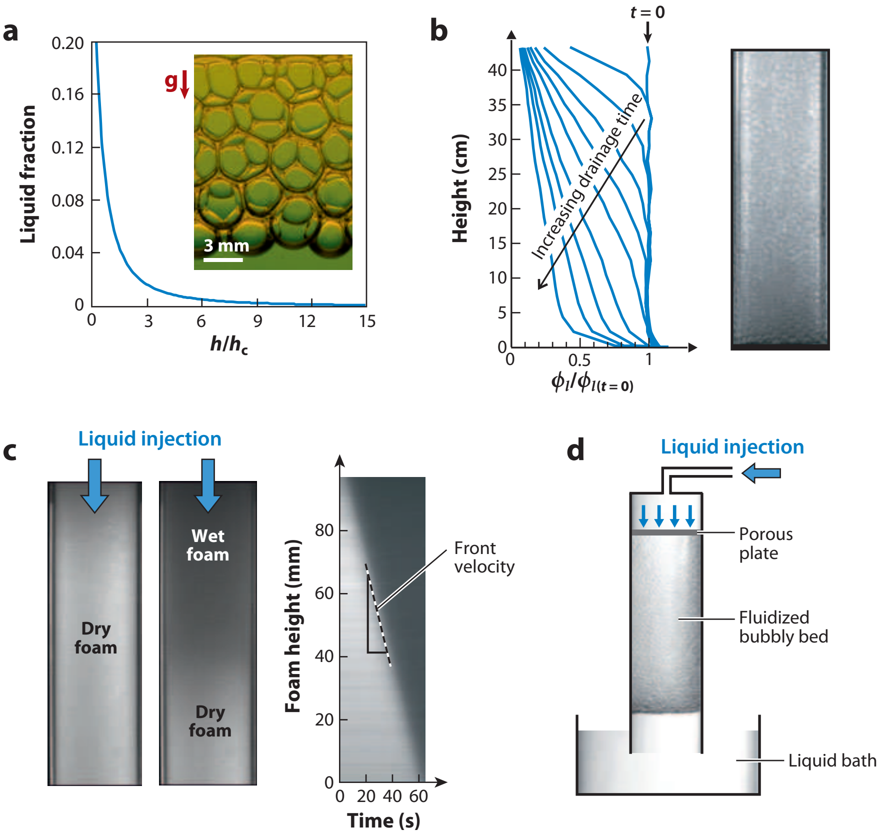


Figure 2

(*a*) Equilibrium liquid fraction profile of a disordered foam under gravity. The height h above the bottom of the foam column is normalized by the characteristic length $h_c = \gamma/(\rho g R)$. The inset illustrates the corresponding liquid distribution in the Plateau border network. Photograph courtesy of H. Caps and D. Terwagne, GRASP, University of Liège. (*b*) Free-drainage experiment. (*Left*) Time evolution of the vertical liquid fraction profile in a foam ($R = 90$ μm) measured using electrical conductivity (time interval between successive curves is 6 min). Data taken from Saint-Jalmes & Langevin (2002). At $t = 0$, the liquid fraction is uniform ($\phi_l \cong 0.15$), but as time goes by, a drainage front propagates downward so that the foam dries up at the top. At long times, the vertical profile reaches the equilibrium showed in panel *a*. (*Right*) Such a profile observed by light transmission. Wet regions appear darker than dry ones. (*c*) Forced-drainage experiment. (*Left*) A liquid front invades an initially dry foam column from the top, as observed by light transmission at two successive times. The volumetric velocity of the perfused liquid is $U \cong 0.05$ mm s^{-1}. (*Right*) Spatiotemporal picture of the experiment, allowing the front velocity to be deduced from the slope of the frontier separating the dark region from the light region. (*d*) Fluidized bubbly bed experiment. The height of the bubbly bed is measured as a function of the liquid flow rate imposed from the top via a porous plate. Under these conditions, liquid fractions up to 80% are obtained.

When a foam column containing a liquid of density ρ equilibrates under gravity, the liquid drains until the variation of osmotic pressure and buoyancy over any vertical distance interval dz reaches a balance $d\Pi = \rho g[1 - \phi_l(z)]dz$, which determines the equilibrium liquid distribution (Princen 1986). The resulting liquid volume fraction profile is presented in **Figure 2*a***. In the dry limit, the equilibrium condition in the presence of a pressure gradient ∇P reduces to $-\nabla P \simeq \nabla(\gamma/R_{Pb})$.

FOAMY COMPLEX FLUIDS

Substituting the usual Newtonian foaming liquids with polymer solutions, oil-in-water emulsions, or suspensions of solid grains greatly modifies foam drainage and rheology. Non-Newtonian interstitial flow can be accounted for in drainage theory (Safouane et al. 2006), but models describing the flow of homogeneous foaming liquids are not applicable to suspensions of particles, oil droplets, or aggregates whose size is comparable to the channel diameters (Koczo et al. 1992, Britan et al. 2009, Carn et al. 2009, Guignot et al. 2010, Louvet et al. 2010, Rouyer et al. 2011). Foams laden with a small amount of noncolloidal particles exhibit strongly enhanced viscoelasticity (Cohen-Addad et al. 2007). This effect can be understood in terms of capillary attraction between the particles, leading to a rigidity percolation phenomenon. Concentrated colloidal suspensions can stop drainage owing to their yield stress, and such foams are therefore very stable (Guillermic et al. 2009, Salonen et al. 2012). However, when these aerated yield-stress fluids are sheared, the interstitial flow is restored, and the drainage velocity is then controlled by the applied shear rate (Turner et al. 1999, Goyon et al. 2010). Remarkably, hydrophobic colloidal particles can stabilize foams in the absence of surfactant molecules, by adsorbing at the interfaces (Aveyard et al. 2003, Binks & Murakami 2006, Cervantes-Martinez et al. 2008). The drainage and rheological properties of these foams have hardly been explored.

Because of their large interfacial energy, liquid foams are intrinsically unstable. They evolve via three coupled processes: Coalescence results from the rupture of films between bubbles. It can be avoided by using large surfactant concentrations, the case we consider in this review. In addition, foams age because of drainage, as described in detail in Section 3. The third aging process results from Laplace pressure differences between neighboring bubbles, which drive diffusive gas exchange between them. As the foam ages (i.e., as the time elapsed since production increases), small bubbles shrink to zero size, and the average bubble size grows. This coarsening occurs at a rate that increases with the gas solubility, the diffusion speed in the liquid, and the gas volume fraction. Coarsening is accompanied by intermittent bubble rearrangements that locally relax interfacial energy. The rate of these events decreases as the foam ages (Durian et al. 1991, Cohen-Addad & Höhler 2001). Recent work has shown that drainage, coarsening, and coalescence, which need to be suppressed in many applications, can be inhibited by the use of complex fluids as foaming liquids (see the sidebar, Foamy Complex Fluids).

2.2. Interfacial Rheology

Surfactant molecules spontaneously adsorb on liquid-gas interfaces and decrease their surface tension γ. A liquid-gas interface with adsorbed surfactants does not behave as a free surface. Dilatation creates gradients of surfactant concentration (**Figure 1*b***), leading in turn to surface tension gradients that pull on the interface to reestablish the equilibrium surfactant concentration. The gradient of γ balances with bulk viscous stress at the surface. This induces an entrainment of the liquid beneath called the Marangoni effect. Surface tension gradients not only relax via surface transport of surfactants, they also fade away owing to exchanges between the bulk and the interface. The surface resistance against deformation can be so strong that it behaves as a deformable, but rigid incompressible interface. Conversely, significant surface flow under applied stress indicates interfacial mobility (Levich 1962).

The rheological behavior of the interface is described by the surface stress tensor related to the surface deformation by a constitutive law. For liquid-like interfaces with Newtonian behavior, the Boussinesq-Scriven law relates stress to the surface rate of deformation, depending on the surface dilatational viscosity, μ_d, and surface shear viscosity, μ_s (Edwards et al. 1991, Sagis 2011). For elastic interfaces, a generalized Hooke's law with surface dilatational elastic modulus E and surface shear modulus G_s describes the relation between stress and surface deformation.

LOH: lauryl alcohol

The surface dilatational modulus is defined as $E = A(\mathrm{d}\gamma/\mathrm{d}A)$, where A is the surface area. It describes the resistance of the interface against surface tension gradients or homogeneous dilatation. The surface dilatational viscosity affects the rate at which these gradients vanish owing to intrinsic surface friction or to surfactant diffusion/adsorption from the bulk (Lucassen-Reynders 1981, Ivanov et al. 2005). If the surface is dilated with a frequency f, its response is described by a complex surface dilatational modulus $E^*(f) = E'(f) + i\,E''(f)$. In the limit of high frequency, the interface behaves as an insoluble monolayer, $E' \to E_\infty$ the limiting elasticity, and $E'' = 2\pi f \mu_d$ is reduced to the small intrinsic viscous contribution to μ_d (Ivanov et al. 2005). In the zero frequency limit, surfactant exchanges between the bulk and surface have time to occur, and the interface does not resist dilation: $E' \to 0$ and $E'' \to 0$. At intermediate frequency, the interface is viscoelastic.

The surface rigidity/mobility strongly depends on the surfactant composition. Surfactants of low solubility such as lauryl alcohol (LOH), fatty acids, and some proteins are known to increase considerably both the shear and dilatational moduli and viscosities, enhancing the surface rigidity (e.g., see Sagis 2011). When foam flows, the interfaces are sheared and dilated. However, because in general $\mu_d \gg \mu_s$ and $E_d \gg G_s$, the coupling between bulk rheological properties and interfacial viscoelasticity tends to be dominated by the dilatational surface properties. In contrast, as liquid flows between the bubbles during drainage, the interfaces are essentially sheared, and the surface shear viscosity μ_s generally dominates the surface stress.

3. FLOW IN FOAMS

3.1. Experimental Investigation of Foam Drainage

Free drainage is the spontaneous evolution of foam liquid fraction due to gravity, from an initial state toward the equilibrium state described in Section 2.1. The free-drainage experiment probes liquid flow in foams, by analyzing the volume of liquid drained at the bottom of a foam column, by following the temporal evolution of the vertical liquid fraction profile (Koehler et al. 2000, Carrier et al. 2002, Saint-Jalmes & Langevin 2002) (**Figure 2*b***), or by measuring bubble velocities at the column wall (Maurdev et al. 2006). To interpret such data quantitatively, one must carefully consider the initial liquid fraction profile (Saint-Jalmes et al. 2000) and bubble coarsening (Hilgenfeldt et al. 2001) as they have a significant influence on the drainage.

Miles et al. (1945) have studied another interesting drainage configuration in which liquid is continuously introduced at the top of the foam column. This procedure yields a steady-state flow in which the amount of liquid held in the foam is determined by stopping the liquid flow and draining the foam column. In the forced-drainage variant experiment, the foam column is left to drain before liquid is perfused at a constant flow rate Q (Weaire et al. 1993). A region of wet foam expands downward with a front velocity u_{front} (**Figure 2*c***). The average vertical interstitial liquid velocity above the front is the same as that of the front, and the fraction of liquid flowing through the perfused foam is simply given by $\phi_l \cong U/u_{front}$, where $U = Q/S$, and S is the column cross-sectional area. In contrast to free drainage, flow conditions are the same throughout the wet region, resulting in a constant liquid fraction whose value is set by the balance between gravity and viscous drag. The foam permeability thus reads $K = \mu\, U/(\rho g)$, with μ the bulk shear liquid viscosity, so that the forced-drainage experiment provides a simple tool to determine $K(\phi_l)$. However, for $\phi_l \geq 0.1 - 0.2$, the quantitative analysis of the front become more difficult because of gas compressibility (Koehler 2012), inertia effects (Stevenson & Li 2010), or internal convective motions (Hutzler et al. 1998). Recently, a fluidized bubbly bed experiment (Rouyer et al. 2010) has been used to extend the study of $K(\phi_l)$ to bubbly liquids (**Figure 2*d***).

SDS: sodium dodecyl sulfate

3.2. Liquid Flow Through Plateau Borders, Nodes, and Films

Schematically, liquid drainage through a bubble assembly can be considered as flow through a collection of channels (Plateau borders) and nodes (their junctions). The Reynolds number based on the typical size of a channel (100 μm) and the average channel liquid velocity $\bar{u}$ (1 mm s^{-1}) is $Re = \rho R_{Pb} \bar{u}/\mu \cong 0.01\text{–}1$. For these small Re, the velocity field $\mathbf{u}$ is expected to obey the Stokes equation. Note that for large bubbles (several millimeters) at liquid fraction $\phi_l > 0.1$, inertia effects are expected (Stevenson & Li 2010).

Several models are proposed to capture the hydraulic behavior of Plateau borders. For Hagen-Poiseuille flow driven by a pressure gradient ∇P in an infinitely long solid pipe of cross-sectional area A, the dimensionless channel permeability is

$$k_{Pb} = \mu \, \bar{u}/(A \nabla P). \tag{2}$$

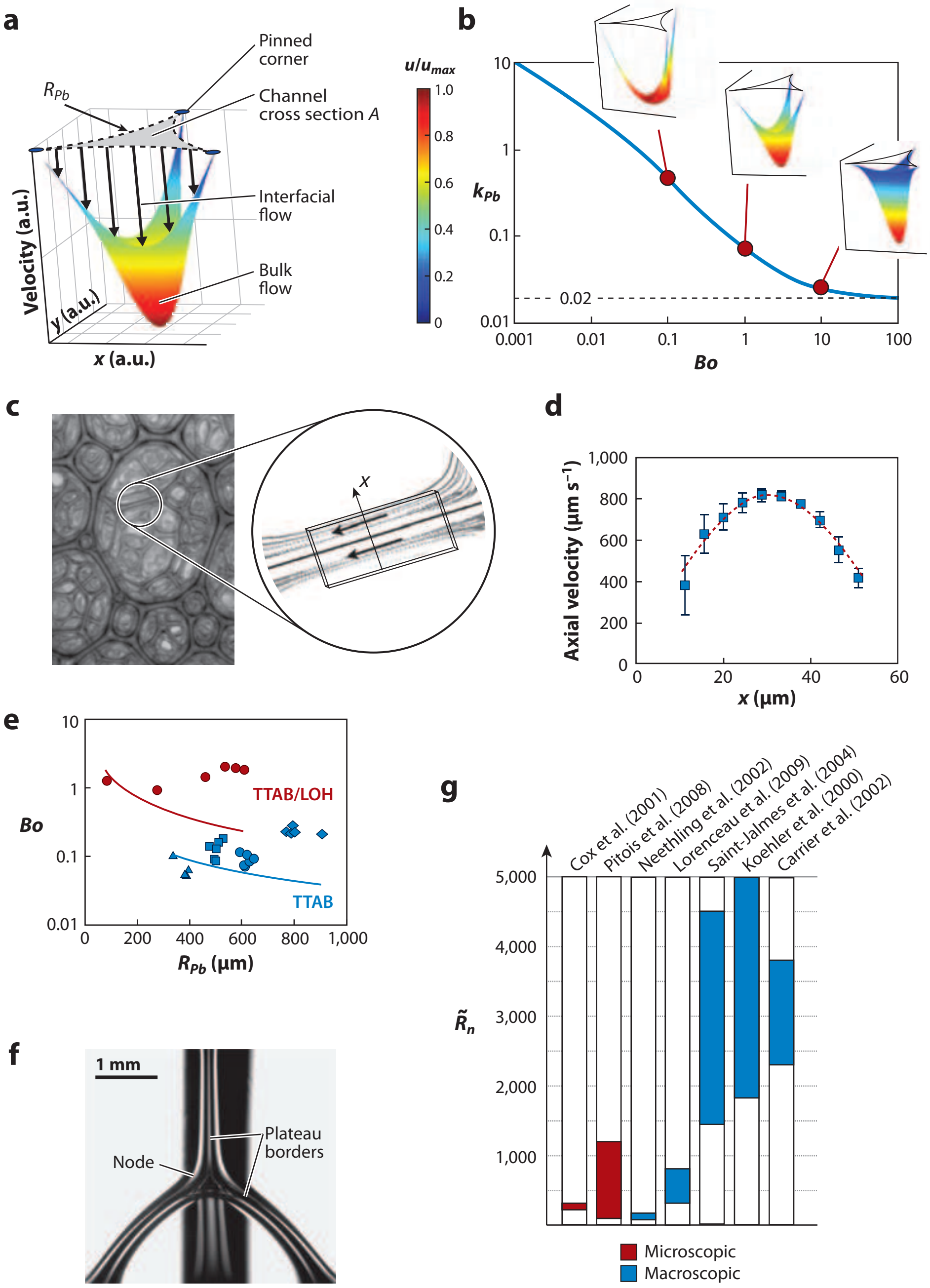
For a Plateau border, A is equal to $(\sqrt{3} - \pi/2)\, R_{Pb}^2$ and the dimensionless permeability $k_{Pb} \cong 0.020$ is calculated (Nguyen 2002, Koehler et al. 2004a). The influence of interfacial mobility has been considered in pioneering work by Leonard & Lemlich (1965a) by balancing the bulk viscous stress with the surface shear viscous stress at the boundary: $\mu\,(\mathbf{n} \cdot \nabla u) = \mu_s \nabla_s^2 u$, where $\mathbf{n}$ is the unit vector normal to the channel surface and ∇_s is the surface gradient. At the three Plateau border corners, the interfaces are assumed to be pinned (**Figure 3*a***). The dimensionless Boussinesq number characterizing the influence of the surface on the flow is

$$Bo = \frac{\mu_s}{\mu R_{Pb}}. \tag{3}$$

Numerical as well as approximate analytical solutions of the Stokes equation with these boundary conditions provide values for $k_{Pb}(Bo)$ (Nguyen 2002, Koehler et al. 2004a). $Bo \gg 1$ corresponds to rigid interface behavior with $k_{Pb} = 0.02$ as above, but as Bo is decreased, k_{Pb} increases significantly owing to the surface shear flow. The length scale $R_{Pb}\, Bo$ measures the region of influence of the pinned corners, as shown in **Figure 3*b***. For small Bo, the interface is entrained by the bulk flow, except in the corner regions, so that the flow is essentially plug-like, and the Plateau border permeability increases tenfold. Numerical studies have evidenced the significant influence on k_{Pb} of the geometry of the corners, accounting for a finite film thickness (Koehler et al. 2004a) or a finite contact angle (Rouyer et al. 2008).

Evidence for the impact of interfacial mobility has been provided by microparticle-tracking experiments using confocal microscopy in dry foams (Koehler et al. 2002, 2004b), as illustrated in **Figure 3*c***. The axial velocity profile measured across a Plateau border is parabolic with boundaries flowing at a velocity that is about half the maximum velocity for sodium dodecyl sulfate (SDS) solutions (**Figure 3*d***). The same experiment performed with a solution of bovine serum albumin revealed a vanishing velocity at the boundary, as expected for large Bo (Koehler et al. 2002). Other experiments have been developed to investigate liquid flow in a single Plateau border (Koczo & Racz 1987, Pitois et al. 2005b). Inspired by microfluidic experiments, the Plateau border apparatus (Pitois et al. 2005b) delivers controlled flow rates through a single vertical Plateau border of a few millimeters in length, while the driving pressure gradient and the channel cross-sectional area are measured. The channel permeability is deduced using Equation 2, and the corresponding Bo is determined from the theoretical curve $k_{Pb}(Bo)$ (**Figure 3*b***). A value of $Bo \cong 0.1$ is found for SDS and tetradecyltrimethylammonium bromide (TTAB) solutions (**Figure 3*e***), corresponding to an extremely low surface shear viscosity $\mu_s \cong 2\ 10^{-8}$ kg s^{-1}. In contrast, $Bo \cong 1$ if LOH is added in sufficient amount. This increase of the surface shear viscosity is indeed expected because of the compact structure formed by LOH molecules solubilized in TTAB or SDS monolayers (Lu et al. 2000). Even though this model is globally consistent in terms of Bo, several features are not captured. In fact, Bo seems to be approximately constant when R_{Pb} changes (**Figure 3*e***) or when the bulk viscosity is increased by the addition of glycerol (Pitois et al. 2005a). These effects are

TTAB: tetradecyltrimethyl-ammonium bromide



observed in interstitial liquid flows through foams, but they are absent in the related problem of falling rivulets between solid plates (Drenckhan et al. 2007).

Interstitial liquid flow induces surfactant concentration gradients at the channel surfaces close to their junctions (see **Figure 4*a***) (Pitois et al. 2009b). The effects of Marangoni stresses are indeed observed in foam drainage, as illustrated in **Figure 4*e*** (Leonard & Lemlich 1965b, Carrier et al. 2002, Pitois et al. 2005b). To model the balance between Marangoni stresses at the channel surfaces and viscous stresses from the bulk, Durand & Langevin (2002) introduced the surface elasticity E_∞ (Durand & Langevin 2002). However, this model predicts a significant decrease of the channel permeability with E_∞ which is incompatible with the values k_{Pb} measured for foams with mobile interfaces. Recently, it has been suggested that Marangoni stresses could be partially relaxed through surfactant recirculation (Pitois et al. 2009b). Schematically, surfactants are entrained by bulk flow along the channel surface, inducing a surface tension gradient along the flow direction. This drives a Marangoni counterflow in the thin channel/film transitional areas (**Figure 4*b–d***). Globally, the permeability k_{Pb} is predicted to decrease with increasing surface elasticity and surface shear viscosity, in qualitative agreement with the data (Pitois et al. 2009b). Moreover, Marangoni flows induce the films to swell up to several micrometers, providing new paths for the liquid (Carrier et al. 2002). A numerical calculation shows that the films' contribution becomes substantial if their thickness h satisfies the relation $h/R \geq \phi_l$ (Koehler et al. 2004a). However, the quantitative contribution of the swollen films to the drainage remains to be elucidated.

Flow in a node is more complex than in a single Plateau border because of the multiple connections with upstream and downstream channels. Few experimental and theoretical works have been devoted to its study. Particle-tracking experiments have shown streamlines, highlighting the laminar character of the flow and the lack of mixing in the nodes (Koehler et al. 2004b). Pressure-drop, ΔP, measurements have been performed (Pitois et al. 2008) in the configuration in which the flow from one vertical upstream Plateau border (with flow rate q) exits via three downstream Plateau borders (**Figure 3*f***). Similar to foam channels, the results show that the flow in a node is significantly influenced by the surfactants used (Pitois et al. 2008). The measured dimensionless hydraulic resistance of the node, $\tilde{R}_n = 1/k_n = (R^3_{Pb}/\mu)(\Delta P/q)$, ranges from 150 for TTAB solutions to 1,200 for TTAB/LOH mixtures (**Figure 3*g***). The value $\tilde{R}_n$ for a TTAB

Figure 3

Liquid flow in a Plateau border. (*a*) Surface plot illustrating the liquid velocity field in an infinitely long Plateau border of radius R_{Pb}. As suggested by the arrows, the channel surface is entrained by the bulk flow except at the three corners at which the channel joins the foam films. Panel *a* courtesy of F. Rouyer. (*b*) Dependence of the dimensionless channel permeability k_{Pb} on Boussinesq number (Equations 2 and 3). In the limit of large *Bo*, the interface provides rigid boundary conditions and $k_{Pb} = 0.02$. As *Bo* diminishes, k_{Pb} increases significantly owing to the surface shear flow. This evolution is illustrated by the three surface plots for $Bo = 0.1$, 1, and 10. Panel *b* courtesy of F. Rouyer. (*c*,*d*) Particle-tracking experiment in an SDS foam perfused with a suspension of 1-μm latex microspheres (Koehler et al. 2004b). Monitoring these tracers in a Plateau border using confocal microscopy reveals flow streamlines. The analysis of the particle positions as a function of time gives the axial velocity profile of the flow. Interfacial mobility is evidenced by the finite velocity measured at the channel's boundaries. (*e*) Influence of interfacial mobility on liquid flow through a single Plateau border, as probed with the Plateau border apparatus (Pitois et al. 2005b). The Boussinesq number is determined for several Plateau border sizes and liquid flow rates, for TTAB (*blue*) and TTAB/LOH (*red*) solutions. Solid lines represent Equation 3, assuming a constant value for μ_s. Data taken from Pitois et al. (2005a). (*f*) Single foam node formed at the junction of four Plateau borders (Pitois et al. 2008). (*g*) Comparison of values for the dimensionless resistance parameter $\tilde{R}_n$ of a foam node published in the literature. Pitois et al. (2008) and Cox et al. (2001) correspond to studies at the microscopic level. The other references correspond to macroscopic studies.

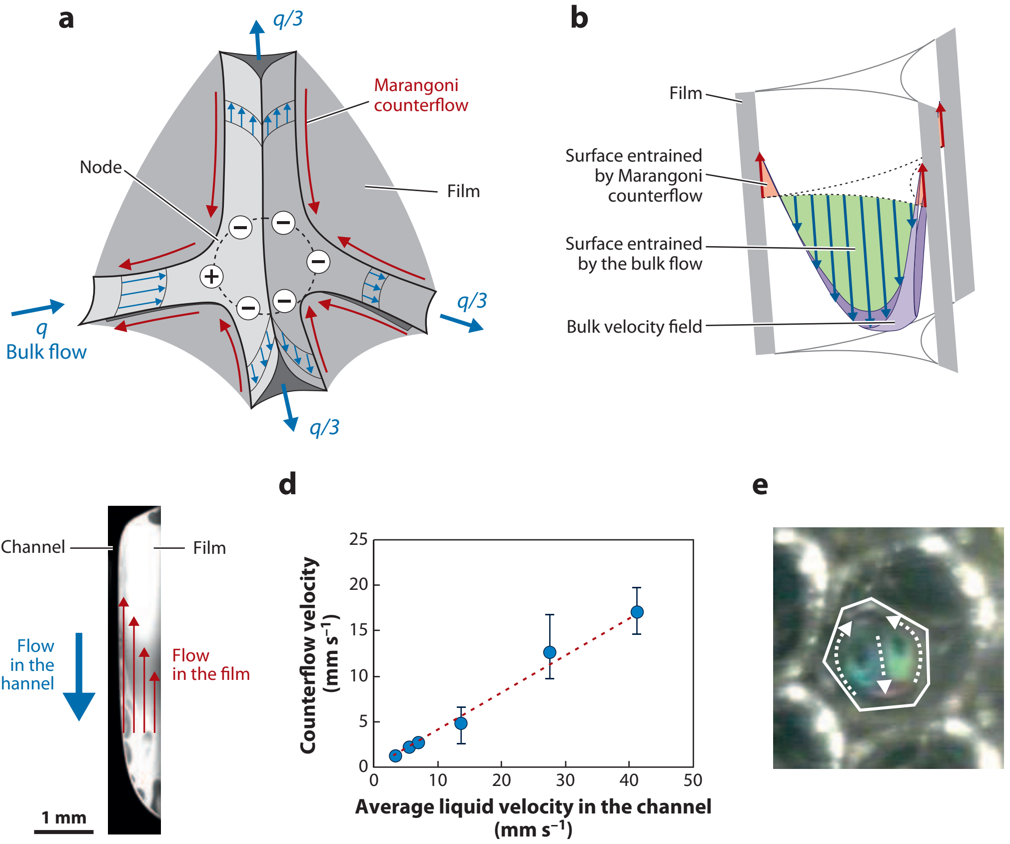


Figure 4

(*a*) Coupled bulk and surface flows close to a node. The injected flow rate q is distributed over three output channels (*blue arrows*). The surface flows at upstream and downstream channels create an excess (*plus signs*) or a lack (*minus signs*) of surfactant molecules compared to the equilibrium surface concentration. Marangoni counterflows in the channel/film transitional regions (*red arrows*) act against this imbalance of surface concentrations. (*b*) Surfactant flow at a Plateau border surface. Surfactants are entrained by bulk flow in the central region of the channel surface, whereas Marangoni counterflows near the corners of the channel cross section entrain surfactant molecules in the opposite direction. This recirculation mechanism reduces the channel permeability k_{Pb}. (*c*,*d*) The channel/film transition region as observed in the Plateau border apparatus (Pitois et al. 2005b). Thin spots driven by Marangoni counterflow are observed (*dark areas in the film*) with velocity proportional to the average liquid velocity in the channel. (*e*) A vertical film in a foam during a perfusion with a solution of sodium dodecyl benzene sulfonate (Carrier et al. 2002). The colored pattern reveals Marangoni counterflows close to the Plateau borders.

solution compares well with that simulated for a foam node with free (totally mobile) interfaces (Cox et al. 2001).

3.3. Multiscale Modeling of Foam Drainage

Foam drainage is modeled as flow through a porous medium composed of Plateau borders and nodes. The interstitial liquid flows under the combined actions of the hydrostatic and capillary pressure gradients. At a scale corresponding to many bubbles, the liquid flow rate per unit area is described by Darcy's law (Koehler et al. 2000), $\mathbf{U} = K(\rho\,\mathbf{g} - \nabla P)/\mu$. One of the main differences between foams and solid porous media is that the foam capillary structure adjusts to the interstitial liquid flow according to Laplace's law. As presented in Section 2.1, the pressure gradient can be expressed as $-\nabla P \cong \nabla(\gamma/R_{Pb})$ in the limit of dry foams. Combining these relations, the continuity equation and the Plateau border geometry described in Section 2.1 give the foam-drainage equation, which describes the dynamics of the foam liquid fraction (Goldfarb et al. 1988, Verbist & Weaire 1994):

$$\frac{d\phi_l}{dt} + \nabla \cdot \left(\frac{K\,\rho}{\mu}\,\mathbf{g} \right) - \nabla \cdot \left(\frac{\gamma\,K}{2\,c\,R\,\mu}\,\phi_l^{-3/2}\,\nabla\phi_l \right) = 0. \tag{4}$$

For free-drainage boundary conditions, it predicts the experimentally observed power-law decrease of ϕ_l with time at a fixed position in space (Koehler et al. 1998, Koehler et al. 2000). Moreover, Equation 4 has a solitary wave solution observed in the case of the forced-drainage experiment (**Figure 2*c***) (Weaire et al. 1997). Capillary forces become important in the front region and set its width: $w_{front} \approx 2.5 \times 10^{-3} \phi_l^{-1/2}$ (Koehler et al. 2000).

Quantitative predictions require an expression of the foam permeability *K*. As a starting point, we consider the channel-dominated regime (Verbist et al. 1996) in which dissipation mainly occurs in the Plateau borders, modeled as a network of randomly oriented channels, a problem first addressed by Saffman (1959). Adapted to foams, the permeability reads $K_{channels} = A k_{Pb}\, \phi_l/3$. In the limit of rigid interfaces, k_{Pb} is equal to 0.02 and $K_{channels} \cong 3.2 \times 10^{-3}\, R^2\, \phi_l^2$. Combining this expression of $K_{channels}$ with Equation 4 gives the foam-drainage equation as first derived by Goldfarb et al. (1988). This model is useful for dry foams stabilized by proteins or specific surfactant mixtures with high surface shear viscosity (Saint-Jalmes & Langevin 2002). However, in many cases, the observed foam permeability is larger than predicted, and one has to consider the effect of interfacial mobility on k_{Pb} (Section 3.2). Moreover, the observed scaling for foam permeability is closer to $\phi_l^{3/2}$ than to ϕ_l^2, as shown in **Figure 5**.

To explain the observed $K \propto \phi_l^{3/2}$ power law, Koehler et al. (1999) proposed the so-called node-dominated regime in which dissipation occurs solely in the nodes, whose volume fraction in the channel network $v_{nodes}/(v_{nodes} + v_{channels})$ is proportional to $\phi_l^{1/2}$ for a dry foam. Balancing the viscous damping force (per unit liquid volume) inside the nodes of the network $\phi_l^{1/2}\,(\mu\,\bar{u}/R_{Pb}^2)$ by the pressure gradient, expressed as $\mu(\bar{u}\phi_l)/K_{nodes}$ using Darcy's law, one obtains the foam permeability $K_{nodes} \sim R^2\,\phi_l^{3/2}$. However, the assumed absence of dissipation in the Plateau borders is hard to reconcile with the known values of μ_s. More general models in which the permeabilities of Plateau borders and nodes are associated (i.e., $1/K = 1/K_{channels} + 1/K_{nodes}$) are generally preferred (Koehler et al. 2000, Carrier et al. 2002, Neethling et al. 2002, Saint-Jalmes et al. 2004, Lorenceau et al. 2009). By setting the respective dissipation parameters for the channels and the nodes, these models allow any power-law exponent in the range 3/2–2 to be predicted (**Figure 5*b*,*c***). However, ad hoc assumptions are generally made to fit experimental data: a vanishing dissipation in the Plateau borders (Koehler et al. 2000), a constant value for k_{Pb} (Neethling et al. 2002) or a constant *Bo* number (Lorenceau et al. 2009), an additional resistance due to the foam films (Carrier et al. 2002), or a complex evolution for the node resistance as a function of *Bo* (Saint-Jalmes et al. 2004). The fitted values for k_{Pb} (or μ_s) and $\tilde{R}_n$ are scattered over a wide range, as shown in **Figure 3*g***, which originates from both the incorrect description of the flow at the microscopic level and the intrinsic difficulty of capturing the foam geometry beyond the dry limit (Lorenceau et al. 2009, Cantat et al. 2010, Koehler 2012). As an alternative to these geometrical models, an approach based on the Carman-Kozeny equation has been applied to foams (Pitois et al. 2009a). The starting point is the specific surface area of the interstitial network of the bubble packing calculated by surface energy minimization (Höhler et al. 2008). It has been shown that for foams with rigid interfaces, the Carman-Kozeny approach reproduces well the channel-dominated regime at low liquid fractions

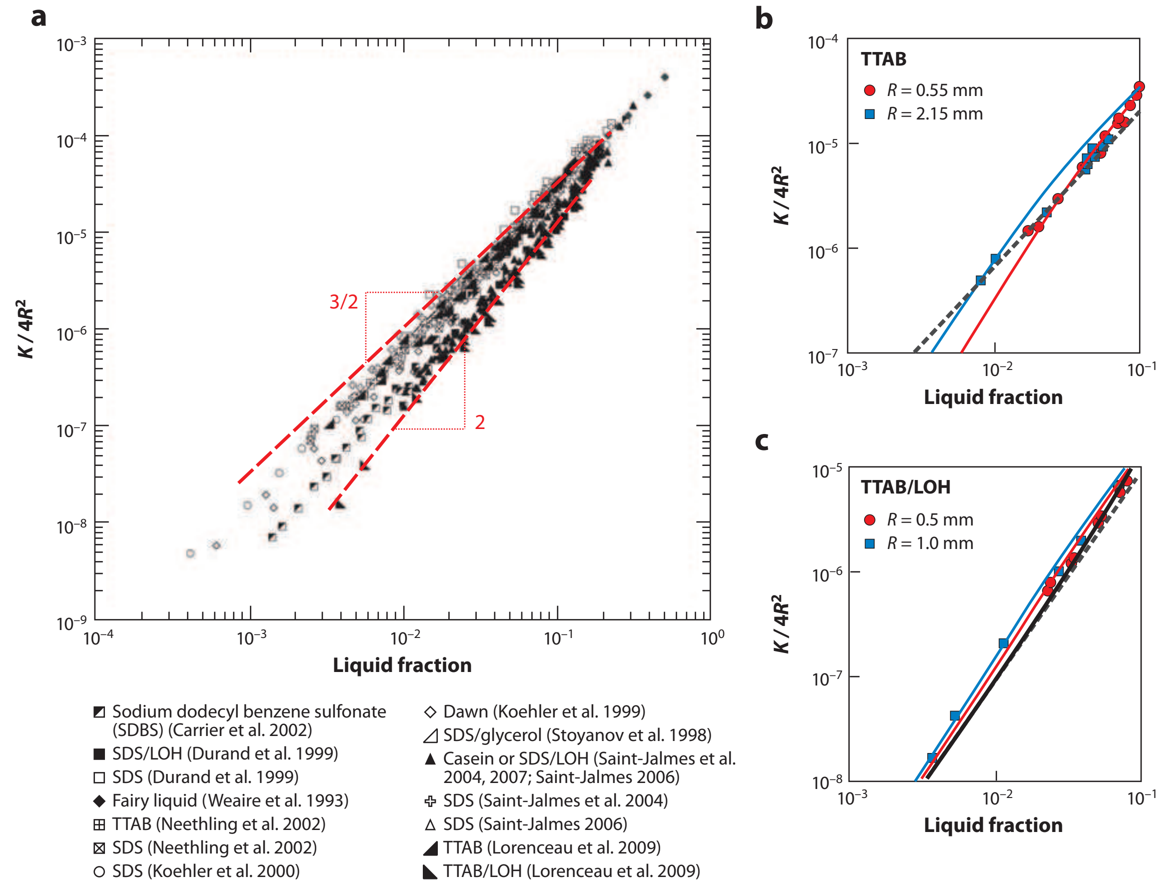


Figure 5

Foam permeability scaled by the square of the bubble diameter as a function of the foam liquid fraction. (*a*) The symbols correspond to experimental data available in the literature for different foaming solutions. The two lines show the slopes of power laws predicted by the node-dominated (3/2) and the channel-dominated (2) models (Section 3.3). (*b*) TTAB solution. Experimental data are shown for (*red circles*) $R = 0.55$ mm and (*blue squares*) $R = 2.15$ mm. The dashed line represents the node-dominated model with $\tilde{R}_n = 2{,}000$, and the red line and the blue line the channel/node model with $\tilde{R}_n = 400$ and $\mu_s = 5 \times 10^{-8}$ kg s^{-1} for $R = 0.55$ mm and $R = 2.15$ mm, respectively. Panel *b* adapted from Lorenceau et al. (2009) and Cantat et al. (2010). (*c*) TTAB/LOH solution. Experimental data for (*red circles*) $R = 0.5$ mm and (*blue squares*) $R = 1.0$ mm. The dashed line represents the channel-dominated model, the black line the Carman-Kozeny model, and the red and the blue line the channel/node model with $\tilde{R}_n = 800$ and $\mu_s = 5 \times 10^{-7}$ kg s^{-1} for $R = 0.5$ mm and $R = 1.0$ mm, respectively. Panel *c* adapted from Lorenceau et al. (2009) and Cantat et al. (2010).

($\phi_l <$ 2%–3%) and agrees with experimental data up to the wet limit (Rouyer et al. 2010). However, this approach cannot be applied to foams with mobile interfaces.

4. FLOWING FOAMS

4.1. Solid-Like or Liquid-Like Mechanical Behavior

Despite being composed only of fluids, aqueous foams exhibit either solid-like or liquid-like mechanical behavior, depending on the applied stress and their liquid content ϕ_l (**Figure 6*a***). Here we consider rheology decoupled from drainage, which means in practice that measurements are

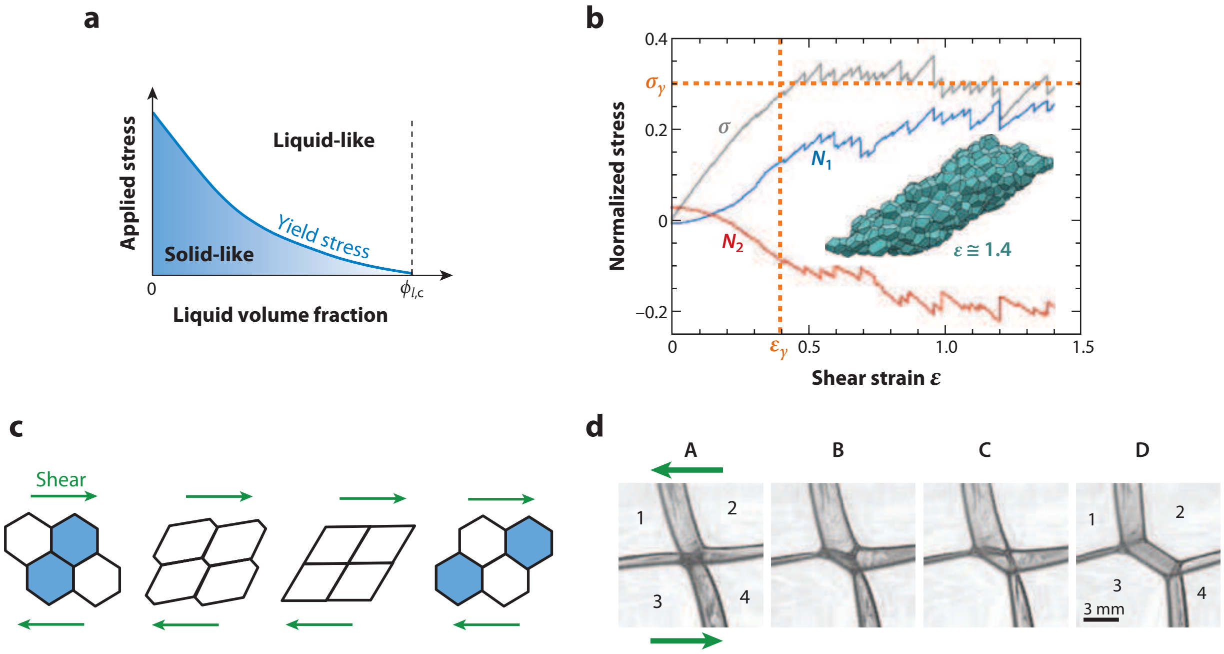


Figure 6

(a) Solid-like or liquid-like mechanical behaviors of foam, depending on the liquid volume fraction and applied shear stress. $\phi_{l,c}$ denotes the liquid volume fraction at the jamming transition. (b) Evolution of the shear stress, the first N_1 and second N_2 normal stress differences with the applied shear strain, for a disordered monodisperse 3D dry foam, simulated under quasi-static conditions using the Surface Evolver software. Stresses are normalized by $\gamma / V^{1/3}$, where V is the bubble volume. For strains below the yield strain ε_y, the response is linear, $\sigma = G_o \varepsilon$ with $G_o \cong 0.5\gamma/R$ (Kraynik & Reinelt 2004). (*Inset*) The foam structure for $\varepsilon \cong 1.4$. Beyond ε_y, the shear stress saturates to the yield stress σ_y, and the foam undergoes a plastic flow. The intermittent jumps of the stresses result from T1 events that would be averaged out in a sample containing a larger number of bubbles. Data taken from Kraynik & Reinelt (2004), courtesy of A. Kraynik. (c) Four structures illustrating how a quasi-statically increasing strain induces a T1 event in a 2D foam (Princen 1983). Some films are lengthened while others are shortened, leading to the formation of a mechanically unstable fourfold junction. At this point, the stress reaches its maximum value. Then bubbles rearrange and relax into a new equilibrium configuration of modified topology, as illustrated by the bubbles highlighted in blue. (d) T1 rearrangement in a dry 3D cluster of four bubbles (*1–4*) induced by a shear strain (in the direction of the *arrows* in the first image) applied so slowly that the duration of the relaxation following the topological switch is independent of the strain rate. Note that the transition from 1 to 4 corresponds to the topological change shown in the last two images of panel *c*. The foaming solution is SLES-CAPB-LOH in water. The time interval between images 1–4 is 1.4 s. Data taken from Biance et al. (2009), courtesy of A.-L. Biance.

performed on a timescale shorter than the drainage time evaluated as $H/u_{front} \approx \mu\,\phi_l\,H/(K\rho g)$, where H is the sample height (Section 3.1). Under a small shear, a foam behaves as a solid. Its elasticity arises from the interfacial energy density ($\approx \gamma/R$) that increases as the bubbles are deformed. The static elastic shear modulus G_o scales as γ/R and vanishes at the jamming transition. For disordered polydisperse foams of mean bubble radius R, G_o is expressed as

$$G_o \cong \frac{\gamma}{R}(1-\phi_l)(\phi_{l,rcp}-\phi_l) \tag{5}$$

with a dimensionless prefactor of order 1 depending on polydispersity (Höhler & Cohen-Addad 2005). The maximum stress that a foam can sustain without flowing defines the static yield stress σ_y, and the corresponding strain is the yield strain ε_y. Both quantities are maximum for a dry foam and tend toward zero at $\phi_{l,c}$ where the foam loses its elasticity. For disordered polydisperse foams, they are given by the empirical laws $\sigma_y = a_1 \frac{\gamma}{R}(\phi_{l,rcp}-\phi_l)^2$ and $\varepsilon_y = a_2(\phi_{l,rcp}-\phi_l)$ with the dimensionless prefactors $a_1 \cong 1 - 2$ and $a_2 \cong 0.2 - 0.5$ (Saint-Jalmes & Durian 1999; Rouyer et al. 2003, 2005; Höhler & Cohen-Addad 2005; Marze et al. 2009).

When a foam yields, plastic flow sets in owing to local bubble rearrangements that relax stress (Kraynik & Reinelt 2004, Cantat & Pitois 2006) as illustrated in **Figure 6*b***. In a dry foam, these topological changes are called T1 events (**Figure 6*c***). At small applied strain rates, bubbles rearrange intermittently and undergo strongly nonaffine motions, whereas for strain rates high enough for rearrangements to overlap in time and space, the flow is more laminar (Gopal & Durian 1999, Sexton et al. 2011). At the macroscopic scale, a foam behaves as a shear-thinning fluid whose relation between shear stress $\sigma(\sigma > \sigma_y)$ and strain rate $\dot{\varepsilon}$ is described by the Herschel-Bulkley phenomenological law:

$$\sigma = \sigma_y + k_c \dot{\varepsilon}^n. \tag{6}$$

The power-law exponent $n < 1$ depends on the interfacial rigidity. The consistency k_c writes dimensionally $k_c \approx \sigma_y \tau^n$, where τ is a characteristic time related to the duration of bubble rearrangements (Gopal & Durian 1999). It allows Equation 6 to be expressed in terms of a dimensionless shear rate $\tau \dot{\varepsilon}$ that measures the ratio of viscous and surface tension effects.

Foam flow behavior depends on the disorder of the structure and on the dynamics at the bubble scale, encompassing the dilatational interfacial rheology that determines the boundary conditions of flow in films and Plateau borders. In addition, slow flows are coupled to aging via coarsening-induced bubble rearrangements. In the following, we review how these processes and couplings come into play in shear start-up, steady flow, creep flow, and viscoelastic relaxations of foams.

4.2. Bubble Rearrangement Dynamics

Bubble rearrangements in which the packing intermittently relaxes into a configuration of lower surface energy can be induced either by the coarsening process or by an applied shear strain. The dynamics of strain-induced rearrangement have been studied in dry two-dimensional (2D) Hele-Shaw foams (Durand & Stone 2006), in soap films held on wire frames (Hutzler et al. 2008), and in dry 3D bubble clusters (Biance et al. 2009). In these cases, the relaxation triggered by quasi-static shear results from an unstable fourfold junction that evolves toward a new equilibrium structure with two threefold junctions (**Figure 6*d***). It is driven by the surface tension pulling on the new film and is hindered by viscous friction due to the bulk liquid viscosity μ or to the dilatational surface viscosity μ_d. Because the film thickness h is generally so small that the Boussinesq number $\mu_d/(\mu h) \gg 1$, surface dissipation dominates. The balance between surface friction and surface tension yields the relaxation time $T \cong \mu_d/\gamma$, with a prefactor of order 1 that is indeed observed experimentally. Typically, we have $\gamma \cong 30$ mN m^{-1} and $\mu_d \cong 10^{-4}$–10^{-1} kg s^{-1}, depending on the rigidity of the interfaces. Thus T can vary by approximately three orders of magnitude from a few milliseconds up to a few seconds (Biance et al. 2009). In this simple model, gradients of surfactant surface concentration are assumed to equilibrate much faster than the relaxation. Otherwise, upon stretching, the new film would have a larger surface tension than its neighbors owing to surface elasticity E_∞, the total capillary force pulling on the border would be reduced, and the relaxation would slow down. Thus the scaling of T is modified by a factor slowly increasing with the dimensionless ratio E_∞/γ (Durand & Stone 2006).

Rearrangement dynamics in coarsening wet foams with rigid interfaces exhibit long durations, typically $T \cong 0.1$–0.4 s, reminiscent of those observed in clusters (Gopal & Durian 1999, Gittings & Durian 2008). In contrast, rearrangements in wet foams with mobile interfaces are much faster: The driving force $\approx \Pi R^2$ is given by the osmotic pressure Π (Equation 1) that pushes the bubbles against each other. As a bubble moves a distance $\approx R$ upon a rearrangement, a volume of liquid is displaced through the interstices of the bubble packing, whose permeability is $K = \alpha R^2$

(see Section 3.3) (Lorenceau et al. 2009). The motion is thus slowed down by a modified viscous Stokes friction $\mu\,(R^2/T)\,(\phi_l/\alpha)$ (Le Merrer et al. 2012). The balance between capillary and viscous forces yields the relaxation time $T \approx \mu\phi_l/(\Pi\alpha)$, diverging at the jamming transition as $\Pi \to 0$. The dimensionless ratio ϕ_l/α is of the order of 10^3–10^4. Experimental results on 3D ordered and disordered foams are in agreement with this scaling with a prefactor ≈0.33. For instance, for foams with mobile interfaces [SLES (sodium lauryl dioxyethylene sulfate)-CAPB (cocoamidopropyl betaine) and glycerol solution], $R = 120\ \mu\text{m}$, $\gamma = 30\ \text{mN m}^{-1}$, $\mu = 1.6\ \text{mPa} \cdot \text{s}$, a duration $T = 20$ ms is measured for $\Pi = 100$ Pa (ϕ_l of the order of 0.1, according to **Figure 1** and Equation 1). These rearrangement dynamics, in which friction is dominated by viscous interstitial flow, resemble those in steadily flowing, concentrated hard-sphere granular suspensions, as demonstrated recently (Cassar et al. 2005, Lespiat et al. 2011, Le Merrer et al. 2012). In contrast, if dissipation is dominated by the flow in the contact facets (of thickness h), the friction force must scale as $\mu(R^2/h)(R/T)$, and the predicted duration becomes $T \propto \mu R^2/(\gamma h)$, as assumed in the bubble model (Durian 1997).

SLES: sodium lauryl dioxyethylene sulfate

CAPB: cocoamidopropyl betaine

To summarize, experiments show that the rearrangement duration varies between a few milliseconds to a few seconds, depending on the bubble size, osmotic pressure, surface tension, liquid viscosity, and surface rigidity. This scaling of local dynamics should be the basis of simulations and mesoscopic models of foam flows.

4.3. Shear Start-Up

Upon quasi-static shear start-up (**Figure 7*a***), the stress overshoots and then settles to a steady-state value. The static yield stress $\sigma_{y\,stat}$ is the maximum stress reached at the onset of bubble rearrangements starting from an initial isotropic equilibrium structure, whereas the dynamic yield stress σ_y is deduced from steady flow in the limit of low strain rates in which the foam structure is intermittently broken down and reformed. In experiments with 3D foams, $\sigma_{y\,stat}$ was found to be roughly 15%–30% higher than σ_y, and this difference was more pronounced for dry foams (**Figure 7*b***) (Khan et al. 1988). Quasi-static simulations predict a stress overshoot for 2D disordered dry foams owing to a reorientation of the films along the principal directions of the strain tensor. However, this tensorial effect is too small to account for the observed overshoot (Raufaste et al. 2010).

Upon stress growth at finite strain rates, experiments with 3D dry foams show that the onset of T1 events is delayed (Rouyer et al. 2003). This effect can be analyzed by considering how each of the three films joining at a Plateau border pulls on it with a net tension that includes dissipative forces. These forces result either from viscous flow in the film (Kraynik & Hansen 1987, Kraynik 1988, Reinelt & Kraynik 1989, Kraynik & Reinelt 1990) or from interfacial resistance, as recently proposed in a 2D model taking into account surface elasticity (Cantat 2011). Shear-induced variations of the film length L create surfactant surface concentration Γ gradients, which in turn lead to variations of the surface tension γ driven by the elasticity E_∞. Gradients of Γ are supposed to relax on a characteristic timescale τ that reflects an adsorption/desorption process or diffusion-limited surfactant exchange with the bulk. This model predicts a stress overshoot that strongly increases with strain rate (**Figure 7*c***) as high strain rates tend to induce large increases of surface tension. The interfacial dissipation delays the onset of T1 events. A strong stress overshoot may induce shear banding as it may lead to an instability in which the system separates into a liquid-like band and a solid-like band coexisting at the same stress (Fielding 2007, Schall & Van Hecke 2010).

4.4. Steady Flow

We now turn to the shear-thinning behavior of foams observed upon steady flow for stresses above the yield stress σ_y (Equation 6). Experiments show that 3D foams and concentrated emulsions

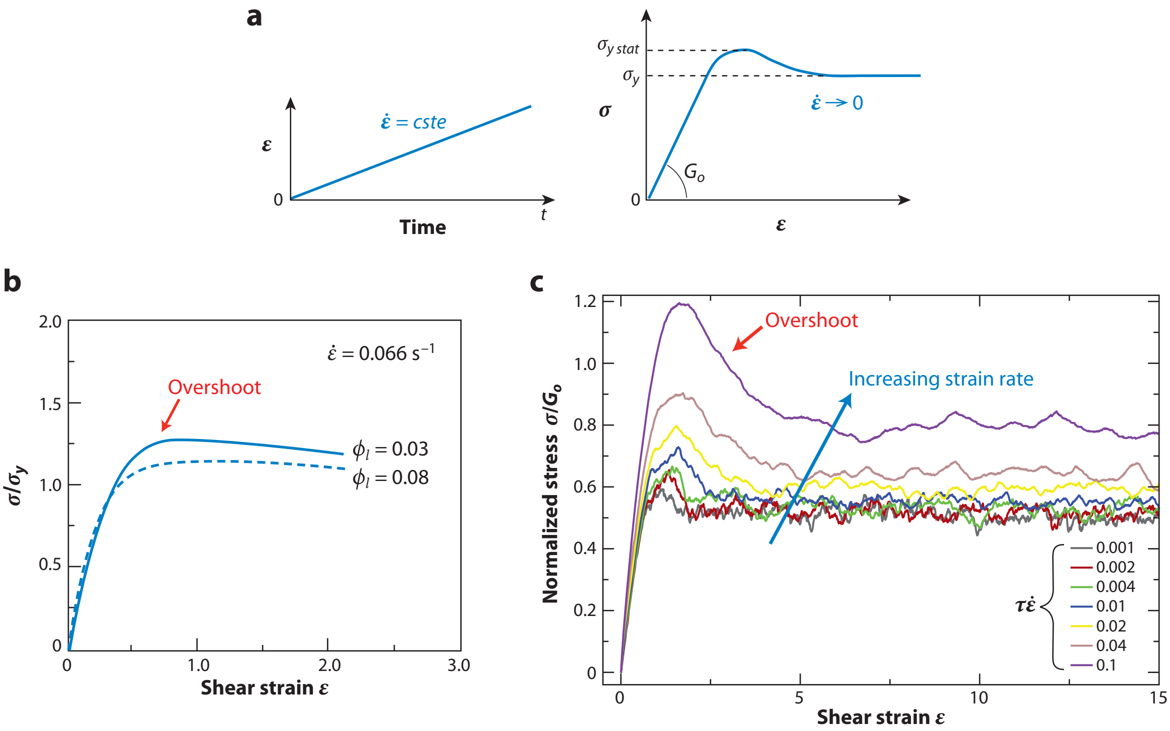


Figure 7

(*a*) Principle of a quasi-static shear start-up experiment. A shear strain $\varepsilon = \dot{\varepsilon}\, t$ with a low shear rate $\dot{\varepsilon}$ is applied to a sample, and the stress σ is measured as a function of ε. Beyond the initially linear elastic response ($\sigma = G_o \varepsilon$), σ grows up to a maximum value $\sigma_{y\,stat}$ and then decreases toward its steady-state limit σ_y. (*b*) Variation of the stress normalized by σ_y upon quasi-static shear start-up for two 3D disordered foams of the same surface tension $\gamma = 23$ mN m^{-1} and average bubble radius $R = 65$ μm. Panel *b* adapted with permission from Khan et al. (1988), copyright 1988, the Society of Rheology. (*c*) Simulated stress-strain variation upon shear start-up for a 2D polydisperse foam with surface elasticity E_∞ for increasing shear rates as indicated. The stress is normalized by the static shear modulus $G_o = 0.52\,\gamma\,\sqrt{\pi/A}$ of a hexagonal foam with the same equilibrium surface tension γ and average bubble area A. τ is the characteristic timescale of surface tension–gradient relaxation. The parameter E_∞/γ is kept constant. Stress overshoots are predicted, even for small strain rates $\tau\,\dot{\varepsilon} \ll 1$. Panel *c* adapted from Cantat (2011). Reproduced with permission of the Royal Society of Chemistry.

behave as nonthixotropic fluids. Moreover, magnetic resonance imaging (MRI) has been used to measure the radial profile of the continuous-phase volume fraction ϕ_l in a wide-gap Couette cell, in which the ratio of gap size to bubble diameter is of the order of 500 or more. In such a heterogeneous stress field, foams and emulsions remain homogeneous without any shear-induced spatial variation of ϕ_l (Ovarlez et al. 2008, 2010).

When sheared in a Couette geometry, any yield-stress fluid exhibits shear localization with a liquid-like zone near the inner rotating cylinder, up to a radius r_c such that $\sigma(r_c) = \sigma_y$ (**Figure 8*a***). The local shear rate can be expressed as $\dot{\varepsilon}(r) = v/r - dv/dr$, where $v(r)$ is the local velocity. For a simple yield-stress fluid, $\dot{\varepsilon}$ is expected to go to zero continuously at $r = r_c$, whereas for a shear-banding material, $\dot{\varepsilon}$ should drop discontinuously from $\dot{\varepsilon}_c$ to zero (Fielding 2007, Dennin 2008). MRI measurements of steady-state velocity profiles reveal that foams and concentrated emulsions behave as simple yield-stress fluids, with no measurable $\dot{\varepsilon}_c$ down to the experimental resolution (typically 0.01–0.3 s^{-1} in foams, as in **Figure 8*b***) (Ovarlez et al. 2008, 2010). By combining MRI profiles and torque measurements for foams and emulsions using the same Couette cell, one can

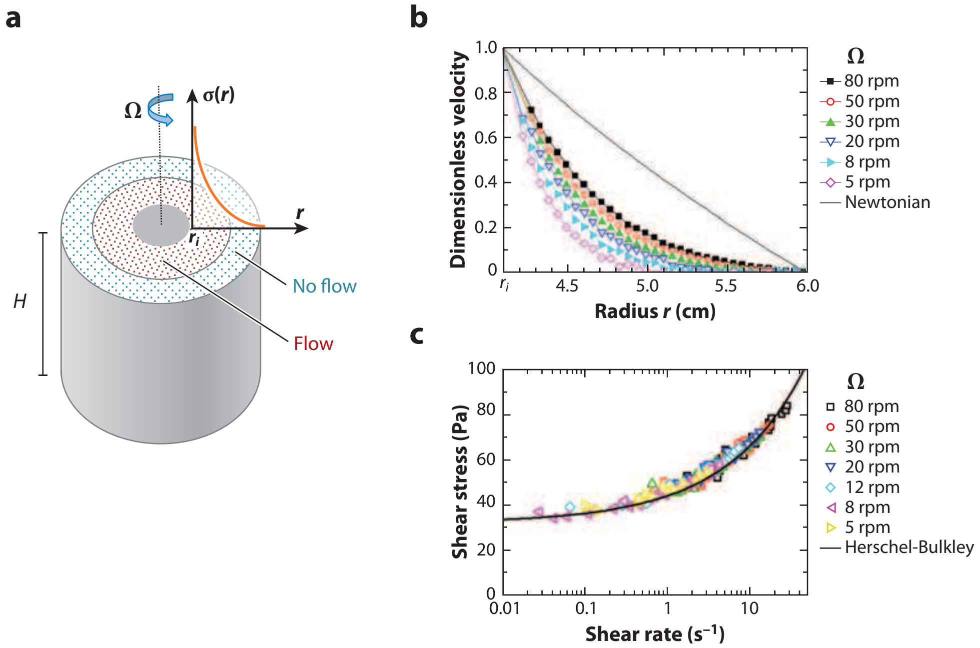


Figure 8

(*a*) Schematic view of a cylindrical Couette cell. The stress varies with the radial position r as $\sigma(r) = \Im/(2\pi H r^2)$, where $\Im$ is the torque exerted by the inner cylinder (radius r_i) rotating with angular velocity Ω. The flow is localized near the rotating cylinder in a region that extends up to a radius r_c, where the stress equals the yield stress. (*b*) Dimensionless velocity profiles $v(r)/v(r_i)$ measured using MRI for a foam steadily sheared in a wide-gap Couette geometry, at different angular velocities Ω. To avoid transient effects, the sample is first sheared at a high Ω so that the entire gap region is flowing. Then Ω is decreased step by step, and the profiles are successively measured. The solid line is the theoretical profile expected for a Newtonian fluid. The foam (SLES CAPB glycerol in water, yielding mobile interfaces) has a liquid fraction $\phi_l = 0.08$, an average bubble radius $R = 22$ μm, and a polydispersity of 55%. Similar velocity profiles are obtained with different bubble sizes, liquid fractions, and interfacial rigidities. Data taken from Ovarlez et al. (2010), adapted with permission by *Europhysics Letters*. (*c*) Local constitutive law measured for the same foam as in panel *b*. Each symbol corresponds to local measurements performed at a given rotational velocity. The solid line represents a Herschel-Bulkley fit (Equation 6) to the data with $\sigma_y = 32$ Pa, $k_c = 12$ Pa $\cdot\, s^n$, and $n = 0.46$. Data taken from Ovarlez et al. (2010), adapted with permission by *Europhysics Letters*.

deduce $\dot{\varepsilon}(r)$ and $\sigma(r)$ for a given radius r, allowing local flow curves to be inferred. They are consistent with a Herschel-Bulkley law (**Figure 8*c***).

We note that long-lived transient shear banding at shear start-up has been observed in emulsions (Becu et al. 2006). In the case of dry 3D foams, which cannot be probed by MRI, shear banding has also been reported at shear start-up, possibly due to the difference between static and dynamic yield stresses (Rouyer et al. 2003). In 2D foams or rafts, shear banding occurs, depending on disorder, polydispersity, and wall friction (Dennin 2008, Wyn et al. 2008, Schall & Van Hecke 2010, Weaire et al. 2010).

We now discuss how shear-thinning behavior can be modeled based on film-level dissipation. In pioneering work, Kraynik & Hansen (1987) considered viscous film traction due to shear-induced stretching. Moreover, a mechanism of liquid withdrawal from or receding into Plateau borders has been predicted, leading to a viscous stress $\sigma_v \sim Ca^{2/3}$, where the capillary number

$Ca = \mu R\dot{\varepsilon}/\gamma$ expresses the ratio of viscous to capillary forces (Schwartz & Princen 1987, Reinelt & Kraynik 1989). As $\dot{\varepsilon}$ increases, entrained films get thicker, the local shear gradients decrease, and so the overall effective foam viscosity diminishes. The macroscopic stress is deduced from the film tensions and orientations, and a Herschel-Bulkley law (Equation 6) is predicted with $n = 2/3$.

A shear-thinning model based on a lubricated sliding motion of bubbles or droplets has been recently proposed for ordered 3D foams and concentrated emulsions (Denkov et al. 2008, Tcholakova et al. 2008). Upon a steady flow, the structure is periodically rearranging, and contact films form and disappear upon transient bubble collisions (**Figure 9*a***). The films thin down under the action of the capillary pressure P_c (Section 2.1), and they are simultaneously sheared as bubbles are entrained by the applied flow. The evolution of the film thickness h is given by the Reynolds equation $\dot{h} = -2P_c h^3/(3\,\mu R_c^2)$, where the contact radius R_c is set by the osmotic pressure (Equation 1). This sets the film thickness $h \propto R\,Ca^{1/2}$, which can be several orders of magnitude larger than the equilibrium one. Assuming that mechanical dissipation arises mainly from the sliding motion with a local shear rate u/h (**Figure 9*a***), the viscous stress exerted on the contact facets is related to the foam viscous stress: $\sigma_v \propto \mu(R\dot{\varepsilon}/h)(R_c/R)^2 \propto (\gamma/R)Ca^{1/2}$. The ϕ_l dependency of the prefactor has been calculated only for moderately wet foams (Tcholakova et al. 2008). By adding σ_v to the yield stress σ_y, a Herschel-Bulkley law (Equation 6) is predicted, $\sigma \propto \sigma_y + (\gamma/R)Ca^{1/2}$. Its exponent $n = 1/2$ is consistent with that found experimentally for concentrated emulsions and for foams with interfaces of low rigidity (Princen & Kiss 1989; Denkov et al. 2005, 2009). Moreover, as illustrated in **Figure 9*b*,*c***, the dependency with bubble size and bulk liquid viscosity is well captured because data obtained for different R and μ all collapse onto a master curve when viscous stress normalized by γ/R is plotted as a function of *Ca*. A dependency of σ_y on $\dot{\varepsilon}$ (Mohan et al. 2011) is not considered in this model.

In contrast to foams with mobile interfaces, shear-thinning exponents $n < 1/2$ are found for foams with rigid interfaces (**Figure 9*d*,*e***). The viscosity effect is no longer captured by the capillary number, as shown by the variation of n with μ. Denkov et al. (2009) established a correlation between the rigidity of the interfaces and the shear-thinning behavior. Interfaces with low rigidity ($|E^*|$ of the order of a few millinewtons per meter) have an exponent $n \cong 1/2$, in contrast to high–surface modulus interfaces ($|E^*|$ of the order of 100–1,000 mN m^{-1}) that have $n < 1/2$. However, a model of the transition criterion between low– and high–surface moduli behaviors is missing. In ordered flowing foam, the interfaces are periodically dilated with a frequency of order $\dot{\varepsilon}$. This leads to the dissipation of mechanical energy in proportion to the surface dilatational loss modulus E'' and an additional contribution to the viscous stress (Denkov et al. 2009). Indeed, recent simulations of disordered 2D foams taking into account surface elasticity predict $\sigma_v \propto \dot{\varepsilon}^{0.6}$, which is similar to the observed power laws (Cantat 2011) (**Figure 9*g***).

The role of disorder for shear thinning has been pointed out by recent simulations (Langlois et al. 2008, Sexton et al. 2011) based on the bubble model (Durian 1995). Bubble interactions are schematically described by a harmonic repulsion and by a linear friction $F_{drag} = bu$, where u is the velocity difference between neighbors. The film thickness h that determines the friction coefficient $b \approx \mu R^2/h$ is assumed to be constant in contrast to the lubrication model above. Despite the Newtonian local friction, a Herschel-Bulkley law with an exponent $n \cong 0.5$ describes the flow macroscopically (Langlois et al. 2008). This behavior is attributed to the disorder of the packing that leads to large velocity fluctuations at low $\dot{\varepsilon}$ (Sexton et al. 2011). Only at high $\dot{\varepsilon}$ do bubbles form lanes in the shear direction, and the flow becomes laminar. Thus for small shear rates, the simple picture of bubbles sliding in ordered lanes, as assumed in a theoretical model of viscous friction (Denkov et al. 2008), may not capture all the physical ingredients.

Using an inclined-plane rheometer (**Figure 9*f***), in which the ratio of shear stress σ to osmotic pressure Π ($\sigma/\Pi = \tan\theta$) is imposed by the angle of inclination θ, Lespiat et al. (2011) have

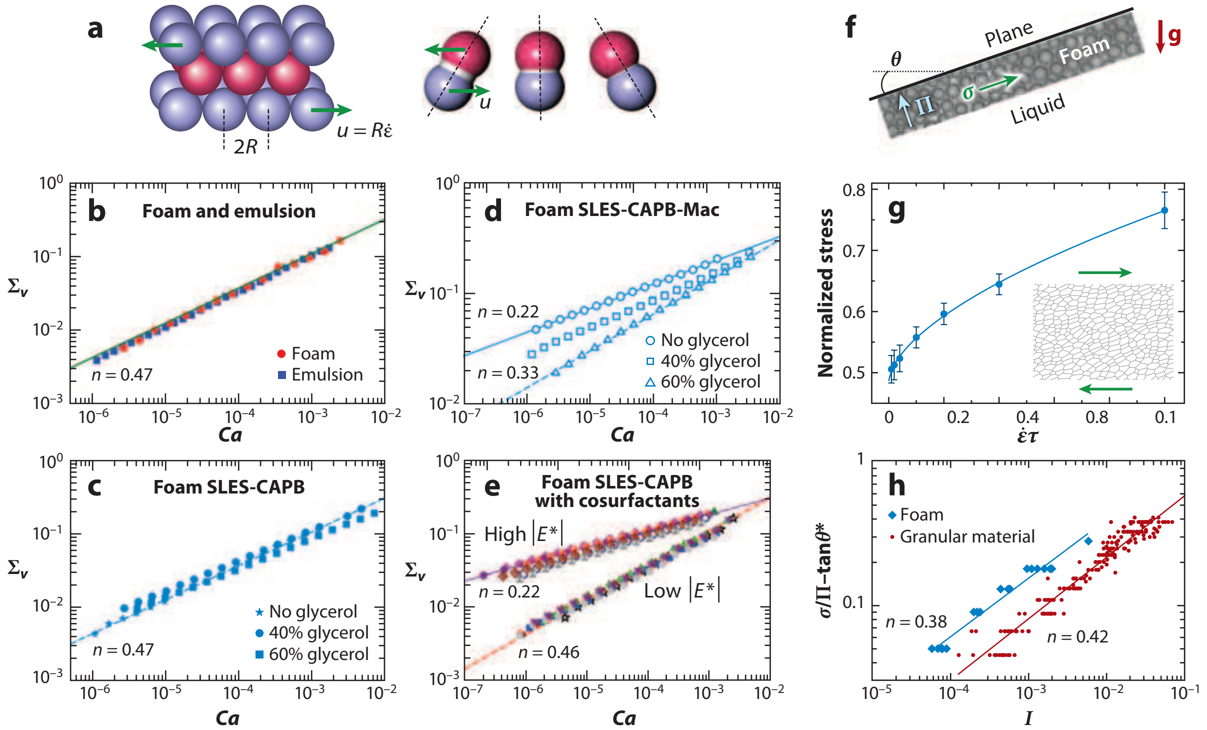


Figure 9

(*a*, *left*) Motion of neighboring bubble layers in a steadily sheared ordered foam. The structure is periodically rearranging as bubble layers are entrained by the applied shear. (*Right*) Film thinning during a bubble collision. u is the velocity difference between two adjacent bubbles in neighboring layers. Panel *a* adapted from Denkov et al. (2008). (*b*) Dimensionless viscous stress $\Sigma_v = \sigma_v R_{32}/\gamma$ versus capillary number $Ca = \mu R_{32}\dot{\varepsilon}/\gamma$, measured for a foam and a concentrated emulsion of the same liquid fraction $\phi_l = 0.10$. The foam is composed of SLES-CAPB surfactants in water (Sauter mean radius $R_{32} = 150$ μm, $\gamma = 30$ mN m^{-1}, and $\mu = 1$ mPa · s; mobile interfaces with $|E^*| \cong 4$ mN m^{-1} measured with the oscillating-drop method at 0.2 Hz, *red circles*). The emulsion consists of hexadecane droplets stabilized in water by a nonionic surfactant ($R_{32} = 4$ μm, $\gamma = 2$ mN m^{-1}, and $\mu = 5$ mPa · s, *blue squares*). The data are fitted by a Herschel-Bulkley law (Equation 6) with $n = 0.47$. Data taken from Denkov et al. (2009). (*c*) Effect of the viscosity of the foaming liquid on the viscous stress of foams with mobile interfaces and liquid fraction $\phi_l = 0.10$. $\Sigma_v = \sigma_v R_{32}/\gamma$ is plotted as a function of Ca, for foams with the same SLES-CAPB surfactants as in panel *b* at increasing glycerol concentration (wt%). The surface tension and the surface dilatational modulus are constant: $\gamma \cong 29$ mN m^{-1} and $|E^*| \cong$ 4–5 mN m^{-1}, measured as in panel *b*. Data taken from Denkov et al. (2009). (*d*) Effect of the viscosity of the foaming liquid on the viscous stress of foams with rigid interfaces and the same liquid fraction $\phi_l = 0.10$. The solution is composed of SLES-CAPB surfactant, glycerol, and a fatty acid [myristic acid (Mac)] that rigidifies the interfaces owing to its low solubility in water: $\gamma \cong 23$ mN m^{-1} and $|E^*| \cong 300 - 350$ mN m^{-1}, measured as in panel *b*. Data taken from Denkov et al. (2009). (*e*) Σ_v versus Ca for SLES-CAPB foams with cosurfactants. Each different cosurfactant is shown with a unique colored symbol. For those significantly increasing $|E^*|$ (such as cetyl trimethyl ammonium chloride, lauric acid, myristic acid, palmitic acid, and lauryl alcohol), $n \cong 0.22$ is found. For cosurfactants that do not rigidify the interfaces, $n \cong 0.46$ is found, as in panel *c*. All the foams have $\phi_l = 0.10$. Data taken from Denkov et al. (2009). (*f*) Schematic drawing of a flowing foam driven by buoyancy along an immersed incline plane. **g** is the gravity acceleration. (*g*) Steady-flow curve predicted by a simulation for a dry disordered 2D foam with surface elasticity. The stress is normalized as in **Figure 7*c***, and the strain rate $\dot{\varepsilon}$ is normalized by the characteristic timescale τ of the relaxation of surface tension gradients. (*Inset*) Foam structure upon shear applied in the direction of the arrows. Data taken from Cantat (2011). Panels *b*–*d*, *e*, and *g* reproduced by permission of the Royal Society of Chemistry. (*h*) Variation of the dimensionless viscous stress $\sigma/\Pi - \tan\theta^*$ with the dimensionless shear rate $I = \mu\dot{\varepsilon}/\Pi$ for steady flow of a foam (*blue diamonds*) and for a granular suspension (*red circles*), both near the jamming point. The foam is made of TTAB in water and has mobile interfaces. The bubble size is $R = 60$ μm with 10% polydispersity. Data taken from Lespiat et al. (2011), with permission by the American Physical Society.

probed the rheology of a polydisperse foam with mobile interfaces near the jamming transition. In this regime, yielding is found to be controlled by the ratio $\sigma_y/\Pi = \tan\theta^* \cong 5°$. For $\theta < \theta^*$, the foam slides as a whole along the plane, whereas for $\theta > \theta^*$, it flows, and a Herschel-Bulkley law is found: $\sigma/\Pi = \tan\theta^* + I^{0.4}$ (Lespiat et al. 2011). For the dimensionless shear rate $I = \mu\dot{\varepsilon}/\Pi$, μ/Π is interpreted as the rearrangement duration (Section 4.2). Moreover, the steady flow of granular suspensions follows a similar friction law as wet foams with mobile interfaces near the jamming transition (**Figure 9*b***), suggesting an analogy between the rearrangement dynamics in both cases.

As foam flows along a smooth wall, it slips (Denkov et al. 2005, 2009; Marze et al. 2008). The viscous stress at the wall σ_w depends on the rigidity of the interfaces that sets the film-level dissipation mechanisms. For mobile interfaces, bubble-wall friction arises from the parietal Plateau borders, and it is described by a Bretherton-type friction law: $\sigma_w \propto Ca_w^{2/3}$. The capillary number $Ca_w = \mu V_o/\gamma$ is defined with respect to the relative velocity V_o between the foam and the wall (Denkov et al. 2005). For rigid interfaces, an additional friction originates from the contact wetting films. As they are sheared, their thickness is set by the balance between a hydrodynamic lift force and the capillary force, and a scaling for the viscous stress $\sigma_w \propto Ca_w^{1/2}$ is observed, which is different from the mobile case (Denkov et al. 2005). These two regimes of friction have been observed experimentally with foams of various constitutions (Denkov et al. 2005, 2009). Wall friction is crucially important for 2D foam flow confined by solid boundaries (Cantat et al. 2004, Cantat & Delannay 2005, Terriac et al. 2006) and has been modeled by the viscous froth model (Kern et al. 2004, Cox 2005, Embley & Grassia 2011).

4.5. Creep Flow and Linear Relaxations

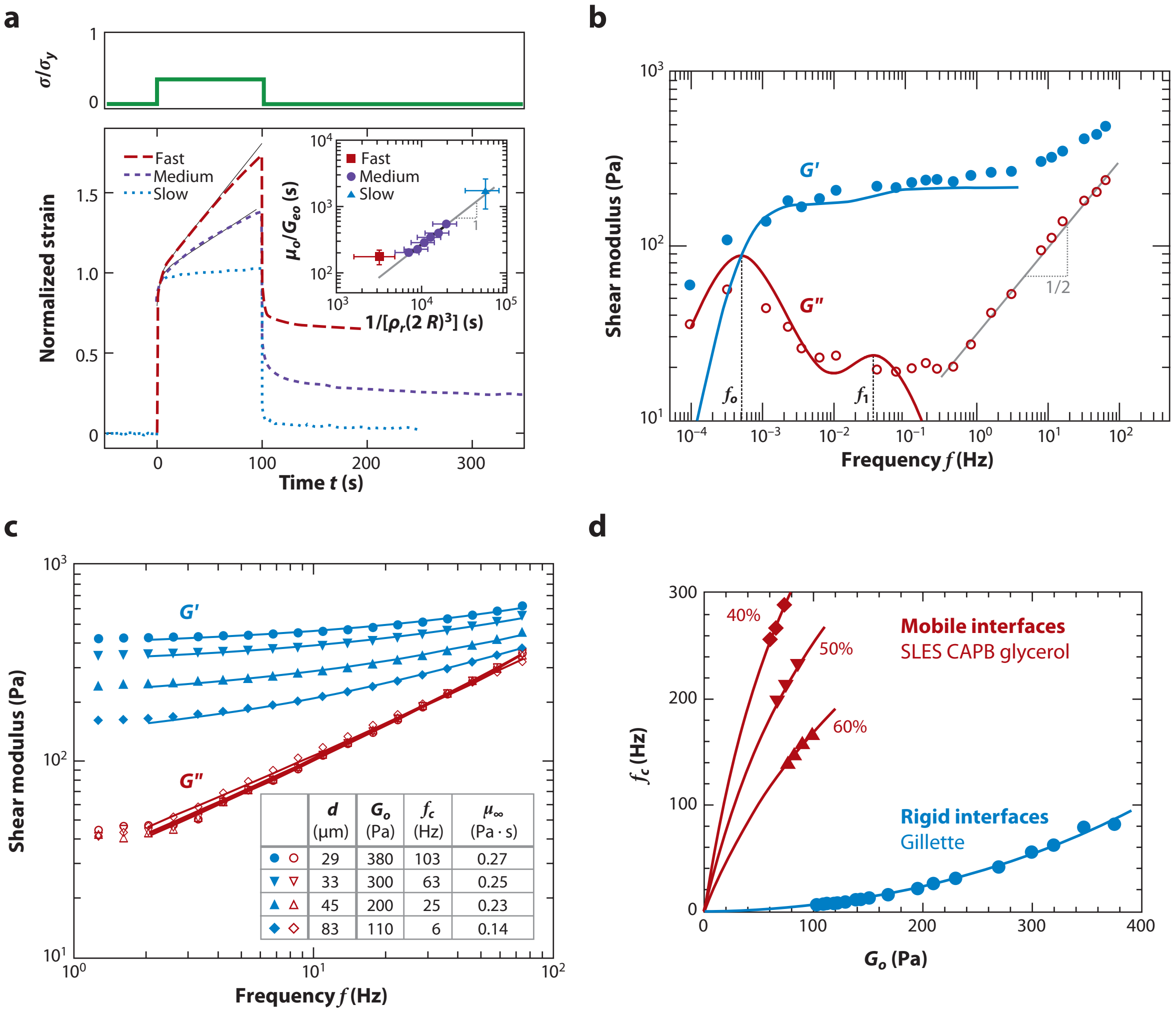
When a coarsening foam is subjected to a small constant shear step stress ($\sigma < \sigma_y$ applied at $t = 0$), it is initially deformed as an elastic material, and after a transient relaxation, it slowly creeps like a highly viscous fluid (**Figure 10*a***). The strain increases with time t as $\varepsilon(t) = \sigma\{1/G_o + t/\mu_o + [1 - \exp(-t/\tau_1)]/G_1\}$, where the parameters G_o, G_1, μ_o, and τ_1 depend on the physicochemical characteristics of the foam (Cohen-Addad et al. 2004, Marze et al. 2005).

Figure 10

(*a*) Creep response of a foam in the solid-like regime. A constant step stress σ ($\sigma < \sigma_y$) is applied during a time interval $\Delta t = 100$ s. The normalized strain response $\varepsilon\, G_{eo}/\sigma$ is plotted as a function of time, for foams with three different coarsening rates. During Δt, the bubble size remains to a good approximation constant, but many rearrangements are induced by the coarsening. (*Inset*) Characteristic time of the mechanical relaxation μ_o/G_{eo}, measured using rheometry, versus the average time interval between coarsening-induced bubble rearrangements in a volume $(2\,R)^3$. This time interval is measured in situ using a multiple coherent light-scattering technique (diffusing-wave spectroscopy). The fitted volume of a rearrangement is $v_r = (6\,R)^3$. The foam surfactants are AOK (sodium α-olefine sulfonate)-LOH surfactant and N_2 gas (fast; *red long dashes*), or a mixture of N_2 and C_6F_{14} vapor (slow; *blue dots*), and Gillette normal regular foam (medium; *short purple dashes*). The liquid fraction is $\phi_l = 0.07$. Data taken from Cohen-Addad et al. (2004). (*b*) Oscillatory response, showing variations of the storage G' and loss G'' shear moduli with frequency. The continuous lines represent the moduli deduced from the creep compliance $J(t)$ using a Laplace transform $G^*(\omega) = (1/i\,\omega)\,\mathrm{L}[J(t)][i\,\omega]$. The variations of G^* at low frequency arise from two relaxation processes of characteristic frequencies, $f_o = G_o/(2\,\pi\,\mu_o)$ and $f_1 = 1/(2\pi\tau_1)$. The straight line of slope 1/2 is a guide for the eyes. The sample is Gillette shaving cream with a liquid volume fraction $\phi_l = 0.08$, bubble size $R = 25$ μm, and surface tension $\gamma = 30$ mN m^{-1}. Data taken from Cohen-Addad et al. (1998, 2004) and Gopal & Durian (2003). (*c*) Fast relaxations in a foam with rigid liquid-gas interfaces, showing variations of G' and G'' with frequency between 1 and 100 Hz for different mean bubble sizes. The sample is Gillette shaving cream with $\phi_l = 0.08$. In contrast to G', the loss modulus G'' is independent of the bubble size. The lines correspond to a fit of $G^*(f) = G_o(1 + \sqrt{if/f_c}) + 2\pi i\mu_\infty f$ for each given average bubble diameter d. Data taken from Krishan et al. (2010), reproduced with permission by American Physical Society. (*d*) Characteristic frequency f_c of fast relaxations versus static shear modulus G_o, for foams with rigid interfaces (Gillette shaving cream) or mobile interfaces (SLES-CAPB) and different glycerol concentrations (in wt%) as labeled. The relaxations slow down as the bulk liquid viscosity increases with the amount of glycerol. The continuous lines are fits to the predictions given in the text. Data taken from Krishan et al. (2010).

For steady creep ($t \gg \tau_1$), the strain increases linearly with time: $\varepsilon(t) \cong \sigma(1/G_{eo} + t/\mu_o)$, as for a Maxwell fluid of shear elastic modulus $G_{eo} = 1/G_o + 1/G_1$ and viscosity μ_o. This has been explained as a consequence of a temporary local loss of elasticity upon coarsening-induced bubble rearrangements. It is described by a mesoscopic homogenization model relating μ_o, G_{eo}, the number of rearrangements ρ_r per unit time and volume, and the effective average volume of the region where the local stress is relaxed upon a rearrangement v_r: $\mu_o = G_{eo}/(\rho_r v_r)$ (Cohen-Addad et al. 2004, Vincent-Bonnieu et al. 2006). In situ measurements of the creep rate and the coarsening-induced rearrangement rate confirm quantitatively the prediction that creep slows down as the coarsening rate decreases (see the inset in **Figure 10*a***). The transient creep relaxation time scales as the ratio of surface dilatational viscosity and surface tension $\tau_1 \approx \mu_d/\gamma$ (Cohen-Addad et al. 2004). For foams with rigid interfaces, the measured values of τ_1 are of the order of 0.1 s, which is in rough agreement with the values of μ_d/γ measured independently (Cohen-Addad et al. 2004).

The creep compliance $\varepsilon(t)/\sigma$ is related to the complex shear modulus $G^* = G' + iG''$ by a Laplace transform, as expected for a linear response (**Figure 10*b***). The peak of G'' is the hallmark of the coarsening-driven relaxation (at $f_o \cong 0.0005$ Hz in **Figure 10*b***). Thus at low frequency, a



	d (μm)	Go (Pa)	fc (Hz)	μ∞ (Pa · s)
● ○	29	380	103	0.27
▼ ▽	33	300	63	0.25
▲ △	45	200	25	0.23
◆ ◇	83	110	6	0.14

foam behaves indeed as a Maxwell liquid, and a scaling law governs the variation of $G^*(f, t_a)$ with coarsening age t_a and frequency f: $G^*(f, t_a) \cong 2\pi i f a(t_a) G_0(t_a)/[1 + 2\pi i f a(t_a) G_0(t_a)]$. The frequency scaling factor is $a(t_a) = \mu_0(t_a)/G_0(t_a)$ (Höhler & Cohen-Addad 2005). As a foam coarsens, the average bubble size increases, and the frequency of rearrangements decreases; therefore, the foam becomes softer [as $G_0(t_a) \sim 1/R$ according to Equation 5] and $a(t_a)$ increases (the peak is shifted toward smaller frequencies according to the mesoscopic model).

In the frequency range $1 \lesssim f \lesssim 100$ Hz, foams exhibit fast relaxations unrelated to the coarsening process, in which the loss modulus G'' increases roughly as $f^{1/2}$ (**Figure 10*c***) (Cohen-Addad et al. 1998, Gopal & Durian 2003, Krishan et al. 2010). Similar variations of the complex moduli have been reported for concentrated emulsions (Liu et al. 1996) and soft pastes (Cloitre 2011). Within this frequency range, for foams with moderate interfacial rigidity, the modulus can be approximately described by $G^*(f) \cong G_0(1 + \sqrt{if/f_c}) + 2\pi i \mu_\infty f$. Similar behavior is predicted for packings of soft spheres in a viscous liquid in which the disorder gives rise to generic collective floppy relaxation modes (Liu et al. 1996, Tighe 2011). μ_∞ is an effective viscosity that becomes dominant at high frequency. In foams, the characteristic relaxation frequency f_c can be set by dissipative processes in the films or in their junctions (Buzza et al. 1995). The scaling of f_c with bubble size $f_c \propto 1/R^2$ found experimentally for foams with rigid interfaces indicates that the dissipation due to shear flow within the liquid films is dominant. For foams with mobile interfaces, a different scaling (omitting prefactors) is observed $f_c \propto 1/(\mu R + \mu_d)$, suggesting that dissipation mainly results from flow at the junctions between the Plateau borders and the films (**Figure 10*d***). In addition, the interfacial viscoelasticity of the gas-liquid interfaces may also contribute directly to the scaling of G^* with frequency.

5. CONCLUDING REMARKS

The permeation of liquid through a foam and the flow of a foam as a whole are both closely linked to the structure of the bubble packing, composed of Plateau borders, nodes, and films. Under static conditions, this structure and its mechanical properties are well understood. The flow behavior is much more complex owing to a coupling between the dynamics of the bulk liquid and the rheological properties of the gas-liquid interfaces. The interfacial behavior can be mobile or rigid, depending on the physicochemistry of the adsorbed surfactant layer. Thus the modeling of permeation in foams and their flow requires multiscale approaches. The coupling with interfacial rheology is specific to foams and emulsions, but the mechanisms of flow on the bubble scale are reminiscent of those encountered in other concentrated dispersions, such as soft pastes and granular suspensions, suggesting fruitful analogies.

FUTURE ISSUES

1. The coupling between flow in channels and in the adjacent films is still not completely understood. How can Marangoni-driven recirculation be modeled quantitatively?
2. Current models of foam permeability are accurate in the dry limit. New approaches are needed to model the behavior close to the wet limit in which channels and nodes are no longer clearly distinguished.
3. Structural disorder and film-level flow coupled to interfacial rheology have been identified as possible origins of mechanical dissipation in foams. How can their interplay be delineated?

4. The coupling between successive bubble rearrangements depends on their spatial and temporal extent. What sets these timescales and length scales? What is their impact on macroscopic flow? How do these dynamics in foams compare with flow dynamics in other soft jammed systems?

DISCLOSURE STATEMENT

The authors are not aware of any biases that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank A.L. Biance, I. Cantat, C. Caroli, S. Costa, P. Grassia, A. Lemaître, M. Le Merrer, K. Okomura, G. Ovarlez, and the CECAM Workshop "Dissipative Rheology of Foams" (Dublin 2012) for stimulating discussions.

LITERATURE CITED

Aveyard R, Binks BP, Clint JH. 2003. Emulsions stabilised solely by colloidal particles. *Adv. Colloid Interface Sci.* 100–102:503–46

Becu L, Manneville S, Colin A. 2006. Yielding and flow in adhesive and nonadhesive concentrated emulsions. *Phys. Rev. Lett.* 96:138302

Biance A-L, Cohen-Addad S, Höhler R. 2009. Topological transition dynamics in a strained bubble cluster. *Soft Matter* 5:4672–79

Binks BP, Murakami R. 2006. Phase inversion of particle-stabilized materials from foams to dry water. *Nat. Mater.* 5:865–69

Britan A, Liverts M, Ben-Dor G, Koehler SA, Bennani N. 2009. The effect of fine particles on the drainage and coarsening of foam. *Colloid Surf. A* 344:15–23

Buzza DMA, Lu C-YD, Cates ME. 1995. Linear shear rheology of incompressible foams. *J. Phys. II* 5:37–52

Cantat I. 2011. Gibbs elasticity effect in foam shear flows: a non quasi-static 2D numerical simulation. *Soft Matter* 7:448–55

Cantat I, Cohen-Addad S, Elias F, Graner F, Höhler R, et al. 2010. *Les Mousses, Structure et Dynamique*. Paris: Belin

Cantat I, Delannay R. 2005. Dissipative flows of 2D foams. *Eur. Phys. J. E* 18:55–67

Cantat I, Kern N, Delannay R. 2004. Dissipation in foam flowing through narrow channels. *Europhys. Lett.* 65:726–32

Cantat I, Pitois O. 2006. Stokes experiment in a liquid foam. *Phys. Fluids* 18:083302

Carn F, Colin A, Pitois O, Vignes-Adler M, Backov R. 2009. Foam drainage in the presence of nanoparticle-surfactant mixtures. *Langmuir* 25:7847–56

Carrier V, Destouesse S, Colin A. 2002. Foam drainage: a film contribution? *Phys. Rev. E* 65:061404

Cassar C, Nicolas M, Pouliquen O. 2005. Submarine granular flows down inclined planes. *Phys. Fluids* 17:103301

Cervantes-Martinez A, Rio E, Delon G, Saint-Jalmes A, Langevin D, Binks BP. 2008. On the origin of the remarkable stability of aqueous foams stabilised by nanoparticles: link with microscopic surface properties. *Soft Matter* 4:1531–35

Cloitre M. 2011. Yielding, flow and slip in microgel suspensions: from microstructure to macroscopic rheology. In *Microgel Suspensions: Fundamentals and Applications*, ed. A Fernandez-Nieves, HWJ Mattsson, DA Weitz, pp. 285–311. New York: Wiley

Cohen-Addad S, Hoballah H, Höhler R. 1998. Viscoelastic response of a coarsening foam. *Phys. Rev. E* 57:6897–901

Cohen-Addad S, Höhler R. 2001. Bubble dynamics relaxation in aqueous foam probed by multispeckle diffusing-wave spectroscopy. *Phys. Rev. Lett.* 86:4700–3

Cohen-Addad S, Höhler R, Khidas Y. 2004. Origin of the slow linear viscoelastic response of aqueous foams. *Phys. Rev. Lett.* 93:028302

Cohen-Addad S, Krzan M, Hohler R, Herzhaft B. 2007. Rigidity percolation in particle-laden foams. *Phys. Rev. Lett.* 99:168001

Cox SJ. 2005. A viscous froth model for dry foams in the Surface Evolver. *Colloid Surf. A* 263:81–89

Cox SJ, Bradley G, Hutzler S, Weaire D. 2001. Vertex corrections in the theory of foam drainage. *J. Phys. Condens. Matter* 13:4863–69

Denkov ND, Subramanian V, Gurovich D, Lips A. 2005. Wall slip and viscous dissipation in sheared foams: effect of surface mobility. *Colloid Surf. A* 263:129–45

Denkov ND, Tcholakova S, Golemanov K, Ananthapadmanabhan KP, Lips A. 2008. Viscous friction in foams and concentrated emulsions under steady shear. *Phys. Rev. Lett.* 100:138301

Denkov ND, Tcholakova S, Golemanov K, Ananthpadmanabhan KP, Lips A. 2009. The role of surfactant type and bubble surface mobility in foam rheology. *Soft Matter* 5:3389–408

Dennin M. 2008. Discontinuous jamming transitions in soft materials: coexistence of flowing and jammed states. *J. Phys. Condens. Matter* 20:283103

Drenckhan W, Ritacco H, Saint-Jalmes A, Saugey A, McGuinness P, et al. 2007. Fluid dynamics of rivulet flow between plates. *Phys. Fluids* 19:102101

Durand M, Langevin D. 2002. Physicochemical approach to the theory of foam drainage. *Eur. Phys. J. E* 7:35–44

Durand M, Martinoty G, Langevin D. 1999. Liquid flow through aqueous foams: from the plateau border-dominated regime to the node-dominated regime. *Phys. Rev. E* 60:R6307–8

Durand M, Stone HA. 2006. Relaxation time of the topological *T*1 process in a two-dimensional foam. *Phys. Rev. Lett.* 97:226101

Durian DJ. 1995. Foam mechanics at the bubble scale. *Phys. Rev. Lett.* 75:4780–83

Durian DJ. 1997. Bubble-scale model of foam mechanics: melting, nonlinear behavior, and avalanches. *Phys. Rev. E* 55:1739–51

Durian DJ, Weitz DA, Pine DJ. 1991. Scaling behavior in shaving cream. *Phys. Rev. A* 44:R7902–5

Edwards DA, Brenner H, Wasan DT. 1991. *Interfacial Transport Processes and Rheology*. Stoneham, MA: Butterworth-Heinemann

Embley B, Grassia P. 2011. Viscous froth simulations with surfactant mass transfer and Marangoni effects: deviations from Plateau's rules. *Colloid Surf. A* 382:8–17

Fielding SM. 2007. Complex dynamics of shear banded flows. *Soft Matter* 3:1262–79

Gittings AS, Durian D. 2008. Statistics of bubble rearrangement dynamics in a coarsening foam. *Phys. Rev. E* 78:066313

Goldfarb I, Kann KB, Schreiber IR. 1988. Liquid flow in foams. *Fluid Dyn.* 23:244–49

Gopal AD, Durian DJ. 1999. Shear-induced "melting" of an aqueous foam. *J. Colloid Interface Sci.* 213:169–78

Gopal AD, Durian DJ. 2003. Relaxing in foam. *Phys. Rev. Lett.* 91:188303

Goyon J, Bertrand F, Pitois O, Ovarlez G. 2010. Shear induced drainage in foamy yield-stress fluids. *Phys. Rev. Lett.* 104:128301

Guignot S, Faure S, Vignes-Adler M, Pitois O. 2010. Liquid and particles retention in foamed suspensions. *Chem. Eng. Sci.* 65:2579–85

Guillermic RM, Salonen A, Emile J, Saint-Jalmes A. 2009. Surfactant foams doped with laponite: unusual behaviors induced by aging and confinement. *Soft Matter* 5:4975–82

Hilgenfeldt S, Koehler SA, Stone HA. 2001. Dynamics of coarsening foams: accelerated and self-limiting drainage. *Phys. Rev. Lett.* 86:4704–7

Höhler R, Cohen-Addad S. 2005. Rheology of liquid foam. *J. Phys. Condens. Matter* 17:R1041–69

Höhler R, Yip Cheung Sang Y, Lorenceau E, Cohen-Addad S. 2008. Osmotic pressure and structures of monodisperse ordered foam. *Langmuir* 24:418–25

Hutzler S, Saadatfar M, van der Net A, Weaire D, Cox SJ. 2008. The dynamics of a topological change in a system of soap films. *Colloids Surf. A* 323:123–31

Hutzler S, Weaire D, Crawford R. 1998. Convective instability in foam drainage. *Europhys. Lett.* 41:461–66

Ivanov IB, Danov KD, Ananthapadmanabhan KP, Lips A. 2005. Interfacial rheology of adsorbed layers with surface reaction: on the origin of the dilatational surface viscosity. *Adv. Colloid Interface Sci.* 114–15:61–92

Kern N, Weaire D, Martin A, Hutzler S, Cox SJ. 2004. Two-dimensional viscous froth model for foam dynamics. *Phys. Rev. E* 70:041411

Khan SA, Schnepper CA, Armstrong RC. 1988. Foam rheology: III. Measurement of shear flow properties. *J. Rheol.* 32:69–92

Koczo K, Lobo LA, Wasan DT. 1992. Effect of oil on foam stability: aqueous foams stabilized by emulsions. *J. Colloid Interface Sci.* 150:492–506

Koczo K, Racz G. 1987. Flow in a Plateau border. *Colloids Surf.* 22:97–110

Koehler S. 2012. Foam drainage. In *Foam Engineering: Fundamentals and Applications*, ed. P Stevenson, pp. 27–58. New York: Wiley

Koehler SA, Hilgenfeldt S, Stone HA. 1999. Liquid flow through aqueous foams: the node-dominated foam drainage equation. *Phys. Rev. Lett.* 82:4232–35

Koehler SA, Hilgenfeldt S, Stone HA. 2000. A generalized view of foam drainage: experiment and theory. *Langmuir* 16:6327–41

Koehler SA, Hilgenfeldt S, Stone HA. 2004a. Foam drainage on the microscale: I. Modeling flow through single Plateau borders. *J. Colloid Interface Sci.* 276:420–38

Koehler SA, Hilgenfeldt S, Weeks ER, Stone HA. 2002. Drainage of single Plateau borders: direct observation of rigid and mobile interfaces. *Phys. Rev. E* 66:040601

Koehler SA, Hilgenfeldt S, Weeks ER, Stone HA. 2004b. Foam drainage on the microscale II. Imaging flow through single Plateau borders. *J. Colloid Interface Sci.* 276:439–49

Koehler SA, Stone HA, Brenner MP, Eggers J. 1998. Dynamics of foam drainage. *Phys. Rev. E* 58:2097–106

Kraynik AM. 1988. Foam flows. *Annu. Rev. Fluid. Mech.* 20:325–57

Kraynik AM, Hansen MG. 1987. Foam rheology: a model of viscous phenomena. *J. Rheol.* 31:175–205

Kraynik AM, Reinelt D. 1990. On the shearing flow of foams and concentrated emulsions. *J. Fluid Mech.* 215:431–55

Kraynik AM, Reinelt DA. 2004. Microrheology of random polydisperse foam. *Proc. 14th Int. Congr. Rheol.* Seoul: Korean Soc. Rheol.

Krishan K, Helal A, Höhler R, Cohen-Addad S. 2010. Fast relaxations in foam. *Phys. Rev. E* 82:011405

Langlois VJ, Hutzler S, Weaire D. 2008. Rheological properties of the soft-disk model of two-dimensional foams. *Phys. Rev. E* 78:021401

Le Merrer M, Cohen-Addad S, Höhler R. 2012. Bubble rearrangement duration in foams near the jamming point. *Phys. Rev. Lett.* 108:188301

Leonard RA, Lemlich R. 1965a. A study of interstitial liquid flow in foam. Part I. Theoretical model and application to foam fractionation. *AIChE J.* 11:18–25

Leonard RA, Lemlich R. 1965b. A study of interstitial liquid flow in foam. Part II: experimental verification and observations. *AIChE J.* 11:25–29

Lespiat R, Cohen-Addad S, Höhler R. 2011. Jamming and flow of random-close-packed spherical bubbles: an analogy with granular materials. *Phys. Rev. Lett.* 106:148302

Levich VG. 1962. *Physicochemical Hydrodynamics*. Englewood Cliffs, NJ: Prentice Hall

Liu AJ, Ramaswamy S, Mason TG, Gang H, Weitz DA. 1996. Anomalous viscous loss in emulsions. *Phys. Rev. Lett.* 76:3017–20

Lorenceau E, Louvet N, Rouyer F, Pitois O. 2009. Permeability of aqueous foams. *Eur. Phys. J. E* 28:293–304

Louvet N, Höhler R, Pitois O. 2010. Capture of particles in soft porous media. *Phys. Rev. E* 82:041405

Lu JR, Thomas K, Penfold J. 2000. Surfactant layers at the air-water interface: structure and composition. *Adv. Colloid Interface Sci.* 84:143–304

Lucassen-Reynders EH. 1981. *Anionic Surfactants.* New York: Marcel Dekker

Marze S, Guillermic RM, Saint-Jalmes A. 2009. Oscillatory rheology of aqueous foams: surfactant, liquid fraction, experimental protocol and aging effects. *Soft Matter* 5:1937–46

Marze S, Langevin D, Saint-Jalmes A. 2008. Aqueous foam slip and shear regimes determined by rheometry and multiple light scattering. *J. Rheol.* 52:1091–111

Marze SPL, Saint-Jalmes A, Langevin D. 2005. Protein and surfactant foams: linear rheology and dilatancy effect. *Colloids Surface A* 263:121–28

Maurdev G, Saint-Jalmes A, Langevin D. 2006. Bubble motion measurements during foam drainage and coarsening. *J. Colloid Interface Sci.* 300:735–43

Miles GD, Shedlovsky GJ, Ross J. 1945. Foam drainage. *J. Phys. Chem.* 49:93–107

Mohan L, Seth JR, Mohan L, Locatelli-Champagne C, Cloitre M, Bonnecaze RT. 2011. A micromechanical model to predict the flow of soft particle glasses. *Nat. Mater.* 10:838–43

Neethling SJ, Lee HT, Cilliers JJ. 2002. A foam drainage equation generalized for all liquid contents. *J. Phys. Condens. Matter* 14:331–42

Nguyen AV. 2002. Liquid drainage in single plateau borders of foam. *J. Colloid Interface Sci.* 249:194–99

Ovarlez G, Krishan K, Cohen-Addad S. 2010. Investigation of shear banding in three-dimensional foams. *Europhys. Lett.* 91:68005

Ovarlez G, Rodts S, Ragouilliaux A, Coussot P, Goyon J, Colin A. 2008. Wide-gap Couette flows of dense emulsions: local concentration measurements, and comparison between macroscopic and local constitutive law measurements through magnetic resonance imaging. *Phys. Rev. E* 78:036307

Pitois O, Fritz C, Vignes-Adler M. 2005a. Hydrodynamic resistance of a single foam channel. *Colloids Surf. A* 261:109–14

Pitois O, Fritz C, Vignes-Adler M. 2005b. Liquid drainage through aqueous foam: study of the flow on the bubble scale. *J. Colloid Interface Sci.* 282:458–65

Pitois O, Lorenceau E, Louvet N, Rouyer F. 2009a. Specific surface area model for foam permeability. *Langmuir* 25:97–100

Pitois O, Louvet N, Lorenceau E, Rouyer F. 2008. Node contribution to the permeability of liquid foams. *J. Colloid Interface Sci.* 322:675–77

Pitois O, Louvet N, Rouyer F. 2009b. Recirculation model for liquid flow in foam channels. *Eur. Phys. J. E* 30:27–35

Princen HM. 1983. Rheology of foams and highly concentrated emulsions: I. Elastic properties and yield stress of a cylindrical model system. *J. Colloid Interface Sci.* 91:160–75

Princen HM. 1986. Osmotic pressure of foams and highly concentrated emulsions. 1. Theoretical considerations. *Langmuir* 2:519–24

Princen HM, Kiss AD. 1987. Osmotic pressure of foams and highly concentrated emulsions. 2. Determination from the variation in volume fraction with height in an equilibrated column. *Langmuir* 3:36–41

Princen HM, Kiss AD. 1989. Rheology of foams and highly concentrated emulsions: IV. An experimental study of the shear viscosity and yield stress of concentrated emulsions. *J. Colloid Interface Sci.* 128:176–87

Raufaste C, Cox SJ, Marmottant P, Graner F. 2010. Discrete rearranging disordered patterns: prediction of elastic and plastic behavior, and application to two-dimensional foams. *Phys. Rev. E* 81:031404

Reinelt DA, Kraynik AM. 1989. Viscous effects in the rheology of foams and concentrated emulsions. *J. Colloid Interface Sci.* 132:491–503

Rouyer F, Cohen-Addad S, Hohler R. 2005. Is the yield stress of aqueous foam a well-defined quantity? *Colloids Surf. A* 263:111–16

Rouyer F, Cohen-Addad S, Vignes-Adler M, Höhler R. 2003. Dynamics of yielding observed in a three-dimensional aqueous dry foam. *Phys. Rev. E* 67:021405

Rouyer F, Lorenceau E, Pitois O. 2008. Film junction effect on foam drainage. *Colloids Surface A* 324:234–36

Rouyer F, Louvet N, Fritz C, Pitois O. 2011. Transport of coarse particles in liquid foams: coupling of confinement and buoyancy effects. *Soft Matter* 7:4812–20

Rouyer F, Pitois O, Lorenceau E, Louvet N. 2010. Permeability of a bubble assembly: from the very dry to the wet limit. *Phys. Fluids* 22:043302

Saffman PG. 1959. A theory of dispersion in a porous medium. *J. Fluid Mech.* 6:321–49

Safouane M, Saint-Jalmes A, Bergeron V, Langevin D. 2006. Viscosity effects in foam drainage: Newtonian and non-Newtonian foaming fluids. *Eur. Phys. J. E* 19:195–202

Sagis LMC. 2011. Dynamic properties of interfaces in soft matter: experiments and theory. *Rev. Mod. Phys.* 83:1367–403

Saint-Jalmes A. 2006. Physical chemistry in foam drainage and coarsening. *Soft Matter* 2:836–49

Saint-Jalmes A, Durian DJ. 1999. Vanishing elasticity for wet foams: equivalence with emulsions and role of polydispersity. *J. Rheol.* 43:1411–22

Saint-Jalmes A, Langevin D. 2002. Time evolution of aqueous foams: drainage and coarsening. *J. Phys. Condens. Matter* 14:9397–412

Saint-Jalmes A, Marze S, Ritacco H, Langevin D, Bail S, et al. 2007. Diffusive liquid propagation in porous and elastic materials: the case of foams under microgravity conditions. *Phys. Rev. Lett.* 98:058303

Saint-Jalmes A, Vera MU, Durian DJ. 2000. Free drainage of aqueous foams: container shape effects on capillarity and vertical gradients. *Europhys. Lett.* 50:695–701

Saint-Jalmes A, Zhang Y, Langevin D. 2004. Quantitative description of foam drainage: transitions with surface mobility. *Eur. Phys. J. E* 15:53–60

Salonen A, Lhermerout R, Rio E, Langevin D, Saint-Jalmes A. 2012. Dual gas and oil dispersions in water: production and stability of *foamulsion*. *Soft Matter* 8:699–706

Schall P, Van Hecke M. 2010. Shear bands in matter with granularity. *Annu. Rev. Fluid Mech.* 42:67–88

Schwartz LW, Princen HM. 1987. A theory of extensional viscosity for flowing foams and concentrated emulsions. *J. Colloid Interface Sci.* 118:201–11

Sexton MB, Mobius ME, Hutzler S. 2011. Bubble dynamics and rheology in sheared two-dimensional foams. *Soft Matter* 7:11252–58

Stevenson P, Li X. 2010. A viscous-inertial model of foam drainage. *Chem. Eng. Res. Des.* 88:928–35

Stoyanov S, Dushkin C, Langevin D, Weaire D, Verbist G. 1998. Effect of the rheology on foam drainage. *Langmuir* 14:4663–65

Tcholakova S, Denkov ND, Golemanov K, Ananthapadmanabhan KP, Lips A. 2008. Theoretical model of viscous friction inside steadily sheared foams and concentrated emulsions. *Phys. Rev. E* 78:011405

Terriac E, Etrillard J, Cantat I. 2006. Viscous force exerted on a foam at a solid boundary: influence of the liquid fraction and of the bubble size. *Europhys. Lett.* 74:909–15

Tighe BP. 2011. A relaxing approach to jamming. *Phys. Rev. Lett.* 107:158303

Turner D, Dlugogorski B, Palmer T. 1999. Factors affecting the stability of foamed concentrated emulsions. *Colloids Surfaces A* 150:171–84

Verbist G, Weaire D. 1994. A soluble model for foam drainage. *Europhys. Lett.* 26:631

Verbist G, Weaire D, Kraynik AM. 1996. The foam drainage equation. *J. Phys. Condens. Matter* 8:3715–31

Vincent-Bonnieu S, Höhler R, Cohen-Addad S. 2006. Slow viscoelastic relaxation and aging in aqueous foam. *Europhys. Lett.* 74:533–39

Weaire D, Barry J, Hutzler S. 2010. The continuum theory of shear localization in two-dimensional foam. *J. Phys. Condens. Matter* 22:193101

Weaire D, Hutzler S. 1999. *The Physics of Foams*. Oxford: Clarendon

Weaire D, Hutzler S, Verbist G, Peters E. 1997. A review of foam drainage. *Adv. Chem. Phys.* 102:315–74

Weaire D, Pittet N, Hutzler S, Pardal D. 1993. Steady-state drainage of an aqueous foam. *Phys. Rev. Lett.* 71:2670–73

Wyn A, Davies T, Cox SJ. 2008. Simulations of two-dimensional foam rheology: localization in linear Couette flow and the interaction of settling discs. *Eur. Phys. J. E* 26:81–89

Annual Review of
Fluid Mechanics

Volume 45, 2013

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Fluid Mechanics* articles may be found at http://fluid.annualreviews.org/errata.shtml

Langmuir **2000,** *16,* 6327−6341

A Generalized View of Foam Drainage: Experiment and Theory

Stephan A. Koehler,* Sascha Hilgenfeldt, and Howard A. Stone

Pierce Hall, Division of Engineering & Applied Sciences, Harvard University, Cambridge, Massachusetts 02138

Received October 5, 1999. In Final Form: February 23, 2000

A new experimental method is presented using fluorescein dye to determine the spatial and temporal variations of the liquid volume fraction in aqueous foams. This method is used for quantitative studies of liquid redistribution (*drainage*) in three types of experiments: forced, free, and pulsed drainage. Characteristic quantities, such as the drainage velocity, show power-law dependences on experimental parameters that are inconsistent with traditional foam drainage models based on Poiseuille-type flow in the liquid-carrying channels (Plateau borders) of the foam. To obtain a theoretical description, the *foam drainage equation* is generalized using an energy argument which accounts for viscous dissipation in both the channels and the nodes (or vertices, which are the junctions of four channels) of the liquid network. Good agreement with results for all three types of drainage experiments is found when using this new model in the limit where the dissipation is dominated by the nodes.

1. Introduction

There are many industrial as well as everyday uses for foams in either liquid or solid form. Aqueous foams are ubiquitous in the kitchen: examples are whipped cream, chocolate mousse (N.B. mousse is the French word for foam), beer foam, and soapy dish water. Polymeric foams have numerous applications such as seat cushions and packaging (e.g. Styrofoam). Recently there has also been interest in foaming metal melts to produce metallic foams, which are useful in many mechanical applications because of their structural stability and ultralight weight. These foams have recently come into use in the aerospace and automotive industries.[1,2]

Many foams are made by introducing bubbles into a liquid which contains a surfactant in order to stabilize the gas/liquid interfaces. The spatial and temporal liquid distribution is governed by the fluid dynamics of the liquid in the foam. Under the influence of gravity liquid will *drain* from the foam and accumulate at the bottom, leaving the top deprived of liquid. Such an inhomogeneity of liquid content may be undesirable in applications such as the aforementioned structural metal foams. It is unclear how the dynamics of fluid flow varies for different kinds of foams: Is there a simple description of fluid transport through foams that is universally applicable to different situations and different foams? Do different foams have different boundary conditions at the gas/liquid interface? In particular, how does the presence of a surfactant affect the boundary conditions? Even the fluid dynamics of aqueous (soap) foams, which are the simplest and easiest system to study, is poorly understood.[3−5] Here we report new results for three types of drainage experiments for aqueous foams, along with a modification to the existing foam drainage model that succeeds in capturing the features of liquid drainage in all of these cases.

Foam drainage is the flow of liquid through the interstitial spaces between the bubbles. The flow is driven by capillarity and (usually) gravity and is resisted by viscous damping.[6−9] The interstitial space, i.e., the liquid volume, can be divided into (i) *films*, which form between two adjacent bubbles and are bounded by almost flat bubble faces, (ii) *channels* (also known as *Plateau borders*), where three films meet, and (iii) *nodes* (also known as *vertices* or *junctions*), where four channels meet (see Figure 1b). On the scale of single bubbles the flow is rather intricate: for example in the nodes flows from different channels merge and subsequently split into flows through other channels. The dynamics is further complicated by the flow along the gas/liquid interface; e.g., are the surfaces rigid or mobile? In spite of this small-scale complexity, on a *macroscopic* scale (i.e. lengths of at least several bubble diameters) simple ordered flow is often observed and it is possible to successfully describe foam drainage using effective-medium models.

There are other processes that can change the macroscopic appearance of a foam, in particular *coarsening*, which refers to the growth of the average bubble size. This can occur either by film rupture (coalescence of adjacent bubbles) or by diffusive coarsening, where large bubbles grow at the expense of smaller bubbles, whose higher capillary pressure makes them lose gas. This latter effect is similar to the process known as Ostwald ripening of metal grains.[10] With a good surfactant, films are very stable and rupture events occur only rarely. Diffusive coarsening will always be important at long enough times; however, in the experiments reported here we have taken countermeasures to avoid coarsening.

---

* To whom correspondence should be addressed. E-mail: skoehler@deas.harvard.edu.

(1) Gibson, L. J.; Ashby, M. F. *Cellular Solids*; Cambridge University Press: Cambridge, U.K., 1997.

(2) Evans, A. G.; Hutchinson, J. W.; Ashby, M. F. *Curr. Opin. Solid State Mater.* **1998**, *3*, 288−303.

(3) Desai, D.; Kumar, R. *Chem. Eng. Sci.* **1982**, *37*, 1361.

(4) Kraynik, A. M. Sandia National Laboratories, Internal Report, 1983; 83-0844.

(5) Koehler, S. A.; Hilgenfeldt, S.; Stone, H. A. *Phys. Rev. Lett.* **1999**, *82*, 4232.

(6) Bikerman, J. J. *Foams*; Springer: New York, 1973.

(7) Leonard, R. A.; Lemlich, R. *AIChE J.* **1965**, *11*, 18.

(8) Weaire, D.; Hutzler, S.; Verbist, G.; Peters, E. *Adv. Chem. Phys.* **1997**, *102*, 315.

(9) Bhakta, A.; Ruckenstein, E. *Adv. Colloid Interface Sci.* **1997**, *70*, 1−124.

(10) Smith, C. S. Grain Shapes and other Metallurgical Applications of Topology. In *Metal Interfaces*; Herring, C., Ed.; American Society for Metals: Cleveland, OH, 1952; p 65.

10.1021/la9913147 CCC: $19.00

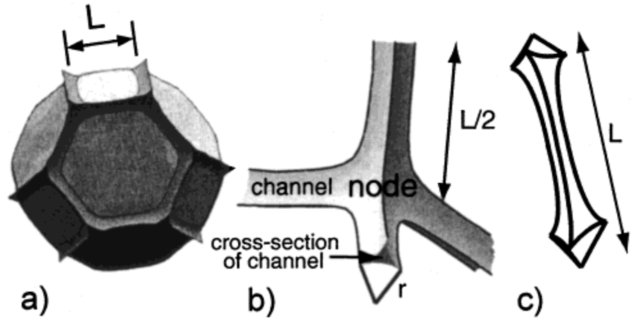


Figure 1. (a) Tetrakaidecahedral bubble with liquid-filled channels and nodes, with edge length L, that composes an idealized Kelvin foam (reproduced from a figure by A. Kraynik, private communication). A liquid network unit composed of one node and four half-channels with corresponding volume fraction $\epsilon = 0.005$ is shown in (b). A "dog-bone"-shaped liquid network unit, one channel with one-quarter-node at each end, is sketched in (c). (a) and (b) have been generated using the Surface Evolver.[27]

In the past, researchers have focused on a variety of aspects of soap foams. The 19th century Belgian scientist J. A. F. Plateau, at a time when he was already blind, originated the study of soap foams, mainly investigating their geometry. He formulated "Plateau's rules", which state that three films must meet at 120° angles and four channels at the tetrahedral angle of $\arccos(-1/3) \approx 109.5°$, and also the eminent minimal surface problem known as the Plateau problem.[11,12] The reader interested in the large body of foam research spanning physical chemistry and engineering applications is referred to the monographs by Mysels et al.[13] and Bikerman[6] and the recent compilation by Prud'homme and Khan[14] as well as the proceedings edited by Sadoc and Rivier.[15]

We are particularly interested in simple dynamical models for foam drainage. Research in this spirit appears to begin with the work of Leonard and Lemlich,[7] who performed foam drainage experiments and developed a model based upon balancing gravity and viscous effects for a Poiseuille-type flow in the channels. Kraynik found analytical solutions to this model for the case of negligible surface tension.[4] The attempt to remove the effects of surface tension recently led Durian and co-workers to perform drainage experiments using a clever geometry of the foam container[16] (earlier publications[17,18] also treat container shape effects on drainage). Reintroducing surface tension into the description of drainage, Goldfarb et al. developed what is now called the *foam drainage equation*.[19] This theory was supported by systematic experiments performed in Weaire's group several years later,[20] who also presented an independent reformulation of the foam drainage equation. However, already Desai and Kumar[3] and Kraynik[4] remarked that the basic assumptions of this standard model lead to slower drainage rates than those observed in their experiments. In recent work, we proposed a modification to the model, based upon the nodes of the liquid network as the principal region of dissipation. This modified model agrees well with data from our forced drainage experiments.[5] Very recently Langevin and co-workers found that changing the surface viscosity of a SDS (sodium dodecyl sulfate) solution by adding dodecanol affects the drainage rates, and for SDS alone their forced drainage results agree with ours.[21] Further evidence that changing the interfacial chemistry significantly alters the drainage rate is provided by Wilde et al.,[22] who report that foams stabilized by proteins drain much slower than foams stabilized by low molecular weight surfactants, such as SDS.

In this paper, we treat the two different models as limiting cases of a generalized theory and demonstrate that the theory which focuses on dissipation in the nodes[5] is in close agreement with data for a variety of drainage experiments. The remainder of this section details the geometry of foams and presents the main ideas of a foam drainage theory, while section 2 describes the experimental setup and methods, along with the main results of the measurements. In sections 3 and 4 a generalized foam drainage equation is derived and discussed and compared to the experimental results. Concluding remarks are given in section 5.

1.1. Foam Geometry. The surfaces of the bubbles of aqueous foams are coated with surfactants which are necessary to stabilize the films against rupture by providing a disjoining pressure (electrostatic and steric in nature) that keeps opposing faces from merging.[23] Outside of the film regions, the dominant force determining the foam geometry is surface tension, which acts to minimize the surface area of the bubbles and results in constant mean curvature.

One simple idealized foam structure is the Kelvin foam,[24] which is a collection of regular tetrakaidecahedral bubbles (cells) with edge length L (see Figure 1a and e.g. Princen[25]). The maximum dimension of a cell is approximately 2.8 L. Typical film thicknesses of the faces of aqueous foams are $\lesssim$ 100 nm,[26] so that most of the liquid resides in the channels and nodes, which are typically between 10 μm and 1 mm wide. Here the contact angle between two channel walls meeting at a face is assumed to be vanishingly small, and the liquid content of the face (and thus the film thickness) is assumed zero. As the amount of liquid in the channels and nodes increases, the mean curvature of the interfaces decreases, and the edges and corners of the polygonal bubbles become more rounded. We quantify the liquid content of a foam by the *liquid volume fraction* ϵ, defined as the volume of liquid in a macroscopic foam region divided by the total volume of the region.

---

(11) Plateau, J. A. F. *Statique expérimentale et théorique des liquides soumis aux seules forces moléculaires*; Gauthier-Villars, Trubner et cie: Paris, 1873.

(12) Isenberg, C. *The Science of Soap Films and Soap Bubbles*; Dover: London, 1992.

(13) Mysels, K. J.; Shinoda, K.; Frankel, S. *Soap Films, studies of their thinning*; Pergamon Press: London, 1959.

(14) Prud'homme, R. K., Khan, S. A., Eds. *Foams, Theory, Measurements and Applications*; Marcel Dekker: New York, 1996.

(15) Sadoc, J. F., Rivier, N., Eds. *Foams and Emulsions*; Kluwer Academic: Dordrecht, The Netherlands, 1997.

(16) Saint-Jalmes, A.; Vera, M. U.; Durian, D. J. *Euro. Phys. J. B* **1999**, *12*, 67−73.

(17) Princen, H. M. *J. Colloid Interface Sci.* **1990**, *134*, 188.

(18) Narsimhan, G.; Ruckenstein, E. Structure, Drainage, and Coalescence of Foams and Concentrated Emulsions. In *Foams, Theory, Measurements and Applications*; Prud'homme, R. K.; Khan, S. A., Eds.; Marcel Dekker: New York, 1996; p 99.

(19) Goldfarb, I. I.; Kann, K. B.; Shreiber, I. R. *Fluid Dyn.* **1988**, *23*, 244.

(20) Verbist, G.; Weaire, D.; Kraynik, A. *J. Phys. Condens. Matter* **1996**, *8*, 3715−3731.

(21) Durand, M.; Martinoty, G.; Langevin, D. *Phys. Rev. E* **1999**, *60*, R6307.

(22) Wilde, P. J.; Mackie, A. R.; Husband, F. A.; Gunning, A. P.; Morris V. J.; Fillery-Travis, A. The Role of Interfacial Structure and Composition on Foam Drainage and Fluid Dynamics. In *Foams and Films: Proceedings of the International Workshop on Foams and Films, Leuven, Belgium*; Weaire, D., Banhart, J., Eds.; MIT-Verlag: Bremen, Germany, 1999; p 59.

(23) Adamson, A. W. *Physical Chemistry of Surfaces*; Wiley: New York, 1990.

(24) Kelvin, L. *Philos. Mag.* **1887**, *24*, 503.

(25) Princen, H. M. *Langmuir* **1986**, *2*, 519.

(26) Wasan, D. T.; Nikolov, A. D.; Lobo, L. A.; Kozco, K.; Edwards, D. A. *Prog. Surf. Sci.* **1992**, *39*, 119.

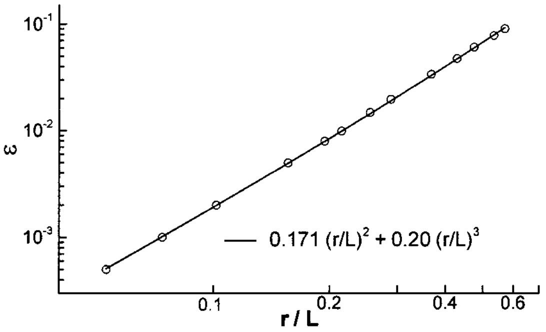


Figure 2. Liquid volume fraction ϵ of the idealized Kelvin foam versus r/L as computed (○) by the Surface Evolver[27] in the absence of gravity. The solid line is the polynomial fit (1) which is in excellent agreement with the numerical data.

To arrive at a geometrical description of the channel network (no faces) we consider the equilibrium situation without gravity, where the mean curvature H is constant. The two principal radii of curvature r_1 and r_2 will change from point to point, obeying $2H = r_1^{-1} + r_2^{-1}$. We define the *characteristic radius of curvature* r by $r^{-1} = 2H$ and plot ϵ as a function of r in Figure 2. The points are computed for the idealized tetrakaidecahedral foam using the Surface Evolver,[27] and the solid line is a fit to the formula

$$\epsilon = \delta_\epsilon \left(\frac{r}{L}\right)^2 + \delta' \left(\frac{r}{L}\right)^3 \qquad \delta_\epsilon \approx 0.171 \qquad \delta' \approx 0.20 \tag{1}$$

Since here we are principally concerned with the dry foam limit, corresponding to $\epsilon \ll 1$, we have $r \ll L$, and the approximation

$$r \approx \delta_\epsilon^{-1/2} L \epsilon^{1/2} \tag{2}$$

is good for liquid volume fractions $\epsilon \lesssim 0.1$.

There is a simple geometrical justification for the polynomial fit (1). Foams with low liquid content ($\epsilon \lesssim 0.05$) have channels that are long, essentially straight, and slender, containing much more liquid than the nodes. With decreasing ϵ the channel length approaches L and the transverse radius of curvature of the channel approaches r from above because the longitudinal radius of curvature becomes large. The cross-sectional area of a channel approaches $\delta_a r^2$, with $\delta_a = \sqrt{3} - \pi/2$, which is the area between three contiguous circles of radius r (see Figure 1b). If one neglects small overlap regions in the junctions, the liquid content of a tetrakaidecahedral bubble is $12\delta_a r^2 L$ because there are 12 complete channels per bubble. The volume of the tetrakaidecahedron[1] is $2^{7/2}L^3$, so the liquid volume fraction due to the channels is $12\delta_a 2^{-7/2}(r/L)^2 \approx 0.171(r/L)^2$, which is the first term on the right-hand side of (1). Higher-order corrections arise from the presence of nodes which (i) diminish the length of the channels by $O(r)$ and (ii) introduce a node volume $\nu_n = \delta_n r^3$, with $\delta_n = O(1)$.[28] This leads to the δ' term in (1), the only genuine fit parameter in the equation.[29] For foams with low liquid content, however, the channels account for almost all of the liquid, and we will neglect node corrections to the volume.

1.2. Foams as Porous Media.

It is useful to consider an analogy between the drainage of a foam and the flow of liquid through a porous medium such as sand or a packed bed of rigid spheres.[30,31] The dimensions and number density of the connected pores determine the permeability of the porous medium to liquid flow. Darcy's law relates the driving pressure gradient $\mathbf{G}$ to the permeability k and the average liquid velocity through the medium $\mathbf{v}$,

$$\mathbf{G} = -\nabla p + \rho \mathbf{g} = \mu \mathbf{v}/k \tag{3}$$

where μ is the viscosity, p is the liquid pressure, and $\rho\mathbf{g}$ is the gravitational force. Permeability has dimensions of length squared and scales with the square of a characteristic pore size. For foams the interstitial space between the bubbles (channels and nodes) plays the role of the pores through which liquid can pass. However the interstitial space (i.e. volume fraction) of a foam is not fixed and bubble deformation accompanies liquid flow. Thus unlike conventional porous media, foams have a permeability that is dynamically coupled to the liquid flux.

Note that eq 3 is *macroscopic* in the sense that all quantities (e.g. the gradient) are averaged over length scales larger than a pore size. Unless explicitly denoted "microscopic", all quantities will be understood as macroscopic, describing foam drainage on an effective-medium level.

1.3. The Generalized Foam Drainage Equation.

We seek to describe the spatial and temporal dynamics of the liquid volume fraction, $\epsilon(\mathbf{x}, t)$. Inertial terms can be neglected, as we estimate that the Reynolds numbers of liquid flow, based upon a typical channel radius, are never larger than 10 and in most cases are much smaller than 1. The discussion below is a modified presentation of Koehler, Hilgenfeldt, and Stone[32] and has many features in common with previous models.[19,33]

A suitable starting point is the equation of mass conservation,

$$\frac{\partial \epsilon}{\partial t} + \nabla \cdot (\epsilon \mathbf{v}) = 0 \tag{4}$$

in which $\mathbf{v}$ (see (3)) is the macroscopic velocity of the flow through the foam. On the left-hand side of (3), the liquid pressure p is given by the Young–Laplace equation

$$p = p_{\mathrm{gas}} - \gamma/r \tag{5}$$

where γ is the surface tension. The foam is assumed monodisperse, so all of the bubbles have the same gas pressure p_{gas}, as well as the same volume, because the weight of the foam is small and compression due to gravity is negligible.[8]

With (2) and (5) the driving force $\mathbf{G}$ from (3) becomes

$$\mathbf{G} = \rho \mathbf{g} + \nabla(\gamma/r) \approx \rho \mathbf{g} + \frac{\gamma \delta_\epsilon^{1/2}}{L} \nabla \epsilon^{-1/2} \tag{6}$$

We use (3) and (6) to solve for $\mathbf{v}$ in terms of ϵ and the permeability $k(\epsilon)$ and substitute this into (4) to arrive at the *generalized foam drainage equation*

---

[27] Brakke, K. Univ. of Minnesota Geometry Center, Surface Evolver Version 2.10c; http://www.geom.umn.edu.

[28] The prefactor δ_n for the node volume can be estimated assuming that the geometrical shape of the nodes resembles that of the interstitial space between hexagonally close-packed spheres (which is the structure of a wet foam with spherical bubbles). The packing fraction of the fcc structure is $\sqrt{2}\pi/6 \approx 0.74$, and each sphere has ${}^{14}/_{4}$ complete nodes. This yields $\delta_n \approx 0.31$.

[29] Phelan, R.; Weaire, D.; Peters, E.; Verbist, G. J. Phys. Condens. Matter **1996**, *8*, L475.

[30] Zick, A. A.; Homsy, G. M. *J. Fluid Mech.* **1982**, *115*, 13.

[31] Larson, R. E.; Higdon, J. J. L. *Phys. Fluids A* **1989**, *1*, 38.

[32] Koehler, S. A.; Hilgenfeldt, S.; Stone, H. A. Foam Drainage: Experiments and a New Model. In *Foams and Films: Proceedings of the International Workshop on Foams and Films, Leuven, Belgium*; Weaire, D., Banhart, J., Eds.; MIT-Verlag: Bremen, Germany, 1999; p 49.

[33] Weaire, D.; Phelan, R. *J. Phys. Condens. Matter* **1996**, *8*, 9519.

$$\mu\frac{\partial\epsilon}{\partial t} + \rho\mathbf{g}\cdot\nabla(k(\epsilon)\epsilon) - \frac{\gamma\delta_{\epsilon}^{1/2}}{L}\nabla\cdot(k(\epsilon)\nabla\epsilon^{1/2}) = 0 \qquad (7)$$

It remains to compute the foam permeability in terms of ϵ. A detailed derivation of a general expression for $k(\epsilon)$ will be presented in section 3. Here, we give a simple argument for the functional form of $k(\epsilon)$ in the two limiting models of foam drainage.

The original foam drainage model is based upon rigid gas/liquid interfaces, which cause Poiseuille-like flow in the channels.[19,20] The viscous damping force (per unit volume) in the channels then is $O(\mu v' r^{-2})$, where v' is a characteristic liquid velocity which is on the order of the macroscopic liquid velocity. For dry foams the node volume is negligible (cf. (1)), and so the viscous damping of the liquid flow is dominated by the channels. Since the Reynolds number is small, the viscous damping equals the driving force from Darcy's law (cf. (6)), and it follows from $\mu v/k(\epsilon) \propto \mu v' r^{-2}$ that $k(\epsilon) \propto r^2 \propto L^2\epsilon$. To differentiate from other models, we call this the *channel-dominated model*.

An alternative boundary condition stipulates that the gas/liquid interface is not rigid but rather freely slipping and stress-free.[5] Although the flow in the nodes may be very complicated due to the mixing and bending of the flow (see Figure 1b), the viscous damping force (per unit volume) inside the nodes is $O(\mu v' r^{-2})$, because r is the only length scale of the nodes. Inside the channels however the viscous damping is negligible, because there the flow is pluglike and the velocity gradients are small. Hence the nodes dominate the dissipation, and we shall call this the *node-dominated model*. Thus the *average* damping force over a network structure is given by $O(\mu v' r^{-2}(r/L))$, because the ratio of the node volume to the volume of the network is $O(r/L)$. Equating the net damping force to the term $\mu v/k(\epsilon)$ of Darcy's law (3) yields $k(\epsilon) \propto rL \propto L^2\epsilon^{1/2}$.

2. Foam Drainage Experiments

In the following we briefly describe the experimental techniques we use to measure the liquid volume fraction of draining foams, which differ from previous work[7,34,35] in several respects. We then review the three different types of foam drainage experiments that we have studied in detail: *forced*, *free*, and *pulsed drainage*.[20] In this section, we present the main results of the measurements. An extensive analysis of the drainage dynamics is deferred to section 4.

2.1. Experimental Procedure. To study foam drainage experimentally we have developed a new measurement scheme and a new procedure to produce a uniform, nearly monodisperse foam that does not show signs of coarsening.[5] The foam is inside a vertical Tygon tube that is 2 m long and has a diameter of 1.3 cm (see Figure 3). The bottom of this *foaming tube* is inserted into a vessel containing the soap solution, and a slow flow of C_2F_6 gas through a blunt-end syringe needle produces a constant supply of monodisperse bubbles, filling the tube at a rate of 0.025 cm/s. At the top the foam flows out of the Tygon tube into a catch basin about 1 m above the experimental region of interest.

The foaming solution is composed of single-distilled water, SDS surfactant (sodium dodecyl sulfate), and fluorescein salt in the ratio of 1 to 5×10^{-3} to 2.5×10^{-4} by weight. The corresponding molarity of SDS is 0.017 M, which is well above the critical micelle concentration (CMC) of this surfactant.[36] The results were unchanged when the soap concentration was either doubled or halved. SDS is a surfactant whose properties are well-known and which is readily available in purified form. The findings reported in the present work support earlier experiments with Dawn dish detergent, tap water, and air[5] and show that the drainage behavior seems to be robust toward these changes. We also conducted some experiments using the surfactant AOS (α-olefinsulfonate), with essentially unchanged results.

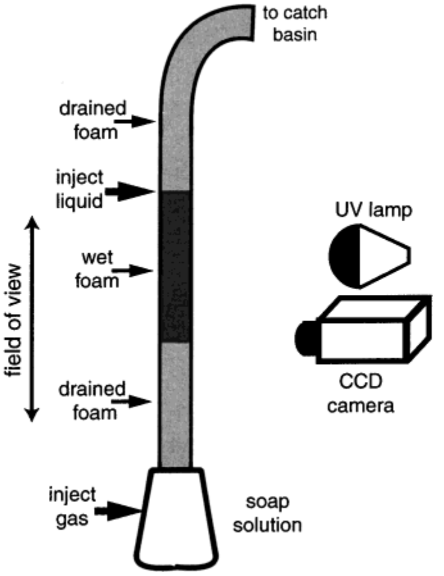


Figure 3. Schematic of the experimental setup, here showing forced drainage. Liquid is injected into the foaming tube at a constant rate above the field of view of the CCD camera. This produces a growing region of wet foam in the otherwise well-drained foam. The CCD camera records the fluorescence of the foam illuminated by UV light from which the liquid volume fraction in the foam can be deduced.

As a result of the slow bubbling, the foam inside the foaming tube is well drained and very *dry* ($\epsilon < 10^{-4}$). We chose C_2F_6 for bubbling, because this gas has a low diffusion coefficient and a low saturation level in water which minimizes coarsening. The bubble size was determined from a close-up photograph of the foam, and the average edge length is $L = 0.15$ cm with a standard deviation of 0.04 cm. The foam does not age or coarsen with time, when observed at a fixed height in the tube, because it is continually being replenished from the bottom. The edge length L does not vary noticeably over the region of observation, which is 76 cm high and starts 40 cm above the foaming solution.

The actual foam drainage experiments record the dynamics of excess liquid added to the well-drained foam. This excess liquid is the same soap solution described above and is injected into the foam through a syringe needle inside the foaming tube. A programmable syringe pump was used for the injection process, and the amount of liquid actually dispensed was independently checked with digital calipers.

2.2. Fluorescence Measurements. The experimental measurements use UV light to excite fluorescence in the liquid part of the foam. The small amount of fluorescein salt dissolved in the foaming solution absorbs the illuminating UV radiation and emits visible (yellow-green) light. The fluorescence intensity, recorded by a CCD digital camera, with a green light filter mounted, is thus a measure of liquid content. Consider a UV light ray that passes through the foam, which has a total path length H in the foam and intensity $I_{\mathrm{UV}}(h)$, where $h \le H$ is the distance traversed. The attenuated intensity is $I_{\mathrm{UV}}(h) = \exp(-\alpha_{\mathrm{liq}}\epsilon h)$, where α_{liq} is the UV absorption coefficient (inverse of absorption length) of the fluorescein solution. We assume that the emission of visible light from the fluorescein occurs isotropically and that for small ϵ the visible light is not reabsorbed noticeably, making the total detected fluorescence intensity

$$I \propto 1 - \exp(-\alpha_{\mathrm{liq}}\epsilon H) \qquad (8)$$

so that a linear relation $I \propto \epsilon$ between detected intensity and

---

(34) Weaire, D.; Pittet, N.; Hutzler, S.; Pardal, D. *Phys. Rev. Lett.* **1993**, *71*, 2670.

(35) Weaire, D.; Findlay, S.; Verbist, G. *J. Phys. Condens. Matter* **1995**, *7*, L217.

(36) Prud'homme, R. K.; Warr, G. G. Foams in Mineral Flotation and Separation Processes. In *Foams, Theory, Measurements and Applications*; Prud'homme, R. K., Khan, S. A., Eds.; Marcel Dekker: New York, 1996; p 511.

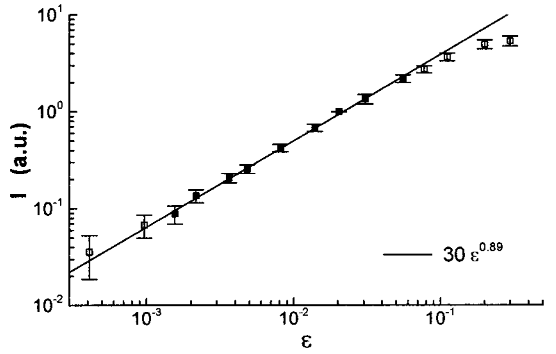


Figure 4. Forced drainage fluorescence intensity plotted against main body liquid volume fraction, determined from eq 10. The line shows the best fit to power-law behavior; see eq 9. The open squares were not included in the power-law fit.

volume fraction is valid if $\alpha_{\mathrm{liq}}\epsilon H \ll 1$, and the intensity saturates ($I \rightarrow I_{\max}$) for $\alpha_{\mathrm{liq}}\epsilon H \gg 1$.

An independent measurement of liquid volume fraction is possible when performing a forced drainage experiment, as described in section 2.3. Figure 4 shows that the relationship between I and ϵ for $\epsilon \lesssim 0.05$ is close to linear in an extended range of liquid volume fractions. A small deviation is, however, evidenced by the best-fit power law

$$I \propto \epsilon^{0.89} \tag{9}$$

which we attribute to a slight nonlinearity of the camera response. From the range of approximate linearity in Figure 4 and using eq 8, we estimate $\alpha_{\mathrm{liq}}^{-1} \approx 0.5$ mm, which is comparable to the liquid path lengths ϵH in our experiments, where H is the diameter of the foaming tube.

Data points at large $\epsilon \gtrsim 0.1$ were not used for the fit because of the saturation of the fluorescence. Also, two data points at extremely small $\epsilon \lesssim 10^{-3}$ were excluded as these very faint signals cannot be detected with reasonable accuracy.

2.3. Forced Drainage. In the forced drainage experiment, a constant flux of the soap solution is added at the top of the foaming tube (about 20 cm above the field of view of the camera), and the liquid invades the completely drained foam from above. The measured volume fraction profiles at five successive times in Figure 5a show a *drainage wave* making its way down the foam tube. It is observed that all forced drainage profiles have a distinct front that does not change in shape with time and moves downward with a constant velocity v_{f}, thus forming a solitary wave. Because the foam is generated at the bottom at a constant rate, it is necessary to add the (very slow) speed at which the foam rises to the measured downward front velocity of the drainage wave.

The forced drainage wave profile consists of three regions (see Figure 5a): the drained region below the traveling wave, which has very little liquid volume content ($\epsilon < 10^{-4}$), the transition (front) region, which has a characteristic width (typically several bubble diameters), and finally the *main body*, which has uniform volume fraction $\epsilon = \epsilon_{\mathrm{main}}$. We choose to measure the front width as the vertical distance along the profile where ϵ drops from 80% to 20% of its main body value. In the experiments, both the front speed v_{f} and width w_{f} are measured as a function of the injected one-dimensional liquid volume flux Φ. Because of conservation of liquid volume, the flux through the foam is simply the product of the average macroscopic liquid velocity v and the main body liquid volume fraction ϵ_{main}; thus

$$\epsilon_{\mathrm{main}} = \Phi/v \tag{10}$$

The liquid velocity of the forced drainage wave is the front velocity, $v = v_{\mathrm{f}}$, because the profile maintains its shape—see Figure 5a. If the front traveled faster than the average liquid velocity, it would broaden, and if it were slower than the average liquid velocity, a steepening of the front would be observed.

Relation (10) allows for an independent calibration of fluorescence intensity measurements as mentioned in section 2.2 and displayed in Figure 4, as the volume flux Φ is a controlled quantity and $v = v_{\mathrm{f}}$ is directly measured without the need for an intensity calibration.

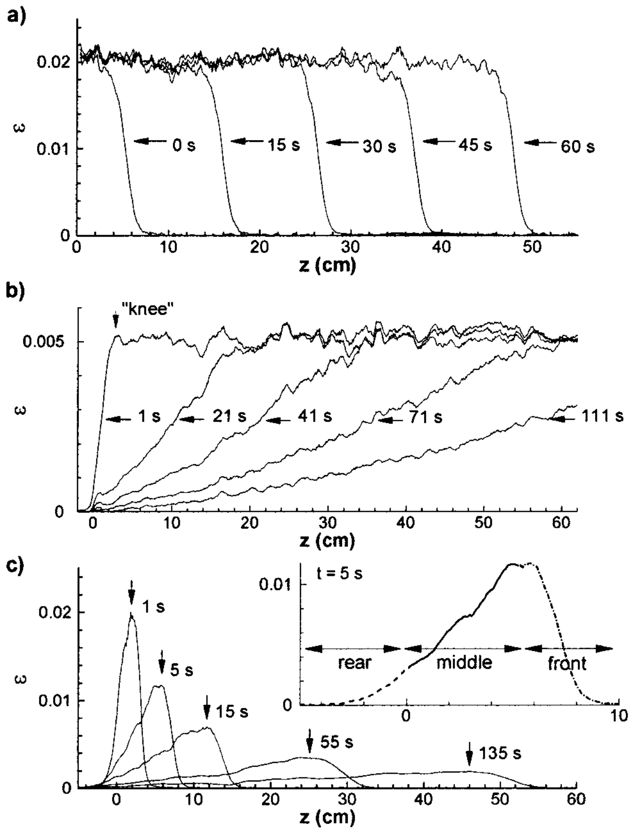


Figure 5. Profiles showing the dynamics of three foam drainage experiments: (a) *forced drainage*, (b) *free drainage*, and (c) *pulsed drainage*. Each profile is labeled by the time (in seconds) when it was taken, and the $+z$-axis is pointing in the direction of gravity. Horizontal arrows in (a) and (b) show the location of the half-maximum of the profiles. The vertical arrow in (b) shows the location of the "knee" of the pulse at $t = 1$ s. In (c) the arrows show the location of the pulse maximum for five successive profiles. All profiles have been averaged over approximately 1 cm (a few bubble diameters) and for pulsed drainage at longer times over several centimeters. The inset of (c) shows the three regions of a pulse to be discussed in section 2.5 (here for $t = 5$ s): (i) the rear, which is above the injection point (dashed line); (ii) the middle region extending from the injection point to the pulse maximum (solid line); (iii) the front, which is the region below the pulse maximum (dashed and dotted line). Note that relation (9) was used to determine ϵ.

Figure 6 shows the front velocity (measured at half-maximum) versus ϵ_{main} for 15 forced drainage experiments with the flux varying over more than four decades from $\Phi = 3.0 \times 10^{-5}$ cm/s to $\Phi = 1.1$ cm/s. The error bars are determined by the standard deviations in the measured front velocity as it moves down the camera's field of view, and the relative error rapidly decreases with increasing volume fraction. The data points demonstrate a power-law dependence of v_{f} on ϵ_{main},

$$v_{\mathrm{f}} \approx c_v \epsilon_{\mathrm{main}}^{d_v} \qquad c_v \approx 7.94 \text{ cm/s} \qquad d_v \approx 0.60 \tag{11}$$

which is shown by the solid line.[37] The dashed line shows the expected dependence of v_{f} on ϵ using the traditional channel-dominated foam drainage equation based upon a rigid gas/liquid interface. Clearly, this model does not describe the data well.

Figure 7 is a log–log plot of the measured front width of the forced drainage profile against ϵ_{main}, which shows a decrease in

---

(37) Note that plotting v against Φ gives a power-law fit $v_{\mathrm{f}}/(\mathrm{cm/s}) = 3.66\ (\Phi/(\mathrm{cm/s}))^{0.37}$. This is in close agreement with experiments using Dawn detergent.[5]

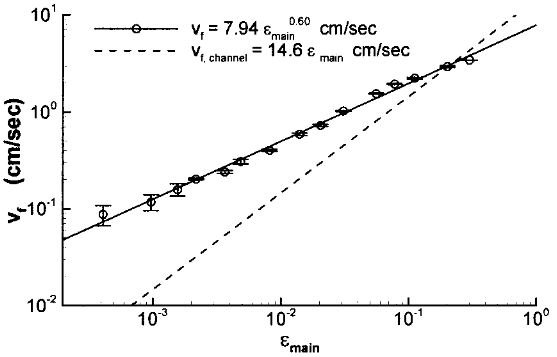


Figure 6. Front velocity v_f, plotted against main body liquid volume fraction ϵ_{main} of the forced drainage wave, determined from eq 10. The solid line is the best fit to power-law dependence (11), and the dashed line is the channel-dominated forced drainage prediction (rigid wall boundary condition).

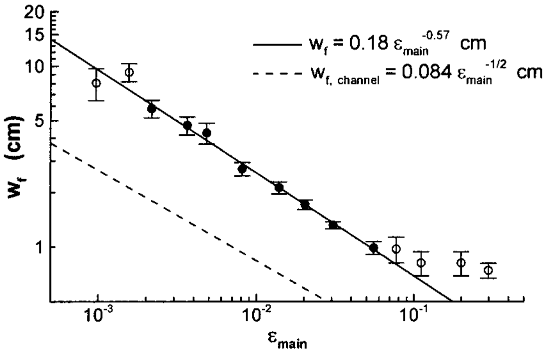


Figure 7. Forced drainage front width plotted against main body liquid volume fraction. The filled circles were used to determine the best fit power law (solid line); see eq 12. The dashed line is the prediction from channel-dominated foam drainage theory.

front width that levels off at $\epsilon \gtrsim 0.05$. The solid data points fall on a straight line, indicating power-law behavior, for which a fit yields

$$w_f \approx c_w\epsilon_{main}^{d_w} \qquad c_w \approx 0.18\text{ cm} \qquad d_w \approx -0.57 \qquad (12)$$

The first two data points in Figure 7 ($\epsilon_{main} \lesssim 10^{-3}$) were not used for the fit, because of the extremely low signal and relatively large error in the fluorescence intensity. The dashed line shows the prediction of channel-dominated theory which captures the power-law behavior of the front width measurements for $\epsilon_{main} \lesssim 0.05$, but the prefactor is less than half that found in (12).

The leveling off of the front width can in part be attributed to the break down of the geometric approximations of section 1.1, because the assumption that the channels contain most of the liquid is no longer valid for large ϵ. Another important consideration is that our experiments study one-dimensional flow along the vertical, which is valid provided that liquid rapidly spreads horizontally as to remove radial variations in ϵ.[38] Any radial variations in the liquid volume fraction will give rise to horizontally acting capillary forces which scale as $1/R$, where $R \approx 0.63$ cm is the tube radius. The capillary force along the vertical scales inversely with the front width, $1/w_f$, and thus is weaker than the horizontal capillary force for $w_f \gtrsim R$. In this case, we expect liquid to be distributed more or less evenly throughout the cross-section of the tube, and one-dimensional flow is a good approximation. For $w_f \lesssim R$, however, capillary forces are too weak to allow for sufficient radial liquid redistribution. Instead of measuring the front of a one-dimensional wave in the vertical direction, the fluorescence intensity then picks up the radial ϵ profile. Figure 7 shows that for large $\epsilon_{main} \gtrsim 0.1$ the front width indeed saturates at $w_f \sim R$.

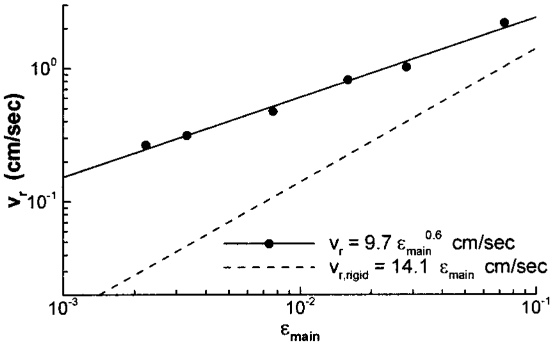


Figure 8. Rear velocity of free drainage versus main body liquid volume fraction. The solid line is the best power-law fit to the data points, and the dashed line is the channel-dominated foam drainage prediction.

2.4. Free Drainage. *Free drainage* is the evolution of the liquid volume fraction of an initially uniform foam of finite height. Figure 5b shows the experimental dynamics with initial uniform (main body) volume fraction $\epsilon_{main} \approx 5 \times 10^{-3}$. A foam of finite height is created by inserting a smaller tube into the foaming tube at the top of the camera's field of view where the injection needle is located. Sucking air through the smaller tube removes foam above the injection point. The origin of the z-axis is the location of the injection needle, which has an experimental uncertainty of about ±1 mm and is about 5 mm (equivalent to a few bubble diameters) below the top of the foam. For the duration of the free drainage experiment, unlike the other two experiments, foaming from the bottom is halted. To achieve the initial condition of a uniform volume fraction, a continuous flux of liquid is injected, as in forced drainage, at the top of the foam, and the front is allowed to reach the bottom of the foam. At that point in time, the foam is uniform, and turning off the liquid supply marks the beginning of the free drainage experiment.

We observe that the average edge length for some free drainage experiments increased somewhat from the forced drainage value $L = 0.15 \pm 0.04$ cm. During the foaming process, the foam is pushed up through the foaming tube and there is friction of the bubbles against the walls. This friction appears to compress the bubbles inside the foaming tube. Stopping the foaming for free drainage, as well as removing foam above the injection point, relieves the compression. This increase in L induces slight changes in the prefactors c_v and c_w of (11) and (12).

Under the influence of gravity the free drainage profiles of liquid volume fraction versus height evolve such that ϵ increases monotonically from the top to the bottom of the foam at all times (see Figure 5b). Liquid continues to drain from the foam, and the profiles approach a steady state at long times.[39]

In contrast to forced drainage, where the front is moving down the foam without changing its shape, the free drainage *rear*, the region above the (uniform) main body, is getting wider with time as it moves downward. We refer to the velocity of the half-maximum point of the rear as the *rear velocity*, v_r. The solid line in Figure 8 shows the best power-law fit of v_r versus ϵ_{main}, and the dashed line shows the expected rear velocity from the channel-dominated foam drainage model.

A distinctive feature of the free drainage profiles is the "knee" (Figure 5b), which indicates the transition region between the main body and the draining foam above. The knee is measured to be moving downward at a velocity of $v_k \approx 1.5v_r$, which will be discussed theoretically in section 4.2.2.

It is useful to track the rate at which the liquid volume fraction decreases with time. However, as is obvious from the profiles in

---

[38] We will show later in section 4.1 that flow through channels experiences a slipping boundary condition and is not Poiseuille-like. However, in the channels lining the foaming tube wall a no-slip boundary condition is partially imposed, which increases the viscous dissipation of these channels and decreases the liquid velocity. We thus expect in the front region that capillarity redistributes liquid from wetter channels in the center of the foaming tube to drier channels at the periphery.

[39] Princen, H. M.; Kiss, A. D. *Langmuir* **1987**, *3*, 36.

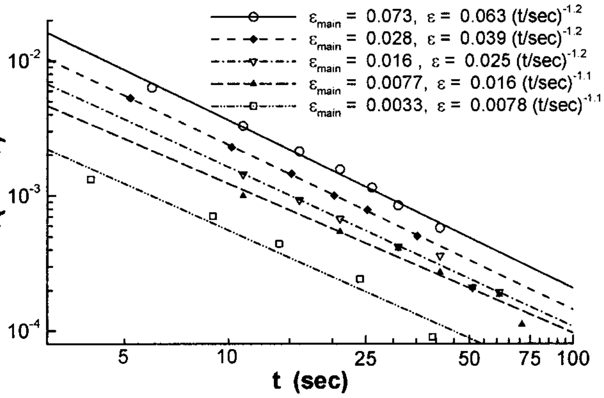


Figure 9. Decrease of liquid volume fraction at the injection point for five free drainage profiles, $\epsilon_{\text{main}} = 3.3 \times 10^{-3}$, ..., 7.3 $\times$ 10^{-2}. The lines show best fits to power-law behavior.

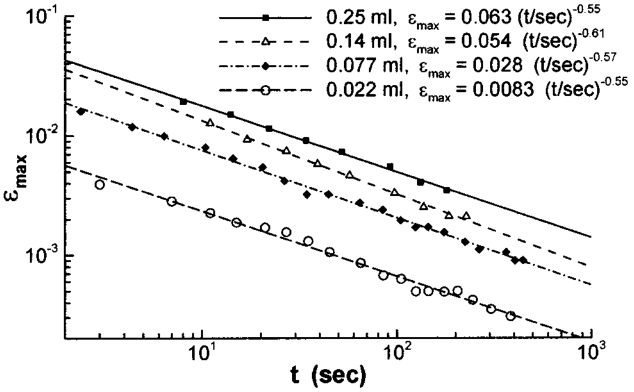


Figure 10. Peak volume fraction of the drainage pulse as a function of time for pulse volumes $V_{\text{liq}} = 0.25$, 0.14, 0.077, and 0.022 mL. Lines show best fits to power-law behavior.

Figure 5b, the dynamics depends on time and position as well. We fix the position at the injection point and plot in Figure 9 the decrease of ϵ with time for four different free drainage experiments with different main body volume fractions ϵ_{main}. Although the data cover only a small range in time, because the liquid volume fraction quickly drops to immeasurably small values, it appears that there is a rough power-law dependence with $\epsilon \propto t^{-1.2}$. We point out that the channel-dominated drainage equation applied to free drainage[40] predicts that $\epsilon \propto t^{-2/3}$ for points close to the top of the foam.

2.5. Pulsed Drainage. *Pulsed drainage* is the evolution of a (small) finite volume of liquid ("pulse") injected into a foam with very low liquid content; the injected volume has an uncertainty of ±0.005 mL. As with free drainage, the position of the injection needle marks the origin for the z-axis. The bolus is injected within 1 s using the syringe pump, and the end of the injection sets the time $t = 0$, which is known within ±1 s. Gravity pulls the injected liquid downward, and capillarity spreads the pulse in all directions, as shown by the liquid volume fraction profiles in Figure 5c. Convenient reference points of the profiles are the peak height (maximum liquid volume fraction) ϵ_{max} and peak position z_{max} of the pulse.

Data for the decrease of ϵ_{max} with time from four experiments with different injection volumes are shown in Figure 10. Figure 11 shows the corresponding motion of z_{max} in time. Both figures span one decade in pulse volume and several decades in time. The larger pulses travel downward faster and by $t \approx 200$ s have moved outside the field of view, which ends at $z \approx 70$ cm. The two smaller pulses stay in the field of view longer; however, their fluorescence signals are much weaker.

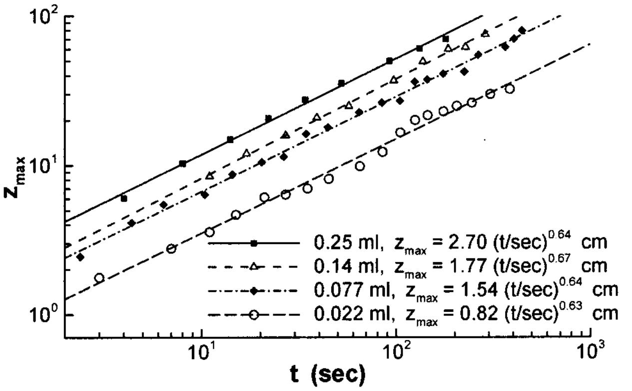


Figure 11. Peak position of the drainage pulse as a function of time for pulse volumes $V_{\text{liq}} = 0.25$, 0.14, 0.077, and 0.022 mL. Lines show best fits to power-law behavior.

The data points fall on straight lines on the log−log plot, and power-law behaviors

$$\epsilon_{\text{max}} = c_\epsilon t^{d_\epsilon} \qquad z_{\text{max}} = c_z t^{d_z} \tag{13}$$

are observed, with $d_\epsilon \approx$ -0.57 and $d_z \approx 0.65$. Again we observe (small) deviations from the channel-dominated foam drainage theory, which predicts $d_\epsilon = -{}^1/_2$ and $d_z = +{}^1/_2$.[40]

We verified that the total liquid content, which is determined from the integral under the ϵ-profiles (cf. Figure 5c), is conserved provided that the complete pulse remains within the field of view of the camera. This also supports the validity of our intensity calibration.

3. Energy Argument for the Foam Permeability

As discussed in section 1.2, our task will be the derivation of a generalized foam permeability $k(\epsilon)$ which accounts for the viscous dissipation in *both* channels and nodes.

In order to derive $k(\epsilon)$, we consider the energy balance over a "dog-bone" foam network unit (Figure 1c) with liquid volume ν. All dissipation in the gas bubbles is neglected, and we only consider dissipation in the liquid. If one balances the rate of work per unit volume of foam done by pressure and body forces against the rate of work done by viscous stresses, one obtains (from the Stokes equation)

$$\int_\nu \mathbf{u}_m \cdot (\rho \mathbf{g} - \nabla_m p_m)\, \mathrm{d}\nu = -\int_\nu \mu \mathbf{u}_m \cdot \nabla_m^2 \mathbf{u}_m\, \mathrm{d}\nu \tag{14}$$

where $\rho\mathbf{g}$ is the gravitational force and p_m and $\mathbf{u}_m$ are the microscopic pressure and velocity field, respectively. The rate of work per unit volume that drives the flow in the network unit (left-hand side of (14)) can be rewritten as

$$\mathbf{u} \cdot \mathbf{G} \nu \tag{15}$$

where $\mathbf{G} = \rho\mathbf{g} - \nabla p$ is the macroscopic driving pressure gradient (3). By dropping the subscript m, we indicate that averages over the network unit have been performed, making the resulting quantities macroscopic. The liquid volume ν consists of two quarter-node contributions $\nu_{\text{n}} \approx \delta_{\text{n}} r^3/2$ and a channel contribution $\nu_{\text{c}} \approx \delta_\epsilon r^2 L$; see discussion after (2). Removing length and velocity scales from inside the integral on the right-hand side of (14) yields

$$\mu u^2 \left(\frac{\delta_{\text{n}} r}{2} \int_{2\nu_{\text{n}}/\delta_{\text{n}} r^3} \left(\frac{\mathbf{u}_m}{u}\right) \cdot r^2 \nabla_m^2 \left(\frac{\mathbf{u}_m}{u}\right) \frac{2\mathrm{d}\nu}{\delta_{\text{n}} r^3} + \right.$$
$$\delta_{\text{a}} L \int_{\nu_{\text{c}}/\delta_{\text{a}} r^2 L} \left(\frac{\mathbf{u}_m}{u}\right) \cdot r^2 \nabla_{\perp m}^2 \left(\frac{\mathbf{u}_m}{u}\right) \frac{\mathrm{d}\nu}{\delta_{\text{a}} r^2 L} +$$
$$\left. \frac{\delta_{\text{a}} r^2}{L} \int_{\nu_{\text{c}}/\delta_{\text{a}} r^2 L} \left(\frac{\mathbf{u}_m}{u}\right) \cdot L^2 \nabla_{||m}^2 \left(\frac{\mathbf{u}_m}{u}\right) \frac{\mathrm{d}\nu}{\delta_{\text{a}} r^2 L} \right) \tag{16}$$

---

(40) Koehler, S. A.; Stone, H. A.; Brenner, M. P.; Eggers, J. *Phys. Rev. E* **1998**, *58*, 2097.

where $\nabla_{\|m}^2$ is the longitudinal component of the Laplacian along the channel, which has a typical scale of L^{-2}, and $\nabla_{\perp m}^2$ is its transverse component with scale r^{-2}. We have indicated explicitly the scales of terms in order to characterize the relative sizes of the sources of dissipation. The first term in (16) is the rate of viscous dissipative work of the flow in the two node regions, while the second and third are respectively the rate of transverse and (extensional and compressional) longitudinal dissipative work in the channel.

Combining (14)−(16) yields

$$\mathbf{G}\cdot\hat{\mathbf{u}} \approx -\mu u\left(\frac{\delta_{\mathrm{n}}}{2\delta_{\mathrm{a}} r L}\int_{\tilde{\nu}_{\mathrm{n}}} \tilde{\mathbf{u}}_m\cdot\tilde{\nabla}_m^2\tilde{\mathbf{u}}_m \,\mathrm{d}\tilde{\nu} + \frac{1}{r^2}\int_{\tilde{\nu}_{\mathrm{c}}} \tilde{\mathbf{u}}_m\cdot\tilde{\nabla}_{\perp m}^2\tilde{\mathbf{u}}_m \,\mathrm{d}\tilde{\nu} + \frac{1}{L^2}\int_{\tilde{\nu}_{\mathrm{c}}} \tilde{\mathbf{u}}_m\cdot\tilde{\nabla}_{\|m}^2\tilde{\mathbf{u}}_m \,\mathrm{d}\tilde{\nu}\right) \quad (17)$$

where ~ denotes dimensionless quantities and $\hat{\mathbf{u}}$ is the unit vector in the direction of flow in the dog-bone network unit. ν in (15) has been approximated by $\nu_{\mathrm{c}} \approx \delta_{\mathrm{a}} r^2 L$, consistent with the small liquid volume fraction limit and the approximations that lead to (2).

A foam is composed of many such interconnected network units with different orientational angles θ between the axis of the unit channel $\hat{\mathbf{u}}$ and the direction of the driving pressure gradient $\hat{\mathbf{G}}$. The average liquid velocity of a region of foam is then given by the average $\mathbf{v} \equiv \langle\mathbf{u}\rangle_\theta$ over all orientations. Averaging (17) over all angles gives

$$\frac{\mathbf{G}}{3} \approx -\mu\mathbf{v}\left(\frac{\delta_{\mathrm{n}}}{2\delta_{a} r L}\int_{\tilde{\nu}_{\mathrm{n}}} \tilde{\mathbf{u}}_m\cdot\tilde{\nabla}_m^2\tilde{\mathbf{u}}_m \,\mathrm{d}\tilde{\nu} + \frac{1}{r^2}\int_{\tilde{\nu}_{\mathrm{c}}} \tilde{\mathbf{u}}_m\cdot\tilde{\nabla}_{\perp m}^2\tilde{\mathbf{u}}_m \,\mathrm{d}\tilde{\nu} + \frac{1}{L^2}\int_{\tilde{\nu}_{\mathrm{c}}} \tilde{\mathbf{u}}_m\cdot\tilde{\nabla}_{\|m}^2\tilde{\mathbf{u}}_m \,\mathrm{d}\tilde{\nu}\right) \quad (18)$$

As expected the average flow velocity is along $\mathbf{G}$. The factor of 3 results from the angular average $\langle\mathbf{G}\cdot\tilde{\mathbf{u}}\rangle_\theta = G\int_0^1\cos^2(\theta)\,\mathrm{d}\cos(\theta) = G/3$.

Thus we have successfully made the transition from a microscopic equation like (14) to a formula on macroscopic scales equivalent to Darcy's law, (3). If one compares (3) and (18), the effective permeability of the foam is

$$k(\epsilon) \approx \left(\frac{3\delta_{\mathrm{n}}}{2\delta_{\mathrm{a}} r L} I_{\mathrm{n}} + \frac{3}{r^2} I_{\mathrm{c}\perp} + \frac{3}{L^2} I_{\mathrm{c}\|}\right)^{-1} \approx \left(\frac{\delta_{\mathrm{n}}\delta_\epsilon^{1/2}}{2\delta_{\mathrm{a}}\epsilon^{1/2}} I_{\mathrm{n}} + \frac{\delta_\epsilon}{\epsilon} I_{\mathrm{c}\perp} + I_{\mathrm{c}\|}\right)^{-1}\frac{L^2}{3} \quad (19)$$

where I_{n}, $I_{\mathrm{c}\perp}$, and $I_{\mathrm{c}\|}$ have been introduced as abbreviations for the absolute values of the corresponding dimensionless dissipative work integrals in (18). The second approximation in (19) uses $r(\epsilon)$ from (2) to arrive at an expression for $k(\epsilon)$. An explicit calculation of these dimensionless integrals in general is quite difficult, and beyond the scope of this work, except for the simple case of no-slip channel walls (i.e. Poiseuille-like flow) as we discuss in the next section. Nonetheless we will treat I_{n}, $I_{\mathrm{c}\perp}$, and $I_{\mathrm{c}\|}$ as dimensionless numbers that for a given foam are independent of ϵ and L.

3.1. Boundary Conditions. The relative contribution of each of the three terms in eq 19 to the permeability of a foam is influenced by the relative liquid volume in the channels and nodes and also is crucially dependent on the boundary conditions at the gas/liquid interface which affect the dimensionless integrals I_{n}, $I_{\mathrm{c}\perp}$, and $I_{\mathrm{c}\|}$.

The original foam drainage model[19,23] is based upon a no-slip (i.e. zero-velocity) boundary condition, making the flow in the channels Poiseuille-like. The second term of (19) then makes the dominant contribution to $k(\epsilon)$ because $r \ll L$, and the permeability is

$$k(\epsilon) = K_1 L^2\epsilon \quad K_1 \equiv (3\delta_\epsilon I_{\mathrm{c}\perp})^{-1} \quad (20)$$

From numerical calculations for flow through a rigid, straight channel with a scalloped-triangle cross section (see Figure 1b) it follows that[5,7] $K_1 \approx 6.3 \times 10^{-3}$. Equation 20 is indeed the proportionality $k(\epsilon) \propto \epsilon$ inferred from the physical argument discussed in section 1.3. As the dissipation in the nodes is negligible here, we call the foam drainage equation resulting from (7) and (20) *channel-dominated*. In its one-dimensional form, (7) for rigid gas/liquid interfaces can be written as

$$\mu\frac{\partial\epsilon}{\partial t} + K_1\rho g L^2\frac{\partial\epsilon^2}{\partial z} - \frac{\gamma\delta_\epsilon^{1/2}K_1 L}{3}\frac{\partial^2\epsilon^{3/2}}{\partial z^2} = 0 \quad (21)$$

where the $+z$ axis points along the direction of gravity.

Two arguments have been made to support the no-slip boundary condition. The first is based upon the assumption that the liquid in the faces is essentially stationary. Any motion of the channel walls results in a surface shear against the rigid faces, which is resisted by the surface viscosity. If the surface (shear) viscosity is large, the surface mobility is low, and the channel boundaries are approximately rigid. Kraynik[4] proposed that the surface shear viscosity μ_{s} should fulfill $\mu_{\mathrm{s}} \gtrsim 10r\mu$ in order to justify the assumption of rigid walls. The surface viscosity is highly dependent on the surfactant species and concentration, and unfortunately, there is considerable uncertainty in the measured values of μ_{s} even for well-defined surfactant systems such as aqueous solutions of SDS.[41] The range of typical values $\mu_{\mathrm{s}} \sim 10^{-4}$–$10^{-2}$ g/s [42] allows for violations of the "Kraynik criterion" for large enough r.

A second argument in favor of a no-slip interface is that the liquid flow through the channels will shear surfactants off the top portion of the channel surface and push them toward the bottom of the channel. If the diffusion and adsorption times for surfactants in the bulk to replenish the surfaces are large, then a surfactant concentration gradient in the direction of the downward flow is created. This sets up a surface (Marangoni) stress that will oppose the downward flow of the surface and so reduce the surface velocity. Again, it is difficult to assess a priori the time scales of exchange of surfactant between surface and bulk.

An alternative physical limit to consider is the case of very small surface stresses,[5] which should be valid for low surface viscosities, mobile faces, and small surfactant concentration gradients along the surface. We remind the reader that the injected liquid in our experiments is the same soap solution used to generate the foam, so that the injection will not automatically set up any surface stresses. If, in addition, the surfactant molecules adsorb and desorb from the surface quickly enough to equilibrate on the time scales of the flow, surfactant concentration gradients should be small and the surfaces would be nearly stress-free. Indeed, when comparing low molecular weight

---

[41] Buzza, D. M. A.; Lu, C.-Y. D.; Cates, M. E. *J. Phys. II* **1995**, *5*, 37−52.

[42] Shah, D. O.; Djabbarah, N. F.; Wasan, D. T. *Colloid Polym. Sci.* **1978**, *256*, 1002.

surfactants such as SDS to surface active proteins, Wilde et al.[22] describe the interfaces of the former as highly mobile.

In the case of mobile interfaces, the flow in the channels is pluglike with some degree of extension/compression toward the top and bottom ends of the channel where the cross-sectional area varies (see Figure 1b,c). Thus, in contrast to the rigid-wall case, $I_{c\perp}$ is not necessarily the dominant term. However, even for a stress-free wall with perfect slip, there will be dissipation in the nodes due to the merging, mixing, and bending of the flow (see the Introduction and Figure 1b; laminar viscous flow in curved channels is referred to as Dean flow[43]). The first term of (19) will then dominate (the third term is negligible because $L \gg r$, as before), and the permeability becomes

$$k(\epsilon) = K_{1/2}L^2\epsilon^{1/2} \qquad K_{1/2} \equiv \frac{2\delta_{\rm a}}{3\delta_\epsilon^{\;1/2}\delta_{\rm n} I_{\rm n}} \tag{22}$$

which is in agreement with the force balance in section 1.3. The permeability coefficient $K_{1/2}$ is determined by the geometry of the node walls and the flow field inside the nodes.[44] We have to caution, though, that experiments varying bubble size suggest a slight dependence of $K_{1/2}$ on L not explicitly indicated in (22). We believe this weak L-dependence may arise either from interfacial rheology or from bubble deformations due to hydrostatic pressure variations. We are not aware of any calculation of the dimensionless dissipation integral $I_{\rm n}$; however an experimentally deduced value is reported in section 4.1. Because of the dominance of node dissipation we call the equation obtained from (7) with the foam permeability (22) the *node-dominated foam drainage equation*

$$\mu\frac{\partial\epsilon}{\partial t} + K_{1/2}\rho g L^2\frac{\partial\epsilon^{3/2}}{\partial z} - \frac{K_{1/2}\delta_\epsilon^{\;1/2}L\gamma}{2}\frac{\partial^2\epsilon}{\partial z^2} = 0 \tag{23}$$

given here in its one-dimensional form projected onto the z-axis. Note that the exponents in (23) differ from those in (21) which is due to the change in the permeability's dependence on ϵ (cf. (20) and (22)).

The approach assuming moving, stress-free walls does not seem unrealistic considering that it has been known for some time that the gas/liquid interfaces in a foam do not have to be rigid. In experiments involving soap films, observers have distinguished between rigid and (simple or irregular) mobile films.[13] In our foam system we visually observe motion in the faces, similar to that described by Leonard and Lemlich.[7]

4. Analysis and Comparison to Experiment

Armed with a generalized theory for foam permeability (section 3) and data from three types of experiments (section 2), we now proceed to discuss the dynamics of foam drainage in more detail. We begin with forced drainage and show that our data supports the node-dominated model. In subsequent sections we use the node-dominated equation (23) to analyze and elucidate the experimental results for free and pulsed drainage from sections 2.4 and 2.5. The dynamics of the channel-dominated theory has been described in a previous publication.[40]

4.1. Forced Drainage. The natural decomposition of the traveling wave profile into a well-drained region, front region, and main body region (cf. section 2.3 and Figure 5a) sets three characteristic scales for a particular experiment: (i) a length scale, proportional to the front width $w_{\rm f}$, (ii) a velocity scale proportional to the front speed $v_{\rm f}$, and (iii) the liquid volume fraction of the main body, $\epsilon_{\rm main}$. These characteristic scales are determined by physical parameters, such as surface tension, density, gravitational acceleration, and viscosity as well as experimental parameters such as liquid flux and bubble size. We will later use these scales to normalize experimental data as well as computational results.

To arrive at a theoretical description of forced drainage, we make the following ansatz for a traveling volume fraction profile, $\epsilon(z, t) = \epsilon(s \equiv z - v_{\rm f}\, t)$; i.e., we transform into a frame of reference moving with the drainage wave speed $v_{\rm f}$. Using this ansatz in the generalized foam drainage equation (7), integrating once with respect to s and using the boundary condition $\epsilon(\infty) = 0$ yields

$$-\mu v_{\rm f} + k(\epsilon)\left(\rho g - \frac{\delta_\epsilon^{\;1/2}\gamma}{2L\epsilon^{3/2}}\frac{{\rm d}\epsilon}{{\rm d}s}\right) = 0 \tag{24}$$

In the main body region of the wave above the front, the liquid volume fraction is constant, $\epsilon(-\infty) = \epsilon_{\rm main}$. The ϵ-derivative in (24) then vanishes, and a simple relationship between permeability and volume fraction results,

$$k(\epsilon_{\rm main}) = \frac{\mu v_{\rm f}}{\rho g} \approx \frac{\mu c_v}{\rho g}\epsilon_{\rm main}^{d_v} \tag{25}$$

where the second equality follows from the experimentally observed power-law behavior (11). Comparison of the measured $d_v \approx 0.6$ with the channel-dominated (i.e. "no-slip") approximation (20), $k(\epsilon) \propto \epsilon$, and the node-dominated (i.e. "no-stress") approximation (22), $k(\epsilon) \propto \epsilon^{1/2}$, shows that the permeability is closer to being dominated by the nodes, at least for the aqueous foams used here. In particular, from (22) and (25) the model predicts that the forced drainage velocity is

$$v_{\rm f} = V_{\rm F}\epsilon_{\rm main}^{1/2} \qquad V_{\rm F} = K_{1/2}\rho g L^2/\mu \tag{26}$$

with $K_{1/2}$ defined in (22). Since we are unable to compute $K_{1/2}$, we determine it from the experimental data for $v_{\rm f}$ (Figure 6). In order to obtain a controlled comparison to node-dominated theory, however, we must not use the original fit (11) with $d_v \approx 0.60$, but instead perform a best fit of the measured data to (26), where $d_v = {}^1/_2$ exactly. The result is

$$v_{\rm f} \approx \tilde{c}_v\epsilon_{\rm main}^{1/2} \qquad \tilde{c}_v \approx 5.12\ {\rm cm/s} \tag{27}$$

which, by comparison to (26), identifies the theoretical velocity scale $V_{\rm F}$ with the measured $\tilde{c}_v$. With $\rho = 1$ g cm^{-3}, $g = 981$ cm s^{-2}, $\mu = 0.01$ g cm^{-1}, and $L = 0.15$ cm, we obtain[45] $K_{1/2} \approx 2.3 \times 10^{-3}$. Using eq 22 and approximating[28] $\delta_{\rm n} \approx 0.3$ gives $I_{\rm n} \approx 400$ for the dimensionless node

---

(43) Leal, G. *Laminar Flow and Convective Transport Processes*; Butterworth-Heinemann: Boston, MA, 1992.

(44) The convention we adopt for the subscript of the permeability coefficient is $k(\epsilon) = K_\chi L^2\epsilon^\chi$; $\chi = {}^1/_2$ for node-dominated drainage, and $\chi = 1$ for channel-dominated drainage.

(45) It is instructive to consider the permeability of an fcc-close-packed bed of spheres ($\epsilon_{\rm fcc} \approx 0.26$), which has the same geometry as a monodisperse foam at the critical volume fraction $\epsilon = \epsilon_{\rm fcc}$ composed of spherical bubbles. In our notation, $k_{\rm fcc} \approx (2a^2)/(9(1 - \epsilon_{\rm fcc})435)$ for spheres of radius a.[30] The equivalent sphere radius for the foam follows from the volume of a cell ($2^{7/2}L^3 = 4\pi a^3/3$) to be $a \approx 1.4L$. Substituting for a yields $k_{\rm fcc} \approx 1.4 \times 10^{-3}L^2$. Similar permeabilities are predicted by both foam drainage models: the node-dominated model (mobile gas/liquid interfaces) predicts $k(\epsilon_{\rm fcc}) \approx 1.2 \times 10^{-3}L^2$, and the channel-dominated model (rigid gas/liquid interface) yields $k(\epsilon_{\rm fcc}) \approx 1.6 \times 10^{-3}L^2$. Although both foam drainage models are strictly valid for dry foams only (i.e. $\epsilon \ll \epsilon_{\rm fcc}$), this calculation shows that the modeled permeability is still reasonable even in the limit of wet foams.

dissipation integral in (19), compared with $I_{c\perp} \approx 310$ for the channel-dominated theory.

Is node-dominated drainage a valid description for all of our different types of experiments? Much work has been devoted to the channel-dominated foam drainage model with rigid walls.[8,40] We will focus in the present work on the other limit of plug-flow in the channels and dissipation in the nodes. Although the data for the foam we study here do not exactly match the predictions of either limit, we will show in the remainder of the paper that the agreement with the node-dominated model of foam drainage is much better.

We continue our analysis of the node-dominated model by shifting our focus from the uniform main body to the front region of the forced drainage wave. In the front region ϵ decreases from ϵ_{main} to 0, so that the derivative (capillary) term of (24) becomes important. Using (22), (24), and (26) yields a nonlinear ODE for $\epsilon(s)$,

$$\frac{\mathrm{d}\epsilon}{\mathrm{d}s} = \frac{2L\rho g}{\delta_\epsilon^{\ 1/2}\gamma}(\epsilon^{3/2} - \epsilon_{\mathrm{main}}^{1/2}\epsilon) \qquad (28)$$

A characteristic (capillary) length scale for a foam is

$$Z_{\mathrm{F}} = \frac{\delta_\epsilon^{\ 1/2}\gamma}{2L\rho g} \qquad (29)$$

Defining the new variables $\tilde{s} = s\epsilon_{\mathrm{main}}^{1/2}/Z_{\mathrm{F}}$ and $\tilde{\epsilon} = \epsilon/\epsilon_{\mathrm{main}}$ reduces (28) to

$$\mathrm{d}\tilde{\epsilon}/\mathrm{d}\tilde{s} = \tilde{\epsilon}^{3/2} - \tilde{\epsilon} \qquad (30)$$

Imposing the boundary condition $\tilde{\epsilon}(-\infty) = 1$ leads to the solution

$$\tilde{\epsilon}(\tilde{s}) = (1 + e^{\tilde{s}/2})^{-2} \qquad (31)$$

which is formally the square of a Fermi function. This shows analytically that the node-dominated foam drainage equation (as well as the channel-dominated one[8]) admits a solitary wave solution with unchanging wave profile. This should not be too surprising considering the similarity of equations like (21) and (23) with the Burgers equation and other well-known PDEs with soliton solutions.[46]

The dimensionless width of the front region Δ_s is defined as the distance between the points in the profile where $\tilde{\epsilon} = 0.8$ and $\tilde{\epsilon} = 0.2$, to match the experimental criterion (see section 2.3). Using (31), we find $\Delta_s \approx 4.7$, making the dimensional front width

$$w_{\mathrm{f}} = \Delta_s Z_{\mathrm{F}}\epsilon_{\mathrm{main}}^{-1/2} = \frac{\Delta_s\delta_\epsilon^{\ 1/2}\gamma}{2L\rho g}\epsilon_{\mathrm{main}}^{-1/2} \qquad (32)$$

We note that any equation based upon (1) and (7) predicts that the front width will scale with $\epsilon_{\mathrm{main}}^{-1/2}$. The measured front width exponent $d_w \approx -0.57$ is in fairly close agreement. A ring-tensiometer measurement gives $\gamma \approx 32$ dyne/cm[47] and substituting into (32) predicts a front width prefactor $c_w \approx 0.21$ cm. For consistent comparison to experiment, we again replace our best fit (12) by a fit with the theoretical exponent $d_w = -^1/_2$ and find

$$w_{\mathrm{f}} \approx \tilde{c}_w\epsilon_{\mathrm{main}}^{-1/2} \qquad \tilde{c}_w \approx 0.256 \text{ cm} \qquad (33)$$

The best-fit prefactor, $\tilde{c}_w$, is much closer to the node-dominated prediction than to the channel-dominated prediction of 0.084 cm. This fit also provides us with an experimental value for the length scale Z_{F}, as $Z_{\mathrm{F}} = \tilde{c}_w/\Delta_s \approx 0.0545$ cm.

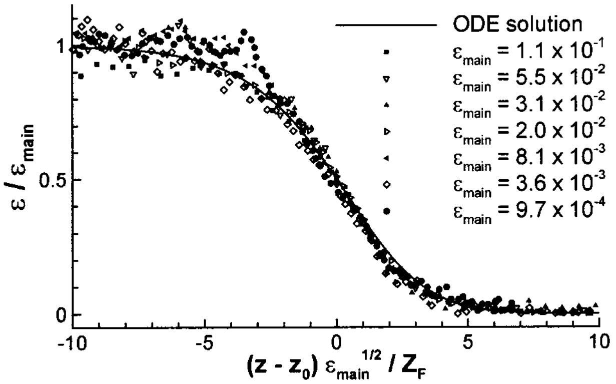


Figure 12. Collapse of the front profiles of forced drainage waves for seven different volume fractions ϵ_{main} (symbols) onto a universal curve. The solid line is the analytical result of (31) which predicts the shape of the wave profile from the node-dominated foam drainage eq 23. The curves have been aligned such that the coordinate $z - z_0 = 0$ coincides with the center of the front region, where we select a value z_0 for each data set.

The analytical solution (31) can be compared with the complete measured profile of the wave front at various ϵ_{main} by rescaling using Z_{F} and ϵ_{main}. Figure 12 shows that all data collapse onto a universal curve, in excellent agreement with the theoretical prediction. Thus, all the spatial and temporal characteristics of forced foam drainage are captured accurately within the node-dominated model.

4.2. Free Drainage. Though the free drainage experiment is conceptually simple, a theoretical description of this process is considerably more involved than for forced drainage. This is largely due to the more complicated boundary conditions for ϵ, namely, zero flux at the top of the foam and $\epsilon = \epsilon_{\mathrm{crit}}$ at the bottom, where the foam is in contact with the soap solution and the bubbles are spherical. The latter condition uses the *critical* volume fraction ϵ_{crit}, which corresponds to the void fraction of close packing of spherical bubbles. In our experiments, the bottom boundary is far below the field of view of the camera, so that we do not have to deal with it explicitly.

It is convenient to divide a free-draining foam into two overlapping regions that grow with time: the rear region which contains the top of the foam and the knee region where the liquid volume fraction plateaus to ϵ_{main} (cf. section 2.4). The knee is moving downward with a constant velocity v_{k}, which experiments show to be greater than the front velocity of forced drainage v_{f} for the same ϵ_{main}. It is also greater than the rear velocity v_{r}, so that the rear region grows over time (see Figure 5b).

4.2.1. Rear Region. At time $t = 0$, a uniform profile $\epsilon(z, 0) = \epsilon_{\mathrm{main}}$ has been established, and the flux of liquid injected into the foam at $z = 0$ is turned off. The no-flux top boundary condition requires that the liquid velocity $v = 0$ at $z = 0$ for $t > 0$ (otherwise the foam would be shrinking in height). For node-dominated drainage this yields

$$\rho g\epsilon^{3/2} - \frac{\delta_\epsilon^{\ 1/2}\gamma}{2L}\frac{\partial\epsilon}{\partial z} = 0 \qquad (z = 0) \qquad (34)$$

which follows from (3), (5), and (22).

---

[46] Grundy, R. E. *IMA J. Appl. Math.* **1983**, *31*, 121–137.

[47] A common value reported in the literature for the surface tension of SDS is 30 dyne/cm.[36]

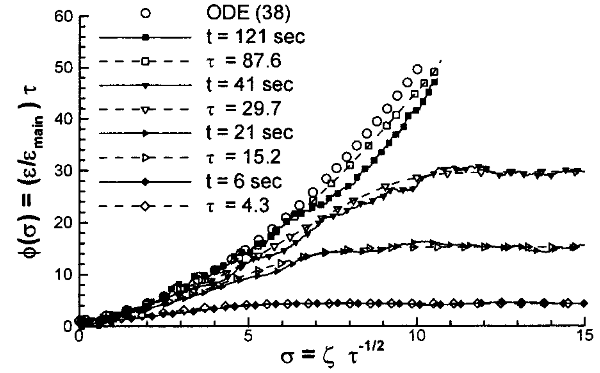


Figure 13. Rear region of simulated and experimental free drainage profiles in rescaled dimensionless coordinates for successive times, starting with a uniform profile. Filled-in symbols are measurements for $\epsilon_{\mathrm{main}} \approx 0.0077$, rescaled using a forced drainage experiment, as explained in the text. All open symbols are simulations: the open circles show the ODE solution (38), and symbols connected by dashed lines are simulations of the full PDE (36) with the no-flux top boundary condition (34). The dimensionless times τ correspond to the measured times t for the experimental profiles. For long times the profiles collapse onto the ODE solution (38).

We define characteristic time and length scales

$$T_{\mathrm{F}} = \frac{\delta_{\epsilon}^{\;1/2}\gamma\mu}{2K_{1/2}(\rho g)^2L^3} \qquad Z_{\mathrm{F}} = \frac{\delta_{\epsilon}^{\;1/2}\gamma}{2\rho g L} \tag{35}$$

which are the same as in the forced drainage problem discussed before (see section 4.1); i.e., $T_{\mathrm{F}} = Z_{\mathrm{F}}/V_{\mathrm{F}} = \tilde{c}_w/(\Delta_s\tilde{c}_v) \approx 0.0106$ s. To nondimensionalize the node-dominated foam drainage equation (23), we transform variables according to $\tilde{\epsilon} = \epsilon/\epsilon_{\mathrm{main}}$, $\zeta = z\epsilon_{\mathrm{main}}^{1/2}/Z_{\mathrm{F}}$, and $\tau = t\epsilon_{\mathrm{main}}/T_{\mathrm{F}}$ and arrive at the dimensionless foam drainage equation

$$\frac{\partial\tilde{\epsilon}}{\partial\tau} + \frac{\partial\tilde{\epsilon}^{3/2}}{\partial\zeta} - \frac{\partial^2\tilde{\epsilon}}{\partial\zeta^2} = 0 \tag{36}$$

In order to solve (36) numerically, we have to impose a second boundary condition on ϵ in addition to (34). Here we are only interested in the top region of the profile far above the foam/liquid interface, so we impose $\tilde{\epsilon}(\infty, t) = 1$ at the bottom of our simulation. Using the IMSL routine DMOLCH we obtain the drainage profiles shown in Figure 13 by the connected open symbols. We have plotted the results in rescaled coordinates, because we can obtain a self-similar solution in this coordinate system as well. In the spirit of our previous work,[40] we note that a transformation of variables according to

$$\tilde{\epsilon} = \tau^{-1}\,\phi(\sigma) \qquad \sigma = \zeta\tau^{-1/2} \tag{37}$$

balances all the terms in (36). This leads to self-similar solutions $\phi(\sigma)$ obeying the ordinary differential equation

$$-\phi - \frac{1}{2}\sigma\phi' + (\phi^{3/2} - \phi')' = 0 \tag{38}$$

with ′ denoting differentiation with respect to σ. The leading-order solution of (38) is

$$\phi(\sigma) \approx (2\sigma/3)^2 \qquad (\sigma \gg 1) \tag{39}$$

which is valid for $\sigma \le 3\tau^{1/2}/2$ (otherwise $\tilde{\epsilon} > 1$). The top no-flux boundary condition (34) transforms into

$$\phi^{3/2} - \phi' = 0 \qquad (\sigma = 0) \tag{40}$$

To solve the ODE (38), we use a shooting algorithm to ensure that the solution asymptotes to the parabolic approximation (39) for the largest $\sigma = \sigma_{\max}$ of the integration range (typically, $\sigma_{\max} \sim 10$). The resulting solution is plotted as open circles in Figure 13. It is approached by the solution of the full PDE at large τ. At any given finite time, however, the rescaled profile reaches its maximum at the knee (to be discussed in the next section) and plateaus at a value of $\phi = \tau$.

The parabolic approximation (39) can be rewritten in $(\tilde{\epsilon}, \zeta, \tau)$ space as

$$\tilde{\epsilon} \approx (2\zeta/3\tau)^2 \tag{41}$$

Note that only for very long times are (39) and (41) good approximations, because only then is their range of validity $(\tau^{1/2} \ll \zeta < 3\tau/2)$ appreciable.

Figure 13 also shows rescaled experimental data (filled symbols) in the self-similar coordinates $\tilde{\epsilon}\tau$ versus $\zeta\tau^{-1/2}$. The length and time scales are determined from the forced drainage experiment with the same ϵ_{main}, as discussed in the previous section. The rescaled experimental data are in very good agreement with the simulations at the corresponding dimensionless times. This figure shows that the rear portion of the experimental profile is indeed parabolic in rescaled coordinates (see eq 39) and not linear, $\epsilon \approx \zeta/2\tau$, as predicted by the channel-dominated foam drainage equation.[40]

Next, we return to the issue raised in section 2.4 pertaining to Figure 8, which shows that the measured rear velocity is considerably higher than that predicted by the channel-dominated model. From eq 41 it is easy to see that the half-maximum at the rear is at a position ζ given by

$$\left(\frac{2\zeta}{3\tau}\right)^2 = \frac{1}{2} \Rightarrow \zeta = \frac{3\tau}{2^{3/2}} \tag{42}$$

It follows that the predicted rear velocity $v_{\mathrm{r}} = (3/2^{3/2})V_{\mathrm{F}}\epsilon_{\mathrm{main}}^{1/2} \approx 1.06\ V_{\mathrm{F}}$. The measured rear velocity of free drainage can be fitted to the node-dominated formula to yield $v_{\mathrm{r}} \approx 6.13\ \epsilon_{\mathrm{main}}^{1/2}$ cm/s, which is in fairly good agreement with the prediction above using $V_{\mathrm{F}} = \tilde{c}_v$; see (27). The predicted rear velocity is about 10% below measurements, which can be attributed to the slightly larger bubbles of the free drainage experiment as described in section 2.4.

Figure 9 previously illustrated that at a fixed position in space near the top of the foam free drainage proceeds as $\epsilon \propto t^{-1.2}$ and not $\epsilon \propto t^{-2/3}$ as predicted by the channel-dominated model.[40] However, this measurement is consistent with the node-dominated foam drainage equation. Close to the top, for $\sigma = \zeta\tau^{-1/2} \ll 1$, it follows from (37) that $\tilde{\epsilon} \approx \tau^{-1}\phi(0)$. For long enough times, then, $\tilde{\epsilon} \propto \tau^{-1}$ is valid for a considerable range of ζ values. Note that the initial decrease in volume fraction is faster, as for times $\tau^{1/2} \ll \zeta$ the parabolic approximation (41) holds, and $\tilde{\epsilon} \propto \tau^{-2}$.

4.2.2. The "Knee".

The main body of the drainage profile with uniform $\epsilon = \epsilon_{\mathrm{main}}$ begins at the knee, which is moving downward in time. Using similar arguments as for the rear velocity (42), we expect the center of the knee (intersection of the rear and main body regions) to be at $\zeta = 3\tau/2$ in dimensionless coordinates. This means that the knee velocity v_{k} is predicted to be $v_{\mathrm{k}} = \sqrt{2}v_{\mathrm{r}}$ (cf. (42)), which agrees well with the measured v_{r} (section 2.4). In

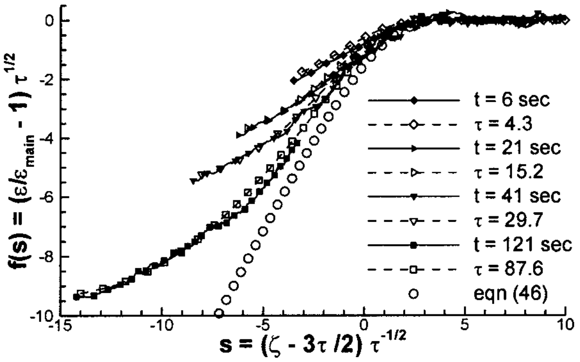


Figure 14. Free drainage profiles in the transition region (about the "knee") for experiment and simulation at successive times in rescaled (f, s) coordinates; see (43). Filled-in symbols are measurements for $\epsilon_{\mathrm{main}} \approx 0.0077$, rescaled according to the preceding forced drainage experiments that create uniform profiles, as described in the text. All open symbols are simulations: the open circles show the ODE solution (46), and connected symbols show simulations of the full PDE (36) with the no-flux top boundary condition (34), at times τ corresponding to experiment.

contrast the channel-dominated foam drainage theory predicts that the knee velocity is *twice* the rear velocity.

For $\zeta \gg 3\tau/2$, $\tilde{\epsilon}$ should saturate to 1, and so we seek a self-similar solution to the dynamics in the transition region about $\zeta \sim 3\tau/2$ of the form

$$\tilde{\epsilon} = 1 + \tau^{-1/2} f(s) \qquad s \equiv (\zeta - 3\tau/2)\tau^{-1/2} \tag{43}$$

Inserting (43) into (36) and expanding $\tilde{\epsilon}^{3/2}$ up to second order in $\tau^{-1/2} f(s)$, we obtain

$$f + sf' + \left(-\frac{3}{4}f^2 + 2f'\right)' + O(\tau^{-1/2}) = 0 \tag{44}$$

In the long-time limit, the error to this approximation becomes vanishingly small. Collapsing the numerical simulations of the PDE in this fashion validates the ansatz, as shown in Figure 14. Integrating (44) once gives

$$sf - \frac{3}{4}f^2 + 2f' = 0 \tag{45}$$

where the integration constant is set to zero because $f(+\infty) = 0$. The solution to (45) is

$$f(s) = -\frac{8e^{-s^2/4}}{3\pi^{1/2}\,\mathrm{erf}(s/2) + C} \tag{46}$$

The constant C should be chosen to match (46) onto the parabolic approximation, $\tilde{\epsilon} \approx (2\zeta/3\tau)^2$ from eq 39. To leading order this requires $f(s) \to -4s/3$ for $s \to -\infty$. From (37) and (43) one obtains $s = \sigma - 3\tau^{1/2}/2$, so that at large enough τ the conditions $\sigma \gg 1$ and $s \ll -1$ can indeed be fulfilled simultaneously. Thus, we find $C = 3\sqrt{\pi}$, and the corresponding solution (46) is plotted as open circles in Figure 14. Again, the solutions of the PDE (36) approach the self-similar solution in rescaled (f,s) coordinates for long times; see Figure 14.

We repeat this rescaling for our experimental data as shown by the solid symbols in Figure 14. As in the previous section, the profiles agree very well with theory. The small mismatches between the simulated and measured profiles can be attributed to uncertainties in the characteristic time and length scales. For the longest time, 121 s, the foam has drained considerably, and the knee no longer is in the field of view of the camera. Thus the rescaled data points do not extend to $s = 0$ as is the case for the profiles at shorter times.

4.3. Pulsed Drainage. We now discuss the one-dimensional dynamics of a pulse with finite liquid volume V_{liq} injected into a foaming tube of cross-sectional area A. We seek a solution to the dimensionless foam drainage equation (36), with the additional constraint that the dimensionless liquid volume be unity, $\int \tilde{\epsilon}\, d\zeta = 1$. This is achieved by choosing the scales

$$T_{\mathrm{P}} = \frac{\delta_{\epsilon}^{\ 3/2}\gamma^3\mu A^2}{8K_{1/2}(\rho g)^4 L^5 V_{\mathrm{liq}}^{\ 2}} \quad Z_{\mathrm{P}} = \left(\frac{\delta_{\epsilon}^{\ 1/2}\gamma}{2\rho g L}\right)^2 \frac{A}{V_{\mathrm{liq}}} \quad \epsilon_{\mathrm{P}} = \frac{V_{\mathrm{liq}}}{AZ_{\mathrm{P}}} \tag{47}$$

The following identities relate the experimentally obtained scaling factors (T_{F}, Z_{P}) of forced and free drainage to the pulsed drainage scaling factors (T_{P}, Z_{P}):

$$T_{\mathrm{P}} = \left(\frac{AZ_{\mathrm{F}}}{V_{\mathrm{liq}}}\right)^2 T_{\mathrm{F}} \qquad Z_{\mathrm{P}} = \frac{AZ_{\mathrm{F}}^{\ 2}}{V_{\mathrm{liq}}} \tag{48}$$

We can compute T_{P} and Z_{P} in this way from forced drainage because the same kind of foam is used for both experiments ($L = 0.15$ cm).

The pulse is composed of three sections (see Figure 5c): (i) The first is the *rear* region, where the pulse profile connects to the dry foam above and which is similar to the rear region of free drainage (cf. section 4.2.1). We formally define this region as the range $\zeta = z/Z_{\mathrm{P}} < 0$, although the solutions we present will remain valid for some interval of positive ζ as well. (ii) Next there is the *middle* region, in which the volume fraction grows from $\zeta = 0$ to the peak of the pulse. (iii) Finally, there is the downward moving *front* region below the peak. The latter has characteristics similar to the advancing forced drainage front (see section 4.1). In each of these regions we discuss asymptotic solutions obtained from the node-dominated foam drainage theory and compare them with our experimental measurements.

4.3.1. Rear Region. Here ϵ is small, so we expect surface tension and gravity to be of equal importance, as in the rear region of free drainage. The same self-similar ansatz (37) used before balances all three terms in the foam drainage equation and transforms it into the ODE (38). We expect the upper region of the pulse to develop in this self-similar fashion, and only toward the pulse's peak should deviations from the self-similar behavior occur. In Figure 15 a simulation (open symbols connected by dashed lines) of an initially Gaussian pulse of unit volume and initial width of unity is plotted in (σ, ϕ)space (the variables of (38)). The collapse of the rear region is very good.

Figure 15 also shows experimental data (filled symbols connected by solid lines) for a $V_{\mathrm{liq}} = 0.077$ mL pulse plotted in (σ, ϕ) space and illustrates the self-similar behavior of the rear of the pulse. Qualitatively the agreement with the simulations is good; however, on a more quantitative level the length scale Z_{P} is too large by a factor of about 1.3−2 to achieve excellent agreement. As the deviations seem to increase over time, they are presumably due to the slight differences between the experimentally measured exponents of pulse dynamics (e.g. the pulse maximum coordinate vs time; cf. Figure 11) and the theoretically predicted values. These discrepancies will be discussed further in the following section. Note also that the only free parameter in our theoretical description is $K_{1/2}$, which has been determined from the forced drainage

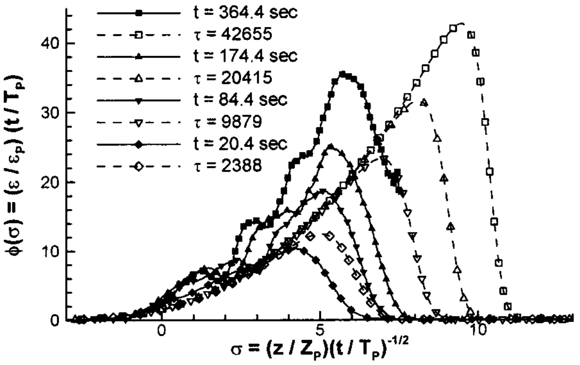


Figure 15. Collapse of the rear regions of simulated and experimental profiles for pulsed drainage using (37). Experimental profiles have filled symbols, and the open symbols show simulations for the corresponding dimensionless times. The liquid volume is $V_{\text{liq}} = 0.077$ mL, and the tube area $A = 1.27$ cm^2. The dips before the peaks of some of the experimental profiles can be attributed to the glaring reflection from the UV lamp off the foaming tube.

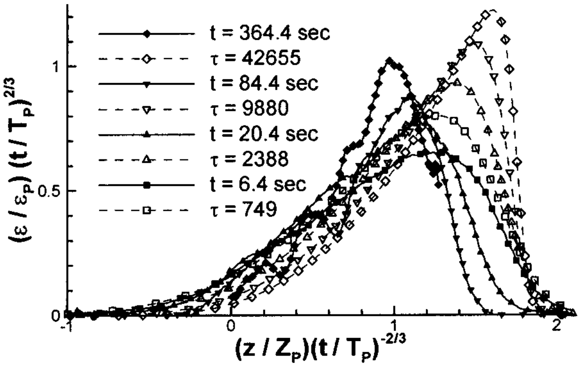


Figure 16. Collapse of pulsed drainage profiles in the middle region where surface tension has been neglected (see (50)). Filled symbols show data for a $V_{\text{liq}} = 0.077$ mL pulse; open symbols are the corresponding simulated profiles.

measurements, so that the theoretical predictions for free and pulsed drainage are parameter-free.

Table 1. Asymptotic (Long-Time) Approach of the Relations (50) of Peak Position and Peak Height Obtained from a Numerical Simulation (See Text)[a]

τ	10^2	10^3	10^4	10^5	10^6	10^7	∞
d_ϵ	−0.51	−0.53	−0.57	−0.61	−0.64	−0.65	$-^2/_3$
d_ζ	0.75	0.74	0.72	0.70	0.69	0.68	$^2/_3$
c_ϵ	0.29	0.33	0.44	0.67	0.93	1.15	1.59
c_ζ	0.74	0.80	0.90	1.10	1.34	1.53	1.89

[a] See (51) for the definitions of d_ϵ, d_ζ, c_ϵ, and c_ζ.

Table 2. Pulse Scales ϵ_P, T_P, and Z_P for Different Injection Volumes V_{liq} Used in the Experiments[a]

V_{liq} (mL)	ϵ_P	T_P (s)	Z_P (cm)	d_ϵ	d_ζ	c_ϵ	c_ζ	$\tau_{\max}$
0.25	13.2	8.1×10^{-4}	0.015	−0.56	0.64	0.74	1.16	6.2×10^5
0.14	4.23	2.5×10^{-3}	0.027	−0.61	0.67	0.93	0.92	2.0×10^5
0.077	1.25	8.5×10^{-3}	0.049	−0.57	0.64	0.72	0.96	5.2×10^4
0.022	0.10	0.11	0.17	−0.55	0.63	0.50	0.80	3.6×10^3

[a] The fifth and sixth columns are the measured exponents of the power-law behaviors for the peak height and peak position (see eq 13). The seventh and eighth columns are the prefactors of the peak volume fraction and position in dimensionless units to be compared with the predictions from Table 1. The last column is the duration of the experiment in dimensionless time; $\tau_{\max} \equiv$ (maximum experimental time)/T_P.

4.3.2. Middle Region. While a rescaling as displayed in Figure 15 may put emphasis on the rear region, it is important to realize that the vast majority of the liquid volume resides in the middle region, whose dominance becomes ever larger as time increases. Thus, the gravitational term of the foam drainage equation dominates and the surface tension term may be neglected. As with free drainage, the parabolic profile $\tilde{\epsilon} \approx (2\zeta/3\tau)^2$ then is an approximate solution. The volume constraint, however, limits the extent of the parabola requiring

$$\int_0^{\zeta_{\max}} \left(\frac{2\zeta}{3\tau}\right)^2 \mathrm{d}\zeta = \frac{4}{27}\frac{\zeta_{\max}^3}{\tau^2} = 1 \tag{49}$$

from which we conclude the power laws

$$\zeta_{\max} \approx 3(\tau/2)^{2/3} \qquad \tilde{\epsilon}_{\max} \approx (2/\tau)^{2/3} \tag{50}$$

for the location and height of the pulse peak, respectively. In accordance with these scalings, we use the coordinates $\zeta\tau^{-2/3}$ and $\tilde{\epsilon}\tau^{2/3}$ to collapse experimental and simulated profiles as shown in Figure 16. In the long time limit, the peak location should asymptote toward $\zeta_{\max}$, and the peak volume fraction should approach $\tilde{\epsilon}_{\max}$. Because the surface tension term was neglected in this approximation, the asymptotic solution of a parabolic profile that abruptly ends at $\zeta_{\max}$ is approached only slowly, which can be seen in the four simulated profiles.

Although the initial liquid volume distribution is not controlled to be Gaussian in experiment as assumed for the simulation, the comparison with the computations is quite favorable and shows that the data collapse in a fashion similar to the simulation. At short times ($t \sim 10$ s) the agreement between theory and experiment is quite good, and at long times ($t \sim 300$ s) the rescaled experimental pulses are within a factor of 2 of the simulated pulses.

To compare the predictions of (50) to experiment in more detail, we attempt to describe the simulated data by the power laws

$$\zeta_{\max} \approx c_\zeta \tau^{d_\zeta} \qquad \tilde{\epsilon}_{\max} \approx c_\epsilon \tau^{d_\epsilon} \tag{51}$$

where d_ζ and d_ϵ are now defined *locally* for a given time τ as the logarithmic derivative of $\zeta_{\max}$ and $\tilde{\epsilon}_{\max}$ with respect to τ, i.e. $d_\zeta = \partial\log(\zeta_{\max})/\partial\log(\tau)$ and $d_\epsilon = \partial\log(\tilde{\epsilon}_{\max})/\partial\log(\tau)$. Only for very large τ do the simulated peak positions and volume fractions approach the behavior predicted by (50) as shown by Table 1. The convergence of the prefactors to the projected values is even slower than that of the exponents.

Table 2 shows that the measured exponent for the peak position d_ζ stays below the asymptotically predicted value of $^2/_3$, while the simulated values in Table 1 are all larger than $^2/_3$ for finite times. This explains the discrepancies seen in Figure 16: as time proceeds, the peak positions, rescaled with the exact theoretical exponent $d_\zeta = {}^2/_3$, drift to the left for the experimental data (with effectively smaller exponents) and to the right for the simulations (which have larger exponents). Since the experimental foam has a permeability of $k(\epsilon) \propto \epsilon^{0.60}$ and not $k(\epsilon) \propto \epsilon^{1/2}$, we expect $d_\zeta = 0.625$.[46] The effect is less pronounced for the scaling of ϵ, where both experiment and theory yield $d_\epsilon > -{}^2/_3$. The shift between experiment and theory in the previous Figure 15 can be explained along the same lines: while all peaks move to the right here as time increases (because of the different rescaling), the measured data do so considerably slower than the simulation. Note that the

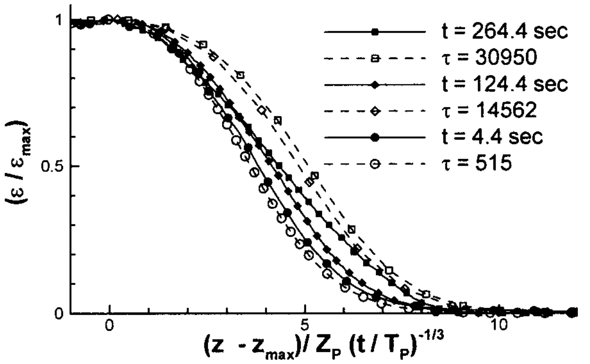


Figure 17. Collapse of the front region of pulsed drainage profiles using (54): experimental data (filled symbols) and simulations (open symbols) at corresponding times. The measurements are for a pulse with $V_{\mathrm{liq}} = 0.077$ mL.

simulations also show that the asymptotic behavior is observed only for very long times (Table 1), and in particular, the peak height requires very long times to reach the asymptote $\tilde{\epsilon}_{\max} \propto \tau^{-2/3}$. Furthermore, in experiments the initial shape of the pulse is not controlled and most likely not Gaussian. With a more irregular initial condition, it may take longer to establish self-similar behavior. Finally, the experimental foam permeability is measured to be $k(\epsilon) \propto \epsilon^{0.60}$ and not $\propto \epsilon^{1/2}$ which leads to $d_\epsilon \approx -0.625$.[46]

The experimental dimensionless prefactors of Table 2, c_ζ and c_ϵ, are generally about a factor of 2 smaller than the prediction of (50). But for times $\tau \sim 10^5$ even the prefactors for simulations starting with a Gaussian profile are only about half the asymptotic value (see Table 1). This emphasizes the extreme slowness of the asymptotics for the middle collapse.

4.3.3. Front Region. Here we treat the advancing front region of the pulse that stretches from $\epsilon_{\max}$ to the dry foam below. In dimensionless variables this region is moving downward as $\zeta_{\max} \approx 3(\tau/2)^{2/3}$ (see (56)), and we make the ansatz that it is spreading at a rate of τ^λ with $\lambda > 0$. We try a self-similar solution of the form

$$\tilde{\epsilon} = (2/\tau)^{2/3} f(s) \qquad s = \tau^\lambda(\zeta - 3(\tau/2)^{2/3}) \tag{52}$$

Substituting into (36) shows that $\lambda = -{}^1/_3$ balances the largest convective term, the gravitational term, and the surface tension term to $O(\tau^{-4/3})$, while two out of the three terms from the time derivative are of magnitude $O(\tau^{-5/3})$ and become negligible at long times. The resulting ordinary differential equation for f is

$$f' - (f^{3/2})' + 2^{-1/3} f'' + O(\tau^{-1/3}) = 0 \tag{53}$$

where the first term stems from the convective time derivative, the second from the advective (gravitational) term, and the third from the diffusive (surface tension) term. We integrate (53) once, which after a substitution of variables, $\tilde{s} = 2^{1/3}s$, and using the boundary condition $f(+\infty) = 0$ gives

$$f(\tilde{s}) - f(\tilde{s})^{3/2} + \frac{\mathrm{d}f(\tilde{s})}{\mathrm{d}\tilde{s}} = 0 \tag{54}$$

which is the same ordinary differential equation that describes the advancing front of forced drainage and whose solution is (31). Figure 17 shows the asymptotic collapse of the bottom of the pulse for both experimental and simulated data using the self-similar ansatz (52). The agreement with simulations is good, although some scatter is present. The solution (31) is indistinguishable from the longest time simulation.

5. Concluding Remarks

We have developed a new method for measuring the liquid volume fraction of foams using fluorescence. It is possible to obtain forced drainage data spanning more than two decades in volume fraction using this technique and extract power-law behavior for the speed and width of the forced drainage wave. For free drainage we measured dynamics over ∼100 s, and for pulsed drainage we were able to track the dynamics up to ∼500 s.

A generalized foam drainage equation has been derived which takes into account the contributions to viscous damping originating from flow in the nodes as well as in the channels. The foam is modeled as a porous medium with a permeability that varies dynamically with the liquid volume fraction. The foam permeability is governed by the contributions to viscous dissipation in the channels and nodes of the liquid network.

There are two limiting cases of this generalized equation: one with a no-slip interface boundary condition, which is the assumption of the original foam drainage equation, and one with a no-stress (ideally mobile) interface. The former, where dissipation is channel-dominated, is consistent with experimental results from publications by Weaire and co-workers.[8,20,34] The latter, for which nodes dominate the dissipation, agrees very well with all the measurements presented here using SDS surfactant, as well as earlier measurements using Dawn dish detergent.[5]

The measured permeability of the foam in the present experiments has a power-law behavior somewhat different from the node-dominated model ($k \propto \epsilon^{0.6}$ rather than $k \propto \epsilon^{1/2}$). However, the deviations are small and the node-dominated limit reproduces the results of our forced, free, and pulsed drainage experiments with good accuracy, whereas the channel-dominated theory fails to explain our data. The only free parameter in our treatment of the node-dominated foam drainage model is the permeability prefactor $K_{1/2}$ which is taken from the measured dependence of the front velocity against ϵ for forced drainage; see eq 26. The complex foam geometry prevents us from directly computing $K_{1/2}$, e.g. using a boundary integral method with no-stress boundary conditions. Allowing additional dissipation in the channels would add a component to the permeability and increase the effective exponent of the permeability somewhat (see eq 19), thus coming closer to the experimental results.

The node-dominated drainage model provides a very good description of the front width of forced drainage as well as the universal shape of the ϵ profile (see (31)), onto which all experimental data can be collapsed after proper rescaling.

Free drainage experiments are well described by node-dominated drainage as well. Rescaling coordinates according to the theoretical descriptions derived from the node-dominated foam drainage equation collapses both the rear and transition ("knee") regions of free drainage profiles (see Figures 13 and 14). In the asymptotic long-time limit of free drainage, self-similar coordinates transform the PDEs into ODEs, which have exact or approximate analytical solutions. Note that here, rather than focusing on the amount of liquid drained out of the foam in free drainage, we describe the full dynamics of the liquid volume fraction profiles.

Likewise, self-similar solutions are found for the three scaling regimes of pulsed drainage profiles—the rear, middle, and front regions. The rear region has similar dynamics to the rear region of free drainage (see (38)), and the front region has a similar shape to the advancing front of pulsed drainage (see (54)). We are able to collapse the experimental data from these regimes onto universal curves for long times (see Figures 15−17). The agreement with theory is good especially when considering that, once $K_{1/2}$ is fixed, no other free parameters are used.

In conclusion, we propose a generalized foam drainage theory of which the existing models are special cases and experimentally find that foam drainage is governed by dissipation in the nodes rather than in the channels, which can be rationalized by assuming a no-stress boundary condition at the gas/liquid interface. In fact, since we are dealing with Newtonian fluids, we would expect that qualitative differences in dynamical behavior, such as different scaling laws, can *only* be attributable to changes in the boundary conditions. Recent experiments by Durand et al.[21] suggest that both limits of the generalized foam drainage equation can be approached for the same system by altering the composition of the surfactant: adding dodecanol to an SDS solution increases the surface viscosity and changes the drainage behavior from node-dominated to channel-dominated. Likewise, performing forced drainage experiments with protein surfactants, we find excellent agreement with the channel-dominated model, which has no free parameters. These results support the validity of a generalized theory of foam drainage, as it has been presented here.

Acknowledgment. We are indebted to A. G. Evans for his guidance and support through the MURI project on Ultralight Materials, ONR Grant No. N00014-1-96-1028. We thank G. McKinley for his helpful experimental suggestions and A. Kraynik, J. Sherwood, J. Eggers, and M. Brenner for many useful discussions.

LA9913147

Experimental probes of surface interactions

Frieder Mugele

- Physical principles of ion adsorption and charge regulation in electric double layers (focus on modeling starting with double layer theory, derivation of PB equation from variational energy minimization)

- Probing surface forces and surface charges using Atomic Force Microscopy (focus on experimental techniques and specific results)

- Physical principles of electrowetting (1h basic principles; 1h application examples)

○ Wednesday :	14h00-15h30
○ Thursday :	09h00-10h30
○ Friday :	09h00-10h00

SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:

ATOMIC FORCE MICROSCOPY

NANOPARTICLES

STRUCTURAL PROPERTIES

SURFACES, INTERFACES AND THIN FILMS

Received
6 January 2014

Accepted
22 April 2014

Published
22 May 2014

Correspondence and requests for materials should be addressed to F.M. (f.mugele@utwente.nl)

* These authors contributed equally to this work.

Direct observation of ionic structure at solid-liquid interfaces: a deep look into the Stern Layer

Igor Siretanu[1]*, Daniel Ebeling[1]*, Martin P. Andersson[2], S. L. Svane Stipp[2], Albert Philipse[3], Martien Cohen Stuart[1], Dirk van den Ende[1] & Frieder Mugele[1]

[1]Physics of Complex Fluids Group and MESA+ Institute, Faculty of Science and Technology, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands, [2]Nano-Science Center, Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark, [3]Van't Hoff Laboratory for Physical and Colloid Chemistry, Debye Institute, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.

The distribution of ions and charge at solid-water interfaces plays an essential role in a wide range of processes in biology, geology and technology. While theoretical models of the solid-electrolyte interface date back to the early 20th century, a detailed picture of the structure of the electric double layer has remained elusive, largely because of experimental techniques have not allowed direct observation of the behaviour of ions, i.e. with subnanometer resolution. We have made use of recent advances in high-resolution Atomic Force Microscopy to reveal, with atomic level precision, the ordered adsorption of the mono- and divalent ions that are common in natural environments to heterogeneous gibbsite/silica surfaces in contact with aqueous electrolytes. Complemented by density functional theory, our experiments produce a detailed picture of the formation of surface phases by templated adsorption of cations, anions and water, stabilized by hydrogen bonding.

Gouy[1], Chapman[2] and Stern[3] laid the foundation for our understanding of the electric double layer by describing the distribution of ions in the vicinity of charged interfaces using Poisson-Boltzmann theory. The classical approach has been refined in many respects, including a variety of sometimes competing microscopic effects, such as preferential binding to specific surface sites[4,5], dispersive ion-substrate interactions[6] and ion correlation effects[7]. More recently, molecular simulations have contributed additional insight, e.g. about the hydration of ions and surfaces. In comparison, common experimental methods such as batch titrations, electrokinetic and surface force measurements provide less direct information on the atomic scale. They integrate laterally over rather large and frequently very heterogeneous surface areas and rely on a large number of assumptions and empirical parameters to fit to theoretical models. Also, along the direction normal to the surface, these techniques average information and attribute it to several of the levels in the electric double layer, based on conceptual model assumptions. It is increasingly recognized[8,9] that quantitative understanding of mineral-fluid interface behaviour is limited because experimental techniques have not been able to capture the complex structure of solid-liquid interfaces with resolution at nanometre scale, parallel and perpendicular to surfaces.

Atomic Force Microscopy (AFM) has recently been advanced to a stage that allows for imaging solid-liquid interfaces at 'true' atomic resolution[10–13]. We have used small amplitude dynamic AFM to explore the surfaces of synthetic nanoparticles of gibbsite (α-Al(OH)$_3$)[14] during exposure to a variety of electrolyte solutions. We chose gibbsite because it can be synthesised reproducibly, to yield suspensions of essentially monodispersed particles. Moreover gibbsite is a good model for some clay mineral surfaces[15]. Sorption of inorganic and organic ions to Al (hydr)oxides, such as gibbsite, and to clay minerals is important for the transport of contaminants and nutrients in the environment and kaolinite, a clay with one Al-OH surface has been reported to play a role in enhancing oil recovery[16,17]. It has long been assumed that the doubly coordinated Al$_2$OH groups on gibbsite basal planes are inactive to deprotonation/protonation reactions and that surface charge and ion sorption are dominated by the singly coordinated aluminol at edges[18]. Recently however, experimental[8,19–21] and numerical studies[21,22] have suggested that missing information about structure at the submicrometre scale and the ratio of edge to basal surface area might have compromised data interpretation.

With small amplitude dynamic AFM, we have collected the required high resolution insight needed for addressing these questions, to directly "see" the structure of the ions adsorbed in the Stern layer and to observe changes in the pattern on the gibbsite basal plane as the contacting solution is changed. We characterised the surfaces at two levels. First, we used AFM spectroscopy at tip-sample distances of a few nanometres during exposure to solutions with a range of concentrations. This provides data on effective surface charge, similar to those obtained by ζ-potential measurements. Next, we recorded atomic scale images at much smaller distances, which gives a direct view of ion distribution within the Stern layer. Finally, using density functional theory, we could confirm the stability of the ordering observed by AFM and gain additional insight into the nature of the bonding and how charge in the Stern layer changes with solution composition.

Results

The gibbsite particles were deposited from a water-ethanol solution (details in Methods Section) on silica wafers that had oxidised in air to produce amorphous SiO_2. The nanoparticles naturally sorb with their {001} basal plane adjacent to the silica surface, exposing a Al-OH surface to the solution. Typically, gibbsite particles attached singly. Lateral dimensions ranged from several 10 to a few 100 nanometres and heights, from 1 to 20 nm (Fig. 1a). All experiments were performed in slightly acidic (pH $\sim$ 6) aqueous electrolyte and we used silicon AFM tips. Tip surfaces had oxidised to amorphous silica so they had the same character as the silica substrate.

We monitored the effective surface charge of the particles by measuring the interaction force between the tip and sample as a function of distance, in frequency modulation force spectroscopy mode[23] (FM-AFM). Two dimensional interaction force maps[24] (colour coded in Fig. 1b) confirm that on the silica substrate, force increases from zero (green) to repulsive values (red) of several hundred picoNewtons at a distance of several nanometers, as expected for two negatively charged surfaces in pure water. Over the gibbsite particles however, attractive force (blue) indicates positive charge. Compared with the silica substrate, there is more lateral variation in the force on the gibbsite particles, indicating a larger degree of heterogeneity. Force decreases toward the particle edges. Force profiles (Fig. 1b) also reveal the location of occasional crystal defects. The local minimum in the attractive force near the center of this specific particle is caused by a twin boundary. This is most easily seen in 2D frequency shift images (Fig. S1b) that show a direct, qualitative measurement of the interaction forces. The minimum force indicates that the effective local surface charge essentially vanishes close to the crystal defect. Typical tip radii of 20–30 nm in the spectroscopy experiments imply a lateral average of a few thousand surface unit cells.

Atomically resolved amplitude modulation images of the basal planes display the periodicity of the gibbsite lattice (Fig. 1c). Close to particle edges, we typically observe a higher density of atomic steps. Frequently, these steps are decorated by adsorbed material (Fig. 1c). Such defects are an important source of charge heterogeneity on gibbsite surfaces. Figure 2a shows line representations of force spectroscopy data for areas such as Figure 1b under several concentrations of NaCl and $CaCl_2$. Each data set was obtained with the same cantilever and sample and care was taken to guarantee that tip shape did not change when solutions were changed (see Supplementary Information). On silica, the force curves (red in Fig. 2a) from a number of sites collapse into a single narrow band for each ion concentration. The interaction curves for gibbsite are more widely spread, with rather weak forces along particle edges (green) and strong attraction in the centre (blue). Our next discussion focusses on the forces in the centre.

The qualitative trends in Figure 2a follow those expected from standard electrostatic screening, i.e. force decreases as salt concentration increases and the absolute force for divalent ions at the same concentration is lower than for monovalent ions. To determine surface charge, we compare the force curves to predictions from DLVO (Derjaguin-Landau-Verwey-Overbeek) theory[23] for electrostatic and van der Waals forces (Supplementary Information). Consistent with expectations, forces measured at small separations lie between the two limiting cases of constant charge and constant potential because of confinement induced charge regulation[25]. However, from the asymptotic regime at large separation, we can readily extract unique values for the effective surface charge, σ_{eff}, for both the tip and sample[24]. For the monovalent salts, σ_{eff} on silica increases with increasing salt concentration, whereas for the divalent salts, it remains constant within experimental error (Fig. 2b). This trend for monovalent salts agrees with the expected enhanced deprotonation of silanol groups on the silica surface: $\equiv \mathrm{SiOH} \rightarrow \mathrm{SiO}^- + \mathrm{H}^+$ as electrostatic screening increases. Fitting the data with a basic Stern model (BSM)[25] yields pKa $\sim$7.5 for silanol deprotonation, in good agreement with literature data[25–29] (black line in Fig. 2b). This supports the effectiveness of our measurement and data analysis procedure. Weakly negative and essentially constant surface charge on the amorphous silica surface in contact with Ca^{2+} and Mg^{2+} has previously been interpreted in terms of cation adsorption[27,30].

On gibbsite, σ_{eff} was positive under all investigated conditions. In solutions of monovalent salts, it increases monotonically from $\sim$0.03 to $\sim$0.1 e/nm^2 as salinity increases. The surface unit cell has an area of $\sim$0.44 nm^2 so these absolute values imply that at most, a few percent of the unit cells carry a net charge. A more intriguing behaviour is observed for the divalent cations. Initially, σ_{eff} increases strongly with increasing salinity, reaching a maximum at 5 to 10 mM and then decreases to negligible values as concentration reaches 100 mM. A slight but consistent specific ion effect was observed in three separate experiments. In $CaCl_2$ solutions, maximum charge is higher and it occurs at somewhat lower concentration than in $MgCl_2$. While the constant increase in σ_{eff} for monovalent salts could be interpreted to result from protonation

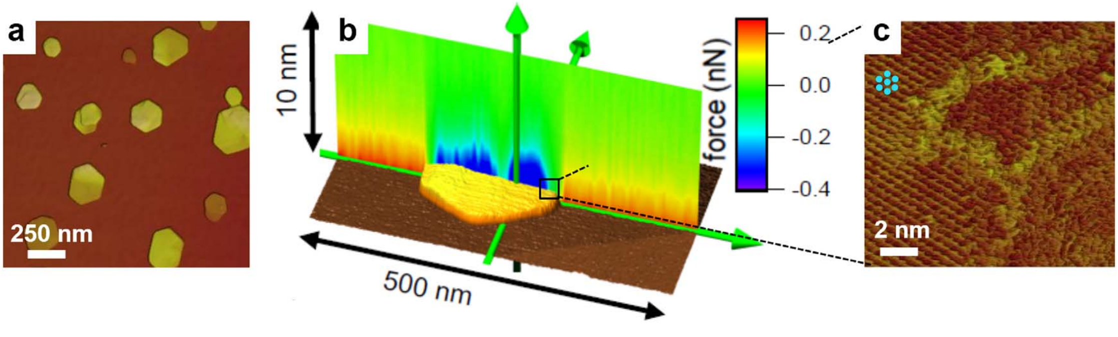


Figure 1 | Atomic force microscopy (AFM) of gibbsite nanoparticles. (a), Topography images of gibbsite on an oxidised silicon wafer. (b), color-coded 2D force field generated from 100 tip-sample interaction curves in 20 mM NaCl at pH $\approx$ 6. (blue: attractive force; red: repulsive force; green: zero force; see scale bar) (c), Amplitude modulation atomic resolution image of a gibbsite particle in ultrapure deionised water. Left part: pseudohexagonal basal plane structure (surface unit cell, a = 0.87 nm, b = 0.50 nm); centre: atomic step disorder on terrace edges; bottom right: edge of the particle.

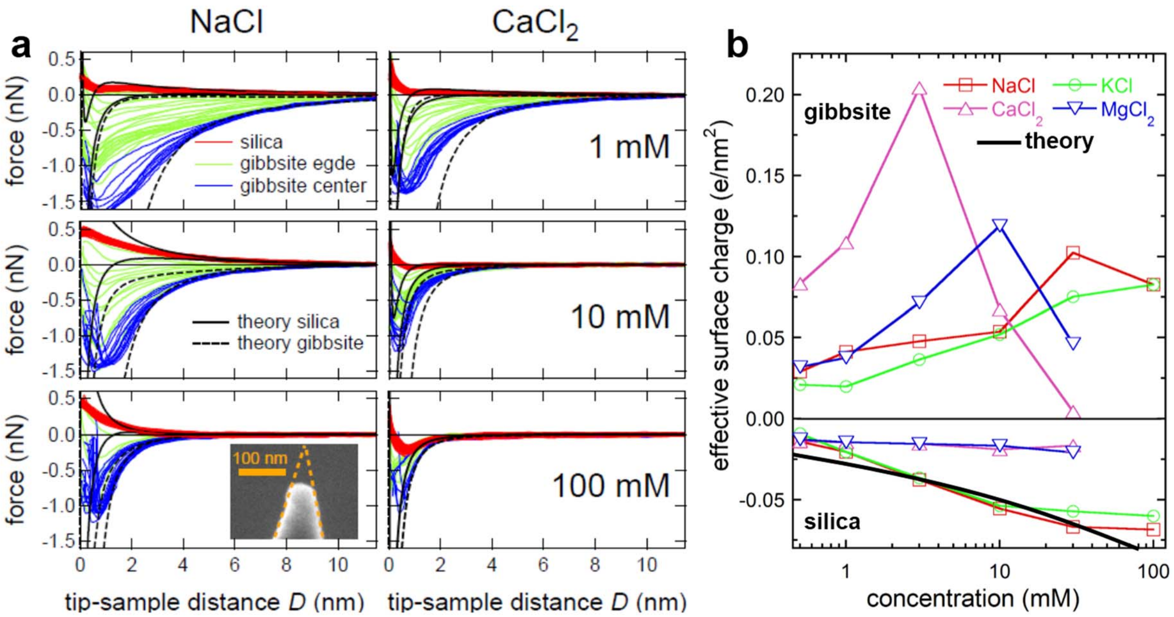


Figure 2 | Electrical properties of amorphous SiO_2 and gibbsite measured with FM-AFM. (a), Force vs distance curves measured over a gibbsite nanoparticle sorbed on oxidised silicon wafers in 1, 10 and 100 mM NaCl and $CaCl_2$ solutions. Red curves: tip on silica substrate. Green: edge of gibbsite particle. Blue: centre of gibbsite particle. Lines (solid: silica; dashed gibbsite): tip sample interaction force according to DLVO theory for constant charge (top) and constant potential (bottom) boundary conditions. Inset: SEM image of AFM tip after the experiment. (CFM Aspire tip, with parameters of silicon cantilever $f_0 = 22.9$ kHz, $c_z = 5.0$ N/m, Q = 9). (b), Surface charge as function of solution composition (pH $\approx$ 6). Symbols: experimental data. Solid black line: best fit assuming deprotonation of silanol groups in monovalent salt solutions (see text for details).

facilitated by improved screening, as we see for silica[9,31,32], the behaviour of divalent cations is more complex. (Fitting the data for the monovalent salts in terms of a simple surface speciation model involving protonation of doubly coordinated Al_2OH groups at low pH yields a pKa value of ~7 and a density of one active group per surface unit cell, reasonably consistent with recent models of the gibbsite surface). The increase and decrease suggests the presence of two separate processes. The first process, dominant at lower salt concentrations, enhances the already positive effect of surface charge. The second reduces it again. Obviously, the first process cannot be driven by electrostatic forces, the second one might be.

At this stage, it is tempting to invoke possible adsorption/desorption reactions to explain Figure 2b. The rather low absolute value of σ_{eff} is consistent with general understanding, that the Al-OH gibbsite basal plane is indeed chemically rather inactive[18–20]. However, atomic force spectroscopy, just as electrokinetic measurements, probes the charge in the diffuse part of the electric double layer. These techniques might be too indirect to deliver a detailed picture of the complex chemical processes that take place at the solid-liquid interface. To overcome this limitation, we imaged the gibbsite surface at atomic resolution under several electrolytes (details in Methods). Figure 3a shows the typical pseudohexagonal pattern of the gibbsite basal plane, imaged under deionized water. The pattern is caused by the arrangement of the octahedral cavities with next neighbour spacing of ~0.5 nm, consistent with dimensions of the surface unit cell with dimensions $a = 0.868$ nm and $b = 0.507$ nm (Fig. 3d), as obtained by x-ray diffraction. Except for an occasional contrast inversion (Fig. S3), which we attribute to loss of true atomic resolution[12], symmetry, contrast and the resolution of the pattern remain unchanged when the water is replaced by solutions of KCl or NaCl. From the absence of changes in surface topography, we conclude that neither the monovalent cations nor Cl^- adsorbs strongly to the gibbsite surface. Ions could be weakly adsorbed and pushed away by the AFM tip, as has been discussed for mica in contact with electrolyte solutions[12,33–35]. Nonadsorption of monovalent ions is completely consistent with protonation as an explanation for the increase in effective surface charge, discussed above.

In stark contrast to behaviour in monovalent salt solutions, gibbsite appearance changes dramatically when the solution is replaced with 10 mM $CaCl_2$ or $MgCl_2$ (Fig. 3b and Fig. S4b). The pseudohexagonal pattern gives way to an array of double rows aligned along the *b* direction (Fig. 3b and e). Each double row consists of alternating bumps. The periodicity along and perpendicular to the double rows is 0.50 nm and 0.87 nm (Fig. 3f), in excellent agreement with the surface atomic structure. There are thus two bumps per surface unit cell, which we interpret to be (possibly hydrated) ions adsorbed from solution.

As we increase the concentration of $CaCl_2$ to 100 mM, we observe a second change in the appearance of the surface. The double rows give way to single rows spaced by one lattice vector along the *b* direction and with one bump per surface unit cell along the *a* direction (Fig. 3c). In between two adjacent rows, a second row of bumps is sometimes seen, typically at much fainter contrast. The same behaviour is observed when gibbsite is exposed to $MgCl_2$ solutions (Fig. S4c). At intermediate concentrations ($\approx$50 mM), we see coexisting domains of double rows and of alternating bright-faint rows (Fig. 4). This suggests two distinct two dimensional adsorbed phases.

At this stage, we can already conclude that the gibbsite basal plane is by no means chemically inactive. Rather than occasional reaction of a few percent of the surface unit cells, as suggested by the low value of σ_{eff} and generally assumed in the literature[5,18–20], our images show that every unit cell accepts at least two adsorbed ions, where the bond is strong enough that it is not pushed away by the tip. The concurrent appearance of the (positive) maximum in σ_{eff} and the double rows in the high resolution images suggests that both phenomena result from adsorption of the same type of ion. Because Cl^- does not affect the surface pattern, even at concentrations of 100 mM NaCl or KCl, we conclude that the double rows must be caused by divalent cation adsorption. The agreement of the measured periodicities of the double row structure with the surface unit cell dimensions suggests bonding to well defined adsorption sites, rather than electrostatic correlation between ions[36]. To identify the adsorption sites, we can

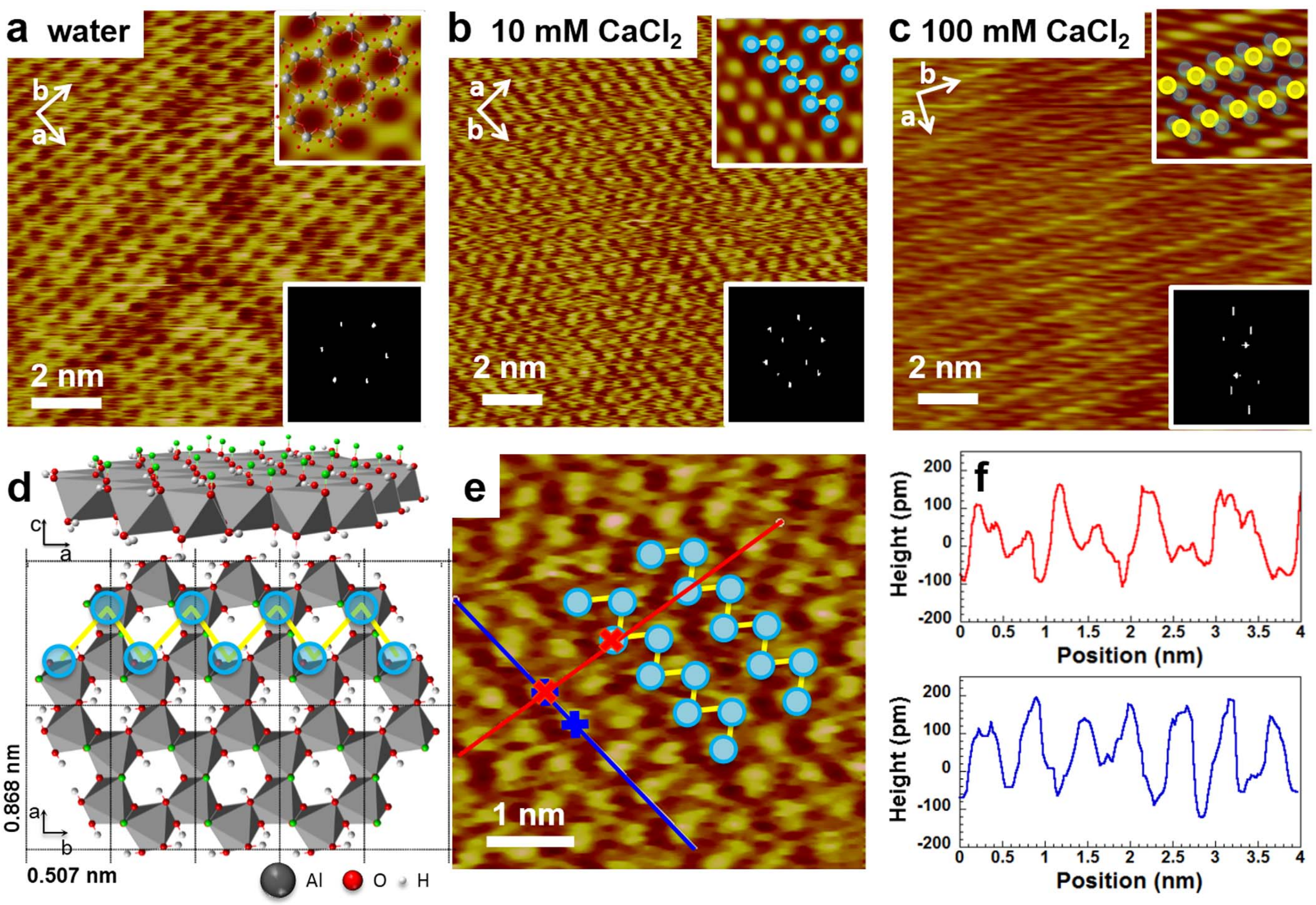


Figure 3 | Atomic resolution AFM images of gibbsite. (a), AFM topographic image of gibbsite basal plane in ultrapure deionised water. Insets: zoomed and Fourier-filtered view with superimposed crystallographic lattice (top); 2D fast Fourier transform of image. (b), same type of data recorded in 10 mM and (c), 100 mM $CaCl_2$ solution. (d), Crystal structure of gibbsite in *ac* and *ab* planes. H atoms pointing perpendicular to the *ab* plane are shown in green. (e), A zoom view of b with schematic indication of position of adsorbates and location of the height profiles in *a* (red) and *b* (blue) directions shown in f. Height profiles in *a* (red) and *b* (blue) directions display periodicities of 0.87 nm and 0.50 nm, respectively.

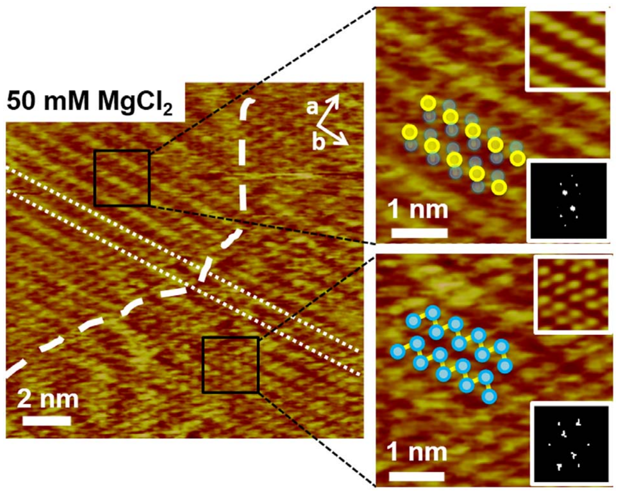


Figure 4 | Gibbsite imaged in 50 mM $MgCl_2$ showing phase-separated domains with double row structure (bottom right) and single row structure (top left) characteristic of low and high salt concentrations. The area to the right of the white dashed line has equivalent height double rows in a zig zag pattern with the same periodicity as Figure 3e and all images obtained under 10 mM $CaCl_2$ or $MgCl_2$ solutions. Left of the dashed line, the rows alternate in height, as observed for all of the surfaces imaged under solutions of 100 mM $CaCl_2$ or $MgCl_2$.

look more closely at the surface structure. The gibbsite surface unit cell has six chemically inequivalent Al_2-OH moieties. Simulations suggest that deprotonation of these sites covers a rather wide range of pK_a[37]. Three of them are located around the central octahedral cavity and point toward the solution (small green dots in Fig. 3d). These OH groups are available for interlayer hydrogen bonding in the bulk gibbsite structure[38] and for hydrogen bonding to adsorbates at the surface[39–41]. Attachment at these sites would produce the observed dimensions and zig zag pattern (Fig. 3d).

The simultaneous decrease of surface charge and change in pattern appearance at higher concentrations suggest adsorption of Cl^- ions. While there is no evidence for Cl^- adsorption on gibbsite, chloride interaction with adsorbed Ca^{2+} and Mg^{2+} could promote attachment. As concentrations increase, both Ca^{2+} and Mg^{2+} form ion pairs with Cl^- so pairing on surfaces is consistent. Chloride adsorption has recently been reported in molecular dynamics simulations of smectite-electrolyte interfaces[42].

The adsorption of two divalent cations per unit cell without any compensation of charge through surface deprotonation or coadsorption of anions corresponds to a hypothetical surface charge of 9.2 e/nm^2. This is inconsistent with the low values of σ_{eff} (Fig. 2b), which correspond to less than one elementary charge per unit cell. Substantial deviations between surface charge determined by macroscopic methods such as titration and values obtained from the diffuse layer, for example by electrokinetic or force measurements, are not uncommon[43]. They are generally attributed to uncertainties in the exact location of the shear plane in electrokinetic measurements and the mobility of weakly adsorbed ions. The mismatch in charge density could originate from surface deprotonation or adsorbed anions, that contribute to the effective surface charge in spectroscopy experiments but that are too mobile to remain localized under high-resolution imaging.

For a more detailed analysis of bonding tendencies and to help explain the surface charge behaviour, we used density functional theory (DFT) to examine the adsorption of Ca^{2+}, Mg^{2+} and Cl^- onto the Al-OH basal plane of gibbsite. We use the COSMO-RS implicit solvent model with periodic boundary conditions to calculate the equilibrium structure of the adsorbed divalent cations for both outer and inner sphere configurations, i.e. with or without water of hydration between the ion and the surface (details in Methods and Supplementary Information). In both cases, stable zig-zag double rows were found. However, only formation of an outer sphere configuration, containing enough hydration water to retain the average bulk ion-water coordination number of six, was exothermic. The formation energies for the divalent ion structures were −118 kJ/mol/$Ca(OH)_2$ and −115 kJ/mol/$Mg(OH)_2$ (Table S1), Figures 5a and S5 show the equilibrium, outer shell configurations that excellently reproduce the experimentally observed double row structure, with alternating adsorption sites. Three of four hydroxyl groups, added to guarantee charge neutrality, act as hydrogen bonding acceptors for surface protons. The fourth OH^- bridges between the two cations. It is interesting that the fourth hydroxyl causes a slight asymmetry in the zig-zag, which is compatible with the experimental data (Supplementary Information, Fig. S6a, where structure from Fig. 5a is superimposed on the AFM image).

Although the surface unit cell is charge neutral, our model offers an interesting explanation of the slight positive surface charge at intermediate salinities. The alternating structure of hydrated divalent cations offers several sites where hydration water and OH^- bridge between two cations. Water adsorbed on similar sites on calcite surfaces is significantly more acidic than bulk water, with pK_a as low as 3 to 4[44]. Additional COMSO-RS DFT calculations for clusters of about 200 atoms, beginning with the converged solution of the periodic calculation, allowed us to calculate pK_a of 10.2 and 4.9 for $H_2O \rightarrow OH^-$ deprotonation for the adsorbed Ca^{2+} and Mg^{2+} structure. These values suggest that the positive charge in the spectroscopy measurements results from partial protonation of hydroxyl that bridges adjacent cations from solution. The pK_a for Mg^{2+} adsorption is lower than for Ca^{2+}, implying that OH^-, and hence the electroneutral configuration, is somewhat more stable for Mg^{2+}, in agreement with the experiments, which show that the maximum charge for the Mg^{2+} structure is always lower than for the Ca structure, Figure 2b and Figure S2.

Finally, we calculated the equilibrium configurations of the adsorbed cations where one Cl^- ion per unit cell replaced the bridging hydroxyl ion. Chloride also bridges adjacent cations, slightly shifted towards the pseudo threefold cavity (Fig. 5b). The vertical position is 210 pm above the plane, averaged over the metal cations (cf. Fig. S6b). This ion exchange disables OH^- protonation and results in a neutral surface structure, which explains the decrease in σ_{eff} at high salinity. The OH^- vs. Cl^- exchange energies are +39 kJ/mol and +47 kJ/mol for the Ca^{2+} and Mg^{2+} structures (Table S2). For $pH = 6$, this implies characteristic concentrations of 30 mM for $CaCl_2$ and 900 mM for $MgCl_2$ to induce the exchange reaction. These values are in very good agreement with the experimental data and even explain the slight shift of maximum σ_{eff} toward higher Mg^{2+} concentrations, compared with Ca^{2+}, in Figure 2b.

In conclusion, the combination of AFM spectroscopy, high resolution imaging and numerical simulations provides unprecedented insight into the complex processes involved in the formation of the electric double layer on mineral surfaces. By resolving the internal structure of the Stern layer we demonstrate a strong affinity for divalent cations of a type of surface that has long been assumed to be chemically inactive. For the specific case of gibbsite, the resulting changes in surface chemistry have important consequences for the role of Al-OH bearing mineral surfaces in modern technologies for enhanced oil recovery.

Methods

Sample preparation. Gibbsite synthesis is described in detail by Wierenga et al.[14] We diluted the gibbsite stock suspension 100 times in a 1 : 1 mixture of ultrapure

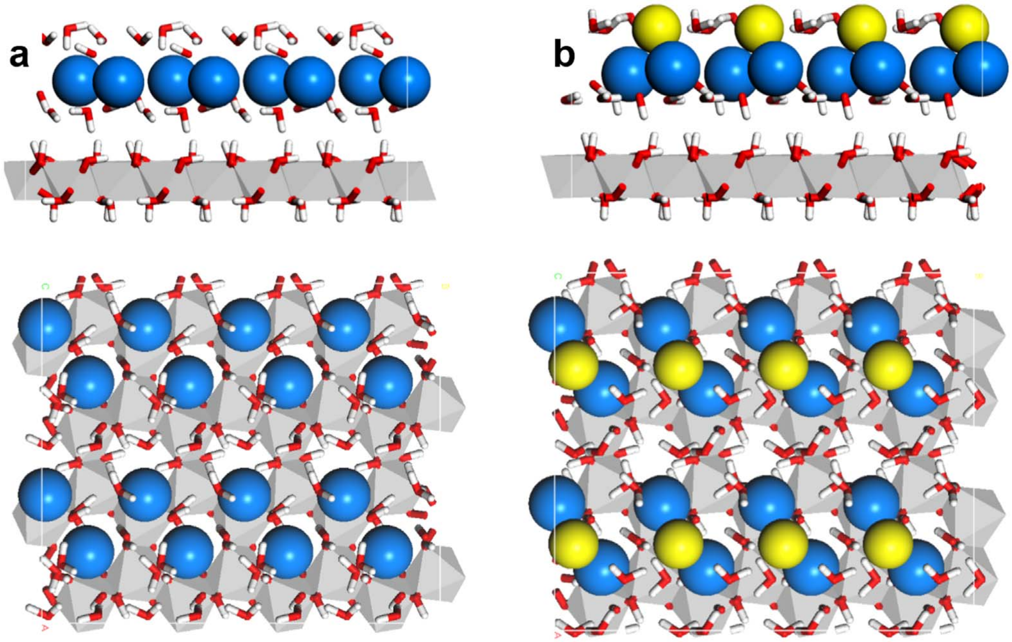


Figure 5 | Equilibrium structure of adsorbed Ca^{2+} (blue) and Cl^- (yellow) on the gibbsite basal plane in contact with aqueous solution predicted by DFT calculations. Red and white: oxygen and hydrogen; gray: Al-O octahedral. (a), Side and top view of the optimized geometry for outer shell adsorption of Ca^{2+} (blue) on gibbsite. A 2 × 2 supercell of our simulation cell is shown for clarity. (b), At higher concentrations of $CaCl_2$. Adsorption plane of Cl^- is 0.22 nm above Ca^{2+}.

deionised water (Milli-Q) and ethanol and we deposited 10 ml on freshly cleaned silica substrates. After 30 s, they were rinsed with copious amounts of deionised water and blown dry with air.

Atomic Force Microscopy (AFM). For atomic resolution, we used a Multimode8 AFM (Bruker Nano) equipped with Nanoscope V controller and an A scanner, operated in tapping mode, mostly with Bruker FASTSCAN-B cantilevers ($f_0 =$ 170 kHz, $c_z = 3$ N/m, $Q = 10$). Controls with Olympus AC-55, Aspire CFM ($f_0 =$ 22.9 kHz, $c_z = 5.0$ N/m, $Q = 9$) and CT130 probes yielded similar results. Before use, tips were cleaned by rinsing with a mixture of ethanol and isopropanol ($\approx 1:1$) and air plasma treatment (Harrick Plasma) for 15–30 min. A standard tapping mode liquid probe holder without O-ring (Bruker Nano) was used for imaging. To minimize drift, the system electronics were allowed to equilibrate for 20–60 min before data were acquired. The AFM was operated in amplitude-modulation mode with free amplitude, A_0, typically less than 2 nm, high scan rate, $\approx$10 Hz, and imaging amplitude set point as high as possible, typically $A/A_0 \geq 0.8$. All images were flattened using Bruker's standard Nanoscope Analysis 1.4 package, including, in some cases slight low pass filtering to improve clarity.

AFM spectroscopy measurements were performed with a Dimension Icon AFM (Bruker Nano) equipped with Nanoscope V controller which does not use a liquid cell exciting "the whole chip", but rather a direct excitation of the cantilever as in usual dynamic AFM in ambient air[45]. Additional drive electronics (QFM-Module, NanoAnalytics GmbH, Germany) was used to operate the system in constant excitation (CE)[46] version of the frequency modulation technique and measure the frequency shift of the oscillating cantilever, since it is known to be more robust, especially for liquid applications, because it does not require an additional feedback loop which keeps the oscillation amplitude constant (as in the constant amplitude mode). Spectroscopy measurements were performed with rectangular silicon cantilevers with conical tips (CFM and CT130, Aspire), using the standard direct drive liquid probe holder and 60 mm glass petri dishes for the samples. The petri dishes, tips and silica sample substrates were rinsed with isopropanol, ethanol and MilliQ water before cleaning with air plasma for 15–30 min. To minimize changes in the tip apex during the spectroscopy measurements, we did not allow the amplitude signal to drop below $\sim$70% of its value far away from the surface (free amplitude ≈ 2 nm) by setting a threshold. Tip-sample forces are calculated from the amplitude and frequency shift vs distance curves as described in elsewhere[47]. The measured interaction forces between tip and sample surface are converted to surface charge using Poisson-Boltzmann theory, taking into account the actual tip geometry[23] (Supplementary Information).

Computational details. Periodic density functional theory (DFT) calculations were performed using the program DMol3 with the COSMO-RS implicit solvent[48], the PBE functional, the DNP basis set and a dispersion correction[49]. We used a 1×2 gibbsite basal plane unit cell with lattice parameters (0.86840×1.01560 nm) defined by x-ray diffraction. Periodic slab calculation included three molecular layers, of which the lowest was frozen during all optimisations. Calculations for predicting pK_a for water binding to the adsorbed cations were performed with a cluster of gibbsite {001} containing 204 atoms. The cluster was terminated with hydrogen at the Al-OH edges to ensure electroneutrality for the structure with adsorbed OH^-. Further description of the experimental and theoretical details is provided in Supplementary Information.

1. Guoy, A. Sur la constitution de la charge electricque a la surface d'un electrolyte. *J. Phys. Theor. Appl.* **9**, 457–468 (1910).
2. Chapman, D. L. A contribution to the theory of electrocapillarity. *Philos. Mag.* **25**, 475–481 (1913).
3. Stern, O. The theory of the electrolytic double-layer. *Z. Electrochem. Angew. P.* **30**, 508 (1924).
4. Hiemstra, T., Venema, P. & Van Riemsdijk, W. H. Intrinsic proton affinity of reactive surface groups of metal (Hydr)oxides: The bond valence principle. *J. Colloid Interface Sci.* **184**, 680–692 (1996).
5. Sposito, G. *The Environmental Chemistry of Aluminum.* Vol. II (CRC Press, Berkeley, 1995).
6. Parsons, D. F., Boström, M., Maceina, T. J., Salis, A. & Ninham, B. W. Why direct or reversed hofmeister series? interplay of hydration, non-electrostatic potentials, and ion size. *Langmuir* **26**, 3323–3328 (2010).
7. Ben-Yaakov, D., Andelman, D., Harries, D. & Podgornik, R. Beyond standard Poisson-Boltzmann theory: Ion-specific interactions in aqueous solutions. *J. Phys.: Condens. Matter* **21**, 424106 (2009).
8. Adekola, F. *et al.* Characterization of acid-base properties of two gibbsite samples in the context of literature results. *J. Colloid Interface Sci.* **354**, 306–317 (2011).
9. Jodin, M. C., Gaboriaud, F. & Humbert, B. Repercussions of size heterogeneity on the measurement of specific surface areas of colloidal minerals: Combination of macroscopic and microscopic analyses. *Am. Mineral.* **89**, 1456–1462 (2004).
10. Fukuma, T., Kobayashi, K., Matsushige, K. & Yamada, H. True atomic resolution in liquid by frequency-modulation atomic force microscopy. *Appl. Phys. Lett.* **87**, 034101 (2005).
11. Voïtchovsky, K., Kuna, J. J., Contera, S. A., Tosatti, E. & Stellacci, F. Direct mapping of the solid-liquid adhesion energy with subnanometre resolution. *Nat. Nanotechnol.* **5**, 401–405 (2010).
12. Loh, S. H. & Jarvis, S. P. Visualization of ion distribution at the mica-electrolyte interface. *Langmuir* **26**, 9176–9178 (2010).
13. Rode, S. *et al.* Atomic-resolution imaging of the polar (0001) surface of $LiNbO_3$ in aqueous solution by frequency modulation atomic force microscopy. *Physical Review B* **86**, 075468 (2012).
14. Wierenga, A. M., Lenstra, T. A. J. & Philipse, A. P. Aqueous dispersions of colloidal gibbsite platelets: synthesis, characterisation and intrinsic viscosity measurements. *Colloids Surf., A* **134**, 359–371 (1998).
15. Verhoeff, A. A. & Lekkerkerker, H. N. W. Direct observation of columnar liquid crystal droplets. *Soft Matter* **8**, 4865–4868 (2012).
16. Somerton, W. H. & Radke, C. J. Role of Clays in the Enhanced Recovery of Petroleum from Some California Sands *J. Pet. Technol.* **35**, 643–654 (1983).
17. Kumar, N., Wang, L., Siretanu, I., Duits, M. & Mugele, F. Salt dependent stability of stearic acid Langmuir-Blodgett films exposed to aqueous electrolytes. *Langmuir* **29**, 5150–5159 (2013).
18. Hiemstra, T., Yong, H. & Van Riemsdijk, W. H. Interfacial charging phenomena of aluminum (Hydr)oxides. *Langmuir* **15**, 5942–5955 (1999).
19. Rosenqvist, J., Persson, P. & Sjöberg, S. Protonation and charging of nanosized gibbsite (α-Al(OH)3) particles in aqueous suspension. *Langmuir* **18**, 4598–4604 (2002).
20. Gan, Y. & Franks, G. V. Charging behavior of the gibbsite basal (001) surface in NaCl solution investigated by AFM colloidal probe technique. *Langmuir* **22**, 6087–6092 (2006).
21. Jodin, M. C., Gaboriaud, F. & Humbert, B. Limitations of potentiometric studies to determine the surface charge of gibbsite γ-$Al(OH)_3$ particles. *J. Colloid Interface Sci.* **287**, 581–591 (2005).
22. Bickmore, B. R., Rosso, K. M. & Mitchell, S. C. *Surface Complexation Modelling, Chapter 9: Is there hope for multi-site complexation (MUSIC) modeling?* Vol. 11 269–283 (Elsevier, 2006).
23. Ebeling, D., Van Den Ende, D. & Mugele, F. Electrostatic interaction forces in aqueous salt solutions of variable concentration and valency. *Nanotechnology* **22**, 305706 (2011).
24. Hölscher, H., Langkat, S. M., Schwarz, A. & Wiesendanger, R. Measurement of three-dimensional force fields with atomic resolution using dynamic force spectroscopy. *Appl. Phys. Lett.* **81**, 4428–4430 (2002).
25. Behrens, S. H. & Borkovec, M. Electrostatic Interaction of Colloidal Surfaces with Variable Charge. *J. Phys. Chem. B* **103**, 2918–2928 (1999).
26. Hiemstra, T., Van Riemsdijk, W. H. & Bolt, G. H. Multisite proton adsorption modeling at the solid/solution interface of (hydr)oxides: A new approach. I. Model description and evaluation of intrinsic reaction constants. *J. Colloid Interface Sci.* **133**, 91–104 (1989).
27. Dove, P. M. & Craven, C. M. Surface charge density on silica in alkali and alkaline earth chloride electrolyte solutions. *Geochim. Cosmochim. Acta* **69**, 4963–4970 (2005).
28. Dishon, M., Zohar, O. & Sivan, U. From repulsion to attraction and back to repulsion: The effect of NaCl, KCl, and CsCl on the force between silica surfaces in aqueous solution. *Langmuir* **25**, 2831–2836 (2009).
29. Labbez, C., Jonsson, B., Skarba, M. & Borkovec, M. Ion-Ion Correlation and Charge Reversal at Titrating Solid Interfaces. *Langmuir* **25**, 7209–7213 (2009).
30. Kosmulski, M. Standard enthalpies of ion adsorption onto oxides from aqueous solutions and mixed solvents. *Colloids Surf., A* **83**, 237–243 (1994).
31. Franks, G. V. & Gan, Y. Charging behavior at the alumina-water interface and implications for ceramic processing. *J. Am. Ceram. Soc.* **90**, 3373–3388 (2007).
32. Bickmore, B. R., Rosso, K. M., Nagy, K. L., Cygan, R. T. & Tadanier, C. J. Ab initio determination of edge surface structures for dioctahedral 2 : 1 phyli-osilicates: Implications for acid-base reactivity. *Clays Clay Miner.* **51**, 359–371 (2003).
33. Kobayashi, K. *et al.* Visualization of hydration layers on muscovite mica in aqueous solution by frequency-modulation atomic force microscopy. *J. Chem. Phys.* **138**, 184704 (2013).
34. Ebeling, D. & Solares, S. D. Amplitude modulation dynamic force microscopy imaging in liquids with atomic resolution: Comparison of phase contrasts in single and dual mode operation. *Nanotechnology* **24**, 135702 (2013).
35. Kilpatrick, J. I., Loh, S. H. & Jarvis, S. P. Directly probing the effects of ions on hydration forces at interfaces. *J. Am. Chem. Soc.* **135**, 2628–2634 (2013).
36. Lyklema, J. Overcharging, charge reversal: Chemistry or physics? *Colloids Surf., A* **291**, 3–12 (2006).
37. Bickmore, B. R., Tadanier, C. J., Rosso, K. M., Monn, W. D. & Eggett, D. L. Bond-valence methods for pKa prediction: Critical reanalysis and a new approach. *Geochim. Cosmochim. Acta* **68**, 2025–2042 (2004).
38. Vyalikh, A., Zesewitz, K. & Scheler, U. Hydrogen bonds and local symmetry in the crystal structure of gibbsite. *Magn. Reson. Chem.* **48**, 877–881 (2010).
39. Tunega, D., Haberhauer, G., Gerzabek, M. H. & Lischka, H. Theoretical study of adsorption sites on the (001) surfaces of 1 : 1 clay minerals. *Langmuir* **18**, 139–147 (2002).
40. Tunega, D., Gerzabek, M. H. & Lischka, H. Ab initio molecular dynamics study of a monomolecular water layer on octahedral and tetrahedral kaolinite surfaces. *J. Phys. Chem. B* **108**, 5930–5936 (2004).
41. Hu, X. L. & Michaelides, A. Water on the hydroxylated (001) surface of kaolinite: From monomer adsorption to a flat 2D wetting layer. *Surf. Sci.* **602**, 960–974 (2008).

42. Bourg, I. C. & Sposito, G. Molecular dynamics simulations of the electrical double layer on smectite surfaces contacting concentrated mixed electrolyte (NaCl-CaCl$_2$) solutions. *J. Colloid Interface Sci.* **360**, 701–715 (2011).
43. Lyklema, J. J. *Fundamentals of Interface and Colloid Science: Solid-Liquid Interfaces.* Vol. 2 (Elsevier, 1995).
44. Andersson, M. P. & Stipp, S. L. S. How acidic is water on calcite? *J. Phys. Chem. C* **116**, 18779–18787 (2012).
45. Beer, S. d., Ende, D. v. d. & Mugele, F. Dissipation and oscillatory solvation forces in confined liquids studied by small-amplitude atomic force spectroscopy. *Nanotechnology* **21**, 325703 (2010).
46. Hölscher, H., Gotsmann, B. & Schirmeisen, A. Dynamic force spectroscopy using the frequency modulation technique with constant excitation. *Phys. Rev. B, PRB* **68**, 1534011–1534014 (2003).
47. Ebeling, D. & Hölscher, H. Analysis of the constant-excitation mode in frequency-modulation atomic force microscopy with active Q-Control applied in ambient conditions and liquids. *J. Appl. Phys.* **102**, 114310 (2007).
48. Klamt, A. & Schüürmann, G. COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans.* **2**, 799–805 (1993).
49. Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **27**, 1787–1799 (2006).

Acknowledgments

The authors thank M. Borkovec and J.J. Lyklema for insightful discussions. We also thank Kislon Voitchovsky for providing useful hints on atomic resolution measurements, L.Verhoeff for providing a gibbsite sample for preliminary AFM experiments and Ian Collins and the BP Pushing Reservoir Limits team for discussion. Financial support was provided through the *Exploratory Research* (*ExploRe*) program of BP plc.

Author contributions

I.S. and D.E. performed the experiments; M.P.A. performed the DFT calculations; A.P. synthesised the gibbsite and developed the sample preparation procedures; M.C.S., S.L.S.S., M.P.A., I.S., D.E. and F.M. contributed to discussion and writing; I.S., D.E., M.P.A. and F.M. designed the research.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Siretanu, I. *et al.* Direct observation of ionic structure at solid-liquid interfaces: a deep look into the Stern Layer. *Sci. Rep.* **4**, 4956; DOI:10.1038/srep04956 (2014).

Supplementary Information

Direct observation of ionic structure at solid - liquid interfaces: A deep look into the Stern Layer

Igor Siretanu[1†], Daniel Ebeling[1†], Martin P. Andersson[2], S. L. Svane Stipp[2], Albert Philipse[3], Martien Cohen Stuart[1], Dirk van den Ende[1] and Frieder Mugele[1]*

[1]Physics of Complex Fluids Group and MESA+ Institute, Faculty of Science and Technology, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

[2]Nano-Science Center, Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark

[3]Van't Hoff Laboratory for Physical and Colloid Chemistry, Debye Institute, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

I. Experimental and theoretical details

Gibbsite synthesis[1]**:** An aqueous solution of aluminium-iso-propoxide (15.2 g/L), aluminiumsec-butoxide (20 g/L) and HCl (0.86 M) was stirred in a closed vessel at room temperature for 10 days. The resulting turbid solution of amorphous aluminium hydroxide suspension was transferred to an autoclave for 72 h at 85 °C and then put into cellophane dialysis tubes and dialyzed against deionized water for 1 week. The procedure yielded a platelet concentration of 5-8 g/L in the final suspension. The BET (N_2) specific surface for the material was 76 m^2/g and it did not change during storage of the stock material. Transmission electron micrographs show that colloids are well defined and fairly monodisperse hexagons with an average diameter of 160 nm. Potentiometric titration and the electrophoretic mobility suggest that edges of gibbsite colloids have an isoelectric point close to pH 7, which differs from that on the faces, i.e. pH~10. To prepare the samples for spectroscopy and AFM measurements, the stock suspension was diluted 100 times in a mixture of ultrapure deionised water (MilliQ) and ethanol (1:1). 10 ml of the suspension are deposited onto freshly cleaned silica substrates. After 30 s residence time, the samples is rinsed with copious amounts of ultrapure water and blown dry with air.

Force spectroscopy: Measurements were made in a range of solutions made with NaCl, KCl, $MgCl_2{\cdot}6H_2O$, $CaCl_2{\cdot}2H_2O$ purchased from Sigma Aldrich, all at puriss, p.a. or ACS reagent grade. We made a stock solution of 1 M for each salt and diluted it for each experiment to concentrations of 0.5, 1, 3, 10, 30, and 100 mM. For each of the 24 cases, 100 interaction curves were measured along a scan line over a single gibbsite platelet. For each salt, a new, freshly cleaned silica substrate and petri dish were used and the series of measurements began with the sample exposed to the 0.5 mM solution. After completing the spectroscopy measurements for the lowest concentration, the cantilever was retracted from the surface by 0.5 mm, without removing it from the liquid. The concentration was increased to 1, then 3, 10, 50 and 100 mM in turn, by adding aliquots of the stock solution. At each concentration, the liquid in the petri dish was gently mixed by flushing liquid in and out with a micropipette. The solution was allowed to equilibrate with the surface for 10 to 15 min. The sample was removed from the petri dish and mounted in the fluid cell of the AFM while it remained wet. Before and after collecting the force measurements, we recorded images of the gibbsite platelet, to make sure it had not drifted out of the scanning area. To get data from the various salt species that could be compared, all measurements were performed with the same cantilever. Each time a solution was replaced, the cantilever, the tip holder and the fluid cell were thoroughly flushed with ultrapure deionised water. The tip-sample force was determined from the amplitude and frequency shift vs distance curves using the method presented by Hölscher et al.[2,3].

The measured interaction forces between tip and sample surface were converted to surface charge with Poisson-Boltzmann theory. To optimise accuracy in the fitting process, several points were considered. Rather than using the classical formula for spherical particles, the actual geometry of the tip, derived from SEM images, was used for integrating the stress on the solid surface[4]:

$$F_{el}(D) = \underbrace{P(D)\pi R^2}_{flat\ circular\ area} + \underbrace{2\pi \int_R^\infty P\left(D + \frac{r-R}{\tan\alpha}\right) r\mathrm{d}r}_{cone} \tag{S1}$$

where $P(D)$ represents the pressure in the gap, D represents the tip-sample distance, R, the radius, α, the half angle of the tip. A similar expression is obtained for van der Waals part of the interaction for the same geometry

$$F_{vdW}(D) = \underbrace{f(D)\pi R^2}_{flat\ circular\ area} + \underbrace{2\pi \int_R^\infty f\left(D + \frac{r-R}{\tan\alpha}\right) r\mathrm{d}r}_{cone} \tag{S2}$$

where $f(D) = A_H/(6\pi D^3)$ describes the van der Waals force per unit area for two flat surfaces and A_H is the Hamaker constant. The total tip-sample interaction force is obtained by adding the contributions from the van der Waals and electrostatic interactions, $F_{ts}(D) = F_{el}(D - z_0) + F_{vdW}(D)$. Here we account for a possible shift by z_0 of the effective surface separation between the electrostatic interactions originating at the Helmholtz plane, whereas the van der Waals interactions start at the sample surface. To account for the thickness of Stern layers on tip and surface, we used $z_0 =$ 0.5-0.6 nm[5]. More details about theoretical background can be found in ref [4].

Using tabulated values for the Hamaker constants $A_{Gbs} = 1.2\times10^{-20}$ J (gibbsite-water-SiO_2)[6,7] and for silica, $A_{SiO2} = 0.65\times10^{-20}$ J (SiO_2-water-SiO_2)[8], we calculated the AFM tip radius by fitting the model (eq. S2) to the experimental force curves for the highest concentration of divalent salts on silica, for which electrostatic interactions can be neglected. This results in a tip radius of ~56 nm, in agreement with 52±5 nm measured from the SEM images after completion of the AFM experiments.

Tips were initially brought into contact with the surface to blunt them slightly. Force distance curves at high concentrations of divalent cations were conducted several times at the beginning and at the end of each measurement to verify that the tip radius did not vary through the measurement.

Keeping this value fixed, the surface charge as a function of the salt concentration is obtained by fitting the full model curves to the experimental data. In doing so, the symmetric part of the system (silica tip - silica surface) is used to determine the tip charge. Standard expressions are used to calculate the Debye screening length based on the salt concentration. **C**onstant **C**harge (cc) and **C**onstant **P**otential (cp) solutions to the Poisson–Boltzmann (PB) equation are calculated. The surface charge is determined by fitting the model force curves to the experimental data for tip-sample distances of a few nm, where the cc and the cp solutions overlap.

The heterogeneous nature of the sample allowed us to obtain a calibration of the tip on the silica for every data set, only a few nanometers from the gibbsite particle. This eliminated the need for independent tip calibration.

DFT: We used the DMol3 program with the COSMO-RS implicit solvent[9,10,11,12] the PBE[13] functional, the DNP basis set and the dispersion correction by Grimme[14] for the periodic density functional theory (DFT) calculations of the gibbsite {001} basal plane. We used a 1x2 unit cell and the gamma point only for all calculations. We constructed the simulation cell to be three molecular layers thick, perpendicular to the {001} face and the lowest layer was frozen in bulk positions throughout all simulations. All COSMO-RS calculations were performed using the parameterization DMOL3_PBE_C30_1301 in COSMOtherm vC30_1301[15]. All presented geometries were optimized to the medium setting for Dmol3.

All chemical reaction energies included entropy contributions from translational and rotational degrees of freedom for non-slab species, calculated using standard expressions for a gas phase pressure of 1 bar, the same condition used to calculate the solvation energies. Vibrational degrees of freedom were excluded, because of the difficulties involved in calculating these contributions for the slab systems. First, the computation is very expensive, and second, DFT relies on the harmonic approximation for determining the vibrational frequencies. The lowest lying modes can be very anharmonic, and contribute the most to the vibrational entropy. It is therefore very hard to assess the accuracy of this contribution and we have therefore disregarded vibrations in our calculations.

We tested several starting geometries. The water coordination number for adsorbed Ca^{2+} and Mg^{2+} was at least 6 for all calculations and both inner and outer shell adsorption of Ca^{2+} and

Mg^{2+} to the gibbsite basal plane were investigated. All surfaces were neutral. To generate inner shell adsorption complexes, we initialized the calculations by placing two partially dehydrated Ca^{2+} or Mg^{2+} ions per unit cell close to the surface. At the same time, we removed the three more reactive protons from the surface and explicitly added one hydroxyl to keep the unit cell electroneutral. In an alternate test, we left the cations fully hydrated, removed three protons from the hydration water and added one hydroxyl ion per unit cell. Both initial configurations relaxed into the stable zigzag configuration, but the calculated reaction energies clearly favoured the outer shell configuration.

The free hydration energy of the ions in our study was calculated as well using the reference state [1 bar gas / 1 mol solvent], to show that the DMol parameterization in COSMO-RS gives reasonable solvation properties for the ions. The solvation of the divalent cations was determined according to the following reaction:

$$M^{2+} + 6H_2O(aq) \rightarrow M(H_2O)_6^{2+}(aq) \qquad \text{(S3)}$$

Here (aq) means COSMO-RS solvation in water. Without the explicit hydration water, the free energies of solvation of the divalent cations are about 500 kJ/mol too weak. Including waters of hydration clearly improved the simulation, which is why we included them in our adsorption reactions as well, both for the free and the adsorbed ions.

The reaction per unit cell of gibbsite, for outer shell adsorption complexes was:

$$gibbsite(001)\left(H_2O\right)_6(aq) + 2Ca\left(H_2O\right)_6^{2+}(aq) + 4OH^-(aq) \rightarrow gibbsite(001)Ca_2\left(H_2O\right)_6\left(OH\right)_4\left(ads,aq\right) + 12H_2O(aq) \qquad \text{(S4)}$$

To explore the configurations for high salinity solutions, we determined the energy required to exchange one hydroxyl ion in each unit cell for a Cl^- ion, i.e.

$$gibbsite(001)Ca_2\left(H_2O\right)_6\left(OH\right)_4\left(ads,aq\right) + Cl^-(aq) \rightarrow gibbsite(001)Ca_2\left(H_2O\right)_6\left(OH\right)_3Cl\left(ads,aq\right) + OH^-(aq) \qquad \text{(S5).}$$

Equivalent reactions were calculated for $MgCl_2$ as well, Table S1 and S2. The most stable geometries for the Mg structures are shown in Figure S7. The COSMO surface for a bare gibbsite surface is shown in Figure S8 as an example of implicit solvation of a periodic structure.

II. Tables:

Table S1: Reaction energies for outer shell adsorption, which results in the zigzag structure from DFT calculations and the COSMO-RS implicit solvent model. Reactions are written using the stoichiometry for the primitive gibbsite unit cell but the energies are reported per cation. Hydration waters for the gibbsite structures are excluded in the reaction formula for clarity.

Reaction	**ΔG (per $M^{2+}(OH)_2$) (kJ/mol)**	**$[M^{2+}]$ for coverage = 0.5, pH=6 (M)**
gibbsite(001) + $2Ca(H_2O)_6^{2+} + 4OH^- \rightarrow$ gibbsite(001)$Ca_2(OH)_4$(ads)+$12H_2O$	-118	$2\cdot10^{-5}$
gibbsite(001) + $2Mg(H_2O)_6^{2+} + 4OH^- \rightarrow$ gibbsite(001)$Mg_2(OH)_4$(ads)+$12H_2O$	-115	$9\cdot10^{-5}$

Table S2: Reaction energies for transforming double row structure (low salt concentrations) into single row structure (high conc.) from DFT calculations and the COSMO-RS implicit solvent model. Reactions are written using the stoichiometry per unit cell. Hydration waters for the gibbsite structures are excluded in the reaction formula for clarity.

Reaction	**ΔG (per Cl^-) (kJ/mol)**	**$[Cl^-]$ for coverage = 0.5, pH=6 (M)**
gibbsite(001)$Ca_2(OH)_4$(ads)+ $Cl^- \rightarrow$ gibbsite(001)$Ca_2(OH)_3Cl$(ads) + OH^-	39	0.060
gibbsite(001)$Mg_2(OH)_4$(ads) + $Cl^- \rightarrow$ gibbsite(001)$Mg_2(OH)_3Cl$(ads) + OH^-	47	1.8

Table S3: Comparison of the experimental free energy of hydration for the ions in this study with our calculations using DFT and COSMO-RS. Energies are in kJ/mol.

Ion	Experimental[16]	Calculated
Mg^{2+}	-1830	-1785
Ca^{2+}	-1505	-1440
OH^-	-430	-515
Cl^-	-340	-369

III. Additional Figures

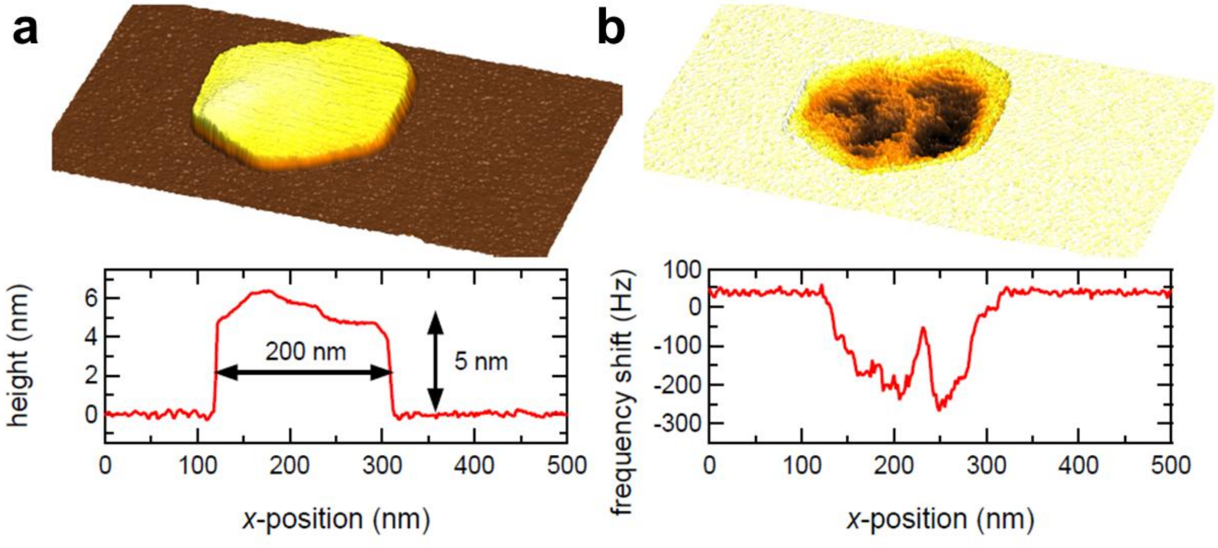


Figure S1: FM-AFM images of a single gibbsite nanoparticle adsorbed to a silica substrate **a,** 500×250 nm^2 FM-AFM topography image **b,** frequency shift image, each with representative cross section through the gibbsite particle. The nonuniform distribution of frequency shift, revealed by the spike, attributed to a grain boundary running through the crystal. Imaging conditions: 20 mM NaCl solution, CFM Aspire rectangular silicon cantilever, conical tip; f_0 = 19.1 kHz, c_Z = 4.2 N/m, Q = 9.5.

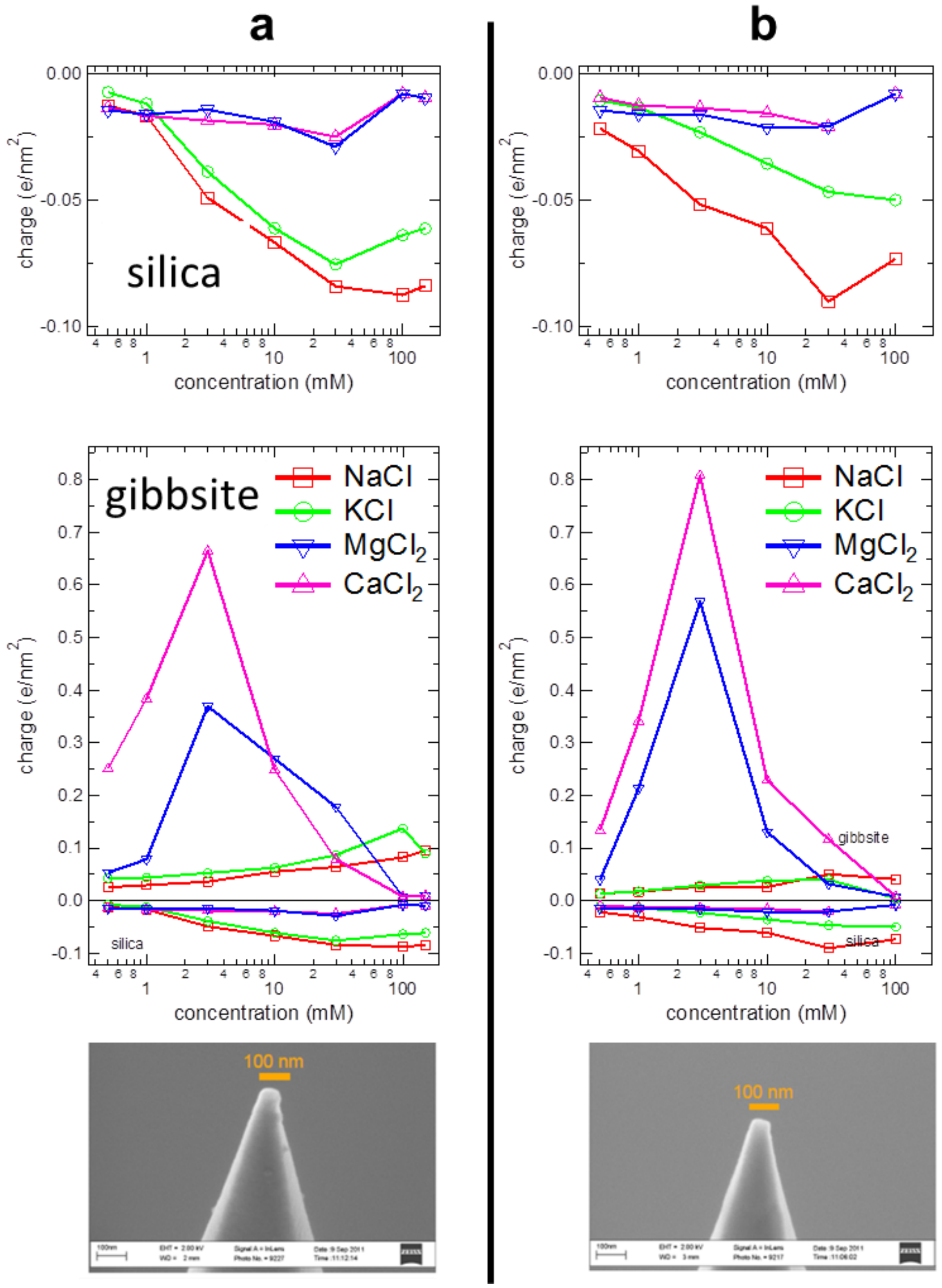


Figure S2: Effective surface charge vs. solution composition for two other tips on silica and gibbsite.

a, Tip parameters and SEM micrographs for Aspire CFM: $f_0 = 19.0$ kHz, $c_z = 4.7$ N/m, $Q = 11$; **b**, and for Aspire CT-130: $f_0 = 68.4$ kHz, $c_z = 48.8$ N/m, $Q = 35$. The images show the tips after use.

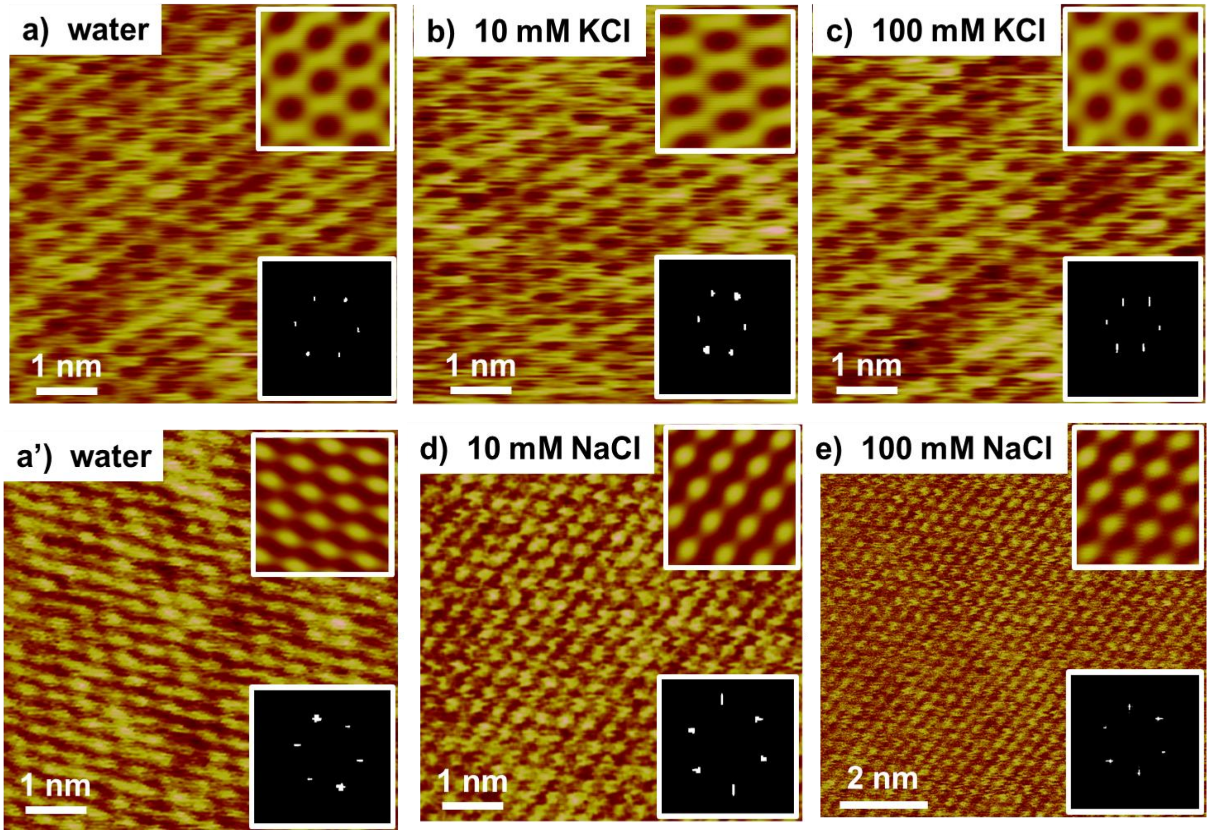


Figure S3: High resolution, noncontact AM-AFM topographic images of gibbsite during exposure to water and the monovalent cations

a,a', AFM images of the gibbsite basal plane obtained with super sharp tips at room temperature. The pseudohexagonal structure has 0.5 nm periodicity; **b-e**, Images taken in 10 mM and 100 mM NaCl and KCl solutions show no changes in topography compared with images taken in water. The pattern, with a periodicity of 0.5 nm is the same in all images.

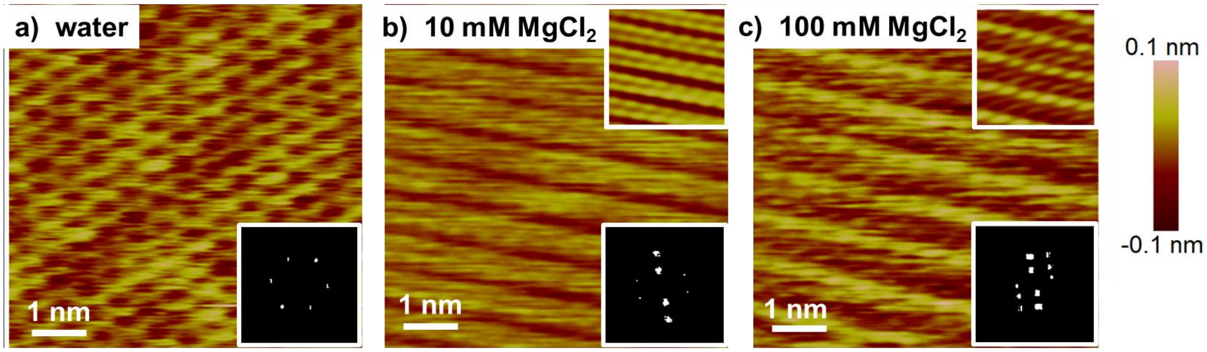


Figure S4: High resolution, noncontact AM-AFM topographic images of gibbsite, taken in water and $MgCl_2$ solutions.

The pattern is quite different than that observed in monovalent solutions.

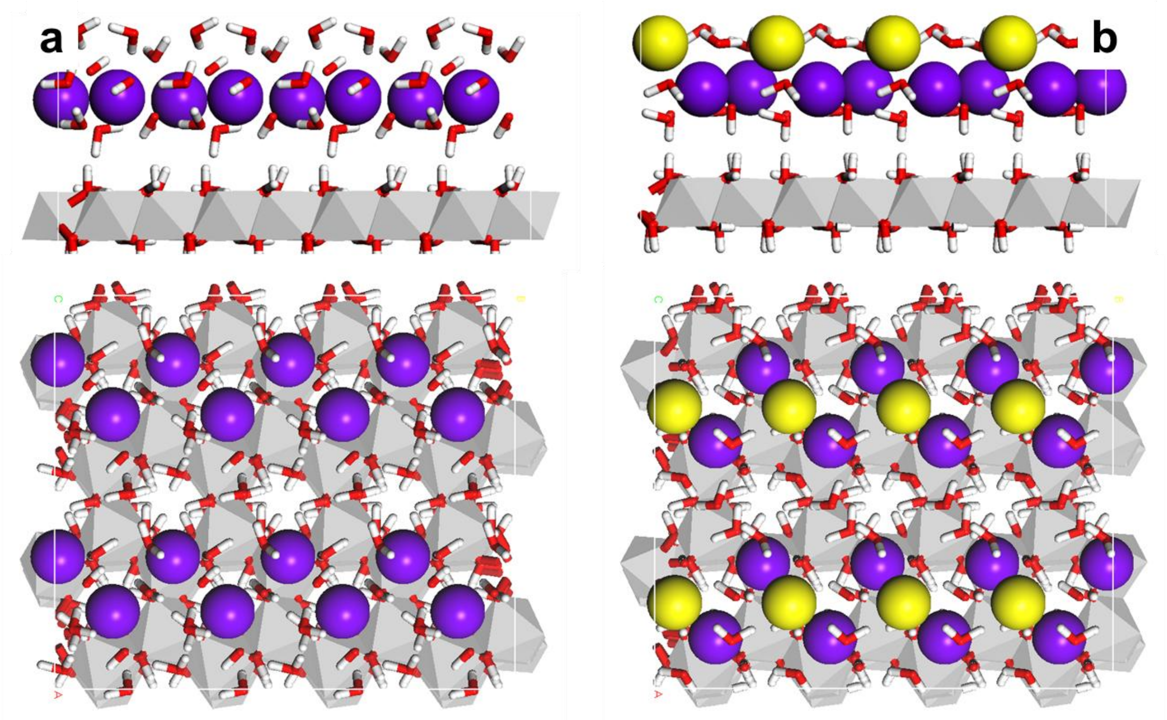


Figure S5: Structural models produced by DFT simulations

a, $Mg_2(OH)_4$ adsorbed on gibbsite, which produces a zigzag structure and **b**, the same surface but with chloride included, i.e. $Mg_2(OH)_3Cl$, which produces single rows We have outlined four of the unit cells that we used in the simulations with a white box, which indicates the periodic boundary conditions.

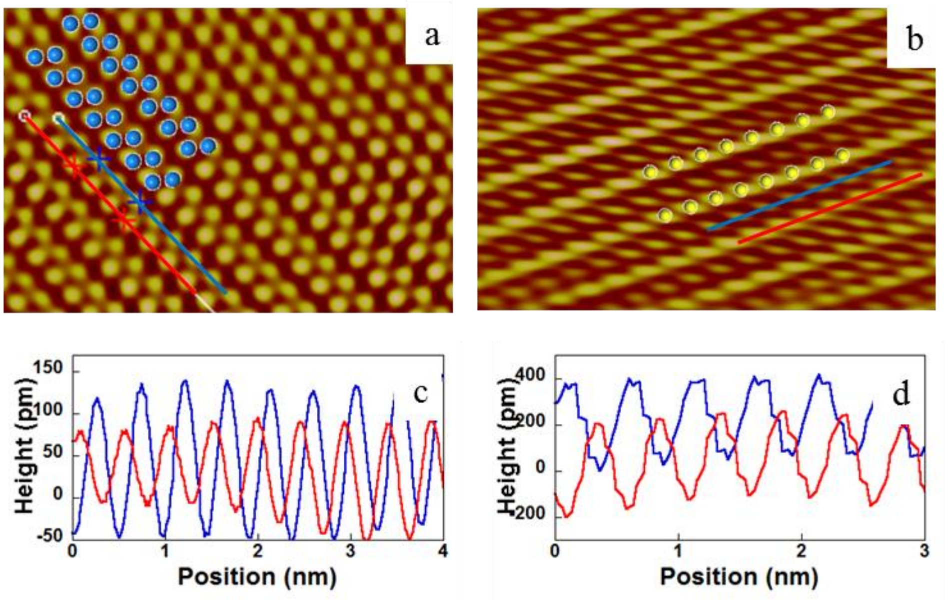


Figure S6: AFM topographic images after processing with Fourier transformation taken in **a**, 10 mM and **b,** 100 mM $CaCl_2$ solution (same as Fig 3b and c) superimposed with structural models showing only adsorbed Ca^{2+} (blue) and Cl^- (yelow) ions from Fig 5a and b. **c,** Height profiles are taken as indicated by solid lines in **a,** reveling the distance between the vertical position of the two Ca^{2+} ions which is around <50 pm as predicted by DFT calculations. **d,** Height profiles corresponding to b, showing vertical position of the Cl^- ions which is around 200 pm above the metal cations in agreement with DFT calculations.

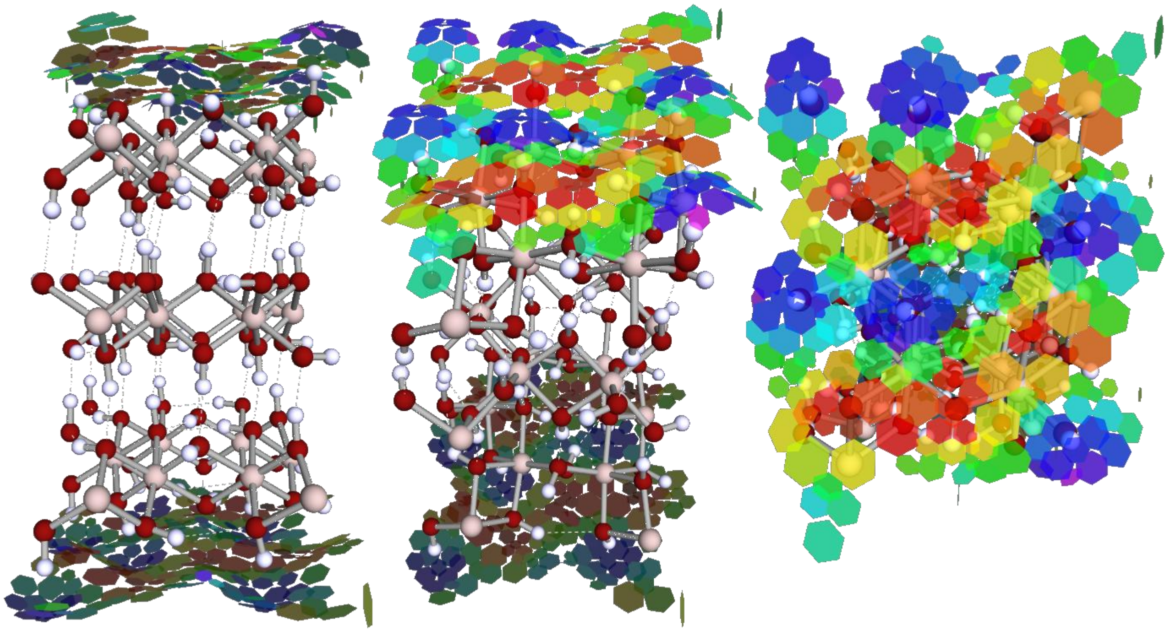


Figure S7: Molecular structure of a (1x2) unit cell of gibbsite seen from various angles, showing the COSMO surface for the periodic structure.

Red indicates negative charge and blue, positive; green and yellow are near neutral. Left, side view; right, top view and middle, tilted.

References:

1 Wierenga, A. M., Lenstra, T. A. J. & Philipse, A. P. Aqueous dispersions of colloidal gibbsite platelets: synthesis, characterisation and intrinsic viscosity measurements. *Colloids Surf., A* **134**, 359-371 (1998).

2 Ebeling, D. & Hölscher, H. Analysis of the constant-excitation mode in frequency-modulation atomic force microscopy with active Q-Control applied in ambient conditions and liquids. *J. Appl. Phys.* **102** (2007).

3 Ebeling, D., Oesterhelt, F. & Hölscher, H. Dynamic force spectroscopy of single chainlike molecules using the frequency-modulation technique with constant-excitation. *Appl. Phys. Lett.* **95** (2009).

4 Ebeling, D., Van Den Ende, D. & Mugele, F. Electrostatic interaction forces in aqueous salt solutions of variable concentration and valency. *Nanotechnology* **22** (2011).

5 Biesheuvel, P. M. Implications of the charge regulation model for the interaction of hydrophilic surfaces in water. *Langmuir* **17**, 3553-3556 (2001).

6 Gan, Y. & Franks, G. V. Charging behavior of the gibbsite basal (001) surface in NaCl solution investigated by AFM colloidal probe technique. *Langmuir* **22**, 6087-6092 (2006).

7 Butt, H. J., Cappella, B. & Kappl, M. Force measurements with the atomic force microscope: Technique, interpretation and applications. *Surf. Sci. Rep.* **59**, 1-152 (2005).

8 Dishon, M., Zohar, O. & Sivan, U. From repulsion to attraction and back to repulsion: The effect of NaCl, KCl, and CsCl on the force between silica surfaces in aqueous solution. *Langmuir* **25**, 2831-2836 (2009).

9 Klamt, A., Eckert, F. & Arlt, W. COSMO-RS: An alternative to simulation for calculating thermodynamic properties of liquid mixtures. *Annu. Rev. Chem. Biomol. Eng.* **1**, 101-122 (2010).

10 Klamt, A., Jonas, V., Bürger, T. & Lohrenz, J. C. W. Refinement and parametrization of COSMO-RS. *J. Phys. Chem. A* **102**, 5074-5085 (1998).

11 Klamt, A. & Schüürmann, G. COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2*, 799-805 (1993).

12 Delley, B. The conductor-like screening model for polymers and surfaces. *Molecular Simulation* **32**, 117-123 (2006).

13 Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865-3868 (1996).

14 Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **27**, 1787-1799 (2006).

15 Eckert, F. & Klamt, A. COSMOtherm Version C3.0, Release 13.01. COSMOlogic GmbH & Co. KG,. (Leverkusen, Germany. , 2013).

16 Marcus, Y. Thermodynamics of solvation of ions. Part 5. - Gibbs free energy of hydration at 298.15 K. *J. Chem. Soc., Faraday Trans.* **87**, 2995-2999 (1991).

www.nature.com/scientificreports

SCIENTIFIC REPORTS

OPEN

Ion adsorption-induced wetting transition in oil-water-mineral systems

Received: 08 February 2015
Accepted: 16 April 2015
Published: 27 May 2015

Frieder Mugele[1], Bijoyendra Bera[1], Andrea Cavalli[1], Igor Siretanu[1], Armando Maestro[1], Michel Duits[1], Martien Cohen-Stuart[1], Dirk van den Ende[1], Isabella Stocker[2] & Ian Collins[2]

The relative wettability of oil and water on solid surfaces is generally governed by a complex competition of molecular interaction forces acting in such three-phase systems. Herein, we experimentally demonstrate how the adsorption of in nature abundant divalent Ca^{2+} cations to solid-liquid interfaces induces a macroscopic wetting transition from finite contact angles (≈10°) with to near-zero contact angles without divalent cations. We developed a quantitative model based on DLVO theory to demonstrate that this transition, which is observed on model clay surfaces, mica, but not on silica surfaces nor for monovalent K^+ and Na^+ cations is driven by charge reversal of the solid-liquid interface. Small amounts of a polar hydrocarbon, stearic acid, added to the ambient decane synergistically enhance the effect and lead to water contact angles up to 70° in the presence of Ca^{2+}. Our results imply that it is the removal of divalent cations that makes reservoir rocks more hydrophilic, suggesting a generalizable strategy to control wettability and an explanation for the success of so-called low salinity water flooding, a recent enhanced oil recovery technology.

The relative wettability of oil and water on porous solids is crucial to many environmental and technological processes including imbibition, soil contamination/remediation, oil-water separation, and the recovery of crude oil from geological reservoirs[1–7]. Good wettability of a porous matrix to one liquid generally implies stronger retention of that fluid and simultaneously easier displacement of the other. In standard 'water flooding' oil recovery, (sea) water is injected into the ground to displace oil from the porous rock, typically at an efficiency $<50\%$. For decades, oil companies have explored adding chemicals such as surfactants and polymers to the injection water to improve the process[8,9]. More recently, it was discovered that the efficiency can also be improved by reducing the salinity of the injection water[10], *i.e.* without adding expensive and potentially harmful chemicals, known as low salinity water flooding (LSWF). Yet, reported increases in recovery vary substantially and the microscopic mechanisms responsible for the recovery increment remain debated[9,11–13]. A wide variety of mechanisms has been proposed to explain the effect, including the mobilization of fines, interfacial tension variations, multicomponent ion exchange, and double layer expansion[10–12,14]. Many of these mechanisms are interrelated and may ultimately result in improved water wettability of the rock but evidence discriminating between them is scarce. The key challenge in identifying the reasons for the success of LSWF lies in the intrinsic complexity of the system and the lack of direct access to its microscopic properties. Here, we experimentally demonstrate for a well-defined model system a consistent scenario leading from ion adsorption at the solid-liquid interface to charge reversal and from there to wettability alteration. We also derive a model that provides quantitative predictions of the experimentally observed contact angles. Our results clarify many previous observations in core flooding experiments, including in particular the relevance of divalent cations, clays, pH, and polar organic species.

[1]University of Twente, MESA+ Institute for Nanotechnology, Physics of Complex Fluids, P.O. Box 217, 7500AE Enschede (The Netherlands). [2]BP Exploration Operation Company Ltd., Chertsey Road, Sunbury-on-Thames, TW16 7LN, (United Kingdom). Correspondence and requests for materials should be addressed to F.M. (email: f.mugele@utwente.nl)

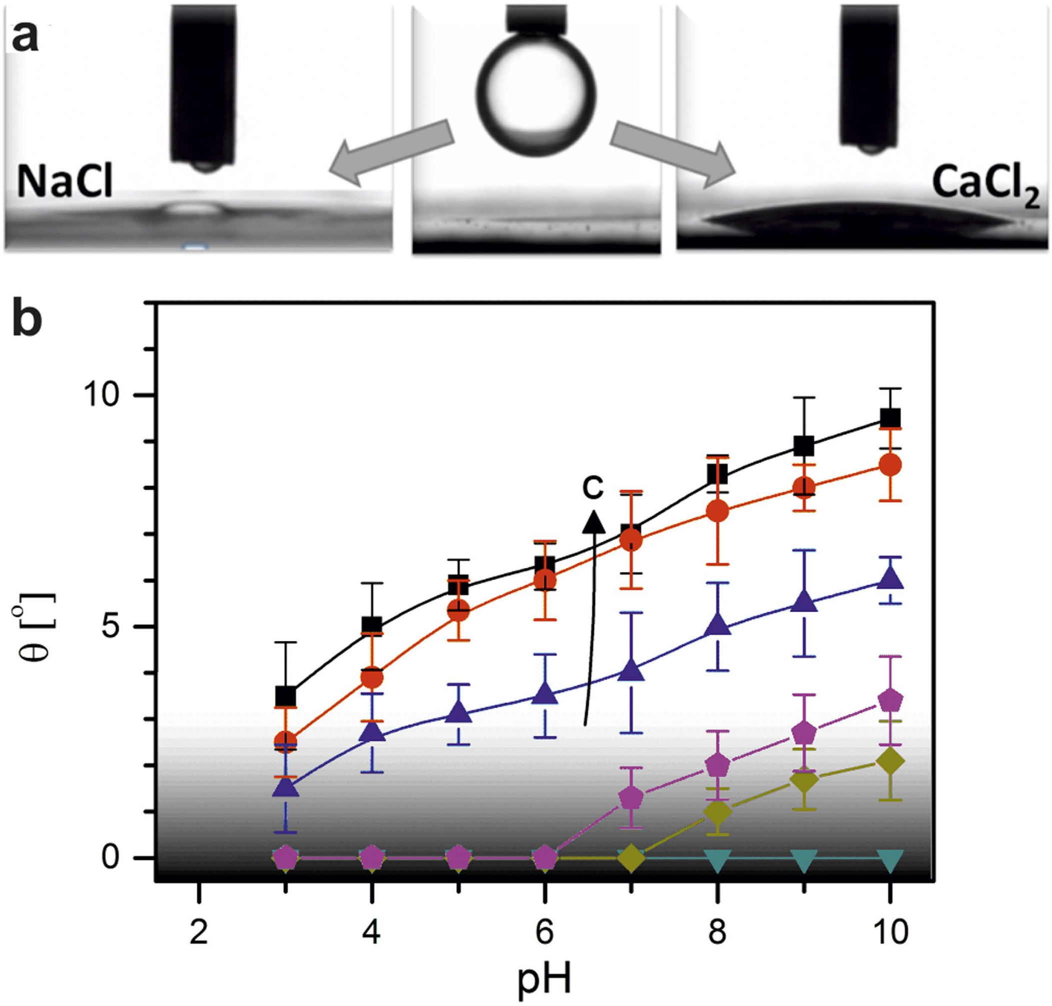


Figure 1. Water wetting on mica in ambient decane for monovalent & divalent salt solutions. (a) Side view of drops of 1 M (pH 7) aqueous solutions of NaCl (left) and $CaCl_2$ (right) immediately after bringing the drop on the needle in contact with the mica surface (ambient fluid: decane; needle diameter: 0.5 mm). NaCl solutions display immeasurably small contact angles, $CaCl_2$ solutions can display a finite contact angle, depending on concentration and pH. (**b**) Symbols: Equilibrium contact angle on mica vs pH for $CaCl_2$ salt solutions of various concentrations; 1,10,30 mM (downward triangles), 50 mM (olive diamonds), 80 mM (purple pentagons), 100 mM (blue triangles), 500 mM (red circles), 1 M (black squares). Solid lines: guides to the eye. The shaded region indicates very low contact angles, which are close or below the sensitivity of the instrument. The arrow with the letter c denotes the direction of increasing salt concentration.

Results & Discussions

Wettability alteration. The rock of common sandstone reservoirs consists of highly polar materials such as quartz and clays that in ambient air are completely wetted by both water and oil. To analyze the competitive wetting of oil and water on these substrates, we measured the macroscopic contact angle of water of variable salt content against decane. We chose flat, freshly cleaved mica and freshly cleaned silica surfaces as model materials, to represent the basic components of sandstone reservoirs. The macroscopic contact angle of water as observed in side view images, Fig. 1a, on mica and silica in ambient decane was found to depend strongly on the composition of the aqueous phase. We varied pH between 3 and 10 and concentrations of NaCl, KCl and $CaCl_2$ from 1 mM to 1 M (see Methods). Aqueous drops containing monovalent cations invariably spread to immeasurably small contact angles ($<2°$); in contrast, drops containing divalent cations displayed finite contact angles on mica for concentrations above ≈ 50 mM and $pH > 4$ (Fig. 1; see also Supplementary Material, movies S1 and S2). On silica, negligible contact angles were found for all pH's and concentrations of all salts investigated, *i.e.* including the ones with divalent cations.

Proposed adsorption mechanism. To identify the origin of the wetting transition on mica, we analyzed the force balance between the decane-water (γ), solid-decane (γ_{so}) and solid-water (γ_{sw})

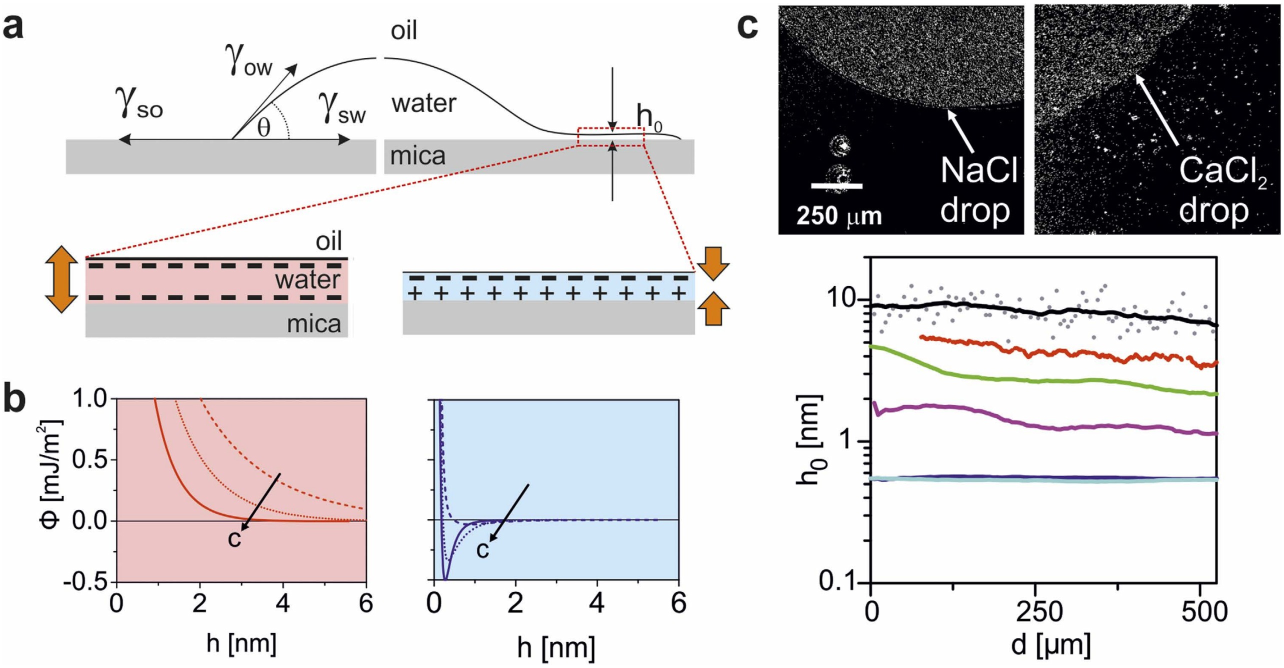


Figure 2. Proposed mechanism of wetting transition through ion adsorption and charge reversal at mica-water interface. (a) Schematic view of force balance, thin film with equilibrium thickness h_0 (top) and surface charge configurations of repulsive (bottom left) and attractive interface potential (bottom right). (**b**) Effective interface potential for surface charges of equal (left, red lines, mica-NaCl solution at pH 6-oil) and of opposite sign (right, blue lines, mica-$CaCl_2$ solution at pH 6-oil), leading to near-zero and finite contact angles, respectively. Lines denote salt concentrations: 1 mM (dashed lines), 10 mM (dotted lines) & 100 mM (solid lines). The arrows with the letter c denote the direction of increasing salt concentration. (**c**) Ellipsometry images (top) and resulting thickness profiles (bottom): film thickness (h_0) vs distance from contact line for aqueous drops for various concentrations of $CaCl_2$: 1 M (dark blue), 500 mM (magenta), 100 mM (green) and 10 mM (red) at pH 6; 1 M NaCl (black; grey symbols indicate scatter of raw data). Light blue: water film thickness in decane before adding aqueous drop.

interfacial tension at the three phase contact line. Under partial wetting conditions, the spreading pressure $S = \gamma_{so} - (\gamma + \gamma_{sw})$is negative and the equilibrium contact angle θ (measured through the aqueous phase) is given by Young's equation (Fig. 2a), $\cos\ \theta = (\gamma_{so} - \gamma_{sw})/\gamma^{15}$. For water in contact with non-polar oils, γ depends very weakly on pH and salt content (Supplementary Figure 1) and hence has a negligible influence on the wettability[16]. γ_{sw} usually decreases as salt content increases due to the spontaneous formation of an electric double layer at the solid-water interface[17]. Because any *reduction* of γ_{sw} can only induce a decrease of θ , the observed increase upon addition of Ca^{2+} and Mg^{2+} ions must be caused by an even stronger decrease of γ_{so}. The latter is plausible if the system forms a nanometer thin aqueous film next to the macroscopic drop with a salinity-dependent thickness h_0 (Fig. 2a). Using imaging ellipsometry we indeed detected such a film, as shown in Fig. 2c. Upon increasing the $CaCl_2$ concentration, h_0 decreased from approximately 8 nm to less than 1 nm. For pure water and for NaCl solutions, ellipsometry measurements revealed that θ is very small but finite despite the apparent spreading in side view images; h_0 was found to be ≈ 10 nm. Given the existence of this nanofilm, we can write the equilibrium tension γ_{so} in terms of oil-water and solid-water interfacial tensions plus an effective interface potential Φ(h) representing the molecular interactions between the solid-water and the water-oil interface as[15] $\gamma_{so} = \gamma_{sw} + \gamma + \Phi(h_0)$. Here, $\Phi(h_0)$ is the equilibrium value of Φ(h) corresponding to the equilibrium film thickness $h = h_0$, such that

$$\cos\ \theta = 1 + \Phi(h_0)/\gamma \tag{1}$$

The ion-induced wettability alteration thus reflects the salt-dependence of Φ(h), Fig. 2b.

Interfacial charge reversal. We decomposed $\Phi(h) = \Phi_h(h) + \Phi_{vdW}(h) + \Phi_{el}(h)$ into contributions from short-range chemical hydration forces $\Phi_h(h) = \Phi_h^0 \exp(-h/\lambda)$, van der Waals forces $\Phi_{vdW}(h) = A\ /\ 12\pi\ h^2$, and electrostatic forces $\Phi_{el}(h)$. While the amplitude Φ_h^0 and the decay length λ of the repulsive hydration forces as well as the Hamaker constant A generally vary weakly with pH and

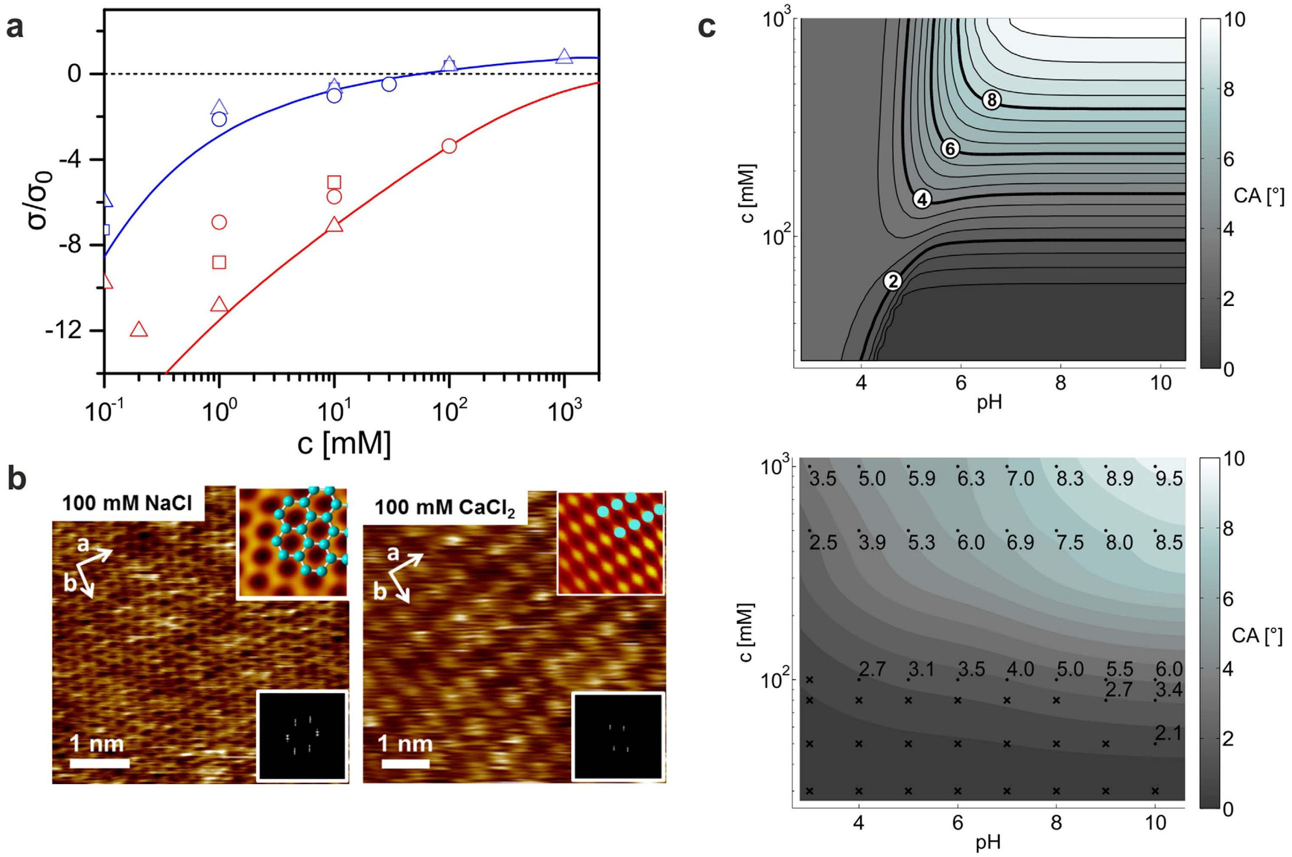


Figure 3. Ion adsorption at mica-water interface. (**a**) Surface Charge calculated from ζ potential measurements (circles) vs. concentration of solutions of NaCl (red) and $CaCl_2$ (blue) at pH 6. Solid lines: surface complexation model predictions. Blue triangles: AFM data from[25]; blue squares[24], red triangles[23], red squares[32]: surface forces apparatus measurements. The charge density is normalized by the characteristic scale σ_0 arising from the Poisson-Boltzmann equation, $\sigma_0 = \varepsilon\varepsilon_0 k_B T \kappa_D / e$, where κ_D is the Debye parameter. (**b**) AFM images of mica-water interface showing the characteristic hexagonal lattice of mica in 100 mM NaCl solution (left), and a rectangular symmetry caused by (presumably hydrated) adsorbed Ca^{2+} ions in 100 mM $CaCl_2$ (right). Insets: filtered zoomed views with overlaid lattice structure (top) and Fast Fourier Transform image of the same data (bottom). **c** Gray scale encoded contact angle vs. pH and $CaCl_2$ concentration. Top: model prediction; bottom: experimental data. Symbols (x: $\theta < 2°$) and numbers: experimental data same as Fig. 1b with interpolated gray scale. Smoothed lines are guides to the eye based on the experimental datapoints.

salt concentration, they are not expected to change sign for the conditions of our experiments[18]. Hence, we conclude that the observed wettability alteration is driven by $\Phi_{el}(h)$. The latter is repulsive and thus favors complete wetting if the charge densities σ_{sw} and σ_{ow} of the solid-water and the oil-water interface, respectively, carry the same sign. Vice versa, surface charges of opposite signs result in attraction and partial wetting. σ_{sw} and σ_{ow} are thus key parameters controlling wettability, as recently recognized in the context of wetting transitions with electrolyte solutions[19,20].

For oil-water interfaces, σ_{ow} is negative for $pH > 3$. The adsorption of ions is rather weak[21,22], as we corroborated using streaming potential measurements with solid eicosane mimicking decane. In streaming potential measurements for NaCl and KCl solutions, negative surface charges prevailed on mica for all conditions investigated, in agreement with surface force measurements[17,23]. For $CaCl_2$, however, a much stronger adsorption was found, Fig. 3a, leading to charge reversal at concentrations beyond ~ 50 mM[24]. Atomic force microscopy (AFM) confirmed this distinct difference between monovalent and divalent cations. While AFM images in pure water and aqueous NaCl and KCl solutions displayed the intrinsic hexagonal appearance of bare mica, a transition to a rectangular pattern was found for ambient $CaCl_2$ solutions, Fig. 3b[25]. Similar to gibbsite-water interfaces[26], we attribute this pattern to a layer of strongly adsorbed, possibly hydrated, divalent cations that reverse the sign of σ_{sw}.

Interaction between interfaces. To quantitatively assess this suggested mechanism, we explicitly calculate the various contributions to the disjoining pressure discussed in the previous section. $\Phi_h(h)$ is characterized by an amplitude $\Phi_h^0 = 10...50\ \mathrm{mJ}/m^2$ and a decay length $\lambda < 1\ \mathrm{nm}$[25,27]. For $\Phi_{\mathrm{vdW}}(\mathrm{h})$ we use a Hamaker constant $A = -\ 0.4 \times 10^{-21}\mathrm{J}$ limited by the experimental constraint that the finite contact angle of NaCl and KCl solutions must not exceed 2°. This negative Hamaker constant implies long range partial wetting, which arises from the fact that water has a lower refractive index than both mica and oil. We obtain the electrostatic contribution Φ_{el} to the disjoining pressure by solving the Poisson-Boltzmann equation for the electrostatic potential Ψ inside the thin film, which reads $\Psi'' = -\ e/\varepsilon\varepsilon_0 \sum_i Z_i c_i^\infty \exp(-\ Z_i e\Psi/k_B T)$. In the equation, e is the elementary charge, ε_0 and ε are the vacuum and relative permittivity of the medium and $k_B T$is the thermal energy in the system. The sum runs over the ions in the solution, with Z_i representing the valence and c_i^∞ the bulk concentration of the i-th specie. Here, we have used the full Poisson Boltzmann expression instead of classical examples[28,29] of a reduced equation, since the zeta potentials in our system are clearly beyond 25 mV. We apply constant charge (CC) boundary conditions, where the surface charges σ_{sw} (at the solid-water interface) and σ_{ow} (at the oil-water interface) are determined from the corresponding surface complexation model (see Methods), by fitting to experimentally measured streaming potentials. Once the electrostatic potential Ψ is known, we find the contribution to the disjoining pressure Φ_{el} by evaluating the standard expression $\Phi_{el} = -\int_\infty^h \left[k_B T \sum c_i^\infty \exp(-\ eZ_i\Psi/k_B T) - \frac{1}{2}\varepsilon\varepsilon_0 (d\Psi/dx)^2\right]_{h'/2} dh'$ [18].

Adding up all the contributions to the disjoining pressure, we find that for sufficiently high Ca^{2+} concentrations, $\Phi(\mathrm{h})$ indeed develops a pronounced minimum at small h_0, corresponding to water contact angles up to 10°, as depicted in Fig. 2b. For Na^+ and K^+, however, a very shallow minimum corresponding to a small but finite contact angle appears, due of the dominance of attractive van der Waals interactions (*i.e.* $A < 0$) for large film values of h.

Using eq. (1), we extracted the contact angle θ from the minima of $\Phi(\mathrm{h})$ for all fluid compositions, Fig. 3c,top. Comparison to the experimental results, Fig. 3c,bottom, shows that the model indeed captures all salient features of the experiments, including in particular the transition from near zero contact angles at low divalent ion concentration and pH to values of $\theta \approx 10^\circ$ at high Ca concentration and pH. For monovalent cations on mica and for all salts on silica, the same calculation invariably results in repulsive electrostatic forces and hence negligibly small contact angles (<2°).

Synergistically enhanced wettability alteration. Most crude oils contain small proportions of surface-active polar components in addition to alkanes. We investigated the impact of these components on the wettability by adding small amounts of stearic acid (S.A.) to the decane. Water drops containing divalent cations, when deposited on mica under decane/S.A. mixture, initially assumed $\theta \approx 10^\circ$, as in absence of S.A. Within seconds, however, θ increased to values of up to 70° (Fig. 4a,b; Movie S3). For drops containing NaCl, θ slightly increased, too, but never exceeded 10°. AFM imaging of the mica surface after removal from all liquids revealed the origin of this strong autophobic behavior: the surface was covered by a stearate monolayer very similar to partially decomposed Langmuir-Blodgett films of the same material reported earlier[30]. Close to the original contact line of the droplet, this layer was dense with occasional holes; farther away, bare mica was seen with occasional islands of monolayer stearate. Ca^{2+} and S.A. thus synergistically enhance the wettability alteration by promoting the self-assembly of hydrophobic Ca stearate monolayers.

In conclusion, these findings demonstrate how divalent cations in combination with clays and acidic components in the oil can control the wettability of oil-water-rock systems in water flooding oil recovery. The observed reduction in the water-mica contact angle in ambient decane of approximately 10°, as a result of removing divalent ions from the water, is itself sufficient to result in several percent of incremental oil recovery[31]. More generally, our results suggest a universal strategy to manipulate wettability by controlling the adsorption of ions to solid-liquid interfaces.

Methods

Experimental System. Anhydrous n-decane (>99%, Sigma Aldrich) is passed five times through a vertical column of Alumina powder ($\mathrm{Al_2O_3}$, Sigma Aldrich, Puriss grade > 98%) to remove any surface-active impurities. The ultrapure water (resistivity 18 MΩ) used to prepare the salt solutions is obtained from a Millipore water treatment system (Synergy UV Instruments). Solutions of various concentration (between 1 mM to 1 M cation concentration) are prepared for NaCl, KCl or $\mathrm{CaCl_2}$ salts (Sigma Aldrich). The pH of the solution is adjusted between 3 and 10 using HCl/$\mathrm{HNO_3}$ and NaOH (0.1 M, Sigma Aldrich). Muscovite mica (B&M Mica Company Inc., USA; initial thickness 340 μm) and oxidized silicon wafers with an amorphous silicon oxide layer (thickness: 30 nm) mimicking silica represent the surface of a solid rock. Mica sheets are cleaved inside the oil phase to obtain a pristine surface during the experiment. Silica surfaces are cleaned using a combination of Piranha solution (followed by extensive rinsing with ultrapure water) and plasma treatment.

Contact angle measurements. The wetting of aqueous drops on mica is characterized using a commercial contact angle goniometer (OCA 20L, Dataphysics Inc.). The measurement is based on sessile-drop method using aqueous drops with a volume of 2 μL placed on solid substrate. The contact angle of the drops is extracted from video snapshots using the tangent-fitting method in data analysis software (SCA

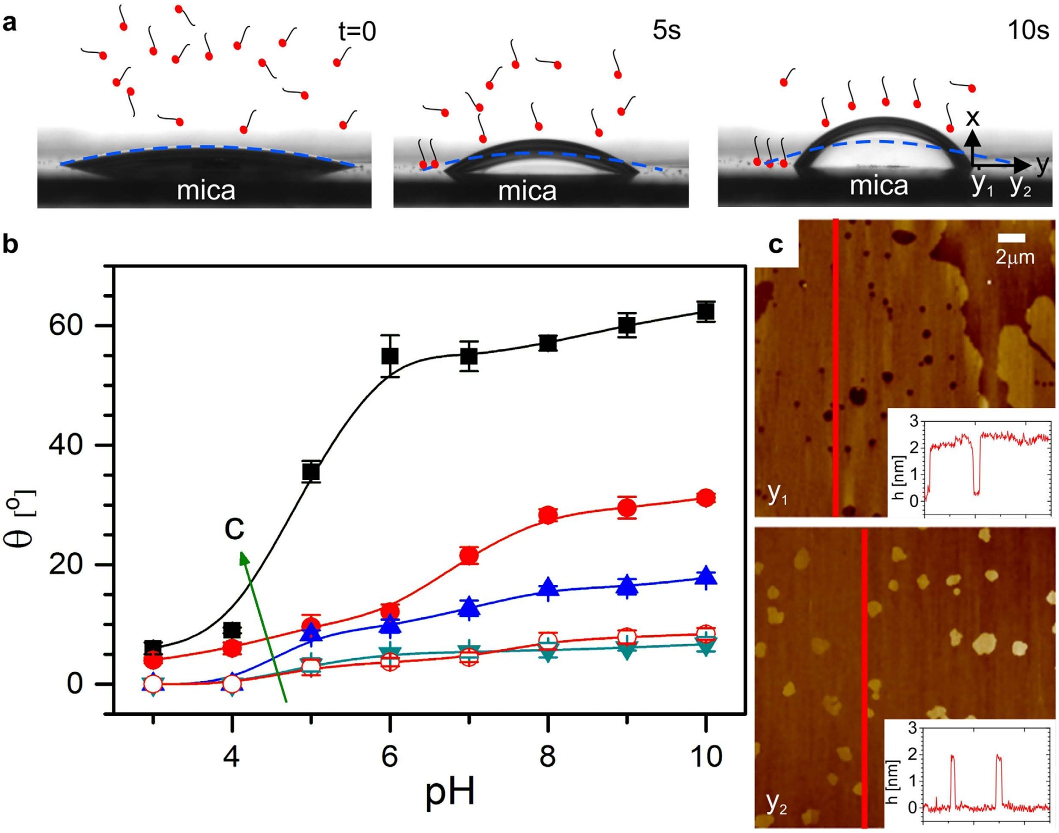


Figure 4. Cation-induced surfactant adsorption on solid substrate in oil. (**a**) Snapshots of drops of 1 M $CaCl_2$ solution (pH = 9) on mica immersed in ambient decane containing 100 μM stearic acid, immediately after deposition (t = 0) and 5 s and 10 s later. Drops display autophobic behavior due to the deposition of organic layers on the substrate. (**b**) Equilibrium contact angle vs. pH for various concentrations of $CaCl_2$: 1 mM (cyan downward triangles), 10 mM (blue upward triangles), 100 mM (red circles), 1 M (black squares) and NaCl: 100 mM (red open circles). The arrow with the letter c denotes the direction of increasing salt concentration. Stearic acid concentration: 100 μM. (**c**) After drop removal and drying AFM images display an almost complete monolayer at a distance of $y_1 = 100\,\mu m$ from the original contact line and an almost bare substrate with occasional stearate islands at $y_2 = 800\,\mu m$. Height profiles, corresponding to the red lines in the AFM images, demonstrate that the thickness of the layer corresponds to the length of a stearate monolayer.

22) provided with the instrument. Contact angles can be determined with a relative accuracy of ±1°. The minimum contact angle that can be determined on reflective surfaces is approximately 1.5°. Before placing the aqueous drops on the substrates, pendant drop measurements are performed to determine the oil/water interfacial tension (IFT). Constancy of the IFT over time ensures that the oil is devoid of residual surface active contaminants after passing the alumina powder column.

Ellipsometry. Thickness measurements of ultrathin wetting films were performed using an imaging ellipsometer (Accurion). The ellipsometer is equipped with custom-built quartz tubes attached to both the source (laser) and the detector arm to enable measurements under liquid at variable angle of incidence. In the case of mica, the bottom side of the substrate was roughened and coated with an index matched epoxy resin to suppress interference. Null ellipsometry experiments were performed. The thickness h_0 of the potentially adsorbed water layer is extracted from the ellipsometric angles Ψ and Δ assuming the bulk refractive index of the adjacent aqueous drop using standard Fresnel coefficients for a three layer system (substrate –water–oil).

Zeta Potential measurement. Surface charge and surface potential of solid/water (or oil/water) interfaces were determined by streaming potential measurements using a ZetaCAD instrument (CAD Instruments, France). The measurement cell consists of two substrates of the solid under investigation (50 mm x 30 mm) at a separation of 100 μm. Measured ζ potentials are converted to (diffuse layer) surface charges using Grahame's equation.

Surface complexation modeling. The surface charge of solid-water interfaces is modelled using standard surface complexation models involving adsorption/desorption reactions of cations X_i ($i = H^+$, Na^+, Ca^{2+}) to surface sites S following the scheme $SX \leftrightarrow S^- + X^+$. Each reaction is characterized by an equilibrium constant K with a corresponding value $pK = -\log K$. The law of mass action relates the cation concentration $[X_i]_s$ at the surface and the surface concentrations {SX} and {S⁻} to the equilibrium constant: $K_i\left\{SX_i\right\} = \left\{S^-\right\}[X_i]_s$. Local concentrations at the surface are related to the corresponding bulk concentrations c_i^∞ by a Boltzmann factor $[X_i]_s = c_i^\infty \exp(-Z_i e\Psi_0/k_B T)$, where Ψ_0 is the potential at the surface and Z_i the valency of species *i*. For the oil-water interface, the primary charge generation mechanism is assumed to be the autolysis of water $H_2O|_s \leftrightarrow H^+ + OH^-|_s$[21]. Additional weak cation adsorption reactions are included, too. The surface charge is then given by the relation $= e\Gamma\frac{(Z_c-1)\frac{C^{Z_c}}{K_C}-1}{1+\frac{[H^+]}{K_H}+\frac{C^{Z_c}}{K_C}}$, where $C^{Z_c} = Na^+, Ca^{2+}$ represent the activity of the ions considered. At large separation, the implicit dependence on Ψ_0 is solved equating this value to the one predicted by the Grahame Equation for monovalent $\sigma_{mono}^2 = 4\epsilon\epsilon_0 c_\infty k_B T\,(\cosh e\Psi_0/k_B T - 1)$ and divalent $\sigma_{di}^2 = 2\epsilon\epsilon_0 c_\infty k_B T\,(\exp -2e\Psi_0/k_B T + 2\exp e\Psi_0/k_B T - 3)$ salts, respectively. We use this procedure to extrapolate the value of the surface charges for all pH and salt concentrations considered. Our choice of the equilibrium constants is based on values from literature: a complete overview of all surface reactions and pK values is provided in the supplementary information, Table S1. In Fig. 3a we observe a good agreement between the values obtained by this approach (full lines) and several experimental measurements of the surface charge of Mica for monovalent and divalent salts.

References

1. Yuan, J. *et al.* Superwetting nanowire membranes for selective absorption. *Nat. Nanotechnol.* **3,** 332–336 (2008).
2. Berg, S. *et al.* Real-time 3d imaging of haines jumps in porous media flow. *Proc. Natl. Acad. Sci. USA.* **110,** 3755–3759 (2013).
3. Murison, J. *et al.* Wetting heterogeneities in porous media control flow dissipation. *Phys. Rev. Appl.* **2,** 034002; DOI: 10.1103/PhysRevApplied.2.034002 (2014).
4. Tian, X., Jin, H., Sainio, J., Ras, R. H. A. & Ikkala, O. Droplet and fluid gating by biomimetic Janus membranes. *Adv. Funct. Mater.* **24,** 6023–6028 (2014).
5. Wong, T.-S. *et al.* Bioinspired self-repairing slippery surfaces with pressure-stable omniphobicity. *Nature* **477,** 443–447 (2011).
6. Kovscek, A. R., Wong, H. & Radke, C. J. A pore-level scenario for the development of mixed wettability in oil-reservoirs. *AIChE Jour*. **39,** 1072–1085 (1993).
7. Zhao, X., Blunt, M. J. & Yao, J. Pore-scale modeling: Effects of wettability on waterflood oil recovery. *Jour. Pet. Sci. Engg.* **71,** 169–178 (2010).
8. Lake, L. W. *Enhanced oil recovery*. (Prentice Hall 1989).
9. Muggeridge, A. *et al.* Recovery rates, enhanced oil recovery and technological limits. *Philos. T. Roy. Soc.A* **372,** 20120320; DOI: 10.1098/rsta.2012.0320 (2014).
10. Tang, G. Q. & Morrow, N. R. Influence of brine composition and fines migration on crude oil/brine/rock interactions and oil recovery. *Jour. Pet. Sci. Engg.* **24,** 99–111 (1999).
11. Aksulu, H., Hamso, D., Strand, S., Puntervold, T. & Austad, T. Evaluation of low-salinity enhanced oil recovery effects in sandstone: Effects of the temperature and pH gradient. *Energ. Fuel* **26,** 3497–3503 (2012).
12. Sheng, J. J. Critical review of low-salinity waterflooding. *Jour. Pet. Sci. Engg.* **120,** 216–224 (2014).
13. Matthiesen, J. *et al.* How naturally adsorbed material on minerals affects low salinity enhanced oil recovery. *Energ. Fuel* **28,** 4849–4858 (2014).
14. Lager, A., Webb, K. J., Black, C. J. J., Singleton, M. & Sorbie, K. S. Low salinity oil revoery - an experimental investigation. *Petrophysics* **49,** 28–35 (2006).
15. deGennes, P. G. Wetting: Statics and dynamics. *Rev. Mod. Phys.* **57,** 827–863 (1985).
16. Liu, M., Xue, Z., Liu, H. & Jiang, L. Surface wetting in liquid–liquid–solid triphase systems: Solid-phase-independent transition at the liquid–liquid interface by lewis acid–base interactions. *Angew.Chem.* **51,** 8348–8351 (2012).
17. Pashley, R. M. & Israelachvili, J. N. DLVO and hydration forces between mica surfaces in Mg^{2+}, Ca^{2+}, Sr^{2+}, and Ba^{2+} chloride solutions. *Jour. Coll. Int. Sci.* **97,** 446–455 (1984).
18. Butt, H. J. & Kappl, M. *Surface and interfacial forces*. (Wiley CH Verlag GmbH & Co. KGaA, 2010).
19. Ibagon, I., Bier, M. & Dietrich, S. Wetting in electrolyte solutions. *J Che. Phys.* **138,** 214703; DOI: 10.1063/1.4807760 (2013).
20. Ibagon, I., Bier, M. & Dietrich, S. Order of wetting transitions in electrolyte solutions. *J Che. Phys.* **140** (17), 174713; DOI: 10.1063/1.4873712 (2014).
21. Beattie, J. K. The intrinsic charge on hydrophobic microfluidic substrates. *Lab Chip* **6,** 1409–1411 (2006).
22. Luetzenkirchen, J., Preocanin, T. & Kallay, N. A macroscopic water structure based model for describing charging phenomena at inert hydrophobic surfaces in aqueous electrolyte solutions. *Phy. Chem. Chem. Phys.* **10,** 4946–4955 (2008).
23. Scales, P. J., Grieser, F. & Healy, T. W. Electrokinetics of the muscovite mica aqueous-solution interface. *Langmuir* **6,** 582–589 (1990).
24. Kekicheff, P., Marcelja, S., Senden, T. J. & Shubin, V. E. Charge reversal seen in electrical double-layer interaction of surfaces immersed in 2:1 calcium electrolyte. *J Che. Phys.* **99,** 6098–6113 (1993).
25. Loh, S.-H. & Jarvis, S. P. Visualization of ion distribution at the mica-electrolyte interface. *Langmuir* **26,** 9176–9178 (2010).
26. Siretanu, I. *et al.* Direct observation of ionic structure at solid-liquid interfaces: A deep look into the stern layer. *Sci. Rep.* **4,** 4956; DOI: 10.1038/srep04956 (2014).
27. Kilpatrick, J. I., Loh, S.-H. & Jarvis, S. P. Directly probing the effects of ions on hydration forces at interfaces. *J. Am. Chem. Soc.* **135,** 2628–2634 (2013).
28. Hogg, R., Healy, T. W. & Fuerstenau, D. W. Mutual coagulation of colloidal dispersions. *T. Faraday Soc.* **62,** 1638–1651 (1966).
29. Parsegian, V. A. & Gingell, D. On the Electrostatic Interaction across a Salt Solution between Two Bodies Bearing Unequal Charges. *Biophys. J.* **12,** 1192–1204 (1972).
30. Kumar, N., Wang, L., Siretanu, I., Duits, M. & Mugele, F. Salt dependent stability of stearic acid langmuir-blodgett films exposed to aqueous electrolytes. *Langmuir* **29,** 5150–5159 (2013).

31. Herminghaus, S. Universal phase diagram for wetting on mesoscale roughness. *Physical Review Letters* **109,** 236102; DOI: 10.1103/PhysRevLett.109.236102 (2012).
32. Zembala, M. & Adamczyk, Z. Measurements of streaming potential for mica covered by colloid particles. *Langmuir* **16,** 1593–1601 (2000).

Acknowledgement

We thank Ann Muggeridge for comments on the manuscript. We acknowledge financial support by the ExploRe research program of BP plc.

Author Contributions

F.M., I.St. and I.C. designed the experiments; B.B., A.M. and I.Sir. carried out experiments; B.B., I. Sir., A.C., M.D., M.C.S. and D.v.d.E. analysed the experiments; F.M. wrote the manuscript with contributions from all other authors.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Mugele, F. *et al.* Ion adsorption-induced wetting transition in oil-water-mineral systems. *Sci. Rep.* **5**, 10519; doi: 10.1038/srep10519 (2015).

Supplementary Information

Ion adsorption-induced wettability alteration in oil-water-mineral systems

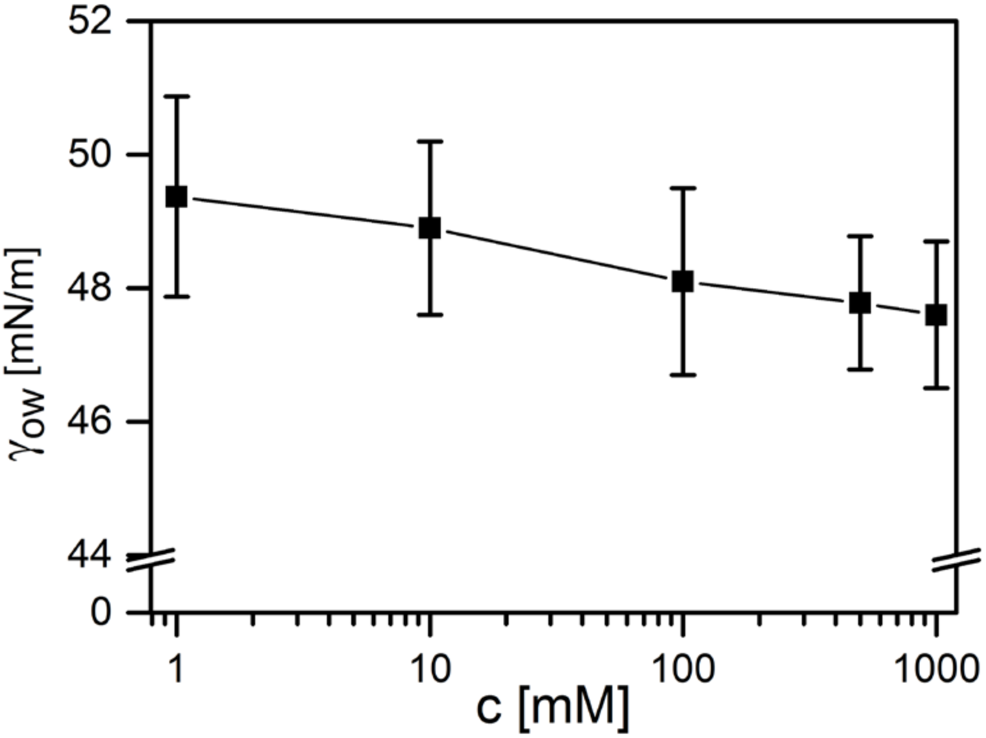
Frieder Mugele[1], Bijoyendra Bera[1], Andrea Cavalli[1], Igor Siretanu[1], Armando Maestro[1], Michel Duits[1], Martien Cohen-Stuart[1], Dirk van den Ende[1], Isabella Stocker[2], and Ian Collins[2]

[1] University of Twente, MESA+ Institute for Nanotechnology, Physics of Complex Fluids, P.O. Box 217, 7500AE Enschede (The Netherlands).

[2] BP Exploration Operation Company Ltd. , Chertsey Road, Sunbury-on-Thames, TW16 7LN, (United Kingdom).

S1. Contribution of Interfacial tension at Oil/Water Interfaces

Oil/water interfacial tension is measured using Pendant Drop method in OCA 20 (Dataphysics Inc.). A pendant drop (volume 15-20 µl) of salt solution is prepared in ambient decane phase, and using tangent fitting to the contour of the drop, interfacial tension is measured. We observe no significant change in interfacial tension with change of pH, but with increasing concentration of salt in the aqueous phase, a small decrease in the interfacial is noticed (Supplementary Figure 1). If the interfacial tensions at solid/water and solid/oil interfaces remain constant, then Supplementary Figure 1 would suggest a decrease in contact angle with increasing salt concentration. But we notice an increase in contact angle with increasing salt concentration, which confirms our hypothesis of existence of a thin wetting film between mica and decane.



Supplementary Figure 1 | Equilibrium interfacial tensions for the decane-water interface. Squares: Interfacial tension vs salt concentration measured with a pendant drop of $CaCl_2$ salt at pH 7 in ambient decane (oil), the error bars show the standard error. Solid line: guide to the eye.

	Mica interface	**Oil interface**
Reactions	$\{SH\} \leftrightarrow \{S^-\} + [H^+]$ $\{SC^{Z_c}\} \leftrightarrow \{S^-\} + \langle C^{Z_c}\rangle$ $\{SC^{Z_c}\} + \{S^-\} + \{SH\} = \Gamma_{\mathrm{mica}}$	$\{SH_2O\} \leftrightarrow \{SOH^-\} + [H^+]$ $\{SOH - C^{Z_c}\} \leftrightarrow \{SOH^-\} + \langle C^{Z_c}\rangle$ $\{SH_2O\} + \{SOH^-\} + \{SOH - C^{Z_c}\} = \Gamma_{\mathrm{oil}}$
Coefficients	$\Gamma_{\mathrm{mica}} = 0.8\ \mathrm{sites/nm^2}$ $PK_{Ca^{2+}} = 1.5$ $PK_{Na^+} = 0.5$ $PK_{H^+} = 5.3$	$\Gamma_{\mathrm{oil}} = 17.3\ \mathrm{sites/nm^2}$ $PK_{Ca^{2+}} = -1$ $PK_{Na^+} = -1$ $PK_{H^+} = 7$
Charge	$\sigma = e\Gamma \frac{(Z_c - 1)\frac{\langle C^{Z_c}\rangle}{K_C} - 1}{1 + \frac{[H^+]}{K_H} + \frac{\langle C^{Z_c}\rangle}{K_C}}$	$\sigma = e\Gamma \frac{(Z_c - 1)\frac{\langle C^{Z_c}\rangle}{K_C} - 1}{1 + \frac{[H^+]}{K_H} + \frac{\langle C^{Z_c}\rangle}{K_C}}$

Supplementary Table 1 | Complexation model coefficients for the reactions at the Mica-water interface (left) and Oil-water interface (right). { } represents surface concentrations, [ ] bulk concentrations and < > the bulk activity of a given specie.

[Movie S1]

Movie S1 | 'Almost' complete wetting with NaCl. Typical behaviour of an aqueous drop (volume 2µl) containing a monovalent salt (NaCl) at any concentration or pH. The drop spreads on the mica (which is in ambient decane) without forming any wedge-shaped drop, thus making impossible to characterize a contact angle with the optical goniometer.

[Movie S2]

Movie S2 | Partial wetting with $CaCl_2$. An aqueous drop (volume 2µl) containing divalent $CaCl_2$/$MgCl_2$ solution (at a concentration higher than 50mM) does not spread on mica (in ambient decane) as in the previous case. A wedge-shaped drop is formed, where a small but definite finite contact angle is possible to measure with the goniometer.

[Movie S3]

Movie S3 | Synergistic contact angle enhancement, with $CaCl_2$ and added stearic acid to the oil phase. An aqueous drop (volume 2µl) containing divalent $CaCl_2$ solution (at a concentration of 1M) is deposited on mica in ambient decane. 0.1mM stearic acid is added to the oil phase. The water droplet forms a finite but small contact angle, which grows rapidly due to the autophobing effect described in the main manuscript.

Nanoscale

PAPER

Cite this: *Nanoscale*, 2015, **7**, 16298

Extracting local surface charges and charge regulation behavior from atomic force microscopy measurements at heterogeneous solid-electrolyte interfaces

Cunlu Zhao, Daniel Ebeling, Igor Siretanu, Dirk van den Ende and Frieder Mugele*

Received 4th August 2015,
Accepted 10th September 2015

DOI: 10.1039/c5nr05261k

www.rsc.org/nanoscale

We present a method to determine the local surface charge of solid–liquid interfaces from Atomic Force Microscopy (AFM) measurements that takes into account shifts of the adsorption/desorption equilibria of protons and ions as the cantilever tip approaches the sample. We recorded AFM force distance curves in dynamic mode with sharp tips on heterogeneous silica surfaces partially covered by gibbsite nanoparticles immersed in an aqueous electrolyte with variable concentrations of dissolved NaCl and KCl at pH 5.8. Forces are analyzed in the framework of Derjaguin–Landau–Verwey–Overbeek (DLVO) theory in combination with a charge regulation boundary that describes adsorption and desorption reactions of protons and ions. A systematic method to extract the equilibrium constants of these reactions by simultaneous least-squared fitting to experimental data for various salt concentrations is developed and is shown to yield highly consistent results for silica-electrolyte interfaces. For gibbsite-electrolyte interfaces, the surface charge can be determined, yet, an unambiguous identification of the relevant surface speciation reactions is not possible, presumably due to a combination of intrinsic chemical complexity and heterogeneity of the nano-particle surfaces.

1. Introduction

In recent years, high resolution imaging and spectroscopy techniques in Atomic Force Microscopy (AFM) have generated unprecedented insights into structure and dissipation in liquids in the vicinity of solid surfaces. Certain organic liquids have attracted specific attention because of their model character and the simplicity of the dominant molecular interaction forces (*e.g.* van der Waals interactions), which gives rise – amongst others – to very pronounced and characteristic oscillatory solvation forces.[1–5] Compared to these systems, water and aqueous electrolytes are much more complex for several reasons including the strongly dipolar character of water molecules, the role of hydrogen bonding, the hydration of surfaces, and the almost unavoidable presence of ions.[6–12] In addition, solid surfaces, including AFM tips, typically acquire finite surface charges upon immersion into water. These surface charges give rise to rather long range electrostatic forces that decay exponentially with a decay length ranging from approximately 1 nm to 100 nm, depending on the salt concentration. Technically, long range electrostatic forces generate a background force that is superimposed onto the more short-ranged chemical forces such as surface and ion hydration forces that play a crucial role in atomic resolution imaging in aqueous environment. More importantly, long range electrostatic forces also provide the physical background field that controls the adsorption of ions, which has been found to have a strong effect not only on the average surface charge but also on the strength of oscillatory hydration forces in water.[7,8] A decent understanding and quantitative characterization of electrostatic interactions is therefore crucial for the interpretation of high resolution AFM experiments in aqueous environment.

In colloid science, the general principles controlling surface charge, ion adsorption, and electrostatic interaction forces are well established. Surface charge and ion adsorption are generally governed by an equilibrium between desorption and adsorption of protons and ions from and to specific sites on the surface.[13,14] To first approximation, the binding energies involved in these processes are governed by short range molecular forces that can be described by the equilibrium constants K (or their counterpart $\mathrm{p}K = -\log K$) of individual adsorption/

Physics of Complex Fluids Group and MESA+ Institute, Faculty of Science and Technology, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands. E-mail: f.mugele@utwente.nl

desorption reactions, or so-called speciation reactions. In addition to the equilibrium constants, the actual fractional coverage of adsorbed/desorbed species for a given situation depends on the *local* concentration of the ions next to the surface. The latter is proportional to the bulk concentration but it is modified by the local electrostatic potential at the surface. Because the latter itself is generated by the surface charge, determining the equilibrium charge density of solid-electrolyte interfaces requires a self-consistent solution of both adsorption/desorption equilibria and the electrostatic potential distribution in the vicinity of the interface. In the classical mean field picture, the solution is obtained by coupling the Poisson–Boltzmann (PB) equation for the distribution of ions and electrostatic potential in the diffuse part of the electric double layer to surface speciation reactions of a variable degree of complexity[15,16] for the adsorbed ions in the Stern part of the electric double layer. Colloidal and AFM force measurements necessarily involve the presence of two solid-electrolyte interfaces in close proximity. Electrostatic forces only arise once the diffuse parts of the double layers overlap. As a consequence, the electrostatic potential and the local concentration of ions that determine surface charge and fractional coverage of the surfaces change upon varying the distance between tip and sample. This phenomenon is known as charge regulation (CR) and was first described in detail by Ninham and Parsegian.[17] Since then, numerous colloidal surface force measurements using the surface forces apparatus (SFA)[18–21] as well as colloidal probe AFM force measurements[22–27] have established that the surface charge typically changes upon approaching two solid surfaces in an ambient electrolyte because of the CR behavior. In particular, the group of Borkovec (see ref. 28 and refs. therein) invested substantial effort to implement charge regulation models in colloid probe AFM force microscopy and to quantify the degree of charge regulation for a wide variety of materials. Because their primary interest was to describe colloidal interaction forces and not the specific surface chemistry, they introduced a so-called constant regulation approach that allows for describing force distance curves and extracting the net surface charges without explicitly specifying the individual surface speciation reactions.

While providing excellent average forces on a mesoscopic scale, colloidal probe AFM force measurements do not provide the lateral resolution that is required to characterize heterogeneous surfaces nor do they provide a sufficiently well-defined confinement geometry to help bridging the gap from the colloidal scale to atomic scale imaging of solid surfaces. To bridge this gap, we recently extended the principle of electrostatic surface characterization from colloidal science to dynamic AFM measurements with sharp cantilever tips and corresponding high lateral resolution of the order of the tip radius (several tens of nanometers).[12,29] Experiments with solutions of chloride salts of Na, K, Ca, and Mg demonstrated that the surface charge of silica and gibbsite surfaces strongly depends on the concentration and valency of the cations. Local surface charge density was extracted based on the asymptotic forces measured at distances much larger than the Debye screening length using solutions of the Poisson–Boltzmann equation under the classical constant charge (CC) or constant potential (CP) boundary condition. In that range, however, the measured forces are inherently small, which limits the accuracy of the measured charge densities. At smaller tip-sample separations, measured forces were found to fall in between the classical CC and CP solutions of the PB equation, indicating the occurrence of charge regulation. In the present work, we implement a data analysis procedure that includes charge regulation to the force measurements presented in ref. 12. Fig. 1(a) presents a typical silica/gibbsite composite sample and the corresponding measured forces for 10 mM NaCl solution at pH 5.8. The novelty of our approach is thus twofold: (i) we demonstrate the applicability and significancy of the charge regulation concept for AFM measurements with sharp tips and correspondingly high lateral resolution. (ii) We go beyond the common constant regulation approach[30,31] and extract directly equilibrium constants of surface speciation reactions using the full non-linear Poisson–

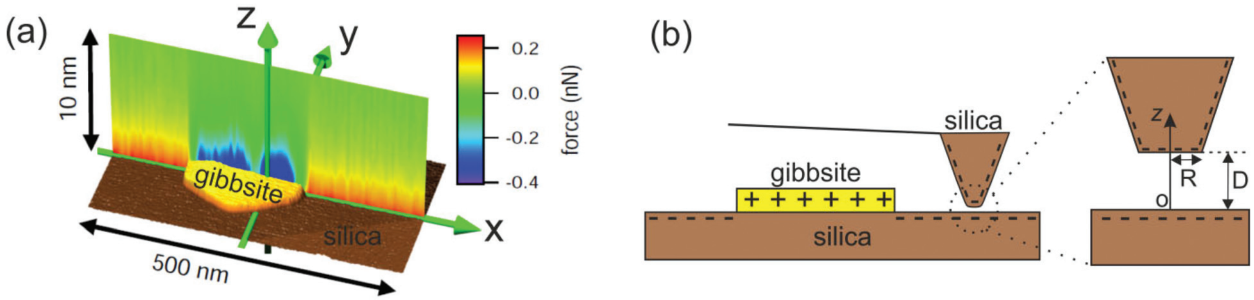


Fig. 1 (a) 3D view of a gibbsite nanoparticle adsorbed onto a silica surface along with force-*vs.*-distance curves (color coded) along a line section through the particle. Repulsive forces (red) upon approaching the silica surface indicate negative surface charge and attractive forces (blue) on the gibbsite particle indicate positive surface charge. Data are acquired with a negatively charged oxidized silicon tip in 10 mM NaCl solution at pH $\approx$ 5.8. Data adapted from ref. 12. (b) Schematic of the sample system (dimensions not to scale) consisting of a gibbsite platelet immobilized on a silica substrate. The magnification of the tip apex in the right part of the figure gives a more detailed view of the used tip geometry which is modeled as a truncated cone with a flat end having radius of $R \approx 30$ nm.

Boltzmann equation and charge regulation boundary condition. We discuss the applicability and reliability of our approach for common silica-electrolyte and for less common gibbsite-electrolyte interfaces.

The manuscript is organized as follows. In section 2, we recapitulate the general charge regulation scheme of coupling the PB description of the ion distribution in the diffuse layer to the surface complexation reactions and discuss specific aspects of its implementation for symmetric and asymmetric material combinations of tip and sample. In section 3, we describe the specific data analysis procedure, which involves the simultaneous optimization of parameters by a least-squared fitting of the charge regulation model to the experimental data obtained for variable fluid compositions. We identify well-defined equilibrium constants for deprotonation and cation adsorption reactions on silica. In section 4 we discuss the consequences of our results for AFM measurements in aqueous electrolytes in general. In particular, we extrapolate our results for silica to a wider range of fluid compositions and discuss possible limitations of the mean field approach inherent to our PB description of the electrolyte.

2 Theoretical framework

2.1 DLVO theory

Tip-sample interaction forces are analyzed in the framework of DLVO theory. In DLVO theory the disjoining pressure between two adjacent surfaces at distance D is decomposed into contributions from van der Waals interaction Π_{vdW} and electrostatic double layer forces Π_{el}.

$$\Pi(D) = \Pi_{\mathrm{vdW}} + \Pi_{\mathrm{el}} \tag{1}$$

Additional contributions to the disjoining pressure due to short range interactions such as hydration forces only become important at tip-sample separations of $\lesssim$1–2 nm. In the present analysis, we disregard these contributions. This implies that our model will only be applicable for tip sample separations beyond 1–2 nm.

Once the disjoining pressure is known, the force on the tip is calculated by integrating Π over the tip surface. For spherical probes as in colloidal probe AFM this is typically done using the Derjaguin approximation (see *e.g.* ref. 32). In our experiments, the AFM tips are slightly flattened leading to a local parallel plate geometry with a rather small contribution from the adjacent cone,[12,29] as sketched in Fig. 1(b). We therefore approximate the total force by

$$F(D) = \pi R^2 \Pi(D) \tag{2}$$

We estimate the absolute uncertainty of the procedure to be of order 10%.[29] Relative trends and the dependence of the force on the fluid composition, however, are not affected by these geometric uncertainties.

van der Waals forces. The contribution due to van der Waals forces is straight forward to analyze, because it can be written as an explicit function of D. Ignoring retardation effects, we can write for two parallel interfaces

$$\Pi_{\mathrm{vdW}}(D) = -\frac{A}{6\pi D^3} \tag{3}$$

where A is the Hamaker constant.

Electric double layer forces. The electrostatic contribution is the more interesting one because it contains the information on the surface chemistry that we are interested in. Yet, this information is contained in the expression for the electrostatic disjoining pressure only in a rather indirect fashion. Formally, we can write Π_{el} as

$$\Pi_{\mathrm{el}}(D) = k_{\mathrm{B}}T\sum_i (c_i(z) - c_{i\infty}) - \frac{\varepsilon\varepsilon_0}{2}\left(\frac{\mathrm{d}\psi}{\mathrm{d}z}\right)^2. \tag{4}$$

Π_{el} consists of a first contribution due to osmotic repulsion caused by local variations of the ion concentration and a second one due to direct electrostatic attraction (Maxwell stress). Here k_{B} is the Boltzmann constant, T is temperature, $\varepsilon\varepsilon_0$ the dielectric permittivity of water. In the first term, the sum runs over all ionic species i in the system. $c_{i\infty}$ is the bulk number concentration of corresponding ions. The solution of eqn (4) depends on the unknown functions $c_i(z)$ and $\psi(z)$, *i.e.*, the concentration profiles of all ionic species and electrostatic potential in the electrolyte at an arbitrary position $d_{\mathrm{s}} < z < D - d_{\mathrm{s}}$ between the two solid surfaces, where d_{s} is the thickness of the Stern layer. Making use of the fact that the ions follow the Boltzmann distribution, *i.e.* $c_i(z) = c_{i\infty}\exp(-Z_ie\psi(z)/k_{\mathrm{B}}T)$, we can calculate the potential distribution $\psi(z)$ by numerically solving the Poisson–Boltzmann (PB) equation

$$\frac{d^2}{\mathrm{d}z^2}\psi(z) = -\frac{e}{\varepsilon\varepsilon_0}\sum_i Z_ic_{i\infty}\exp\left(-\frac{Z_ie\psi(z)}{k_{\mathrm{B}}T}\right) \tag{5}$$

between the substrate surface and the tip using a standard Runge–Kutta algorithm. e represents the elementary charge, and Z_i is the valency of corresponding ions. Eqn (4) and (5) imply that the Π_{el} depends directly only on the field distribution and ion distribution in the diffuse part of the double layer. The surface chemistry that we are actually interested in enters the problem only *via* boundary conditions of eqn (5), which are determined by σ^{I} and σ^{II}, the net surface charge densities of tip and sample. Alternatively, the potentials $\psi(d_{\mathrm{s}})$ and $\psi(D - d_{\mathrm{s}})$ that are related to the surface charges *via* Gauss' law can be specified to solve eqn (5). Once $\psi(z)$ is known as a function of $\psi(d_{\mathrm{s}})$ and $\psi(D - d_{\mathrm{s}})$, we calculate the total charge in the diffuse layer, σ_{d}, by evaluating the integral

$$\sigma_{\mathrm{d}}(\psi(d_{\mathrm{s}}), \psi(D - d_{\mathrm{s}})) = \int_{d_{\mathrm{s}}}^{D-d_{\mathrm{s}}} \frac{e}{\varepsilon\varepsilon_0}\sum_i Z_ic_{i\infty}\exp\left(\frac{Z_ie\psi(z)}{k_{\mathrm{B}}T}\right)\mathrm{d}z \tag{6}$$

However, $\psi(d_{\mathrm{s}})$ and $\psi(D - d_{\mathrm{s}})$ (or equivalently σ^{I} and σ^{II}) are not known *a priori* and need to be determined self-consistently as part of the solution procedure. For an electric double layer at a single solid electrolyte interface this equation reduces to Grahame's equation[14] that relates the total diffuse layer charge density and the potential drop in a double layer.

2.2 Surface charge and charge regulation

As discussed above, the surface charge is controlled by adsorption and desorption equilibrium of protons and salt ions from the solution at the interface. We first recapitulate the problem for a single surface reaction (*i.e.* deprotonation) at the interface following the scheme originally described by Ninham and Parsegian[17] to explain the coupling between surface chemistry and diffuse layer physics using the so-called Gouy–Chapman model of double layer. Subsequently, we generalize this scheme to the actual situation of our experiments that involves several surface reactions including the adsorption of salt ions and the formation of a Stern layer.

Single deprotonation reaction. We consider a surface site SH that can deprotonate to produce a negatively charged site S^- following the simple chemical reaction

$$S^- + H^+ \rightleftharpoons SH \tag{7}$$

The reaction is characterized by an equilibrium constant K_{1H} with a corresponding pK value $pK_{1H} = -\log K_{1H}$. For the silica surfaces to be described below, SH would be simply a silanol group SiOH. The location of the equilibrium of the chemical reaction eqn (7) follows the law of mass action

$$\{S^-\}[H^+]_0 = K_{1H}\{SH\} \tag{8}$$

where curly brackets, { }, indicate surface concentrations and square brackets, [ ], indicate volume concentrations. $[H^+]_0$ is the local proton concentration at the S sites, *i.e.* directly at the surface. The total Γ density of surface sites S is fixed by the geometry and chemistry of the surface, leading to a conservation law

$$\{S^-\} + \{SH\} = \Gamma \tag{9}$$

Eqn (8) and (9) form a set of linear equations for the surface concentration $\{S^-\}$ and $\{SH\}$ that we can formally rewrite as a matrix equation

$$\begin{pmatrix} 1 & 1 \\ [H^+]_0 & -K_{1H} \end{pmatrix} \begin{pmatrix} \{S^-\} \\ \{SH\} \end{pmatrix} = \begin{pmatrix} \Gamma \\ 0 \end{pmatrix} \tag{10}$$

which can be solved for $\{S^-\}$ and $\{SH\}$. Physically, this approach is equivalent to treating the protons adsorbed to the fixed density of surfaces site S as a lattice gas of non-interacting particles with a chemical potential μ_s. This potential is equal to the chemical potential of a reservoir with a concentration $[H^+]_0$. The gain in chemical potential upon adsorption is $\Delta\mu_0 = k_B T \ln K_{1H}/[H^+]_\infty$.[33]

The above procedure results in a surface charge density that is given by

$$\sigma_0 = \sum_i q_i \{X_i\} \tag{11}$$

where $q_i = eZ_i$ is the charge of the surface group of species X_i. With eqn (7) as the only chemical reaction, eqn (11) thus reduces to

$$\sigma_0 = -e\{S^-\} = -\frac{e\Gamma}{1 + \frac{[H^+]_0}{K_{1H}}}. \tag{12}$$

Eqn (12) provides the surface charge as a function of the local proton concentration $[H^+]_0$ at the surface. $[H^+]_0$ deviates from the bulk concentration $[H^+]_\infty$ (which is fixed by the pH = $-\log[H^+]_\infty$ of the solution) because of the unknown electrostatic potential on the surface, ψ_0. $[H^+]_0$ is assumed to follow a Boltzmann distribution

$$[H^+]_0 = [H^+]_\infty e^{-e\psi_0/k_B T} \tag{13}$$

Together, eqn (12) and (13) lead to an expression $\sigma_0 = \sigma_0(\psi_0)$, *i.e.* an equation that connects the potential at the surface to the surface charge *via* the surface chemistry. This relation is the counterpart of eqn (6), which expresses the charge in the diffuse layer as a function of the potential at the surface. Together they assure charge neutrality, *i.e.* the charge on the surfaces of tip and sample have to be compensated by the charge in the diffuse layer.

$$\sigma^{I}(\psi_0{}^{I}) + \sigma^{II}(\psi_0{}^{II}) + \sigma_d(\psi(d_s), \psi(D - d_s)) = 0 \tag{14}$$

For the simple case of a symmetric system with tip and surface both made of the same material (*e.g.* silica in our experiments), we can write down the same potential-charge relations, eqn (12), for both surfaces. Using the simple Gouy–Chapman model of the electric double layer, we assume that the ionizable groups are located directly at the surface and that the Poisson–Boltzmann description of the diffuse layer extends all the way to surface. Hence, we identify $\psi_0{}^{I} = \psi(0)$ and $\psi_0{}^{II} = \psi(D)$. (In this case there is no Stern layer, so $d_s = 0$.) Solving eqn (14) we thus obtain the self-consistent distribution of the electrostatic potential, the salt ions in the solution and the ad/desorption of protons on the surface. The results are inserted into eqn (4) and (2) to calculate the total force, which is then expressed as a function of the parameters that characterize the charging reaction of the surface, *i.e.* the site density Γ and the equilibrium constant K_{1H}. Because Γ is usually fixed by the crystallography of the surface, K_{1H} is the parameter of primary interest that is determined by fitting to the experimental data.

Adsorption of several ions and Stern layer formation. The approach described above can be readily generalized to situations with several surface reactions, such as the adsorption of a cation of valency Z_c to a deprotonated surface site S^-

$$S^- + C^{z_c+} \rightleftharpoons S^-C^{z_c+} \tag{15}$$

with an equilibrium constant K_C. Additional possible reactions include for instance the adsorption of a proton (viz., protonation), and additional anion and cation adsorption reactions to either charged or uncharged sites, each accompanied by its own equilibrium constant.

$$\begin{array}{ll} \mathrm{SH} + \mathrm{H}^{+} \rightleftharpoons \mathrm{SH}_2^{+} & \text{with } K_{2\mathrm{H}} \\ \mathrm{SH}_2^{+} + \mathrm{A}^{z_a-} \rightleftharpoons \mathrm{SH}_2^{+}\mathrm{A}^{z_a-} & \text{with } K_{\mathrm{A}} \\ \mathrm{SH} + \mathrm{C}^{z_c+} \rightleftharpoons \mathrm{SHC}^{z_c+} & \text{with } K_{\mathrm{HC}} \\ \mathrm{SH} + \mathrm{A}^{z_a-} \rightleftharpoons \mathrm{SHA}^{z_a-} & \text{with } K_{\mathrm{HA}} \\ \ldots & \end{array} \tag{16}$$

If lateral interactions between adsorbed species are ignored, as usual in first order approximations, the location of the chemical equilibria is determined for each ionic species separately by a law of mass action using the relevant local concentration of the respective ion in the electrolyte.

While ions are treated as point-like in the Poisson–Boltzmann description of the diffuse layer, their finite size is usually taken into account when considering the position of the adsorbed ions in the Stern layer. Fig. 2 illustrates the Stern models of electric double layer for both surfaces considered in current study. Sophisticated implementations of this idea involve several planes away from the actual surface. Here we consider two planes the "0" plane and the "s" plane, which divides the double layer into a Stern layer and the diffuse layer. The (de-) protonation reaction takes place at the "0" plane, while electrolyte ions are adsorbed at the "s" plane.

This approach results in a generalization of the matrix eqn (10)

$$\begin{pmatrix} 1 & 1 & 1 & \ldots \\ [\mathrm{H}^{+}]_0 & -K_{1\mathrm{H}} & 0 & \ldots \\ [\mathrm{C}^{z_c+}]_{\mathrm{s}} & 0 & -K_{\mathrm{C}} & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{pmatrix} \begin{pmatrix} \{\mathrm{S}^{-}\} \\ \{\mathrm{SH}\} \\ \{\mathrm{S}^{-}\mathrm{C}^{z_c+}\} \\ \ldots \end{pmatrix} = \begin{pmatrix} \Gamma \\ 0 \\ 0 \\ \ldots \end{pmatrix} \tag{17}$$

where the triple dots indicate additional possible surface reactions. Similarly, identifying all possible charged complexes on the surface results in a generalization of eqn (11). Finally, solving the linear matrix eqn (17) leads to a generalized form of eqn (12). To evaluate that expression, the local concentrations of each ion at its specific adsorption plane must be calculated using the Boltzmann distribution with the local electrostatic potential at that plane.[15]

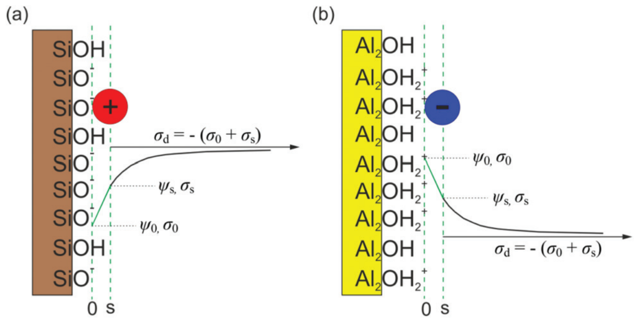


Fig. 2 Schematic of the basic Stern layer model of electric double layer on (a) silica and on (b) gibbsite. Protonated and deprotonated silanol and aluminol groups are located at the "0" plane. Adsorbed cations (Na^+, K^+) on silica and Cl^- anions on gibbsite from the solution are located at the "s" plane to neutralize partially the deprotonated silanol and protonated doubly coordinated aluminol sites. The region between "0"and "s" planes is the so-called Stern layer, and potential drop inside it is linear because it does not contain any space charge. Beyond the "s" plane is the diffuse part of electric double layer, and potential drop is exponential because of non-zero space charge density in the diffuse layer.

2.3 Specific implementation

Silica surfaces in contact with NaCl and KCl solutions. In general, the correct identification of the relevant surface reactions in systems involving several components is rather challenging and – given the indirectness of force and electrokinetic measurements – involves substantial uncertainties. Silica in contact with aqueous solutions of NaCl or KCl of moderate concentration is an ideal model system, arguably the best characterized one in the literature. In this case, we can restrict the reactions to the deprotonation of silanol groups, eqn (7) with SiOH as SH and SiO^- as S^- sites, and the adsorption of a single monovalent cation species, see eqn (15), SiO^-Na^+ or SiO^-K^+ as S^-C^+. So, we consider only the 3 × 3 matrix equation explicitly written in eqn (17). With reference to the double layer structure of silica shown in Fig. 2(a), SiOH and SiO^- sites are located in the "0" plane, and cations are adsorbed at the "s" plane. The charge densities at the "0" and the "s" planes are then given by

$$\sigma_0 = -e(\Gamma - \{\mathrm{SiOH}\}) = -e\{\mathrm{SiO}^{-}\} - e\{\mathrm{SiO}^{-}\mathrm{C}^{+}\} \tag{18}$$

$$\sigma_{\mathrm{s}} = e\{\mathrm{SiO}^{-}\mathrm{C}^{+}\} \tag{19}$$

Because the space charge density between the "0" plane and the "s" plane vanishes, the potential drop in the Stern layer is linear and can be expressed as

$$\psi_0 - \psi_{\mathrm{s}} = \frac{\sigma_0}{C_{\mathrm{s}}} \tag{20}$$

where C_s is the capacitance of the Stern layer.

Together, σ_0 and σ_s define the net or 'effective' charge of each surface in the AFM experiments, *i.e.* we can write $\sigma^{I} = \sigma_0^{I} + \sigma_s^{I}$ for the tip and $\sigma^{II} = \sigma_0^{II} + \sigma_s^{II}$, respectively. These two expressions are inserted in eqn (14) to obtain the solution. Because the charge in the diffuse layer compensates the net charges σ^{I} and σ^{II} and the Poisson–Boltzmann equation is only applied between the *s* planes of the two surfaces, we identify $\psi(d_s) = \psi_s^{I}$ and $\psi(D - d_s) = \psi_s^{II}$ in eqn (14) in the presence of a Stern layer with adsorbed ions. If tip and sample are of the same material (*e.g.*, AFM silica tip over silica substrate), we use the additional simplification $\sigma^{I} = \sigma^{II}$ and $\psi(d_s) = \psi(D - d_s)$. The specific chemical reactions and parameter values required for the data analysis are summarized in Table 1.

Gibbsite surfaces in contact with NaCl and KCl solutions. While the surface complexation on silica surfaces as described above is well established and widely accepted, the origin of the surface charge on the basal plane of gibbsite is less clear. The basal plane of gibbsite has only a doubly coordinated surface group, Al_2OH. Crystallographically, there are six different Al_2OH groups per surface unit cell of gibbsite. The classic mul-

Table 1 Parameters of used in force analysis based on the CR-complemented DLVO theory. Note: the pK values of surface reactions are optimized to fit the experiments, while other parameters are measured or kept to literature values

Surfaces		
Parameters	Silica	Gibbsite basal plane
"0" plane reaction	$SiOH \rightleftharpoons SiO^- + H^+$ with pK_{1H}	$Al_2OH_2{}^+ \rightleftharpoons Al_2OH + H^+$ with pK_{2H}
"s" plane reaction	$SiO^-C^{z_c+} \rightleftharpoons SiO^- + C^{z_c+}$ with pK_C	$Al_2OH_2{}^+A^{z_a-} \rightleftharpoons Al_2OH_2{}^+ + A^{z_a-}$ with pK_A
Stern layer capacitance C_s	2.9 F m^{-2} (ref. 59)	1.49 F m^{-2} (ref. 38)
Site density of surface group Γ	8 nm^{-2} (ref. 59)	13.8 nm^{-2} (ref. 38 and 59)
Hamaker constant A	0.65×10^{-20} J (ref. 26 and 29)	1.2×10^{-20} J (ref. 29 and 39)
Stern layer thickness d_s	0.4 nm (ref. 35 and 60)	
AFM Tip diameter $2R$	52 ± 5 nm (SEM measurement)	
pH of solution	5.8 (measurement)	

tisite complexation (MUSIC) method[33–35] treats all of them the same and predicts that these surface groups are electrically neutral in the pH range of 4 to 10. Bickmore *et al.*[36] treated all OH groups individually in an advanced model that incorporates the bond-valence theory and reveals the molecular structure of the surface *ab initio*. According to those calculations one of the Al_2OH groups per surface unit cell can be protonated with a pK value around 5. Jodin *et al.*[37] considered the possibility of surface relaxation (*e.g.*, the bending of the Al–O–H angle) in their bond-valence MUSIC calculations, and yielded the pK value for protonation of the basal doubly coordinated surface group in the range of 2 to 4. The latter two calculations concur with some experiments[38–40] about the reactivity of doubly coordinated groups. Based on these studies, we tentatively describe our data on gibbsite by the following surface speciation reactions.

$$Al_2OH_2{}^+ \leftrightharpoons Al_2OH + H^+ \quad \text{with } K_{2H} \tag{21}$$

$$Al_2OH_2{}^+Cl^- \leftrightharpoons Al_2OH_2{}^+ + Cl^- \quad \text{with } K_A \tag{22}$$

The charge densities at the "0" and "s" planes are then given by

$$\sigma_0 = e\{Al_2OH_2{}^+\} + e\{Al_2OH_2{}^+Cl^-\} \tag{23}$$

$$\sigma_s = -e\{Al_2OH_2{}^+Cl^-\} \tag{24}$$

The structure of the electric double layer near gibbsite is sketched in Fig. 2(b). Because the charge density between the "0" plane and the "s" plane vanishes, the electrostatic potential ψ_s at the "s" plane can be calculated again from the potential and charge density at "0" plane using eqn (20).

For the calculation of the charge distribution and the forces, we now use the $\sigma^I = \sigma_0{}^I + \sigma_s{}^I$ for the silica tip as obtained in the preceding section and combine it with $\sigma^{II} = \sigma_0{}^{II} + \sigma_s{}^{II}$ for the gibbsite surface. The characterization of the surface charge and surface chemistry of the AFM tip, first using a silica substrate, is thus a necessary prerequisite to characterize the substrate of interest.

3 Results and discussion

3.1 Least-squared fitting

The force *vs.* distance curves calculated using the theoretical model depend on a number of parameters, including the radius of the AFM tip R, the Hamaker constant A, the site density Γ, the capacitance of the Stern layer C_s, and the equilibrium constants K_i of the surface speciation reactions considered. The last are the primary parameters of interest here. Therefore, we use reasonable estimates for the former ones based on tip calibration measurements and literature values as described in Table 1. Only the equilibrium constants K_i are used as free parameters to optimize the agreement between experimental data and calculated model curves. We define a merit function

$$Q(K_1, K_2, \ldots) = \frac{1}{\sum\limits_{j=1}^{N} \left(F_t(D_j) - F_{\exp}(D_j)\right)^2} \tag{25}$$

where F_t and $F_{\exp}$ denote the theoretically calculated and the experimentally measured force value at the distance D_j. Best fit values for the fit parameters are calculated by maximizing Q within a reasonably chosen range of values for the K_i's (or the corresponding pK_i's) under consideration. Such ranges are chosen based on literature data and refined manually in the course of the fitting procedure. To increase the speed of the optimization procedure, we evaluate $F_t(D)$ on a reduced number of points, $N = 15$, chosen equidistantly within the range $2 < D < 15$ nm and determined $F_t(D_j)$ by interpolating between the adjacent $F_t(D)$ values. Empirical tests showed that the use of a larger number of evaluation points had no significant effect on the fit quality and the resulting optimum parameter values, while significantly reducing the speed of the fit process because the optimization involves the solution of the non-linear Poisson–Boltzmann equation at each distance. The minimum separation of 2 nm was chosen to minimize the influence of short range forces such as hydration forces that are not included in the physical model. The maximum value of 15 nm is based on the rapid decay and the dominance of noise beyond that distance for the conditions of the present set of experimental data.

3.2 Surface charge and optimization of p*K* values

Very good fits of individual force curves are easily obtained upon optimizing the merit function Q for each individual fluid composition both on silica and on gibbsite surfaces (Fig. 3). For data acquired on silica, the deviation between model curve and experimental data can be reduced below the symbol size in Fig. 3 for the entire parameter range of interest, *i.e.* for $2 < D < 15$ nm. The optimized model curves that include the CR boundary condition (solid lines in Fig. 3) describe the experimental data of a significantly wider range than the approximate solutions for constant potential (CP) and constant charge (CC) solutions (dashed lines in Fig. 3). For gibbsite, the description of the data by the CR solution is also much better than for the CC and CP solutions. Yet, the quality is not quite as good as in the case of silica and deviations are seen already at tip sample separations of ≈2 nm.

The fit curves shown in Fig. 3 result from a simultaneous global optimization of the data for all salt concentrations. If optimized individually for each concentration, much better fits than shown in the figure can be obtained. Yet, such an approach would be inconsistent with our modeling that assumes concentration-independent equilibrium constants for each reaction. However, it turns out that good agreement between model curve and experimental data for any single salt concentration is not sufficient to determine a unique set of equilibrium constants. Rather, a whole range of combinations of fit parameters provides fits of similar quality. Fig. 4(a) illustrates this observation for a specific data set, where deprotonation of the silanol group and adsorption of Na^+ ions were taken into account in the modeling. This representation of the merit function Q clearly demonstrates that the optimum values of pK_{1H} and pK_{Na} are highly correlated. Fits of equally excellent quality can be obtained for the wide range of parameters shown by the narrow ridge forming the maximum of $Q(pK_{1H}, pK_{Na})$. These results clearly show that a reliable measurement of p*K* values based on force curves for a single fluid composition is impossible when several surface reactions occur simultaneously.

In fact, this result does not come as a surprise. As the theoretical analysis described in the preceding section showed, the electrostatic part of the disjoining pressure, eqn (4), experienced by the AFM tip depends on the surface chemistry only *via* the boundary conditions, $\psi(d_s)$ and $\psi(D - d_s)$ or *via* the corresponding charge density of the diffuse layer, *i.e.* *via* the global charge neutrality condition, eqn (14). Any combination of surface chemical reactions with adequate fractional adsorption that generates the same surface charge gives rise to the same force in the AFM measurement and can thus not be distinguished, as shown in Fig. 4(b). For the present situation, deprotonation and adsorption of monovalent Na^+ (or K^+) cations give rise to the same surface charge density and thus

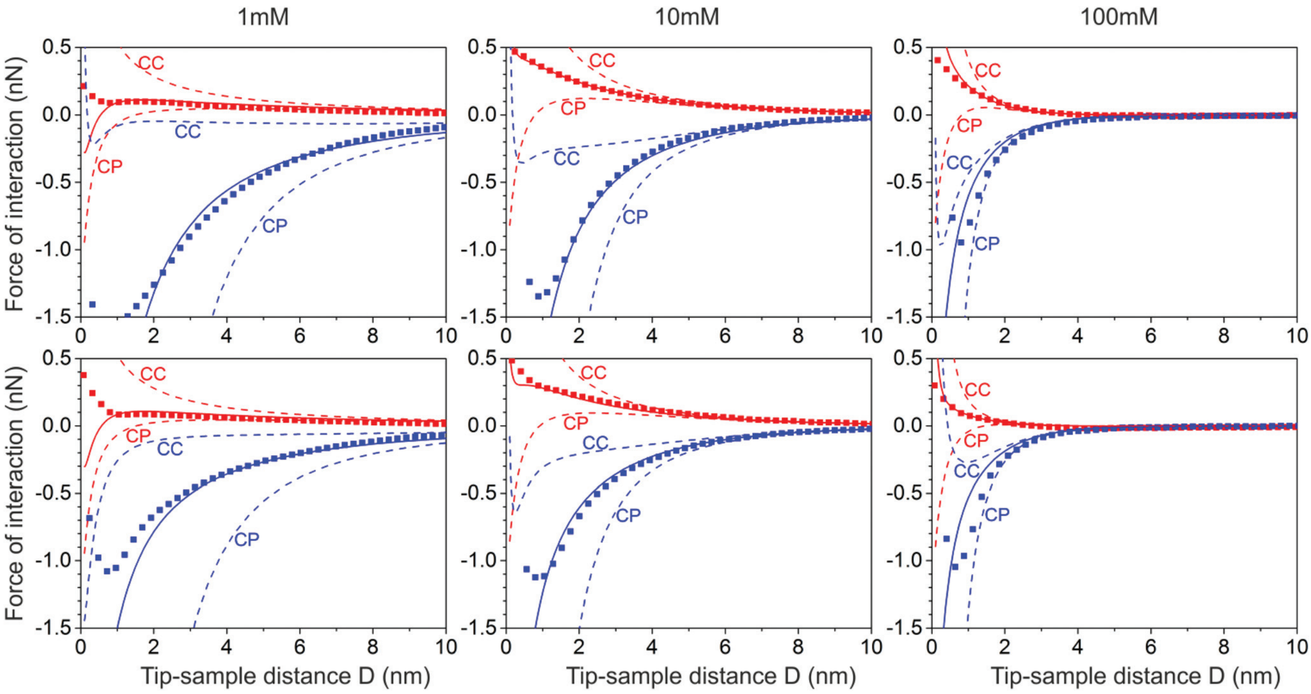


Fig. 3 A comparison of experimental tip-sample force curves (taken from ref. 12) with theoretical force curves for the silica-silica (coded in red) and silica-gibbsite(coded in blue) interaction in aqueous solutions of two monovalent salts (NaCl: first row, KCl: second row) under three different concentrations (1 mM: left column, 10 mM: middle column and 100 mM: right column). Symbols: laterally averaged AFM forces; solid lines: force prediction using charge regulation (CR) with globally optimized p*K* values (see Fig. 5); dashed lines: constant charge (CC) and constant potential (CP) force predictions.

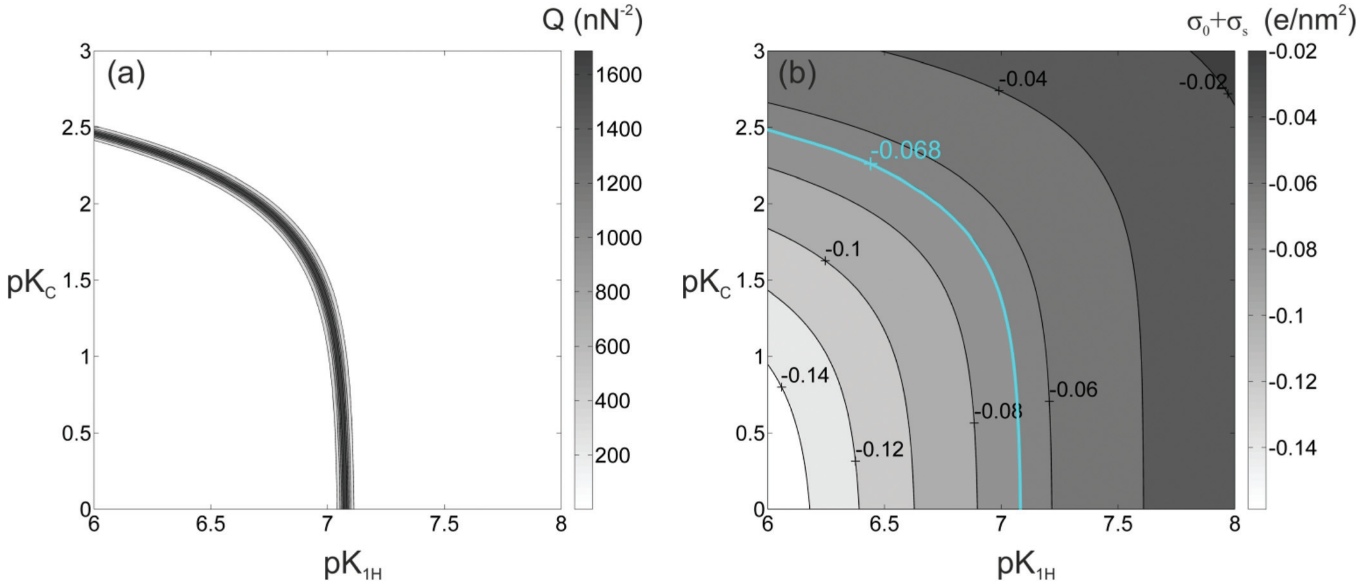


Fig. 4 Correlation of fit parameters pK_{1H} and pK_C on silica surface for a single salt concentration (10 mM NaCl). (a) Contour plot of the fit quality merit function Q (inverse of squared error). (b) Contour plot of net surface charge, $\sigma_0 + \sigma_s = -\sigma_d$, in the limit $D \to \infty$, which illustrates that the quantity probed by the AFM experiment is the diffuse layer charge. Any pK pairs, (pK_{1H}, pK_C), falling on the cyan thick solid line produce a same diffuse layer charge density of $-0.068e/\text{nm}^2$, and thus a same force.

to the same force. The small difference arising from the different locations in the "0" plane and in the "s" plane is apparently insufficient to create a significant difference in the forces. Specifically, Fig. 4a shows that the experimental data can be explained by assuming exclusively (de-) protonation with a value of $pK_{1H} \approx 7.05$ with negligible Na^+ adsorption with arbitrary $pK_{Na} < 0.5$. If pK_{Na} is assumed to be larger than 0.5, pK_{1H} has to be decreased slightly to compensate for the weak adsorption of Na^+.

Similar results are obtained for other concentrations and for KCl, both on silica and on gibbsite surfaces (data not shown). For each combination of materials, the correlation curves on silica have the same qualitative shape. A differently shaped family of correlation curves is obtained on gibbsite. Yet, the exact location of the curves in the parameter space depends on the salt concentration: for higher salt concentrations the contribution of adsorbing Na cations is more pronounced, for lower salt concentrations the effect of deprotonation is dominant. The fact that the exact location of the correlation curves depends on the salt concentration is a direct indication that the adsorption of Na^+ ions does indeed contribute to the surface charge on silica. Because the surface chemistry should be same for all conditions, we can further constrain the equilibrium constants by requiring that one consistent set of pK values – pK_{1H} and pK_C for silica, and pK_{2H} and pK_A for gibbsite – should be obtained by simultaneous optimization of Q for all salt concentrations investigated. Fig. 5 shows the merit function Q_{all} for both silica and gibbsite surfaces determined from all the concentrations in the range of 1 mM to 100 mM (*i.e.*, 1, 3, 10, 30, 100 mM). Forces obtained at the lowest (0.5 mM) salt concentration were excluded from the analysis because the measured forces were overall too low for reliable fitting within the distance range of interest.

Fig. 5a shows the merit function for a superposition of data acquired for the silica surface at different NaCl concentrations. Indeed, it yields a well-defined combination $pK_{1H} = 6.9 \pm 0.3$ and $pK_{Na} = 1.65 \pm 0.1$ for which the model curves simultaneously describe all experimental force curves. From the measurements with KCl solutions the corresponding optimum values are $pK_{1H} = 6.9 \pm 0.3$ and $pK_K = 2.0 \pm 0.2$, as shown in Fig. 5b. Note that the value for pK_{1H} obtained for the two different salts coincides as one should expect. The deprotonation of silanol groups is one of the most widely studied surface reactions in the literature and the reported pK_{1H} values (obtained with potentiometric titration) typically fall in the range 7 to 7.5.[33,41–44] Our result includes the lower part of this range. The exact values are known to depend on the origin and the specific preparation conditions of the silica surfaces. For instance, the surface chemistry of our oxidized surfaces of bulk silicon both on the tip and on the sample surface is expected to differ from bulk amorphous silica such as the one precipitated from silanes in a Stöber reaction that is frequently used in colloidal studies of silica. In addition, we note that deviations of the order of 0.5 pK units may also result from uncertainties in other experimental parameters such as the exact geometry and radius of the AFM tip and physical assumptions related to the Poisson–Boltzmann mean field approach.

Comparison between NaCl and KCl solutions shows a somewhat stronger adsorption for K^+ than for Na^+ ions on silica. This difference, which was already visible in our

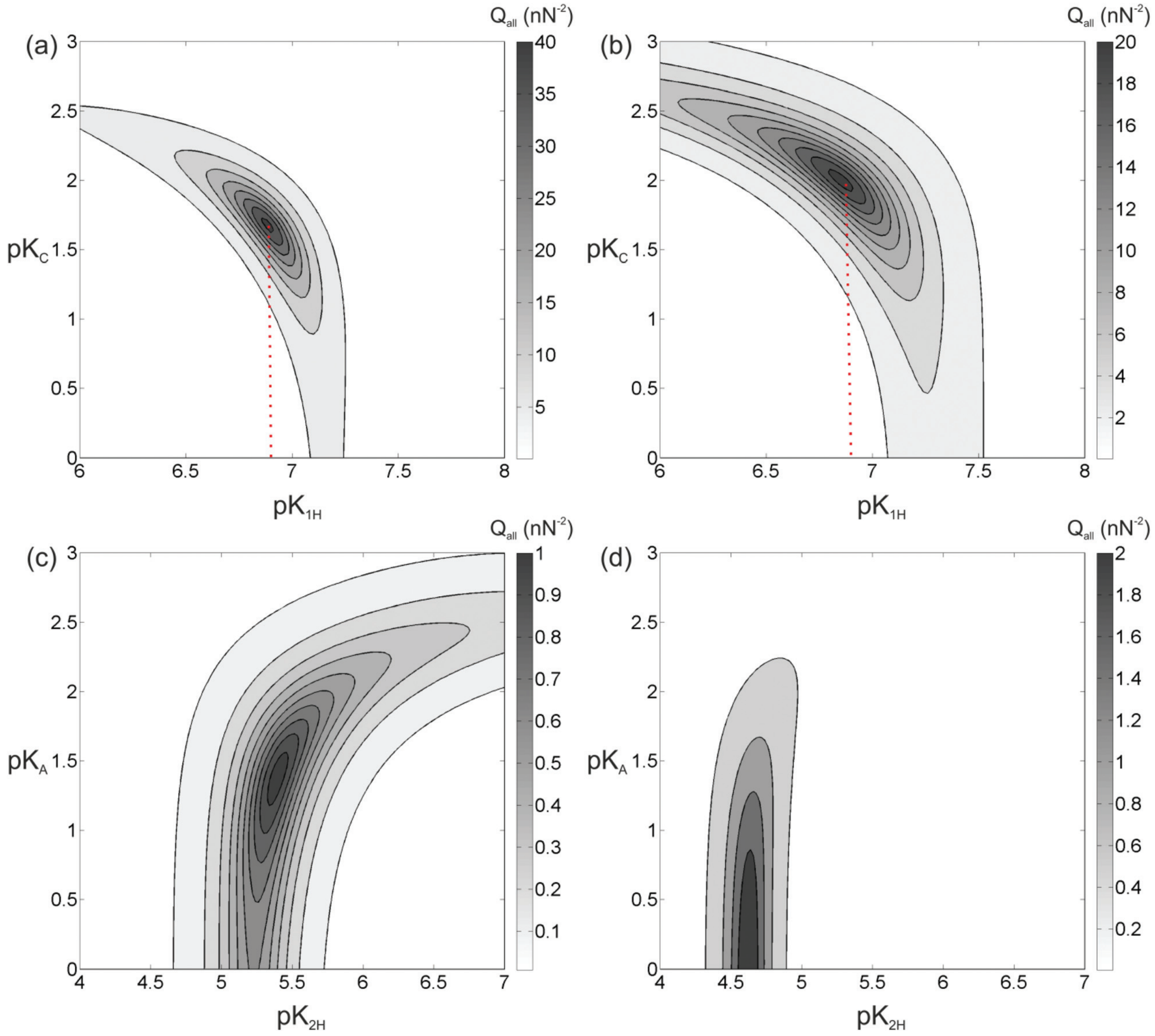


Fig. 5 Contour of the inverse of squared error (Q_{all}) in the parameter space defined by pK values of surface reactions. (a) Q_{all} in parameter space (pK_{1H}, pK_C) for silica in NaCl solution. (b) Q_{all} in parameter space (pK_{1H}, pK_C) for silica in KCl solution. (c) Q_{all} in parameter space (pK_{2H}, pK_A) for gibbsite in NaCl solution. (d) Q_{all} in parameter space (pK_{2H}, pK_A) for gibbsite in KCl solution. For silica surface, local maxima in (a) and (b) provide best estimates of pK_{1H} = 6.9 ± 0.3, pK_{Na} = 1.65 ± 0.1, pK_K = 2.0 ± 0.2. For gibbsite surface, local maxima in (c) and (d) provide best estimates of pK_{2H} = 5.4 ± 0.3/pK_{Cl} = 1.4 ± 0.6 for NaCl and pK_{2H} = 4.6 ± 0.2/pK_{Cl} < 0.7 for KCl, respectively.

approximate analysis of the data based on the constant charge model,[12] is consistent with earlier reports in the literature based on SFA,[21] colloidal probe AFM,[26] and electrokinetic measurements.[45] It is consistent with the idea that the slightly larger K^+ ions are more polarizable than Na^+ and hence experience a somewhat stronger dispersion attraction towards the surface.[46] Simultaneously, the hydration shell of K^+ is somewhat less strongly bound. This may lead to somewhat weaker hydration repulsion between the hydrated ion and the hydrated silica surface.

Similar results can be obtained for gibbsite surfaces. From the plots, we extract optimum values of pK_{2H} = 5.4 ± 0.3 and pK_{Cl} = 1.4 ± 0.6 for the measurements with NaCl (Fig. 5c) and pK_{2H} = 4.6 ± 0.2 and pK_{Cl} < 0.7 for the measurements with KCl (Fig. 5d). The quality of the analysis for gibbsite is less satisfying than in the case of silica discussed above. First, the maximum value of Q_{all} for the optimum combination of pK values is more than an order of magnitude lower, showing that the quality of the fits is not as good as for silica. Second, if the surface speciation reactions, eqn (21) and (22), chosen to model the data are correct, the values of pK_{2H} and pK_{Cl} should actually be the same for both NaCl and KCl solutions. The analysis shows that this is not quite the case. Given the fact that the experimental data are obtained in the course of the same measurement as the very satisfying results on silica, we conclude that the quality of the experimental data as such cannot be the cause of these deviations. This suggests that the specific surface speciation model chosen to describe the data is not appropriate. In fact, we tested a few other possibilities of surface complexation, such as the last two reactions given in eqn (16). None of them yielded more consistent results than the reactions chosen here. Possibly, the actual surface chemistry is in fact much more complex and requires, *e.g.* the involvement of hydration water. The latter was found to be necessary to understand the adsorption of divalent Mg^{2+} and Ca^{2+} cations onto the same surface, as atomically resolved images of these ions on gibbsite in combination with extensive density functional theory (DFT) calculations showed.[12] For those ions, the DFT calculations suggested that six partially deprotonated molecules of hydration water play a crucial role for the observed effective surface charge. For monovalent cations of Na^+ and K^+, the same type of calculation did not reveal a comparable ordered arrangement of adsorbed ions. Yet, it is clear, also from molecular dynamics simulations,[11,47] that rather complex configurations of cations and water molecules can appear at clay-electrolyte interfaces that may be too complex to be captured by a few simple surface complexation reactions as eqn (21) and (22). Moreover, the observed heterogeneity of the force curves on the gibbsite particles points to an intrinsic heterogeneity of the surface charge of the particles that might be caused by intrinsic structural and/or chemical defects on the surface. In this context it is worth noting that the absolute value of the surface charge densities in the present experiments is rather low in all cases, typically of the order of $0.1e/\mathrm{nm}^2$ (see Fig. 6). For a typical tip-sample interaction area of the order of 500 nm^2, this means that the tip typically probes no more than a few tens of charges on each surface. Surface defects carrying – say – 10 elementary charges therefore already produce substantial deviations that undermine the idealized approach of the surface speciation reactions discussed in section 2.3. These considerations also highlight the remarkable degree of homogeneity of the silica surfaces.

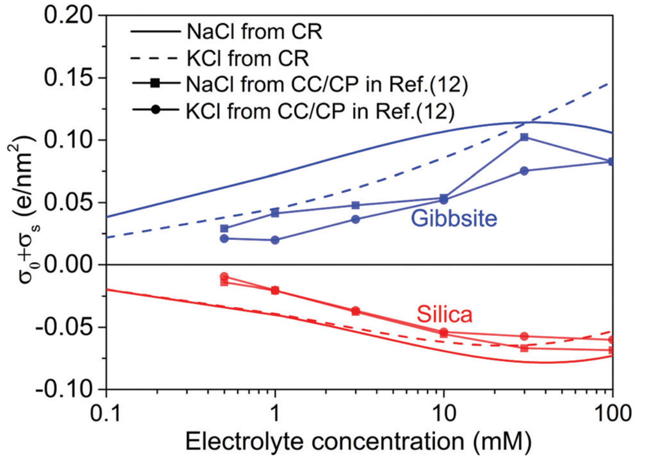


Fig. 6 The unregulated (no interaction between two surface, *i.e.*, $D \to \infty$) diffuse layer charge density, $\sigma_0 + \sigma_s$, as a function of electrolyte concentration for two monovalent electrolytes used in the current study. For comparison purposes, the results determined from CC/CP boundaries in ref. 12 is also included in the plot.

4 Discussion

4.1 Surface charge density

Fig. 6 shows the effective surface charge density $\sigma_0 + \sigma_s$ of silica and gibbsite in both NaCl and KCl solutions as calculated based on the surface speciation reactions with optimized p*K* values for individual solid-electrolyte interfaces, *i.e.* extrapolated to infinite tip-sample separation. On both surfaces, the absolute value of the surface charge density is found to increase with increasing salt concentration. This is due to the fact that increased concentration improves the electrostatic screening and thereby reduces the cost in free energy involved in the creation of surface charge.

Next to the solid and dashed lines representing the results with current charge regulation model, the graph also shows as symbols the results from approximate analysis of the tails of the force–distance curves with constant charge/constant potential model reported earlier in ref. 12. The earlier data display the same trends as the present more sophisticated analysis, yet, that analysis clearly underestimated the absolute values of the charge density. Considering the fact that the charge regulation model produces much better fits of the experimental force curves than the constant charge/constant potential model (see Fig. 3) does, it is thus reasonable to believe that the surface charge extracted from charge regulation model is indeed more reliable.

Note that the surface charge densities obtained for gibbsite should be trusted notwithstanding the uncertainties discussed in the preceding section. As our discussion of Fig. 4 showed, good fits of the force curves imply a correct measurements of the surface charge density. Yet, they do not guarantee the correct identification of the surface chemistry. For the rest of the discussion, we will focus on silica surfaces and explore the consequences of the specific surface speciation reactions that we can trust.

4.2 Charge regulation and local fluid composition

Using the optimized combination of p*K* values, we can analyze the composition of the fluid and the coverage of specific species on the silica surface as a function of the tip-sample separation. Fig. 7 summarizes some in general terms well-known (see *e.g.* ref. 14) basic results for the distribution of protons that arises from the self-consistently determined potential $\psi(z)$ using the Boltzmann distribution. If the separation of tip and sample is large compared to the Debye length, the system displays a typical behavior of the so-called thin double layer limit. The surface charges are screened in a diffuse layer extending a distance of the order of the Debye screening length. In the middle of the gap, the electrolyte has its bulk composition. Under these reference conditions, charge regulation has no effect, as shown by the overlapping

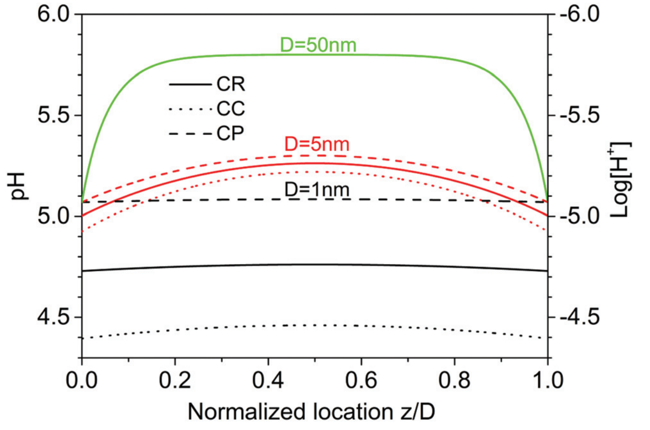


Fig. 7 Effect of charge regulation on pH/proton concentration distribution between tip and silica sample within 10 mM NaCl solution for tip-sample separations of D = 50 nm (green), 5 nm (red) and 1 nm (black) corresponding to weak, intermediate and strong double layer overlap (Debye length: $1/\kappa$ = 3 nm).

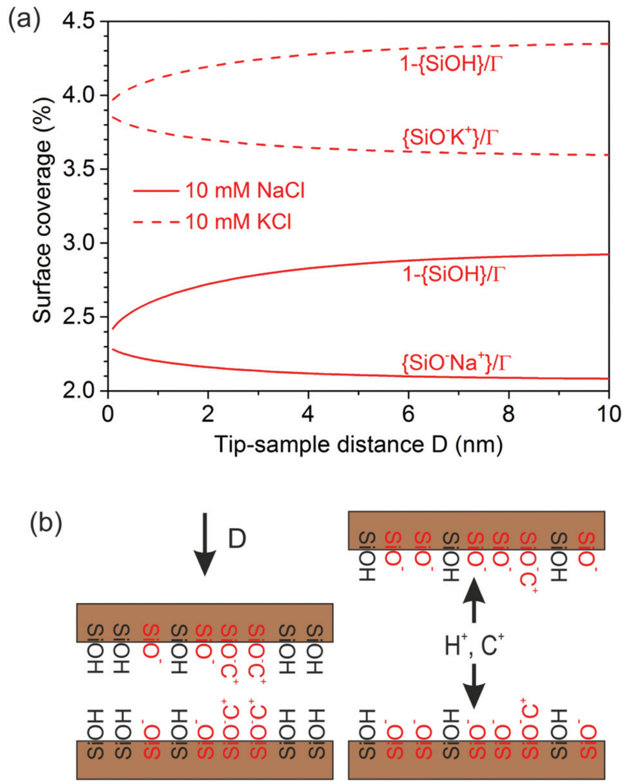


Fig. 8 (a) Fraction of total deprotonated and ion-occupied surface sites in solutions of NaCl and KCl at 10 mM for silica/silica interaction. (b) Schematic illustration of proton and ion transfer during the course of charge regulation as the AFM tip move towards the silica sample. Upon decreasing the tip-sample distance both protons and cations migrate from the solution to the silica surfaces.

blue curves on the top of Fig. 7. For tip-sample gaps of the order of the Debye length or less, the effect of charge regulation becomes apparent. The local concentration of protons increases with decreasing tip-sample separation. Charge regulation then leads to a partial re-adsorption of protons to the solid surface, as a comparison between the dotted lines for the CC boundary and the solid lines for the CR boundary shows. A CP boundary condition (dashed lines) would lead to even more pronounced re-protonation of the surface. At the smallest separation (black curves), the proton concentration is almost constant all across the gap, as expected for the thick double layer limit. Note that the Na^+ and the K^+ ions follow exactly the same Boltzmann distribution, except for the generally different limiting concentration in the bulk.

It is also interesting to consider explicitly the consequences for the composition of the surface. It turns out in the first place that the total degree of deprotonation, $\Gamma - \{\mathrm{SiOH}\}$, is rather low as expected for the conditions of our experiments at pH ≈ 6, see Fig. 8(a). Given the typical site density of 8 sites per nm^2, this low degree of deprotonation implies that the average separation between charged sites on the surface is of the order of a few nanometers, which is comparable to the diameter of supersharp AFM tips that are typically used for high resolution AFM imaging. From that perspective, it is not surprising that supersharp tips hardly feel the presence of such small degrees of surface charge. One may also question, whether the use of a continuous surface charge density is still appropriate under such conditions. The second notable feature in Fig. 8(a) is that most of the deprotonated silanol groups on the surface directly adsorb a cation from the solution. Addition of salt thus promotes the replacement of surface-bound protons by cations. This observation holds for all conditions shown in Fig. 3, and is slightly more pronounced for the slightly more strongly adsorbing K^+ ions than for Na^+. The majority of the surface charge is thus compensated directly in the Stern layer rather than in the diffuse part of the double layer. This conclusion is consistent with X-ray reflectivity studies,[48,49] optical measurements,[50,51] and recent molecular simulations.[47,52] One consequence of this observation is that the surface charge as determined from an AFM (or SFA) force measurement, which is based on the ion distribution in the diffuse layer, is always lower than the charge density determined by a titration measurement that measures the total number of protons or ions adsorbing to or desorbing from a surface.[53]

As a final remark, Fig. 8(a) also shows that the coverage of the various species on the surface does not depend very strongly on the tip-sample separation. As expected, both protons and cations condense onto charged SiO^- sites as tip and sample are brought closer together and the total residual surface charge decreases, as shown schematically in Fig. 8(b). Yet, the total variation between infinite separations and a minimum separation $D \approx 1$ nm, down to which the model is reasonably applicable, is typically of the order of several percent. This implies that the chemical composition of the surface in this range is not dramatically altered by the presence of the AFM tip. This conclusion is crucial for the interpretation of AFM measurements in general.[7,11,12,54–57] It illustrates that the structure of a solid–electrolyte interface as probed by AFM in high resolution spectroscopy or in high resolution imaging mode is hardly affected by the presence of the tip. While charge regulation has a strong effect on the net surface charge and hence the measured forces, as shown in Fig. 3, the fraction of adsorbed ions still provides a fairly good representation of a single interface in contact with a bulk electrolyte for all the conditions studied here.

Some caveats apply. Obviously, this conclusion holds within the limitations of the present mean field Poisson–Boltzmann model. Under conditions of atomic resolution imaging, where short range chemical and hydration forces play an important role, the picture may be altered. Moreover, the present considerations are limited to a symmetric system, in which tip and sample are made of the same material. If the material of tip and sample behave very differently, *e.g.* one acting as a proton donor and the other one as a proton acceptor, the effect of bringing of close proximity may be more pronounced. Yet, our results obtained so far all suggest that the tip can be reasonably well considered as a moderate perturbation of a state that is overall governed by the properties of the individual solid–electrolyte interface.

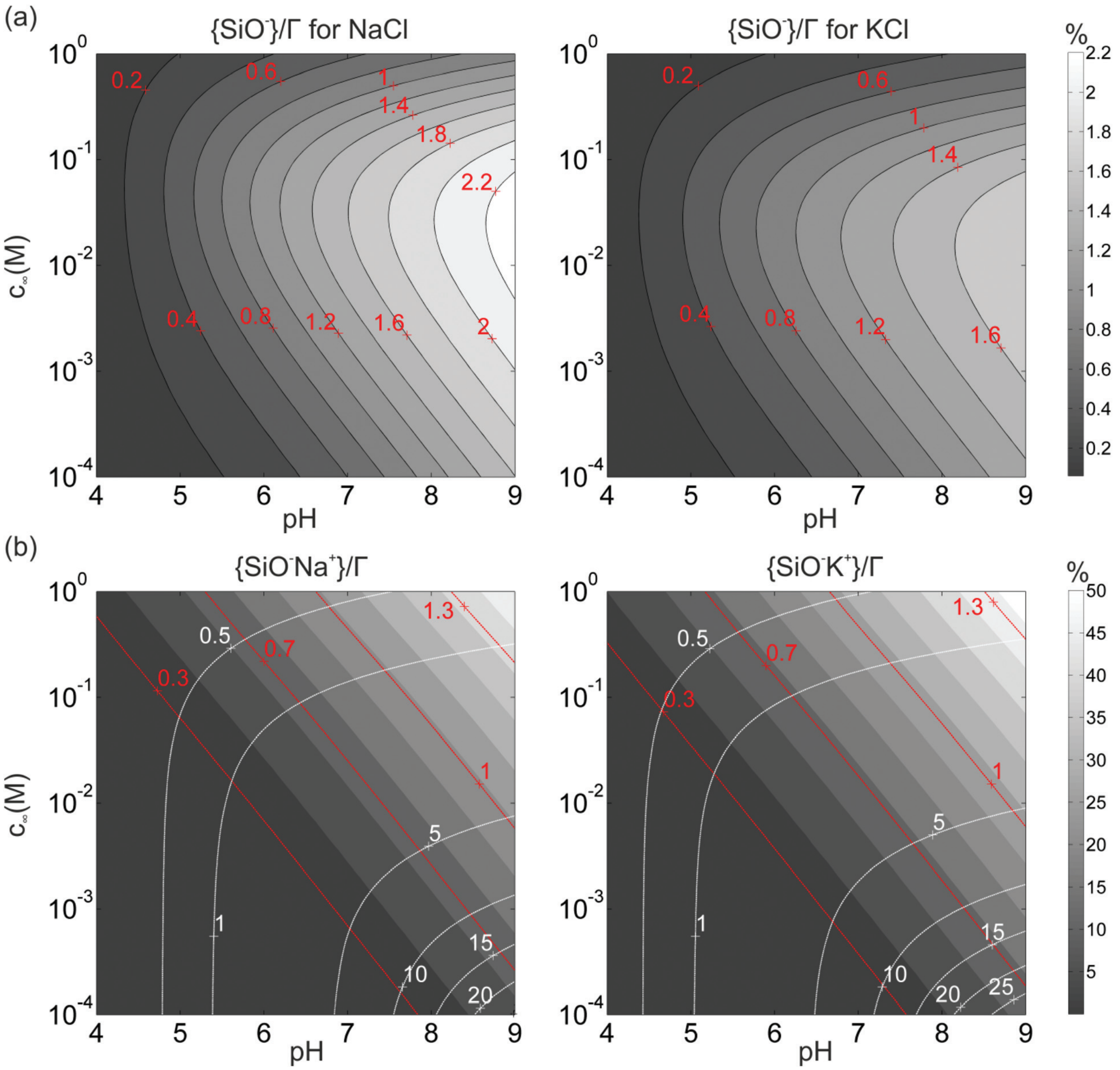


Fig. 9 Calculated fraction of surface complexes for individual silica-electrolyte interfaces as a function of pH and salt concentration. (a) $\{\mathrm{SiO^-}\}/\Gamma$ in NaCl (left) and KCl (right) solutions. (b) $\{\mathrm{SiO^-Na^+}\}/\Gamma$ complexes (left) and $\{\mathrm{SiO^-K^+}\}/\Gamma$ complexes (right) in the corresponding salt solutions. The pK values are those globally optimized in Fig. 5. Plot (b) also includes the contour lines of two dimensionless parameters, $\sqrt{\{\mathrm{SiO^-C^+}\}}/\kappa$ (white dash lines) and $2a\sqrt{\{\mathrm{SiO^-C^+}\}}$ (red dash lines). (Note the difference in gray scale in a) and b).).

4.3 Extrapolation to variable pH and limitations of Poisson–Boltzmann approach

The good consistency of the analysis presented so far for silica surface, encourages us to extrapolate our data to a broader range of fluid compositions, including in particular conditions of variable pH. Substituting the globally optimized pK values of deprotonation and Na^+ and K^+ adsorption into our numerical scheme, we calculate the concentration of surface species in the limit $D \to \infty$ for silica surface, and the results are shown in Fig. 9. The calculations display several expected qualitative trends. For both NaCl and KCl, the fraction of free deprotonated SiO^- groups monotonically increases with increasing pH at all concentrations (Fig. 9a). In contrast, the behavior as a function of the salt concentration at fixed pH is non-monotonic because two opposing processes compete. On the one hand, the increasing salt concentration improves the screening of electric fields and thereby reduces the energetic cost for the system to increase the surface charge by deprotonation. This trend prevails for low salt concentrations. This process competes, however, with the formation of SiO^-C^+ complexes on the surface according to eqn (15). At higher salt concentrations, the latter process dominates and causes a decrease of the fractional coverage of SiO^-, along with an increase of $\{SiO^-C^+\}$. As expected, $\{SiO^-C^+\}$ increases monotonically both with increase pH and with increasing salt concentration, see Fig. 9b.

This extrapolation of the model predictions to a broader range of fluid compositions also allows for a systematic discussion of the limitations of the applicability of the Poisson–Boltzmann treatment applied in this study. Various deviations from the simple Poisson–Boltzmann picture have been discussed in the literature.[58] The most important corrections include the breakdown of the mean field approach due to direct electrostatic correlations between adjacent ions and the neglect of the finite radius a of the ions. A criterion for the validity of the mean field approach can be determined by comparing the average separation between adjacent adsorbed cations C^+ on the surface to the Debye screening length. The white solid lines in Fig. 9b show iso-lines of fixed values of the ratio $\sqrt{\{SiO^-C^+\}}/\kappa$. If this ratio is not too large, many screening charges are found between adjacent surface charges and hence the mean field picture is expected to hold – and *vice versa*. Fig. 9b thus shows that the mean field approach primarily becomes questionable at high pH for low salt concentrations, *i.e.* for conditions of poor screening but nevertheless high degrees of deprotonation.

Similarly, we can consider the ratio between the diameter $2a$ of the (hydrated) ions and the average distance between adsorbed ions, *i.e.* $2a\sqrt{\{SiO^-C^+\}}$. The corresponding red iso-lines in Fig. 9b show that neglecting the finite ion radius is acceptable provided that the pH and salt concentration are not simultaneously high. For the specific conditions of our experiments (pH $\approx$ 6 and $c_\infty = 5 \times 10^{-4}$ to 0.1 M) the two criteria are indeed decently fulfilled, as Fig. 9b shows.

5 Conclusions

We used AFM force measurements with sharp tips to probe the surface chemistries/charging behavior of two interacting solid surfaces in aqueous solutions of NaCl and KCl of variable concentration. In agreement with earlier studies using colloidal probe force microscopy, we find that the measured force–distance curves between tip and sample can be described quantitatively down to tip-sample separations of 1–2 nm by taking into account charge regulation (CR), *i.e.* adsorption/desorption of protons and salt ions as induced by the proximity of tip and sample. For silica surfaces, our systematic approach of simultaneously analyzing experimental data obtained for a variety of salt concentrations allows for quantifying the equilibrium constants for the deprotonation of silanol groups and for the adsorption of Na^+ and K^+ cations. Given the fact that the vast majority of AFM experiments are carried out with tips made of oxidized silicon, our results can be used to quantify the charge of most AFM tips in some of the most common aqueous electrolyte solutions. Our measurements on the gibbsite surface illustrate the usefulness of this knowledge. Although the identification of the correct surface speciation reactions fails on the more complex gibbsite surfaces, the ability to quantify the charge density on the AFM tip nevertheless enables the measurements of the net surface charge density and thus provides a useful characterization of the material. In contrast to earlier colloidal probe AFM measurements, this is now possible in AFM measurements with sharp tips and a lateral resolution of the order of the tip diameter (in nanometers).

Acknowledgements

We thank David Andelman for interesting discussions. This work was financially supported by the Exploratory Research (ExploRe) program of BP plc.

References

1 A. Maali, T. Cohen-Bouhacina, G. Couturier and J.-P. Aimé, *Phys. Rev. Lett.*, 2006, **96**, 086105.
2 G. B. Kaggwa, J. I. Kilpatrick, J. E. Sader and S. P. Jarvis, *Appl. Phys. Lett.*, 2008, **93**, 011909.
3 S. H. Khan, G. Matei, S. Patil and P. M. Hoffmann, *Phys. Rev. Lett.*, 2010, **105**, 106101.
4 S. de Beer, D. van den Ende and F. Mugele, *Nanotechnology*, 2010, **21**, 325703.
5 F. Liu, S. de Beer, D. van den Ende and F. Mugele, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 2013, **87**, 062406.
6 J. J. Kuna, K. Voitchovsky, C. Singh, H. Jiang, S. Mwenifumbo, P. K. Ghorai, M. M. Stevens, S. C. Glotzer and F. Stellacci, *Nat. Mater.*, 2009, **8**, 837–842.
7 T. Fukuma, *Sci. Technol. Adv. Mater.*, 2010, **11**, 033003.

8 J. I. Kilpatrick, S.-H. Loh and S. P. Jarvis, *J. Am. Chem. Soc.*, 2013, **135**, 2628–2634.

9 J. M. Black, D. Walters, A. Labuda, G. Feng, P. C. Hillesheim, S. Dai, P. T. Cummings, S. V. Kalinin, R. Proksch and N. Balke, *Nano Lett.*, 2013, **13**, 5954–5960.

10 E. T. Herruzo, H. Asakawa, T. Fukuma and R. Garcia, *Nanoscale*, 2013, **5**, 2678–2685.

11 M. Ricci, P. Spijker and K. Voïtchovsky, *Nat. Commun.*, 2014, **5**, 4400.

12 I. Siretanu, D. Ebeling, M. P. Andersson, S. L. S. Stipp, A. Philipse, M. C. Stuart, D. van den Ende and F. Mugele, *Sci. Rep.*, 2014, **4**, 4956.

13 R. J. Hunter, *Zeta Potential in Colloid Science: Principles and Applications*, Academic Press, 1981.

14 H. J. Butt, K. Graf and M. Kappl, *Physics and Chemistry of Interfaces*, Wiley, 2006.

15 J. A. Davis, R. O. James and J. O. Leckie, *J. Colloid Interface Sci.*, 1978, **63**, 480–499.

16 J.-F. Boily, J. Lützenkirchen, O. Balmès, J. Beattie and S. Sjöberg, *Colloids Surf., A*, 2001, **179**, 11–27.

17 B. W. Ninham and V. A. Parsegian, *J. Theor. Biol.*, 1971, **31**, 405–428.

18 J. N. Israelachvili and G. E. Adams, *J. Chem. Soc., Faraday Trans. 1*, 1978, **74**, 975–1001.

19 R. M. Pashley and J. N. Israelachvili, *J. Colloid Interface Sci.*, 1984, **97**, 446–455.

20 V. E. Shubin and P. Kékicheff, *J. Colloid Interface Sci.*, 1993, **155**, 108–123.

21 J. P. Chapel, *Langmuir*, 1994, **10**, 4237–4243.

22 H. J. Butt, *Biophys. J.*, 1991, **60**, 1438–1444.

23 I. Larson, C. J. Drummond, D. Y. C. Chan and F. Grieser, *J. Am. Chem. Soc.*, 1993, **115**, 11885–11890.

24 P. Kékicheff and O. Spalla, *Phys. Rev. Lett.*, 1995, **75**, 1851–1854.

25 R. Pericet-Camara, G. Papastavrou and M. Borkovec, *Macromolecules*, 2009, **42**, 1749–1758.

26 M. Dishon, O. Zohar and U. Sivan, *Langmuir*, 2009, **25**, 2831–2836.

27 I. Popa, P. Sinha, M. Finessi, P. Maroni, G. Papastavrou and M. Borkovec, *Phys. Rev. Lett.*, 2010, **104**, 228301.

28 F. J. Montes Ruiz-Cabello, G. Trefalt, P. Maroni and M. Borkovec, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 2014, **90**, 012301.

29 D. Ebeling, D. van den Ende and F. Mugele, *Nanotechnology*, 2011, **22**, 305706.

30 S. L. Carnie and D. Y. C. Chan, *J. Colloid Interface Sci.*, 1993, **161**, 260–264.

31 R. Pericet-Camara, G. Papastavrou, S. H. Behrens and M. Borkovec, *J. Phys. Chem. B*, 2004, **108**, 19467–19475.

32 H.-J. Butt, B. Cappella and M. Kappl, *Surf. Sci. Rep.*, 2005, **59**, 1–152.

33 T. Hiemstra, J. C. M. De Wit and W. H. Van Riemsdijk, *J. Colloid Interface Sci.*, 1989, **133**, 105–117.

34 T. Hiemstra, P. Venema and W. H. V. Riemsdijk, *J. Colloid Interface Sci.*, 1996, **184**, 680–692.

35 T. Hiemstra, H. Yong and W. H. Van Riemsdijk, *Langmuir*, 1999, **15**, 5942–5955.

36 B. R. Bickmore, C. J. Tadanier, K. M. Rosso, W. D. Monn and D. L. Eggett, *Geochim. Cosmochim. Acta*, 2004, **68**, 2025–2042.

37 M.-C. Jodin, F. Gaboriaud and B. Humbert, *J. Colloid Interface Sci.*, 2005, **287**, 581–591.

38 J. Rosenqvist, P. Persson and S. Sjöberg, *Langmuir*, 2002, **18**, 4598–4604.

39 Y. Gan and G. V. Franks, *Langmuir*, 2006, **22**, 6087–6092.

40 G. V. Franks and Y. Gan, *J. Am. Ceram. Soc.*, 2007, **90**, 3373–3388.

41 M. L. Hair and W. Hertl, *J. Phys. Chem.*, 1970, **74**, 91–94.

42 K. R. Marshall, G. L. Rochester and C. H. Simpson, *J. Chem. Ind.*, 1974, **19**, 775–776.

43 N. Sahai and D. A. Sverjensky, *Geochim. Cosmochim. Acta*, 1997, **61**, 2801–2826.

44 S. H. Behrens and D. G. Grier, *J. Chem. Phys.*, 2001, **115**, 6716–6721.

45 G. V. Franks, *J. Colloid Interface Sci.*, 2002, **249**, 44–51.

46 B. W. Ninham and V. Yaminsky, *Langmuir*, 1997, **13**, 2097–2108.

47 I. C. Bourg and G. Sposito, *J. Colloid Interface Sci.*, 2011, **360**, 701–715.

48 Z. Zhang, P. Fenter, L. Cheng, N. C. Sturchio, M. J. Bedzyk, M. Předota, A. Bandura, J. D. Kubicki, S. N. Lvov, P. T. Cummings, A. A. Chialvo, M. K. Ridley, P. Bénézeth, L. Anovitz, D. A. Palmer, M. L. Machesky and D. J. Wesolowski, *Langmuir*, 2004, **20**, 4954–4969.

49 C. Park, P. A. Fenter, K. L. Nagy and N. C. Sturchio, *Phys. Rev. Lett.*, 2006, **97**, 016101.

50 M. Porus, C. Labbez, P. Maroni and M. Borkovec, *J. Chem. Phys.*, 2011, **135**, 064701.

51 L. Wang, C. Zhao, M. H. G. Duits, F. Mugele and I. Siretanu, *Sens. Actuators, B*, 2015, **210**, 649–655.

52 O. Kroutil, Z. Chval, A. A. Skelton and M. Předota, *J. Phys. Chem. C*, 2015, **119**, 9274–9286.

53 J. Lyklema, *Fundamentals of Interface and Colloid Science: Liquid-Fluid Interfaces*, Elsevier Science, 2000.

54 K.-i. Fukui, H. Onishi and Y. Iwasawa, *Chem. Phys. Lett.*, 1997, **280**, 296–301.

55 R. García and R. Pérez, *Surf. Sci. Rep.*, 2002, **47**, 197–301.

56 S.-H. Loh and S. P. Jarvis, *Langmuir*, 2010, **26**, 9176–9178.

57 K. Voïtchovsky, *Nanotechnology*, 2015, **26**, 100501.

58 D. Ben-Yaakov, D. Andelman, D. Harries and R. Podgornik, *J. Phys.: Condens. Matter*, 2009, **21**, 424106.

59 T. Hiemstra, W. H. Van Riemsdijk and G. H. Bolt, *J. Colloid Interface Sci.*, 1989, **133**, 91–104.

60 J. Newman and K. E. Thomas-Alyea, *Electrochemical Systems*, John Wiley & Sons, 2004.

Chem Soc Rev

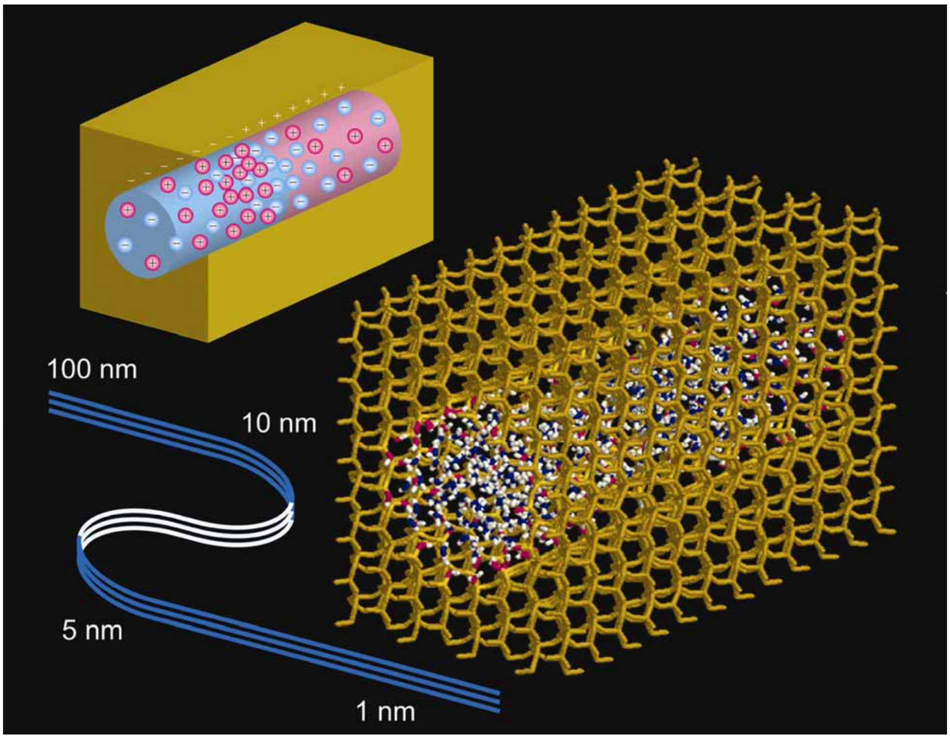
This article was published as part of the

From microfluidic application to nanofluidic phenomena issue

Reviewing the latest advances in microfluidic and nanofluidic research

Guest Editors Professors Albert van den Berg, Harold Craighead and Peidong Yang

Please take a look at the issue 3 table of contents to access other reviews in this themed issue



Soft wetting

Jacco Snoeijer

- Elastocapillary interfaces
- Soft wetting statics
- Soft wetting dynamics

○	Thursday :	11h00-12h30
○	Friday :	11h00-12h30
○	Friday :	16h00-17h30

Capillarity Driven Instability of a Soft Solid

Serge Mora,* Ty Phou, and Jean-Marc Fromental
Laboratoire des Colloïdes, Verres et Nanomatériaux, UMR 5587, Université Montpellier 2 and CNRS, Place Eugène Bataillon, F-34095 Montpellier Cedex, France

Len M. Pismen
Department of Chemical Engineering and Minerva Center for Nonlinear Physics of Complex Systems, Technion—Israel Institute of Technology, Haifa 32000, Israel

Yves Pomeau
Department of Mathematics, University of Arizona, Tucson, Arizona, USA

(Received 20 July 2010; published 17 November 2010)

We report the observation of a Plateau instability in a thin filament of solid gel with a very small elastic modulus. A longitudinal undulation of the surface of the cylinder reduces its area thereby triggering capillary instability, but is counterbalanced by elastic forces following the deformation. This competition leads to a nontrivial instability threshold for a solid cylinder. The ratio of surface tension to elastic modulus defines a characteristic length scale. The onset of linear instability is when the radius of the cylinder is one-sixth of this length scale, in agreement with theory presented here.

DOI: 10.1103/PhysRevLett.105.214301 PACS numbers: 46.32.+x, 46.25.−y, 47.20.Dr, 83.10.Ff

The Rayleigh-Plateau instability (RPI) [1–4] results from the tendency of a given volume of liquid to reduce its area at constant volume. The area of a geometrical cylinder is obviously not at a minimum, this one being reached, as is well known, for a sphere. A long-wave modulation of the surface of a cylinder is unstable: near the maxima of the radius, the Laplace pressure, which is dominated by the azimuthal curvature, decreases and pushes the fluid outward, the converse being true for the minima of the radius where the increase of the Laplace pressure pushes the fluid inward.

Although surface tension exists in solids as well, its effect on the pattern formation is believed to be unobservable at macroscopic scale because surface energy is negligible compared to elastic energy of deformation. Perhaps the only phenomenon where surface tension plays a role is crystal faceting [5] where it does not compete with elastic forces. The RPI may be, however, relevant in soft solids, such as those found very often in biology, and so may play a role in such biological processes as generation of filaments, formation of beaded forms in myelinated nerve fibers, etc. It also provides a unique method for fabricating undulating cylinders with a wavelength controlled by physical parameters. Such cylinders with undulation period in the range of optical wavelength could show, for instance, forbidden wave bands for the propagation of light.

In solids, contrary to liquids, the energy has a volume part that changes alongside modulation of the external surface. As is often noticed, the balance between the two kinds of energies, capillary and elastic, depends on a quantity with the dimension of length $l = \gamma/\mu$, where γ is surface tension and μ is the elastic shear modulus. In usual solids, this is a very small length scale: because of its origin in atomic interactions, one expects l to be of the order of the range of atomic interactions, about a fraction of a nanometer. Therefore the capillary effect, in the numerator of the small length scale, should be typically negligible. Nevertheless, in a very soft solid like a gel just above the percolation threshold, this length scale can be macroscopic. The reason is that the complex molecular structure of such materials reduces by many orders of magnitude the "typical" value of the shear modulus computed from the standard molecular parameters, such as the size of an atom and the energy of a covalent bond. Measured values for γ are few tens of mN/m and for μ measured values are few tens of Pa (very small by comparison with ordinary materials); therefore, l, as well as the expected typical length scale for elastic RPI, may go up to the millimeter range, well above any microscopic length scale of this kind of material.

To demonstrate RPI in a solid, we have used a standard agar gel (Merck KGaA, Germany) dissolved in purified water. Small amounts of methylene blue were added to agar solutions to aid observations. Agar is known to dissolve in boiling water and to form a gel upon cooling to about 35 °C [6,7]. Upon cooling, the shear modulus first rapidly increases and then stabilizes (Fig. 1). A cooled gel behaves as an incompressible elastic solid over a wide frequency range (at least from 300 to 10^{-3} Hz; Fig. 1). We have carefully checked rheological properties of agar gels and measured the shear modulus at various concentrations (from 0.5% to 0.16%).

The experiments are carried out as follows. Liquid solutions of agar (90 °C) with various concentrations are first

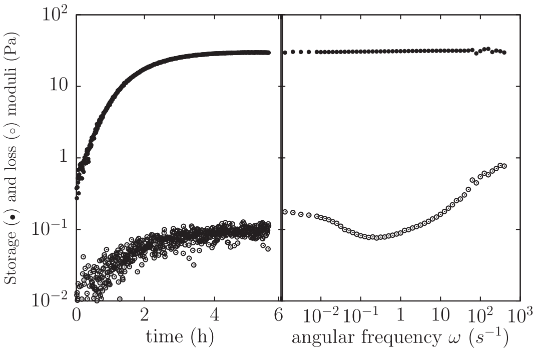


FIG. 1. Linear rheological properties of a 0.18% agar gel hydrogel. The curves are obtained by dynamic oscillatory shear tests, using a strain controlled rheometer (ARES-RFS from TAInstruments) in Couette geometry. Left: Evolution of the storage and loss moduli as functions of time. Right: Storage and loss moduli as functions of the angular frequency 5 h after cooling.

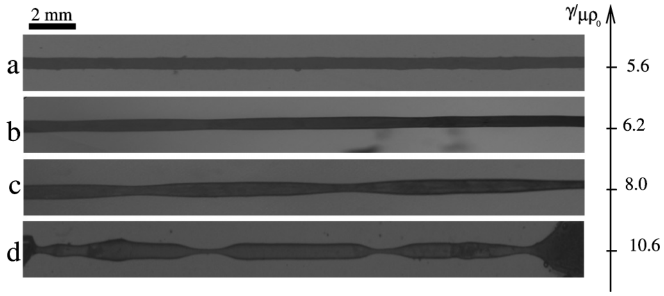


FIG. 2. Equilibrium shape of agar gel cylinders for different values of the shear modulus. Radius is $\rho = 240\ \mu$m, surface tension is $\gamma \simeq 36.5$ mN/m. Shear modulus varies from 12 to 27 Pa. Note the RPI instability for values of $\gamma/(\mu\rho)$ larger than 6.2.

injected in cylindrical moulds made of cellular polystyrene. These moulds are fabricated using two cuboid pieces of cellular polystyrene ($3 \times 0.5 \times 0.5$ cm^3). One of the larger faces of each piece is heated just above the glass transition temperature of polystyrene. The two hot sides are then assembled and a 3 cm long copper wire of a desired diameter is inserted between them. This wire is removed after cooling at room temperature, leaving a hollow cylinder of the same dimensions within the polystyrene block. We have checked by optical microscopy that the roughness of the surface is less than 4 μm. The mould is then preheated in order to prevent partial gelation before the liquid is completely injected.

After injecting the agar solution and cooling for 5 h at room temperature, the mould is dissolved in liquid toluene. Total dissolution takes about 3 min. The strand of agar gel is then released in toluene. The agar gel—toluene surface tension γ is roughly equal to the water-toluene surface tension; this value is used in further computations. To prevent the agar cylinder from wrapping up, the two ends are fixed in a frame before releasing. This yields strands measuring about 2 cm long floating in toluene. Depending on the mould, their radius lies in the range 150–260 μm.

Depending on the initial strand radius ρ_0 and the shear modulus μ of the agar gel, the growth of a surface instability takes place during the mould dissolution. The final steady pattern is seen after dissolution is complete. Strands of agar gel with a high concentration and/or a large radius retain a cylindrical shape after the mould dissolution [Fig. 2(a)]. Strands with a low concentration and a small radius are systematically breaking into two during the dissolution. For intermediate strands, surface undulations develop just after dissolution and remain permanently [Figs. 2(b)–2(d)]. When an unstable filament is gently stretched in the middle, and then released, it recovers its length and shape, thereby demonstrating that the undulation pattern is stable. On the contrary, if pure water is injected into the moulds instead of the agar gel, the released strand breaks, as expected, into separate spherical droplets.

We used toluene saturated with water to prevent shrinking. This makes a fundamental difference between our experiments and those reported by Matsuo and Tanaka [8]. In their case, the instability is driven by diffusion of the gel solvent into the miscible outer fluid. The slowly developing instability they observe cannot be linked to a RPI, because there is no sharp interface and so no surface tension in their experiment.

Within the setup we used, the resolution for the amplitude of the modulations is about 15 μm. To obtain the critical elastic modulus (at a fixed radius) below which cylinders remain straight, the amplitude is plotted as a function of the elastic modulus and fitted by the power law

$$f(\mu) = \alpha(\mu - \mu_c)^\beta \tag{1}$$

with adjustable parameters α, β, and μ_c (Fig. 3); μ_c is the shear modulus at the instability threshold. In this way, we succeed in separating unambiguously the cases where a cylinder is either stable or not. Figure 4 summarizes the experimental stability data in the $\mu - \rho_0$ plane. The plane is divided into two areas, one corresponding to stable straight cylinders and the other to unstable ones.

Just above the threshold [Fig. 2(b)], the instability leads to a varicose shape. Farther away from the threshold, the shape becomes more complicated, with large constant-radius areas interrupted by constrictions [Fig. 2(d)]. In the following, we focus on the physics near the threshold [Figs. 2(a) and 2(b)]. The analysis far beyond the instability threshold requires a nonlinear theory that will be the subject of future work.

Suppose that the surface of a cylinder is perturbed by a small axisymmetric modulation from a constant radius ρ_0 to $\rho(z) = \rho_0 + \lambda(z)$, where z is the coordinate along the axis and $\lambda(z) \ll \rho_0$ (Fig. 5). The mean curvature κ of the surface changes from $1/\rho_0$ to $1/\rho_0 - \lambda(z)/\rho_0^2 - \lambda''(z)$. This yields a Laplace pressure contribution $\gamma\kappa$ to be added to the boundary conditions (bc) for the normal stress on the surface of the cylinder.

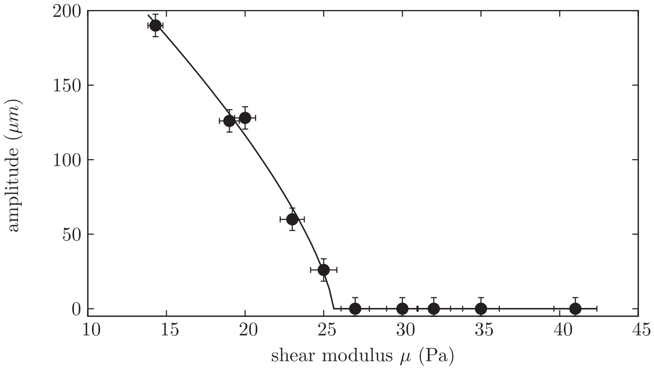


FIG. 3. The modulation amplitude as a function of the shear modulus for a fixed radius of the strands ($\rho = 240$ μm). The solid line is the best fit according to Eq. (1) with $\mu_c = 25.6$ Pa, $\alpha = 34.1$ μm, and $\beta = 0.71$.

Our material can be considered as incompressible, so that its only relevant Lamé coefficient is the shear modulus μ. Using a variational formulation, we write the elastic energy in the cylindrical coordinates r, z as

$$\mathcal{E} = 2\pi \int dz \int_0^\rho dr r \Big[\mu\Big(u_{r,r}^2 + u_{z,z}^2 + \frac{u_r^2}{r^2} + \frac{1}{2}(u_{r,z} + u_{z,r})^2\Big) - p\Big(u_{r,r} + \frac{u_r}{r} + u_{z,z}\Big)\Big], \tag{2}$$

where u_r, u_z are the radial and axial displacements, the indices preceded by a comma denote respective partial derivatives, and p is the Lagrange multiplier imposing the incompressibility condition

$$u_{r,r} + \frac{u_r}{r} + u_{z,z} = 0. \tag{3}$$

By variation with respect to the displacements u_r and u_z, one gets the Cauchy-Poisson equations in cylindrical coordinates:

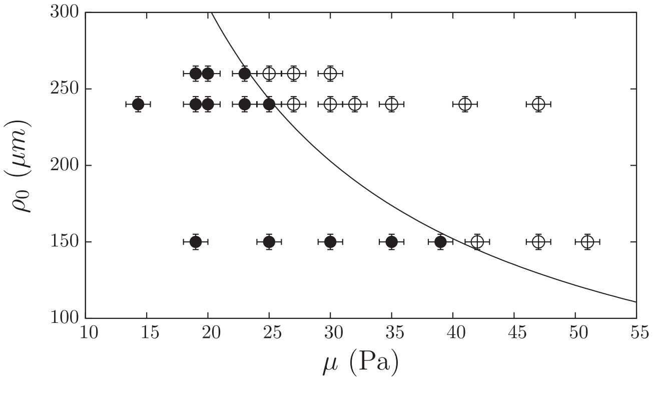


FIG. 4. A series of agar cylinders with different shear modulus has been investigated for each radius. The empty symbols correspond to stable cylinders and filled symbols to unstable cylinders acquiring a varicose shape. The parametric plane $\rho_0 - \mu$ is separated into two domains by the theoretical curve $\gamma/(\mu\rho_0) = 6$ (without fitting parameters) derived in this Letter.

$$\mu[2u_{z,zz} + (u_{r,z} + u_{z,r})_{,r} + r^{-1}(u_{r,z} + u_{z,r})] - p_{,z} = 0, \tag{4}$$

$$\mu(u_{r,zz} - u_{z,zr}) - p_{,r} = 0. \tag{5}$$

The latter equation has been rearranged using the incompressibility condition (3) and the identity

$$u_{r,rr} + \frac{u_{r,r}}{r} - \frac{u_r}{r^2} = \Big(u_{r,r} + \frac{u_r}{r}\Big)_{,r}.$$

The two boundary conditions on the free surface $r = \rho(z)$ express the continuity of stress (including the Laplace capillary pressure), supplemented by two conditions of smoothness at $r = 0$. The Laplace pressure comes from the variation of the capillary energy equal to the area of the perturbed cylinder times surface tension γ. Assuming cylindrical symmetry, this energy reads

$$\mathcal{A} = 2\pi\gamma \int \rho(z)\sqrt{1 + \rho_{,z}^2}\,dz.$$

The variation of the capillary energy caused by changing the shape of the surface of the cylinder reads

$$\delta\mathcal{A} = 2\pi \int \delta\rho(z)\Big(\sqrt{1 + \rho_{,z}^2} - \frac{\rho_{,zz}\rho(z)}{(1 + \rho_{,z}^2)^{3/2}}\Big)dz.$$

This expression has to be added to the contribution $\delta\mathcal{E}_b$ to the variation of $\mathcal{E}$, which comes from the boundary term after integration by parts:

$$\delta\mathcal{E}_b = 2\pi \int \rho(z)[\delta u_r(2\mu u_{r,r} - p) + \mu\delta u_z(u_{r,z} + u_{z,r})]dz,$$

where the integrand is evaluated at $r = \rho(z)$.

Writing now that $\delta\rho(z) = \delta u_r$ at $r = \rho(z)$, and requiring $(\delta\mathcal{E}_b + \delta\mathcal{A})$ to vanish for any possible δu_r and δu_z, one finds the following bc for the tangential and normal stress on the surface:

$$u_{r,z} + u_{z,r} = 0, \tag{6}$$

$$-\gamma\Big(\frac{u_r}{\rho_0^2} + u_{r,zz}\Big) + 2\mu u_{r,r} - p = 0. \tag{7}$$

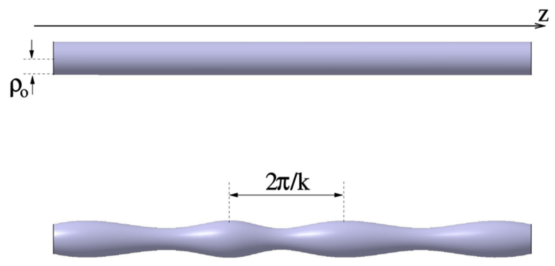


FIG. 5 (color online). Sketch of a straight cylinder before and after a varicose perturbation. The energy increment may be of either sign, depending on the radius, the surface tension, and the shear modulus.

In the last condition, the uniform equilibrium pressure inside the unperturbed straight cylinder given by the standard Laplace value, γ/ρ_0, has been subtracted from p. The problem of linear stability amounts to finding a nontrivial solution of Eqs. (4), (5), and (3) with the bc (6) and (7), imposed at $r = \rho_0$.

In a standard way, we assume a harmonic z dependence of any physical quantity of the order of the perturbation $\sim e^{ikz}$ with a wave number k. Then one can express from Eq. (4) $p(r, k)$ as a function of $v = u_r$ and its derivatives, and use the result in (5) to get a fourth-order equation for v:

$$(\mathcal{L} - k^2)^2 v = 0, \qquad \mathcal{L} = \frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{1}{r^2}. \tag{8}$$

Omitting terms diverging at $r = 0$, the general solution of Eq. (8) is found by using the Wronskian method and can be presented as

$$v(r) = \alpha I_1(kr) + \beta k^2 \Big[K_1(kr) \int_r^{\rho} I_1^2(kr')r'dr' - I_1(kr) \times \int_r^{\rho} r' I_1(kr') K_1(kr')dr' \Big], \tag{9}$$

where I_a and K_a are the modified Bessel function of first and second kind of order a, respectively, and α and β are integration constants. Using Eqs. (3)–(5), the bc (6) and (7) at $r = \rho$ can be expressed in terms of v and its derivatives.

These bc involve the third derivative of v at most, consistent with the fact that Eq. (8) is of the fourth order. Functions singular at $r = 0$ have been excluded by the particular choice of solution in Eq. (9). Using the general solution given by Eq. (9) in the bc (6) evaluated at $r = \rho_0$ and taking note that the integrals in (9) vanish at this point yields $\beta = 2\alpha$. Upon this substitution, we find from the bc (7) that the nontrivial solution exists at the critical value

$$\gamma_c = \frac{2\mu\rho_0}{1 - \rho_0^2 k^2}\Big[\frac{2k\rho_0 I_0(k\rho_0)}{I_1(k\rho_0)} - 1\Big]. \tag{10}$$

Instability occurs at $\gamma > \gamma_c$. It first appears in the long-scale mode $k \to 0$ at $\gamma_c(0) = 6\mu\rho_0$. The instability limit diverges at $k \to 1/\rho_0$.

The curve representing the equation $\gamma = 6\mu\rho_0$ is plotted in the $\mu - \rho_0$ plane in Fig. 4. The toluene-agar gel surface tension is taken as the measured value $\gamma = 36.5$ mN/m. This curve, directly following from the theory with no adjustable parameters, well matches the boundary between the two domains (for stable or unstable cylinders) detected experimentally. The finite wavelength observed in experiments may be either an indication of a slight subcriticality of the instability or a manifestation of nonlinear effects, necessarily present when the instability is observed at the macroscopic scale.

This above exact result for γ_c has been viewed as an approximation by Barrière *et al.* [9] who have used this approach to explain pattern formations during shrinkage of polymer gels reported by Matsuo and Tanaka [8]. In these experiments, the swollen gel in the cylinder becomes surrounded by a shrunken skin of macroscopic thickness. Assuming that the skin acts as an effective surface tension, Barrière *et al.* have found that the orders of magnitude might be in the right range to explain surface instability observed by Tanaka and Matsuo [9]. In these experiments the instability was, however, driven by a skin action modeled by a surface tension rather than by capillarity. Inferring surface tension out of a macroscopic concentration gradient may be problematic: according to the Kirkwood-Buff formula [10], surface tension requires an anisotropic stress tensor in the transition region. Such an anisotropy certainly exists near the interface between the water of our gel and toluene because of their immiscibility, but does not *a priori* exist for two miscible phases, like those of Tanaka and Matsuo. The main distinctive features of our experiments are (i) far smaller elastic moduli of agar gels and (ii) absence of a macroscopic skin at the agar gel surface. The last point is particularly important, and the good agreement between theory and our experiments for the onset value is a strong argument in favor of capillary effects as the cause of the observed instability.

In conclusion, we have given experimental evidence of Rayleigh-Plateau instability in a cylinder of soft solids. Its onset is well described by theory. Contrary to the RPI in liquids, the instability evolves to a steady wavy pattern along the cylinder. From the experiments we conjecture that, beyond a second critical (nonlinear) threshold, the final state is a set of disconnected droplets of solid, like in the case of a fluid cylinder.

L. M. P. acknowledges support by the Human Frontier Science Program (Grant No. RGP0052/2009-C).

---

*smora@univ-montp2.fr

[1] J. Plateau, *Statique Expérimentale et Théorique des Liquides Soumis aux Seules Forces Moléculaires* (Gauthier Villars, Paris, 1873), Vol. II.

[2] Lord Rayleigh, Proc. London Math. Soc. **s1-10**, 4 (1878).

[3] Lord Rayleigh, Proc. R. Soc. London **29**, 71 (1879).

[4] J. Eggers, Rev. Mod. Phys. **69**, 865 (1997).

[5] L. Landau and E. Lifschits, *Statistical Physics* (Pergamon, Oxford, 1980), 3rd ed., Chap. 155, pp. 520–523.

[6] N. Stanley, *Agar* (CRC Press, Boca Raton, FL, 2006), 2nd ed., Chap. 7, pp. 217–231.

[7] R. Lapasin and S. Pricl, *Rheology of Industrial Polysaccharides: Theory and Applications* (Blackie Academic Professional, Glasgow, 1995).

[8] E. Matsuo and T. Tanaka, Nature (London) **358**, 482 (1992).

[9] B. Barrière, K. Sekimoto, and L. Leibler, J. Chem. Phys. **105**, 1735 (1996).

[10] J. Kirkwood and F. Buff, J. Chem. Phys. **17**, 338 (1949).

ANNUAL REVIEWS

Annual Review of Fluid Mechanics

Statics and Dynamics of Soft Wetting

Bruno Andreotti[1] and Jacco H. Snoeijer[2]

[1]Laboratoire de Physique de l'Ecole Normale Supérieure (LPENS), CNRS UMR 8023, Ecole Normale Supérieure, Université PSL, Sorbonne Université, and Université de Paris, 75005 Paris, France; email: andreotti@lps.ens.fr

[2]Physics of Fluids Group, Faculty of Science and Technology, and Mesa+ Institute, University of Twente, 7500AE Enschede, The Netherlands

Annu. Rev. Fluid Mech. 2020. 52:285–308

The *Annual Review of Fluid Mechanics* is online at fluid.annualreviews.org

https://doi.org/10.1146/annurev-fluid-010719-060147

Keywords

wetting and spreading, elasticity, elastocapillarity, contact angles, gels, brushes

Abstract

The laws of wetting are well known for drops on rigid surfaces but change dramatically when the substrate is soft and deformable. The combination of wetting and the intricacies of soft polymeric interfaces have provided many rich examples of fluid–structure interactions, both in terms of phenomenology and from a fundamental perspective. In this review we discuss experimental and theoretical progress on the statics and dynamics of soft wetting. In this context we critically revisit the foundations of capillarity, such as the nature of solid surface tension, the microscopic mechanics near the contact line, and the dissipative mechanisms that lead to unexpected spreading dynamics.

1. INTRODUCTION: FROM RIGID TO SOFT WETTING

A liquid drop sitting on an ordinary solid does not induce any significant deformation of the surface. The angle made at the contact line is selected by surface energies according to Young's law. By contrast, a liquid drop floating on another liquid will strongly deform the interface, with contact angles selected by Neumann's law. Soft compliant substrates, typically reticulated polymer networks, are in between these two extreme cases: They do deform under the effect of capillary forces, but in contrast to liquids, they exhibit an elastic resistance.

The systematic exploration of soft wetting phenomena is fairly recent, spurred by improving technology to tune the properties of soft matter and by progress on elastocapillary phenomena in general—the latter received an extensive review by Bico et al. (2018). The mechanics of extremely soft materials originates from a competition between bulk elasticity and surface effects (Mora et al. 2010, Style et al. 2017), providing a new playground for material design. Applications are numerous and range from adhesives (Autumn et al. 2000, Boesel et al. 2010, Jagota & Hui 2011) to slippery surfaces (Lafuma & Quéré 2011, Wong et al. 2011, Schellenberger et al. 2015, Solomon et al. 2016), highly stretchable synthetic materials (Grandgeorge et al. 2018), and the biomechanics of cells and soft tissues (Manning et al. 2010).

Here, we focus on the wetting of soft elastic substrates, for example by liquid drops. Apart from being of intrinsic interest, drops offer a unique way to study static and dynamic deformations of soft interfaces owing to their nanometrically sharp contact line forcing. Over the years, experiments such as those shown in **Figure 1** have progressively revealed the salient features of the wetting ridge below the contact line. This ridge dramatically alters the macroscopic spreading dynamics, as the moving ridge induces strong viscoelastic dissipation inside the substrate (Carré et al. 1996, Long et al. 1996, Karpitschka et al. 2015, Zhao et al. 2018a). It leads to phenomena such as viscoelastic braking and dynamical depinning, which have no counterparts on rigid surfaces. Further complexities are encountered when the polymeric substrate swells by absorbing the wetting liquid (Cohen Stuart et al. 2006, Kajiya et al. 2011, Dupas et al. 2014, Boulogne et al. 2015) and when dangling chains are present at the surface.

These soft wetting phenomena are not captured by the same laws as rigid wetting (Andreotti & Snoeijer 2016, Style et al. 2017), and one is forced to critically revisit the foundations of capillarity, most notably, (*a*) the nature of surface tension of soft solids and the underlying microstructure of reticulated polymers, (*b*) the force balance near the contact line and the wetting boundary conditions, (*c*) moving contact lines and dissipation mechanisms inside the substrate, and (*d*) the resulting macroscopic motion of droplets. The purpose of this review is to address these fundamental aspects of wetting on soft surfaces by discussing recent experimental and theoretical progress.

Neumann's law: vectorial balance of the three surface tensions at the contact line, traditionally used for liquid subphases

Wetting ridge: localized elastic deformation of the substrate below the contact line

Viscoelastic braking: slowing down of contact-line motion due to dissipation inside a viscoelastic substrate

Surface tension: excess surface force per unit length; often called "surface stress," but here we strictly reserve the term "stress" for the bulk force per area

2. BASIC CONCEPTS

2.1. Soft Polymeric Materials

For ordinary solid materials, externally applied stresses change the interatomic distances and thereby increase the internal energy. The resulting elasticity is of enthalpic origin and leads to elastic moduli of typically 10–100 GPa. Conversely, for reticulated polymer networks that are not in a glassy state, an applied strain puts chains in a less probable conformation, increasing the free energy only for entropic reasons. These networks can therefore undergo large reversible deformations at relatively small applied stress and present small elastic moduli, proportional to the thermal energy $k_B T$ and to the number of chains per unit volume [pieces of polymer between cross-links or between points of entanglement (Watanabe 1999)]. As such, the stiffness of the network can be varied over orders of magnitude via the density of cross-links or entanglements. A single polymer

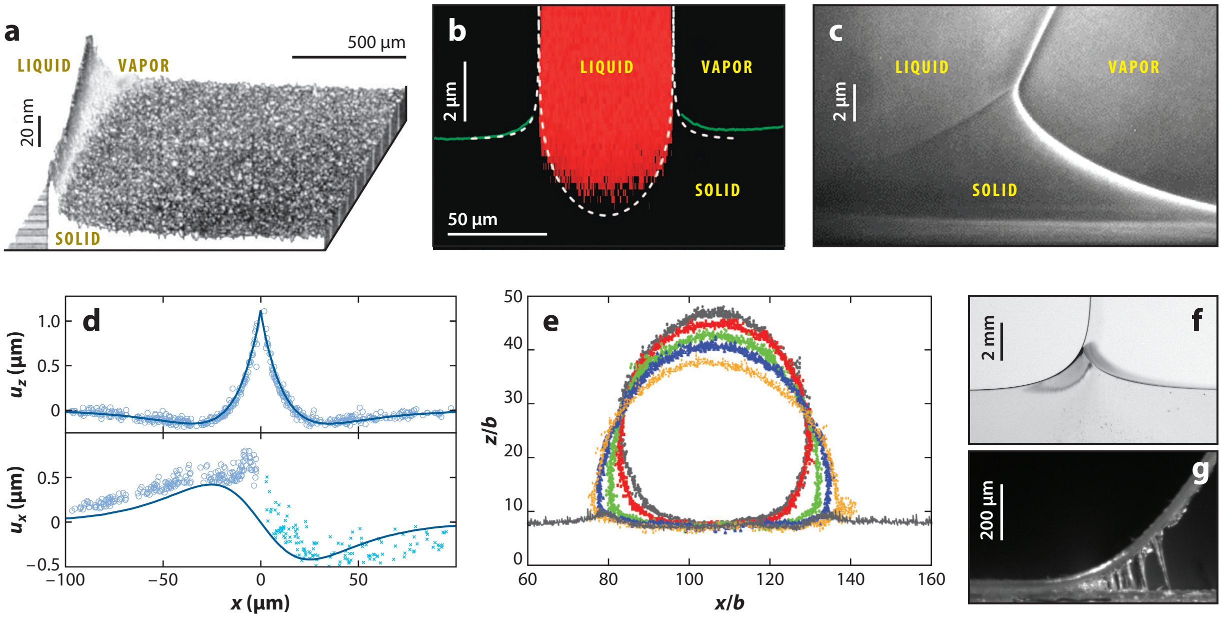


Figure 1

(*a–d*) Historical overview of wetting ridges observed on soft PDMS (polydimethylsiloxane) gels. The experimental methods used are (*a*) white-light interferometry, (*b*) laser scanning confocal microscopy, (*c*) X-ray, and (*d*) confocal microscopy. (*e*) Wetting ridges in molecular dynamics simulations of a polymer drop on a brush, for varying wettability. (*f*) Peeling of a tape from a PDMS gel (Perrin et al. 2019). (*g*) Peeling of a tape from a pressure-sensitive adhesive—note the fibril formation at the peeling front (Villey et al. 2015). Panels adapted with permission from (*a*) Carré et al. (1996), copyright 1996 Springer Nature; (*b*) Pericet-Camara et al. (2008), copyright 2008 American Chemical Society; (*c*) Park et al. (2014); (*d*) Jerison et al. (2011), copyright 2011 American Physical Society; and (*e*) Leonforte & Mueller (2011), copyright 2011 AIP Publishing.

chain has a mechanical response that depends on its length N: In the simple Rouse model, its spring constant scales as k_BT/Nb^2, where b is the monomer length, and its relaxation time τ scales as $\zeta b^2N^2/k_BT$, where ζ is the monomeric friction coefficient controlling the diffusion (Boese & Kremer 1990).

Before entering their wetting behavior, we first provide a very brief description of the different types of polymeric substrates that can be encountered. For a complete view on the statistical physics involved, we refer the reader to textbooks by de Gennes (1979), Doi & Edwards (1988), Rubinstein (2003), and Binder & Kob (2011).

The adhesive properties of polymers strongly depend on their molecular architecture, as illustrated by peeling experiments. **Figure 1*f*** shows a reversible adhesive with a smooth viscoelastic ridge (Perrin et al. 2019), while **Figure 1*g*** involves a pressure-sensitive adhesive that leads to the formation of elongated fibrils (Villey et al. 2015). The latter polymer has strong adhesive bonds (Deplace et al. 2009) and ubiquitous dangling ends that need to be pulled from the matrix during debonding (**Figure 2*a***), in contrast to the former (**Figure 2*b***). The energy required to fracture the two interfaces, Γ, is proportional to the chain length and to the surface density of chains, Σ (**Figure 2*e***) (Creton et al. 1992), and can therefore be orders of magnitude larger than the capillary adhesion energy due to van der Waals interactions (de Gennes 1989, Raphael & De Gennes 1992).

The network structure also determines the viscoelastic bulk rheology. Of particular importance for soft wetting are polymer gels (**Figure 2*b***). These exhibit a fractal structure, with chain lengths spanning from the length of the prepolymer to the size of the sample. By consequence,

Review in Advance first posted on
August 19, 2019. (Changes may
still occur before final publication.)

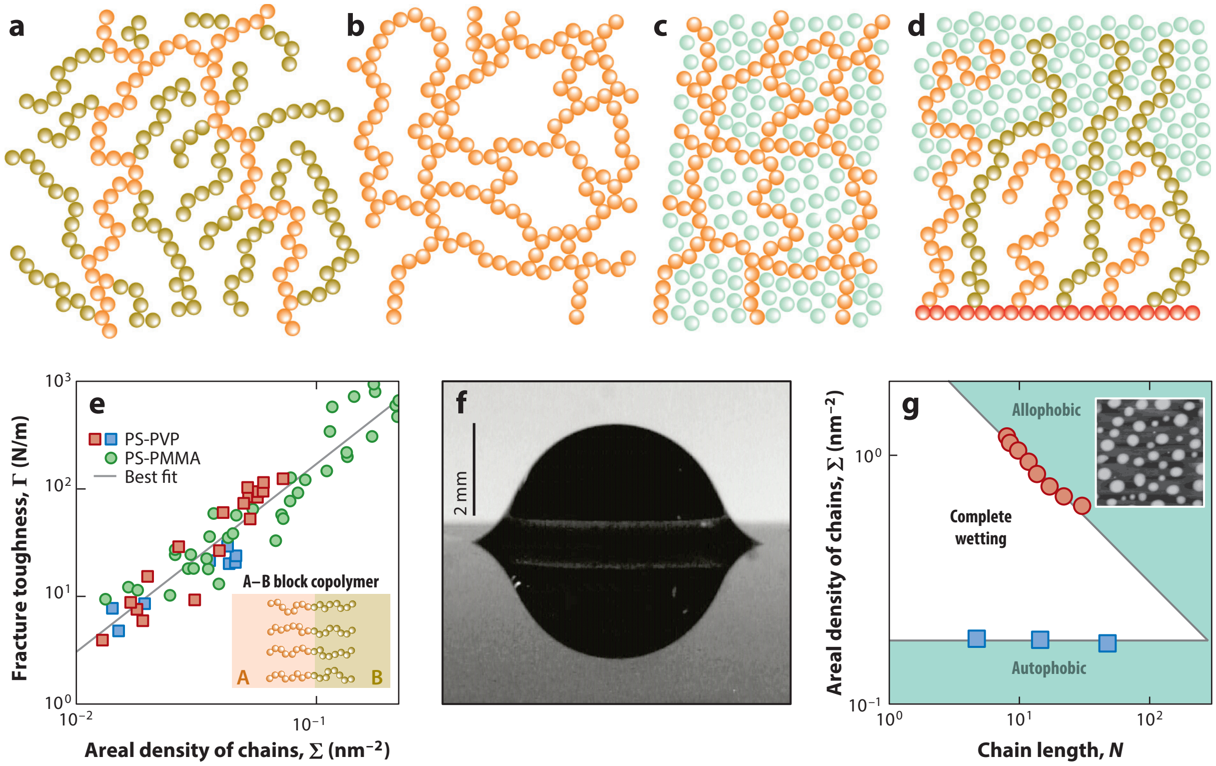


Figure 2

(*Top row*) Schematics of different architectures of polymeric soft solids: (*a*) a pressure-sensitive adhesive, characterized by many dangling chains (*yellow*) connected to a backbone network (*red*); (*b*) a polymeric gel, characterized by a multiscale network without dangling chains; (*c*) a swollen gel with a liquid phase; and (*d*) a polymer brush, here partly swollen. (*e*) The fracture toughness Γ of interfaces reinforced with a block copolymer plotted as a function of the effective areal density of chains, Σ. (*f*) A hydrogel sphere (with shear modulus $G = 61$ Pa) on a silicon wafer, totally wetted by water. (*g*) Wetting phase diagram of a melt of polystyrene of length N in contact with a brush of polystyrene chains end-attached to a substrate with a grafting density Σ. Solid lines are guides for the eyes through experimental points. Abbreviations: PS-PMMA, poly(styrene-*b*-methyl methacrylate); PS-PVP, poly(styrene-*b*-2-vinylpyridine). Panels adapted with permission from (*e*) Creton et al. (1992), copyright 1992 American Chemical Society; (*f*) Chakrabarti et al. (2018), copyright 2018 American Chemical Society; and (*g*) Maas et al. (2002), copyright 2002 American Chemical Society.

Storage $G'(\omega)$ and loss $G''(\omega)$ moduli: the in-phase and out-of-phase stresses of a material under oscillatory strain, respectively; viscoelastic media exhibit both a reversible (storage) and a dissipative (loss) response

the relaxation spectrum spans orders of magnitude in frequency ω. This is reflected by the storage G' and loss G'' moduli, scaling as ω^n with an exponent $n \approx 0.5$ (Winter & Chambon 1986) close to that given by the Rouse model (Onogi et al. 1970). This viscoelastic response determines the substrate's dissipation in dynamical wetting experiments (de Gennes 1996, Long et al. 1996). When providing extra cross-links with respect to the gel point, the material exhibits a finite (static) shear modulus G at low frequency.

The substrate properties change dramatically when the network is swollen by a solvent (**Figure 2*c***). A water drop placed on a hydrogel imbibes into the porous substrate, and one cannot always sharply define a contact line (Kajiya et al. 2011). A hydrogel sphere placed on a rigid surface, which is completely wetted by water, exhibits a finite contact angle (**Figure 2*f***) (Chakrabarti et al. 2018). Polymer gels may even be swollen by a melt phase of the same polymer: Small fractions of un-cross-linked chains can alter the adhesive and dynamical properties (Jensen et al. 2015; Pham et al. 2017; Hourlier-Fargette et al. 2017, 2018). Swelling is also a key aspect of brushes (**Figure 2*d***), which are polymer chains tethered to a rigid substrate either covalently or by adsorption (Alexander 1977, de Gennes 1980, Milner et al. 1988). Brushes exhibit intricate wetting behaviors (Cohen Stuart et al. 2006, Leonforte & Mueller 2011, Mensink et al. 2019) and can even be autophobic with respect to their own melt (**Figure 2*g***) (Maas et al. 2002).

Evidently, the physical chemistry of soft polymeric interfaces constitutes a vast area of research with numerous applications (Fleer et al. 1993). In the following, we primarily restrict ourselves to the statics and dynamics of wetting in the simplest case, where the polymer network is not swollen and where it does not present brush-like dangling chains. Part of the analysis will assume that pinning (contact angle hysteresis) is absent. This idealized situation can indeed be closely approached experimentally, e.g., by the frequently used PDMS (polydimethylsiloxane) gel substrates (**Figure 1**), although many other systems are touched on as well.

Shuttleworth effect: the difference between surface tension Υ and surface energy γ, which arises whenever γ depends on the surface strain, ϵ

Surface strain: change in relative length of a surface element, measured with respect to the substrate's reference state

2.2. Capillarity: Liquid Versus Elastic Interfaces

From a macroscopic thermodynamic perspective, interfaces are characterized by a surface energy, γ. This represents the excess free energy per unit area of an interface (Rowlinson & Widom 1982, de Gennes et al. 2002). The order of magnitude of the surface energy is given by $\gamma \sim k_B T/a^2$, where the scale a is a typical microscopic length. Mechanically, this gives rise to a surface tension Υ_{ij}, which is the surface-analog of the stress tensor. It represents the excess force per unit length in the interface (Marchand et al. 2011); for a liquid, this tension is isotropic, $\Upsilon_{ij} = \gamma\delta_{ij}$. In what follows, the discussion concerns the two-dimensional case, for which we can stick to a scalar description of Υ. For an extensive review on mechanical implications of surface tension in soft solids, we refer the reader to Style et al. (2017).

Unlike liquids, for elastic interfaces the surface tension Υ and the surface energy γ are not equal, owing to the Shuttleworth effect (Shuttleworth 1950, Muller & Saul 2004). The origin of this difference is illustrated in **Figure 3**. **Figure 3*a*** shows that γ is the reversible energy associated with the creation of a solid–vapor interface by separating two solid blocks (i.e., not accounting for the fracture). This procedure is carried out at constant surface strain ϵ, although we remark that in general the surface energy can depend on strain, i.e., $\gamma(\epsilon)$. **Figure 3*b*** shows another way to create

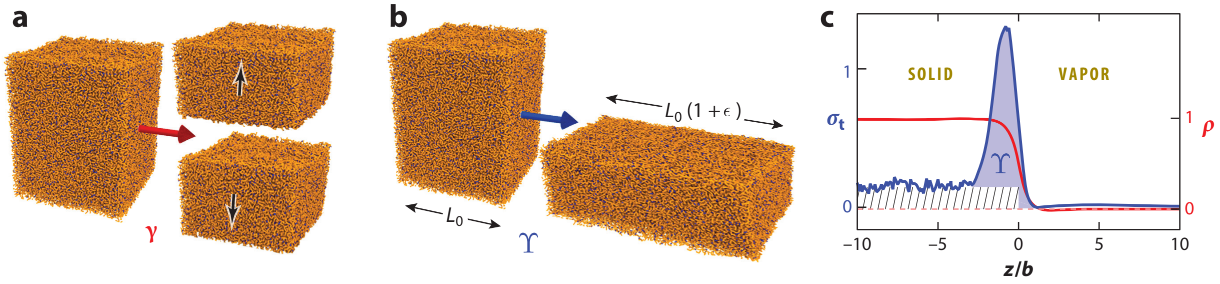


Figure 3

(*a*) The creation of two solid–vapor interfaces by dividing a solid block and bringing the interfaces far away from each other, while keeping the surface strain ϵ constant. The excess energy per unit surface area γ equals the reversible work done during the quasi-static separation. (*b*) The creation of a solid–vapor interfacial area by stretching the elastic solid. The change in interfacial energy equals the work done by the surface tension Υ and gives the Shuttleworth equation (Equation 1). (*c*) Molecular dynamics simulations of a reticulated polymer. (*Red line*) Profile of the monomer density ρ, indicating the location of the solid–vapor interface (the position is expressed in terms of monomer size *b*). (*Blue line*) Profile of the stress anisotropy, $\sigma_t = \sigma_{xx} - \sigma_{zz}$, across a stretched film. The effect of surface tension manifests itself as the peak of σ_t, located inside the liquid–vapor interface (the integral gives Υ). The hatched region is the bulk elastic stress. Data from Liang et al. (2018a,b).

Review in Advance first posted on
August 19, 2019. (Changes may
still occur before final publication.)

a solid–vapor interface, by stretching the interface length by $\delta L = L_0 \delta\epsilon$. The associated work per unit length, $\Upsilon \delta L$, involves surface tension Υ. Equating this work to the increase in surface energy, $\delta(L\gamma) = \delta\,[L_0(1+\epsilon)\gamma]$, one obtains the Shuttleworth equation,

$$\Upsilon(\epsilon) = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\,[(1+\epsilon)\gamma(\epsilon)] = \gamma + (1+\epsilon)\frac{\mathrm{d}\gamma}{\mathrm{d}\epsilon}. \qquad 1.$$

The second term involves the derivative $\gamma' = \mathrm{d}\gamma/\mathrm{d}\epsilon$ and emerges due to the change in surface energy during stretching. This effect is not present for a simple liquid–vapor interface, for which we write $\Upsilon_{\mathrm{LV}} = \gamma_{\mathrm{LV}} = \gamma$.

Interfacial width: capillary forces are not perfectly localized but spread out over a molecular region of size *a* between two bulk phases

Elastocapillary length: length scale arising from the balance of surface tension and the shear modulus, setting the typical scale of wetting ridges

It is of interest to examine the effect of surface tension at the nanoscale, inside the interfacial region (Weijs et al. 2013). The density profile across a soft interface exhibits a smooth transition over the molecular distance $a \sim 10^{-9}$ m, (**Figure 3*c***) (Liang et al. 2018a). Capillarity manifests itself as an anisotropy of stress components, in the direction tangential (σ_{xx}) and normal (σ_{zz}) to the interface. The stress anisotropy, $\sigma_t = \sigma_{zz} - \sigma_{xx}$, represents the excess tangential stress, or tension, localized in the interfacial zone. This effect is well known for liquids (Kirkwood & Buff 1949, Nijmeijer et al. 1990), but the concept equally applies for reticulated polymer networks. This is evidenced by the peak in σ_t (**Figure 3*c***), whose integral gives the interfacial force per unit length: the macroscopic surface tension, Υ. Hence, the excess tangential stress inside the interface, $\sigma_t \sim \Upsilon/a$, can be estimated as $k_{\mathrm{B}}T/a^3 \sim 10^7 - 10^8$ Pa. In crystalline and glassy solids, this stress is negligible with respect to elasticity, which is of enthalpic origin. By comparison, Υ/a is typically orders of magnitude larger than the entropic elastic modulus of a soft-polymer network, $G \sim k_{\mathrm{B}}T/(Nb^3)$, owing to the large number of monomers N between cross-linkers. Since the monomer size b and the interfacial width a are both of molecular size, a crude estimate of the ratio $\gamma/G \sim Nb^3/a^2 \sim Na$ gives the so-called elastocapillary length.

We remark that **Figure 3*c*** was obtained from a molecular simulation of a simple cross-linked polymer network without solvent or dangling chains, for which one expects a liquid-like interfacial structure. The presence of dangling chains extends the interfacial region to the typical chain length, which requires a specific description at a mesoscopic scale.

2.3. Scales of Elastic Wetting

In a continuum perspective, elastocapillary phenomena can be classified in terms of length scales (Schroll et al. 2013, Bico et al. 2018). The elastocapillary length γ/G separates the small scales dominated by capillarity from the large scales dominated by elasticity. By tuning the substrate stiffness, γ/G can be varied over orders of magnitude, from submolecular up to millimeter scales. This length must be compared to the nanometric interface width a and to the macroscopic parameters such as the drop size R and the thickness e of the elastic substrate.

Figure 4*a*, subpanel ***i***, illustrates the rigid limit ($\gamma/G \ll a$). The excess stress γ/a inside the liquid–vapor interface is then negligible compared to G and cannot induce any deformation below the contact line. Subpanel ***ii*** corresponds to the intermediate case ($a \ll \gamma/G \ll R$). Surface tension dominates at small scales and a sharp wetting ridge is formed (Jerison et al. 2011, Limat 2012, Marchand et al. 2012b). On the scale of the drop, however, elasticity is still dominant and the liquid angle remains unaffected. Subpanel ***iii*** illustrates the limit $\gamma/G \gg R$ where elasticity plays no role and the drop takes the shape of a liquid lens (Style & Dufresne 2012, Lubbers et al. 2014). An extensive discussion of the boundary conditions at the contact line, in particular of the applicability of Young's law and Neumann's law, follows in Section 3.

Figure 4*b* shows a drop on a thin membrane, in the regime $\gamma/G \ll e \ll R$. The membrane is curved over the bending elastocapillary length, $\sqrt{B/\gamma}$, which is based on the bending rigidity,

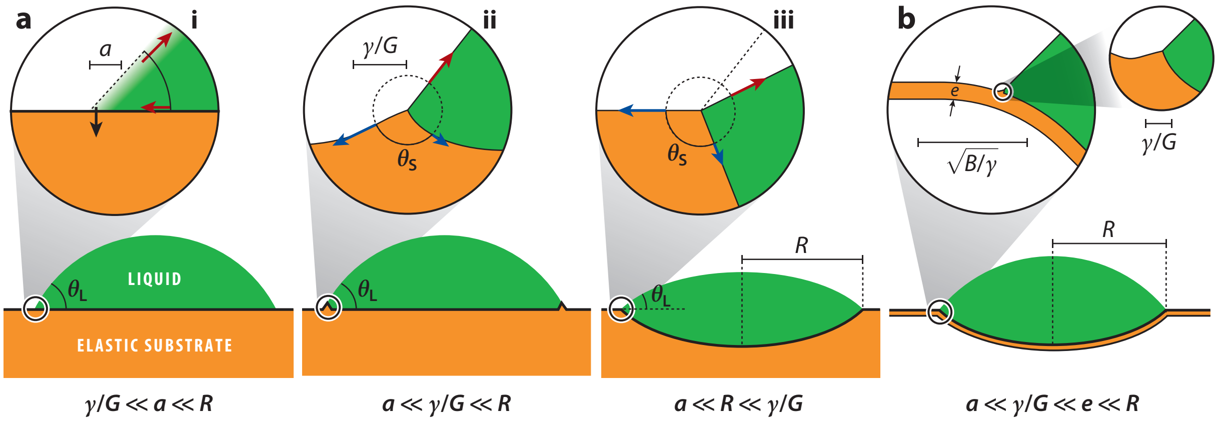


Figure 4

Scales of elastocapillarity. (*a*) Drops on thick elastic substrates of different stiffnesses, tracking the change of the liquid angle θ_L and of the solid angle θ_S (*insets*). The elastocapillary length γ/G must be compared to the width of the interface a and the drop size R. (*i*) The rigid limit, with contact angles governed by Young's law (see Marchand et al. 2011 for the force balance on the liquid wedge). (*ii*) Moderately soft substrates. The wetting ridge remains small compared to the drop size and θ_L still satisfies Young's law. However, θ_S is selected by Neumann's law, which expresses the balance of surface tensions applied to the circular system. (*iii*) The soft limit, where elasticity can be neglected on all scales. The drop resembles a liquid floating lens. (*b*) Typical case for a thin membrane of thickness e. The capillary-induced bending remains smooth at the bending length scale, $\sqrt{B/\gamma}$.

$B \sim Ge^3$. In that case, the wetting ridge is a negligible feature, and one enters the realm of elastocapillary bending and stretching of fibers and sheets (Roman & Bico 2010, Duprat et al. 2012, Schroll et al. 2013, Schulman & Dalnoki-Veress 2015, Bico et al. 2018, Davidovitch & Vella 2018).

3. THE WETTING BOUNDARY CONDITIONS

From a macroscopic perspective, interfaces can be treated as perfectly sharp. The effects of wetting then arise as boundary conditions at the contact line. Here we discuss the energetic derivation of these boundary conditions on elastic surfaces, in the absence of swelling by the wetting liquid. This is complemented by equivalent mechanical interpretations, which allow one to clarify the rigid-to-soft transition, and by a discussion of recent experiments.

3.1. Rigid Substrate Limit: Young's Law

To derive the classical Young's law on a rigid solid, one starts from the free energy $\mathcal{F}$ of the liquid–vapor interface of profile $h(x)$ whose contact line location is defined by the position $x = r$,

$$\mathcal{F} = \int_{-\infty}^{r} \left[\gamma(1+h'^2)^{1/2} + \gamma_{SL}\right] \mathrm{d}x + \int_{r}^{\infty} \gamma_{SV}\, \mathrm{d}x. \qquad 2.$$

The solid–liquid and solid–vapor surface energies respectively are γ_{SL} and γ_{SV}. Minimizing $\mathcal{F}$ with respect to variations of the interface profile $\delta h(x)$ and the contact line position δr yields, after integrating by parts,

$$\delta\mathcal{F} = \left[\frac{\gamma}{[1+h'(x)^2]^{1/2}} + \gamma_{SL} - \gamma_{SV}\right]_{x=r} \delta r - \int_{-\infty}^{r} \gamma\kappa(x)\,\delta h(x)\,\mathrm{d}x, \qquad 3.$$

Review in Advance first posted on August 19, 2019. (Changes may still occur before final publication.)

where $\kappa = h''/(1+h'^2)^{3/2}$ is the interface curvature (Snoeijer & Andreotti 2008). The equilibrium condition $\delta\mathcal{F} = 0$ provides both the normal stress balance at the interface, $\gamma\kappa$ being the Laplace pressure, and the boundary condition at the contact line $x = r$, namely Young's law,

$$\gamma \cos\theta_{\mathrm{L}} = \gamma_{\mathrm{SV}} - \gamma_{\mathrm{SL}}. \qquad 4.$$

Any additional macroscopic force such as gravity or electrostatics acts as an additional normal stress along the interface but does not affect the boundary condition; the boundary condition can be interpreted as a horizontal force balance on an infinitesimal wedge of liquid near the contact line (**Figure 4*a***, subpanel ***i*** inset).

3.2. First Boundary Condition: Neumann's Law for Contact Angles

When the substrate is soft enough to be deformed by the liquid, the elastic energy stored inside the substrate must be included in $\mathcal{F}$. **Figure 5*a*** defines the curvilinear coordinate s along the interface, the tangential vector $\mathbf{t}$ and normal vector $\mathbf{n}$. The stress balance at the interface reads

$$\overline{\overline{\boldsymbol{\sigma}}} \cdot \mathbf{n} - \overline{\overline{\mathbf{T}}} \cdot \mathbf{n} = \frac{\partial}{\partial s}\left(\Upsilon_{\mathrm{S}}\,\mathbf{t}\right), \qquad 5.$$

where $\overline{\overline{\boldsymbol{\sigma}}} \cdot \mathbf{n} \equiv \boldsymbol{\sigma}$ is the elastic traction and $\overline{\overline{\mathbf{T}}} \cdot \mathbf{n}$ is the liquid traction; these are due to the stress tensors in the substrate $\overline{\overline{\boldsymbol{\sigma}}}$ and in the liquid $\overline{\overline{\mathbf{T}}}$, respectively. The surface tension of the solid is denoted Υ_{S}.

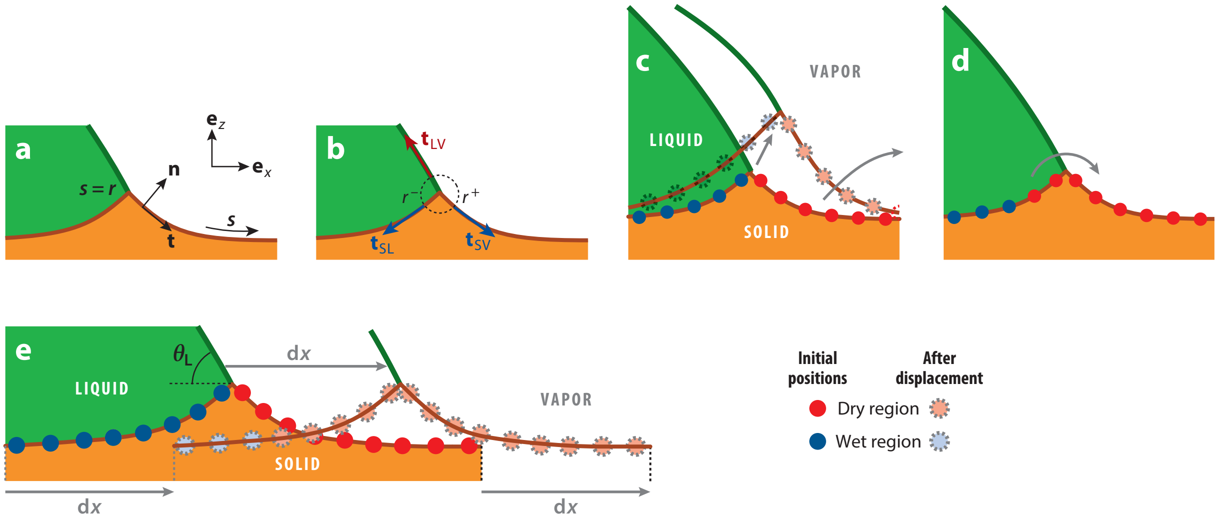


Figure 5

(a) Definition of the unit vectors $\mathbf{e}_x$ and $\mathbf{e}_z$ and the curvilinear coordinate s, running along the solid interface. The contact line is located at $s = r$ so that r_+ and r_- are respectively the limits on both sides. (b) Tangent unit vectors $\mathbf{t}^+ = \mathbf{t}_{\mathrm{SV}}$ and $\mathbf{t}^- = \mathbf{t}_{\mathrm{SL}}$ at both sides of the contact line. (c–d) Wetting at equilibrium requires the energy to be minimal with respect to all possible types of contact line displacement (Snoeijer et al. 2018). (c) Horizontal and vertical displacement in the lab frame. Equilibrium gives the first boundary condition (Equation 7), i.e., Neumann's law. (d) Relative motion of the substrate, while the contact line remains fixed in the lab frame. The colored points indicate the exchange of material points from the dry (*red*) to the wetted region (*blue*). Equilibrium gives the second boundary condition (Equation 9) on the chemical potential. (e) The liquid contact angle θ_{L} can be inferred from a global displacement dx of the contact line. The region close to the contact line remains unchanged, and the changes in (elastic and capillary) energy occur at the edge of the indicated contour.

Review in Advance first posted on
August 19, 2019. (Changes may
still occur before final publication.)

The boundary condition at the contact line is again obtained from a variation of the contact line position. **Figure 5*c*** shows that this involves both horizontal and vertical displacements (Snoeijer et al. 2018). By consequence, there is now a vectorial boundary condition,

$$\gamma \mathbf{t}_{\mathrm{LV}} + [\Upsilon_{\mathrm{S}}\, \mathbf{t}]_{r^-}^{r^+} = \int_{r^-}^{r^+} \mathrm{d}s\, \boldsymbol{\sigma}. \qquad 6.$$

This can be interpreted as the integral of Equation 5 over an infinitesimal zone across the contact line (**Figure 5*b***), using the perfectly localized liquid traction, $\overline{\overline{\mathbf{T}}} \cdot \mathbf{n} = \gamma \mathbf{t}_{\mathrm{LV}} \delta(s - r)$ (Limat 2012). When the substrate is forced into a wedge, the elastic stress is only logarithmically singular (Lubbers et al. 2014), also at large elastic deformation (Singh & Pipkin 1965). Given the weak stress singularity, the integral over $\boldsymbol{\sigma}$ vanishes and Equation 6 becomes

$$\gamma \mathbf{t}_{\mathrm{LV}} + \Upsilon_{\mathrm{SV}} \mathbf{t}_{\mathrm{SV}} + \Upsilon_{\mathrm{SL}} \mathbf{t}_{\mathrm{SL}} = 0. \qquad 7.$$

Figure 5*b* illustrates this result, known as Neumann's law, as clearly observed for the wetting of soft gels (**Figure 1**). This law is normally used for liquid phases (de Gennes et al. 2002) but also applies to sufficiently soft elastic substrates (Limat 2012, Marchand et al. 2012b, Style & Dufresne 2012, Style et al. 2013a). The weak elastic singularity of the (Cauchy) stress persists when prestretching the substrate, which therefore remains, contrary to recent suggestions (Masurel et al. 2018), integrable and does not contribute to Neumann's law (Snoeijer et al. 2018).

After these macroscopic considerations, it is instructive to consider the crossover from rigid to soft substrates from Equation 6. Microscopically, the liquid traction is not sharply localized in space but is spread out over the nanometric width a (**Figure 3*c***). In the stiff limit, $\gamma/G \ll a$, the substrate remains flat and one recovers Young's law (Equation 4). In that case, Equation 6 can be used to compute the integral of elastic traction—integrating not over an infinitesimal zone, but over the contact line width, a. In the vertical direction this gives $\gamma \sin\theta_{\mathrm{L}}$ to balance the upward pull of the droplet. This approach was already taken on by Rusanov (1975), Shanahan (1987), and White (2003) when computing the shape of the ridge for small deformation. Conversely, in the soft limit $\gamma/G \gg a$ where the solid forms a sharp wedge, the integrated elastic force per unit length T_{el} can be estimated as

$$T_{\mathrm{el}} \sim \int_{r-a}^{r+a} \sigma\, \mathrm{d}s \sim Ga \log\left(\frac{\gamma}{Ga}\right). \qquad 8.$$

Molecular dynamics simulations indeed confirmed such an elastic correction to Neumann's law (Liang et al. 2018b). **Figure 6** reports T_{el} for both a droplet and a rigid particle, in contact with a cross-linked polymer network for a range of stiffnesses, here fitted with Equation 8. For experiments on soft wetting, γ/G is typically above the micron scale, for which the elastic correction to Neumann's law is less than one percent.

3.3. Second Boundary Condition: Surface Strain Discontinuity at the Contact Line

In contrast to the rigid limit, the substrate's elasticity allows for relative motion while the contact line remains stationary in the lab frame (**Figure 5*d***). This motion involves an exchange of surface material across the contact line, indicating that the contact line is not pinned. This implies a second

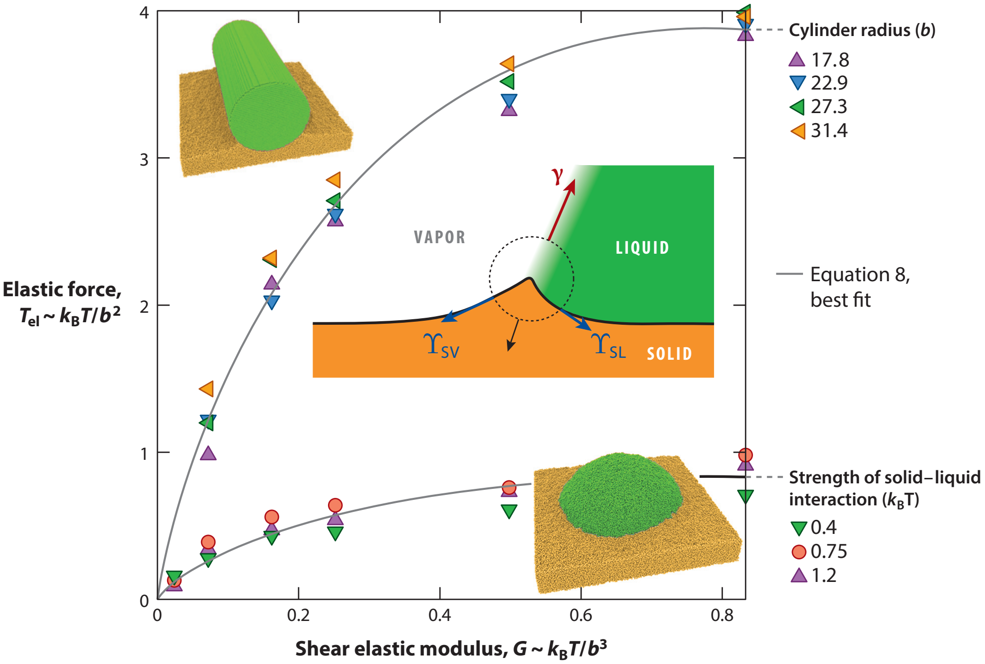


Figure 6

Rigid-to-soft transition on the microscopic scale. (*Inset*) The force balance on a control volume surrounding the contact line. The generalized Neumann's law involves the three surface tensions, γ, Υ_{SV}, and Υ_{SL}, and a residual elastic force per unit length, T_{el}. (*Main graph*) Elastic force T_{el} as a function of the shear elastic modulus G, measured in molecular dynamics simulations by Liang et al. (2018b). Data are represented in units of the thermal energy k_BT and the monomer size b. Data are fitted by Equation 8, showing that T_{el} vanishes in the continuum limit $aG/\gamma \to 0$, where a is the molecular size and G is the shear modulus. Data from Liang et al. (2018b).

boundary condition (Snoeijer et al. 2018),

$$\mu_{SV} = \mu_{SL}, \qquad \text{with} \quad \mu = (1+\epsilon)^2\gamma'(\epsilon) + f_{el} + \int_r^{r^\pm} (\boldsymbol{\sigma}\cdot\mathbf{t})\,\mathrm{d}s. \qquad 9.$$

Here, f_{el} is the (surface) density of elastic energy that is associated to changing the material point below the contact line; it was shown to vanish in linear elasticity (Snoeijer et al. 2018) but possibly contributes at large strain. The integral over $\boldsymbol{\sigma}$ again vanishes in the soft limit ($\gamma/G \gg a$). The condition $\mu_{SV} = \mu_{SL}$ can be viewed as the equality of chemical potential that governs the exchange of material across the contact line. This is in direct analogy with equality of chemical potential across a liquid–vapor interface that regulates the discontinuity in density. Here, Equation 9 serves as a boundary condition for the surface strain ϵ, which in general can be discontinuous across the contact line.

In the rigid limit $\gamma/G \ll a$, any relative motion already implies a contact line displacement. The conditions of Equations 6 and 9 are then obtained from the variation of the same degree of freedom. In this rigid limit, the equality $\mu_{SV} = \mu_{SL}$ predicts the existence of a horizontal elastic force per unit length at the contact line, $\int(\boldsymbol{\sigma}\cdot\mathbf{t})\,\mathrm{d}s = \mathbf{e}_x\cdot\int\boldsymbol{\sigma}\,\mathrm{d}s = \gamma'_{SV} - \gamma'_{SL}$; the same follows from Equation 6 in combination with Young's law (Weijs et al. 2013). This tangential force has indeed been observed in molecular simulations (Seveno et al. 2013).

THE MEMBRANE LIMIT

The membrane limit appears when the layer thickness e and γ/G are negligible with respect to the other scales (**Figure 4*b***). The substrate is locally flat, and the elastic energy can then be integrated across the thickness. For small strains, the membrane limit gives an elastocapillary membrane energy per unit area, $\gamma(\epsilon) + Y\epsilon^2/2$, where Y is the membrane stiffness. Hence, Equation 9 simplifies to $\mu = \gamma' + Y\epsilon$, and the condition $\mu_{\mathrm{SL}} = \mu_{\mathrm{SV}}$ implies a strain discontinuity, $\epsilon_{\mathrm{SV}} - \epsilon_{\mathrm{SL}} = (\gamma'_{\mathrm{SL}} - \gamma'_{\mathrm{SV}})/Y$ (Weijs et al. 2013, Neukirch et al. 2014, Andreotti & Snoeijer 2016). This discontinuity is indeed observed in experiments and molecular simulations using an elastic Wilhelmy plate (**Figure 7*a***). The liquid contact angle with respect to the membrane follows from Young's law based on the total membrane energy, and this gives $\gamma\cos\theta_{\mathrm{L}} = \gamma_{\mathrm{SV}} - \gamma_{\mathrm{SL}} + Y(\epsilon_{\mathrm{SV}}^2 - \epsilon_{\mathrm{SL}}^2)/2$. Hence, the Shuttleworth effect implies an ϵ^2-correction of the liquid angle (Weijs et al. 2013, Neukirch et al. 2014). The equations for the contact angle and the strain discontinuity provide the boundary conditions for the locally flat membrane (**Figure 3*b***). The actual contact angle in the lab frame is obtained by solving the global membrane problem (Schroll et al. 2013, Davidovitch & Vella 2018).

3.4. Selection of the Liquid Contact Angle

The most prominent feature of wetting is the liquid angle θ_{L} with respect to the reference solid surface. It is not selected locally by the Neumann condition but emerges from the global, macroscopic elastocapillary problem of Equation 5 subjected to the two boundary conditions of Equations 6 and 9. The scenario for thick elastic layers is summarized in **Figure 4*a***. For a drop of size R, the liquid angle undergoes a transition from Young's to Neumann's law—but this time on macroscopic scales, crossing over at $\gamma/G \sim R$.

Of particular experimental relevance is the regime $\gamma/G \ll R$, where the wetting ridge appears as a small feature on an otherwise flat substrate. In that case, θ_{L} can be derived from a global displacement of the contact line, as indicated in **Figure 5*e***, comparing the surface and elastic energies at the edges of the contour, far away from the contact line. At the contact line, this global displacement involves a superposition of the motions in **Figure 5*c*,*d***. Whether there is a correction to Young's law for θ_{L} depends on the elastic energy stored on both sides of the contact line, far from it. For free-standing membranes, the Shuttleworth effect can lead to a strain discontinuity, which gives an elastic correction to Young's law of order ϵ^2 (see the sidebar titled The Membrane Limit). For membranes glued to a rigid support, however, any jump in strain near the contact line will be screened by the finite membrane thickness. Hence, the strains far away from the contact line are equal, and θ_{L} follows from Young's law.

3.5. Is There a Shuttleworth Effect for Polymeric Solids?

The Shuttleworth effect is well established in crystalline materials (Muller & Saul 2004). For nonglassy polymer networks, however, it is less obvious that a change in the surface strain ϵ would alter the molecular structure of the interface, which is usually thought of as being close to that of an incompressible liquid. Hence, it is not clear a priori that soft polymeric substrates exhibit a strong dependence of surface energy γ on the applied strain ϵ. Although different numerical and theoretical works have been devoted to this question (Weijs et al. 2013, Liang et al. 2018b, Masurel et al. 2018), it is important in this review to focus on experimental results.

Experimental evidence for a strong Shuttleworth effect in soft wetting is provided in **Figure 7*a***, which shows the elastic displacement inside a thin elastomeric Wilhelmy plate, partially immersed (Marchand et al. 2012a). The observed discontinuity of strain across the contact line implies

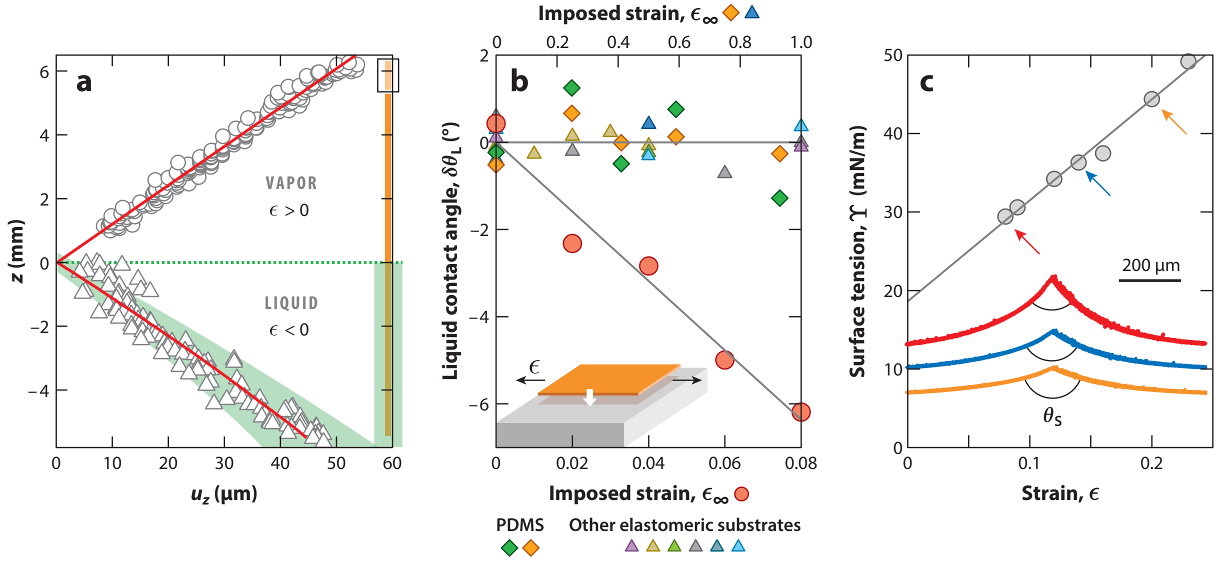


Figure 7

Soft wetting experiments investigating the Shuttleworth effect. (*a*) The elastic Wilhelmy plate, where an extensible rod (of elastocapillary length $\gamma/G \sim 1$ μm) is partially immersed in a liquid. The data represent the vertical displacements along the rod, showing a discontinuity of strain, $\epsilon = \mathrm{d}u_z/\mathrm{d}z$, across the contact line. Data from Marchand et al. (2012a). (*b*) Change of the liquid contact angle $\delta\theta_L$ versus imposed strain ϵ_∞. (*Red circles*; *lower axis*) θ_L on a stiff glassy polymer exhibits a strong dependence on strain. (*Other symbols*; *upper axis*) θ_L for drops on various soft elastomers exhibits no dependence on strain. Triangles correspond to various elastomeric substrates used in experiments by Schulman et al. (2018), and diamonds correspond to soft PDMS (polydimethylsiloxane) (advancing and receding angles) used in experiments by Snoeijer et al. (2018). (*c*) Dependence of solid surface tension Υ on strain ϵ on PDMS, as inferred from the change in the solid angle θ_S of the wetting ridge (*inset*). Data from Xu et al. (2017).

$\gamma'_{SV} - \gamma'_{SL} = 43 \pm 10$ mN/m, which is comparable to the relevant surface energies. A strong contact angle hysteresis was present in these experiments.

A recent series of experiments has systematically investigated the contact angles obtained upon externally stretching the substrate (Schulman et al. 2018). **Figure 7*b*** reports changes in the liquid contact angle $\delta\theta_L$ as a function of the imposed strain ϵ_∞ (with negligible contact angle hysteresis). These experiments were carried out in the regime where θ_L follows Young's law, so that they directly investigate the strain dependence of $\gamma_{SV} - \gamma_{SL}$. For stiff glassy polymers, $\gamma/G \ll a$, there is a clear change of contact angle, and hence a strong Shuttleworth effect, observed for four different liquid–glass combinations (only one data set is shown). The other data correspond to drops on a broad variety of soft elastomers with $\gamma/G \gg a$. These exhibit no variation of θ_L, even for strains up to $\epsilon_\infty = 1$. This implies that $\gamma_{SV} - \gamma_{SL}$ does not depend on strain ϵ. This robust outcome, found for seven different liquid–elastomer combinations (Schulman et al. 2018, Snoeijer et al. 2018), can be interpreted as evidence either that there is no measurable Shuttleworth effect or that there is a fundamental reason why γ_{SV} and γ_{SL} share the same dependence on ϵ. The second boundary condition (Equation 9), which needs to be satisfied in the absence of contact line pinning, indeed requires $\gamma'_{SV} = \gamma'_{SL}$ in the regime of small deformations (Snoeijer et al. 2018). However, rigorous results for large deformations are currently lacking, and it is unknown why many polymeric gels would obey this property.

In another series of experiments on PDMS, the solid angle θ_S was measured as a function of the imposed strain ϵ_∞ (Xu et al. 2017). **Figure 7*c*** shows that θ_S becomes shallower with strain—even though θ_L remains constant. The lack of dependence of θ_S with drop size and substrate thickness consolidates the interpretation of these experiments using Neumann's law (Equation 7) without any residual elastic tension, as mentioned in the discussion of Equation 8 above. The change of θ_S then implies a strong dependence of Υ on ϵ. This method also allows for a direct measurement of the tensorial nature of surface tension when applying anisotropic stretching (Xu et al. 2018).

A definite interpretation of these experiments will require the development of a fully nonlinear modeling approach and a consistent treatment of the Shuttleworth effect, including the second boundary condition (Equation 9). This needs to be complemented with a calibration of the material properties in the regime of large elastic deformations to account for the influence of stretching on the ridge shape.

4. DYNAMICAL ELASTOCAPILLARITY

4.1. Viscoelastic Braking

Pioneering studies by Shanahan & Carré (1995), Carré et al. (1996), and Long et al. (1996) revealed a dramatic slowing down of contact line motion when the substrate becomes exceedingly soft. The motion of the wetting ridge induces a time-dependent deformation of the substrate, leading to strong viscoelastic dissipation that opposes rapid motion. Typical velocities of millimeter-sized drops sliding under the influence of gravity can be as low as 10–100 nm/s (Karpitschka et al. 2016a) due to strong viscoelastic braking. A direct illustration of this effect occurs when drops move more rapidly over thin layers, as these induce less dissipation inside the substrate (Zhao et al. 2018a).

Figure 8*a* shows experimental data for a variety of dynamical wetting experiments. When the contact line moves with a velocity U, the contact angle changes from the equilibrium contact angle according to $\delta\theta_L \sim U^n$, with an exponent n close to 0.5. Linear viscoelasticity precisely predicts this behavior and relates the exponent n to that of the loss modulus, $G''/G \sim (\omega\tau)^n$ (Long et al. 1996). The contact line speed U excites the viscoelastic solid at a frequency $\omega \sim U/\ell$, where ℓ is the width of the ridge; at small velocities, ℓ is on the order of γ/G. A mechanical analysis on the scale of the wetting ridge shows that the contact line motion induces a rotation of the wetting ridge (Karpitschka et al. 2015), shown in the inset of **Figure 8*a***. The associated dissipation in the solid follows as

$$\frac{1}{2}\int \mathrm{d}x\mathrm{d}y \quad \overline{\overline{\boldsymbol{\sigma}}} : \overline{\overline{\dot{\boldsymbol{\gamma}}}} \quad \sim \ell^2\, G''\, \omega \sim \ell^2\, G \left(\frac{U\tau}{\ell}\right)^{n} \frac{U}{\ell} \qquad 10.$$

and is balanced with the injected power $\sim\gamma U \delta\theta$ due to capillary forces. When moving at a velocity U, the contact angle therefore changes from the equilibrium contact angle according to

$$\delta\theta_L \sim \left(\frac{U\tau}{\gamma/G}\right)^{n}. \qquad 11.$$

This is in excellent agreement with experiments, both for advancing and receding motion (**Figure 8*a***). Importantly, the dissipation is only integrable for $n \leq 1$ (Long et al. 1996). For $n = 1$, the standard moving contact line singularity for viscous liquids is recovered (Bonn et al. 2009, Snoeijer & Andreotti 2013). For an analogy between viscoelastic braking and dynamical adhesion, see the sidebar titled Dynamical Adhesion.

The argument can be made rigorous by deriving the shape of the wetting ridge from Equation 5, where the elastic traction $\boldsymbol{\sigma}$ is computed using a Green's function formalism (Johnson 1985). This Green's function approach has been extensively used to compute static wetting ridges (Jerison et al. 2011, Limat 2012, Marchand et al. 2012b, Style & Dufresne 2012, Bostwick et al. 2014, Lubbers et al. 2014) but can be extended to the dynamical case to account for the frequency

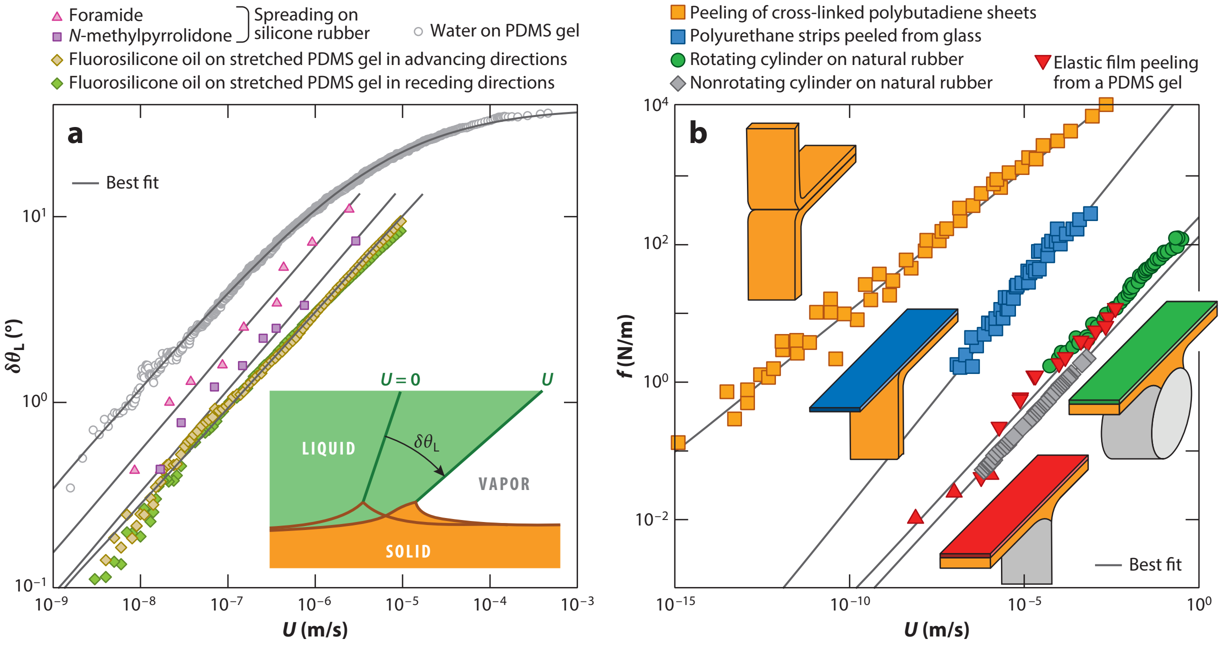


Figure 8

Energy dissipation inside a moving wetting ridge. (*a*) Variation of the liquid angle $\delta\theta_L$ as a function of the contact line velocity U in different liquid–solid wetting experiments: formamide (*red triangles*) and *N*-methylpyrrolidone (*blue squares*) spreading on silicon rubber (Shanahan & Carré 1995), water on a PDMS (polydimethylsiloxane) gel (*white circles*) (Karpitschka et al. 2015), and fluorosilicone oil on a stretched PDMS gel in the advancing (*orange diamonds*) and receding (*green diamonds*) directions (Snoeijer et al. 2018). The solid gray lines show the best fits by power laws, and the red curve includes phenomenologically the saturation at large velocity. (*b*) Relation between the dissipative force f and velocity U in solid–solid adhesion experiments: peeling of cross-linked sheets of polybutadiene (*orange squares*) (Gent 1996); polyurethane strips peeled from glass (*blue squares*) (Maugis & Barquins 1978); a cylinder moving on natural rubber, allowed to rotate (*green circles*) (Charmet et al. 1995) or not (*gray diamonds*) (Robbe-Valloire & Barquins 1998); peeling of an elastic film from a PDMS gel (*red triangles*) (Perrin et al. 2019). The solids lines show best fits by power laws. Data provided courtesy of the cited authors and replotted here.

dependence of the storage and loss moduli. For small velocities, this linear response framework gives the same result (Equation 11) for the ridge rotation (Karpitschka et al. 2015). Hence, the liquid angle passively follows the ridge rotation and maintains a Neumann balance even in the dynamical case.

At larger velocities, however, the size of the wetting ridge decreases. This is because the effective stiffness increases at frequencies beyond τ^{-1}, giving rise to a dynamical elastocapillary length, $\ell \sim \gamma/|G' + iG''|$. The volume over which dissipation occurs is then diminished, leading to a saturation of Equation 11 at large velocities; $\delta\theta_L$ approaches a constant value (**Figure 8*a***) at large velocities. The saturation is indeed observed above the viscoelastic velocity $\gamma/(G\tau)$ (Karpitschka et al. 2015), providing further direct evidence that the wetting dynamics is governed by substrate viscoelasticity.

4.2. Stick-Slip Dynamics

When a droplet is forced to move at high velocity, a remarkable stick-slip dynamics is observed (Pu & Severtson 2008; Pu et al. 2010; Kajiya et al. 2013, 2014; Karpitschka et al. 2015; Park et al. 2017; van Gorcum et al. 2018). **Figure 9*a*** shows the liquid angle as a function of time from a dip

DYNAMICAL ADHESION

Soft adhesives are polymeric layers that are used to bind rigid structures together; they must stick on most surfaces by simple contact under a low normal stress (Creton 2003). Their adherence strength is strongly dependent on the peeling velocity. Most pressure-sensitive adhesives exhibit a large nonlinear viscoelastic dissipation during debonding due to the formation of fibrils from cavitation (Teisseire et al. 2007, Villey et al. 2015) (**Figure 1*g***), as observed using probe tack tests (Nase et al. 2008, Vilmin et al. 2010, Chopin et al. 2018). Polymer chains across the fracture are stretched until they store an elastic energy comparable to the covalent binding energy (Lake & Thomas 1967). A bond-breaking event dissipates all this stored energy, so that the fracture energy Γ is proportional to the length N of the chain and to the cross-link density Σ. Cavitation and fibrillation typically take place when the elasto-adhesive length Γ/G is larger than the adhesive thickness e (Amouroux et al. 2001, Deplace et al. 2009, Creton & Ciccotti 2016).

Conversely, reversible adhesives can be peeled without bulk cavitation or plastic deformations and have been the subject of recent interest in biomimetics (Autumn et al. 2000, Ghatak et al. 2004, Boesel et al. 2010, Jagota & Hui 2011). During fracture, the front exhibits a wetting ridge similar to that observed in soft wetting (**Figure 1*f***), and the resulting dynamics indeed bears a strong resemblance. **Figure 8*b*** reviews the relation between the dissipative force per unit length and the contact line velocity for different fracture tests and different reversible adhesives. They indeed all exhibit a power law dependence that originates from the viscoelastic rheology (Schapery 1975, Newby et al. 1995, de Gennes 1996).

A deep connection between soft wetting and (reversible) adhesion has indeed been established over the last years, primarily in the static regime. Upon increasing the elastocapillary length, now given by γ/G, one can continuously go from the classical JKR (Johnson, Kendall, and Roberts) theory in solid mechanics to the normal case of liquid wetting (Salez et al. 2013, Style et al. 2013c, Hui et al. 2015, Cao et al. 2014). Surface tension in fact provides the the small-scale regularization of the crack singularity observed in JKR (Johnson et al. 1971, Johnson 1985, Liu et al. 2014, Karpitschka et al. 2016b) and in dynamical situations controls the size over which dissipation takes place (Perrin et al. 2019).

coating experiment (Kajiya et al. 2013). Importantly, the stick-slip motion is not associated with permanent defects of the substrate (Bonn et al. 2009, Snoeijer & Andreotti 2013) but here results dynamically from the self-induced wetting ridge. In solid friction, stick-slip occurs as an instability due to a decrease of the friction force when the velocity increases (Baumberger & Caroli 2006). The instability disappears at large velocity due to structural aging (Rice & Ruina 1983). Velocity weakening is also the mechanism invoked to explain the stick-slip observed during the peeling of adhesive tapes (Cortet et al. 2007, De Zotti et al. 2019), and a similar shear weakening leads to the formation of shear bands in complex fluids (Divoux et al. 2016).

In the dynamical wetting of soft solids, the friction force results from the viscoelastic loss due to the ridge motion. A velocity weakening mechanism would therefore relate stick-slip dynamics to the rheology. Indeed, stick-slip appears when the frequency of excitation due to the contact line motion becomes comparable to the rheological cross-over frequency τ^{-1} at which the storage and loss moduli become comparable. For paraffin gels, stick-slip is observed in an intermediate range of velocities (Kajiya et al. 2013, 2014). At high and low speeds, one recovers a continuous spreading, as in solid friction, corresponding to a purely viscous and a purely elastic response, respectively. For PDMS gels, however, a single transition from continuous to stick-slip motion has been reported (Karpitschka et al. 2015, Park et al. 2017, van Gorcum et al. 2018). The lack of continuous motion at high speed can be traced back to the PDMS rheology at high frequency, for which G' and G'' remain comparable for all relevant frequencies beyond τ^{-1}.

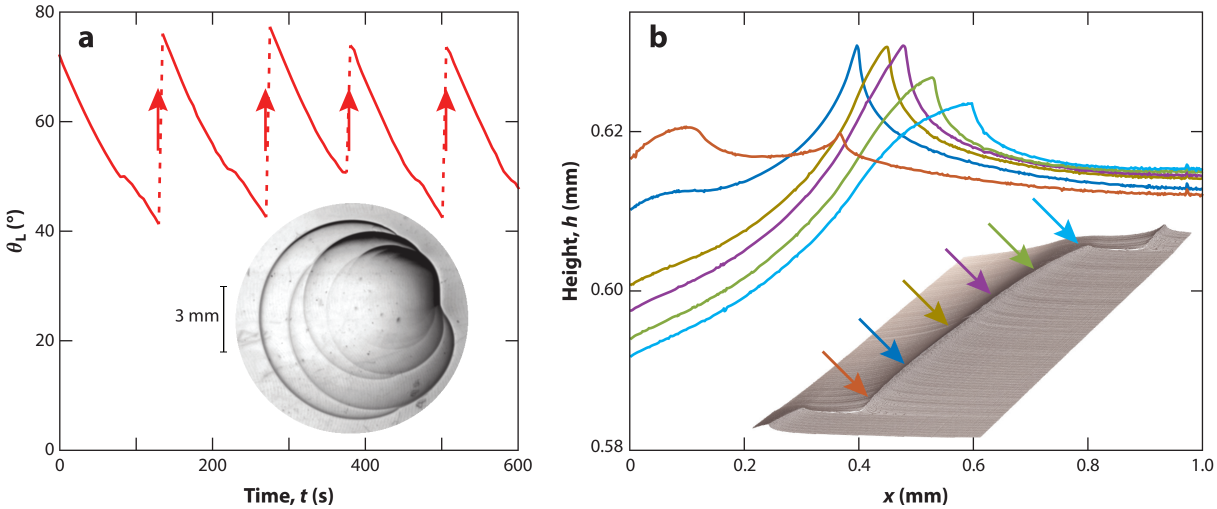


Figure 9

Stick-slip motion during forced spreading. (*a*, *inset*) Picture of a SBS–paraffin gel surface after the spreading of a drop whose stick-slip has left multiple circular patterns. Data from Kajiya et al. (2013). (*a*) Time evolution of the liquid angle θ_L in a dip-coating experiment, performed with the same substrate. The contact angle slowly decreases as it sticks to the substrate and suddenly increases during a slip event. Data from Kajiya et al. (2014). (*b*) Wetting ridge profile as a function of time during a stick-slip cycle, as the contact line is forced over a PDMS gel. Note that the tip angle of the wetting ridge increases when approaching the depinning event. (*Inset*) Space-time diagram. Abbreviations: PDMS, polydimethylsiloxane; SBS, poly(styrene-*b*-butadiene-*b*-styrene). Data from van Gorcum et al. (2018).

In the case of PDMS, however, friction does not decrease with velocity. Instead of a single wetting ridge that accelerates during slip events, direct visualizations have shown that the contact line depins from its own wetting ridge and surfs it (Park et al. 2017, van Gorcum et al. 2018). The depinning is governed by the Gibb's inequality that is normally used for sharp edges on rigid surfaces (van Gorcum et al. 2018) and is related to the increase of the solid angle of the ridge θ_S with contact line velocity (**Figure 9*b***). This change of θ_S points either to time-dependent surface stresses or to dynamical effects beyond a simple static Neumann balance. Although for purely elastic media the wetting ridges decay after depinning, for some systems plastically deformed ridges remain after stick-slip events (Pu et al. 2010), confirming the generality of the depining scenario.

4.3. Manipulating Droplet Motion

When exposed to a humid environment, the deformability of the substrate strongly affects the condensation of droplets (Sokuler et al. 2010) (**Figure 10*a***). Similar features are observed on lubricant-infused surfaces (Kajiya et al. 2016), where instead of a gel layer the drop is in contact with a viscous lubricant that is maintained in a porous substrate (Schellenberger et al. 2015, Keiser et al. 2017). Both cross-linked polymer networks and lubricant surfaces can exhibit vanishing contact angle hysteresis, making them of particular interest for applications such as antifouling and food packaging (Wong et al. 2011, Solomon et al. 2016). In addition, the use of soft coatings can be highly effective for the suppression of splashing (Howland et al. 2016).

Once several drops are in close proximity, the substrate deformation leads to intricate interactions between adjacent drops (**Figure 10*b***). This inverted Cheerios effect resembles the capillary interaction of solid particles at liquid interfaces (known as the Cheerios effect), except that the roles of liquid and solid are reversed. The interactions can be manipulated from attractive to

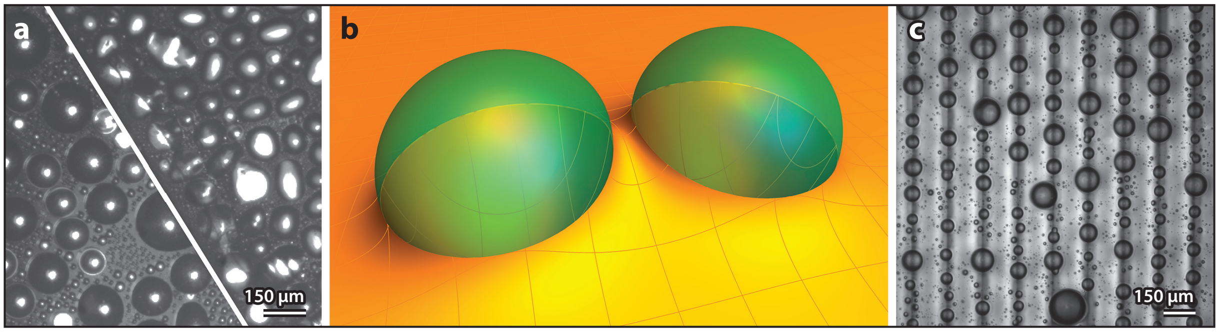


Figure 10

(*a*) Condensation on two PDMS (polydimethylsiloxane) surfaces cooled simultaneously under identical experimental conditions, the left one being five times softer than the right (Sokuler et al. 2010). (*b*) The inverted Cheerios effect: a three-dimensional calculation of the elastocapillary meniscus between two drops, leading to interactions (Karpitschka et al. 2016a). (*c*) Droplet durotaxis. Glycerol droplets, deposited on a flat surface with stiffness gradients, move toward the thickest regions of the substrate (Style et al. 2013b). Panel *a* adapted with permission from Sokuler et al. (2010), copyright 2010 American Chemical Society.

repulsive by tuning the geometry of the substrate (Karpitschka et al. 2016a). A similar approach to manipulate droplet motion is achieved by imposing gradients in the mechanical properties of the substrate. Droplets exhibit durotaxis (Style et al. 2013b), resembling that of biological cells (Saez et al. 2007, Trichet et al. 2012), and tend to migrate toward soft regions (**Figure 10*c***). Such a migration was observed for sliding drops, exhibiting a curved trajectory in the presence of gradients of dissipation, bending toward thicker regions (Zhao et al. 2018a).

4.4. Dynamical Wetting of Brushes and Swollen Gels

The continuum vision of the dynamical elastocapillary problem is perfectly adapted to the case of a totally reticulated polymer gel without dangling ends or residual melt content. In this final section, we report dynamical experiments on other systems, such as brushes and swollen gels. **Figure 2*g*** showed that brushes are highly effective in changing the wettability of the substrate (Maas et al. 2002) and can reduce contact angle hysteresis to within a few percent of a degree (Lhermerout et al. 2016). For dynamical wetting, the dissipation inside the substrate can again be the dominant factor, now due to the stretching and relaxation of the polymer chains. **Figure 11*a*** shows that the dissipative force (quantified by $\delta\theta$) in a PDMS brush is proportional to the contact line velocity, rescaled by the brush thickness e and the relaxation time τ (Lhermerout et al. 2016). This opens up the promising perspective of using dynamic contact lines to probe the nanorheology of polymeric materials (Restagno et al. 2002, Garcia et al. 2016).

The swelling of a polymer gel by un-cross-linked chains may substantially modify the contact line dynamics: A capillary-induced phase separation can occur that extracts free chains from the cross-linked polymer network, which then accumulate at the surface and alter the wetting velocity (Hourlier-Fargette et al. 2017, 2018). Also without phase separation, the dynamics can be dominated by poroelastic effects, competing with viscoelasticity (Johnson 1982, Hong et al. 2008, Doi 2009). After removing the pulling force, the relaxation of wetting ridges may indeed be limited by poroelastic diffusion (Zhao et al. 2018b, Berman et al. 2019).

Figure 11*b* reports an experiment motivated by the food industry, where water drops spread on a soluble glassy polymer (Dupas et al. 2014). At low velocity, the change of the contact angle is due to hydration of the substrate ahead of the contact line, resulting from the evaporation and

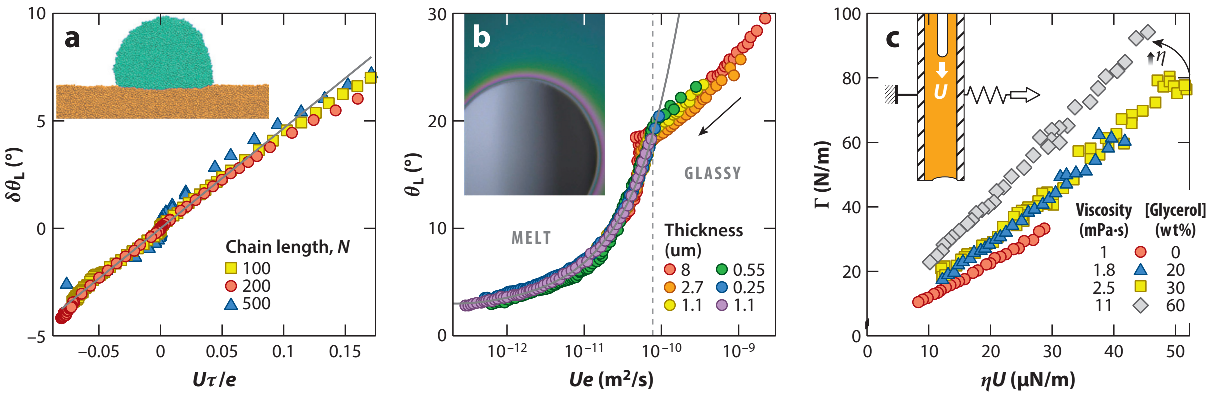


Figure 11

(*a*) Dissipation induced by a liquid contact line moving on PDMS (polydimethylsiloxane) pseudo-brushes, collapsing data for varying chain lengths. The graph shows the variation of the liquid contact angle $\delta\theta_L$ with the velocity U, rescaled by the thickness e of the brush and the Rouse relaxation time τ. (*Inset*) Droplet on a brush in a molecular dynamics simulation (Mensink et al. 2019). (*b*) Liquid angle θ_L as a function of the contact line velocity U for water droplets spreading on maltodextrin of different thicknesses e, from 100 nm to 8 μm. (*c*) Fracture energy release rate Γ as a function of front velocity U for gels with the same gelatin concentration ($c = 5$ wt%) and various glycerol contents and therefore different viscosities η. Panels adapted with permission from (*a*) Lhermerout et al. (2016); (*b*) Dupas et al. (2014), copyright 2014 American Physical Society; and (*c*) Baumberger et al. (2006b), copyright 2006 Springer Nature.

subsequent condensation on the substrate. The water concentration depends on the product of the speed U and the layer thickness e and indeed leads to a collapse of data onto a master curve (**Figure 11*b***). Above a critical velocity, the diffusion is too slow and the substrate remains in a glassy state. Note that the scaling with e is different in panels *a* and *b*, reflecting the different dissipative mechanisms at play. Further complexity is encountered when the droplet wetting the hydrogel contains solutes. A surfactant-laden drop can induce fracture of the gel (Bostwick & Daniels 2013, Grzelka et al. 2017), while the addition of particles can offer strategies for the deposition of particles to avoid the coffee stain effect, and even to tune the contact angle (Boulogne et al. 2015, 2016, 2017).

We have already mentioned the large fracture energy Γ for debonding pressure-sensitive adhesives, associated with pulling connecting chains from the matrix. Similar effects arise in the fracture of swollen physical gels, in which cross-linking is due to weak reversible bonds. **Figure 11*c*** shows that the crack velocity under a given load decreases when the solvent viscosity increases. In this case, the dissipation is dominated by viscoplastic chain pull-out (Raphael & De Gennes 1992), hence the dependence on chain/solvent friction (Baumberger et al. 2006a). This illustrates that further dissipative mechanisms should be anticipated when exploring soft wetting in a broader class of soft-polymer substrates.

FUTURE ISSUES

1. To turn soft wetting experiments into fully quantitative probes for (surface) rheology, researchers should include in future models large elastic deformations (van Brummelen et al. 2017, Masurel et al. 2018, Wu et al. 2018) and a fully consistent implementation of the Shuttleworth effect.
2. What are the physicochemical conditions determining the appearance of the Shuttleworth effect?

3. Under what conditions do soft substrates exhibit hysteresis? When can droplets induce viscoplastic deformations inside the substrate? Could these effects be included as a process zone at the wetting ridge, as in fracture mechanics?
4. What mechanisms could lead to the dynamical variation of θ_S (**Figure 9**), which initiates the depinning in the stick-slip regime?
5. Similar to lubricant-infused surfaces, the combination of surface topography and functionalization by polymers offers new routes toward surface engineering.
6. Contact line forces provide a unique mechanical probe of biological matter, while drops on gels offer a model system to study principles on mechanobiology (Schwarz & Safran 2013, Charras & Sahai 2014, Humphrey et al. 2014).

DISCLOSURE STATEMENT

The authors are not aware of any biases that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We are indebted to our close collaborators, in particular, S. Karpitschka, E. Rolley, and K. Davitt, and we thank T. Baumberger, S. de Beer, A. Pandey, and T. Salez for their valuable feedback on the manuscript. We are grateful to the scientific community present during the Lorentz Center workshop "Capillarity of Soft Interfaces" (Andreotti et al. 2016). We acknowledge financial support from ERC (European Research Council) Consolidator Grant number 616918 (to J.H.S.) and from ANR (French National Agency for Research) grant SMART (to B.A.).

LITERATURE CITED

Alexander S. 1977. Adsorption of chain molecules with a polar head a scaling description. *J. Phys.* 38:983–87

Amouroux N, Petit J, Léger L. 2001. Role of interfacial resistance to shear stress on adhesive peel strength. *Langmuir* 17:6510–17

Andreotti B, Bäumchen O, Boulogne F, Daniels KE, Dufresne ER, et al. 2016. Solid capillarity: When and how does surface tension deform soft solids? *Soft Matter* 12:2993–96

Andreotti B, Snoeijer JH. 2016. Soft wetting and the Shuttleworth effect, at the crossroads between thermodynamics and mechanics. *EPL* 109:66001

Autumn K, Liang YA, Hsieh ST, Zesch W, Chan WP, et al. 2000. Adhesive force of a single gecko foot-hair. *Nature* 405:681–85

Baumberger T, Caroli C. 2006. Solid friction from stick–slip down to pinning and aging. *Adv. Phys.* 55:279–348

Baumberger T, Caroli C, Martina D. 2006a. Fracture of a biopolymer gel as a viscoplastic disentanglement process. *Eur. Phys. J. E* 21:81–89

Baumberger T, Caroli C, Martina D. 2006b. Solvent control of crack dynamics in a reversible hydrogel. *Nat. Mater.* 5:552–55

Berman JD, Randeria M, Style RW, Xu Q, Nichols JR, et al. 2019. Singular dynamics in the failure of soft adhesive contacts. *Soft Matter* 15:1327–34

Bico J, Reyssat É, Roman B. 2018. Elastocapillarity: when surface tension deforms elastic solids. *Annu. Rev. Fluid Mech.* 50:629–59

Binder K, Kob W. 2011. *Glassy Materials and Disordered Solids: An Introduction to their Statistical Mechanics*. Singapore: World Sci.

Boese D, Kremer F. 1990. Molecular dynamics in bulk *cis*-polyisoprene as studied by dielectric spectroscopy. *Macromolecules* 23:829–35

Boesel LF, Greiner C, Arzt E, del Campo A. 2010. Gecko-inspired surfaces: a path to strong and reversible dry adhesives. *Adv. Mater.* 22:2125–37

Bonn D, Eggers J, Indekeu J, Meunier J, Rolley E. 2009. Wetting and spreading. *Rev. Mod. Phys.* 81:739

Bostwick JB, Daniels KE. 2013. Capillary fracture of soft gels. *Phys. Rev. E* 88:042410

Bostwick JB, Shearer M, Daniels K. 2014. Elastocapillary deformations on partially-wetting substrates: rival contact-line models. *Soft Matter* 10:7361–69

Boulogne F, Ingremeau F, Dervaux J, Limat L, Stone HA. 2015. Homogeneous deposition of particles by absorption on hydrogels. *EPL* 112:48004

Boulogne F, Ingremeau F, Limat L, Stone HA. 2016. Tuning the receding contact angle on hydrogels by addition of particles. *Langmuir* 32:5573–79

Boulogne F, Shin S, Dervaux J, Limat L, Stone HA. 2017. Diffusiophoretic manipulation of particles in a drop deposited on a hydrogel. *Soft Matter* 13:5122–29

Cao Z, Stevens MJ, Dobrynin AV. 2014. Adhesion and wetting of nanoparticles on soft surfaces. *Macromolecules* 47:3203–9

Carré A, Gastel JC, Shanahan MER. 1996. Viscoelastic effects in the spreading of liquids. *Nature* 379:432–34

Chakrabarti A, Porat A, Raphaël E, Salez T, Chaudhury MK. 2018. Elastowetting of soft hydrogel spheres. *Langmuir* 34:3894–900

Charmet JC, Verjus C, Barquins M. 1995. Sur la dimension du contact et la cinétique de roulement d'un cylindre long et rigide sous la surface plane et lisse d'un massif de caoutchouc souple. *C.R. Acad. Sci. IIb* 321:443–50

Charras G, Sahai E. 2014. Physical influences of the extracellular environment on cell migration. *Nat. Rev. Mol. Cell Biol.* 15:813–24

Chopin J, Villey R, Yarusso D, Barthel E, Creton C, Ciccotti M. 2018. Nonlinear viscoelastic modeling of adhesive failure for polyacrylate pressure-sensitive adhesives. *Macromolecules* 51:8605–10

Cohen Stuart MA, de Vos WM, Leermakers FAM. 2006. Why surfaces modified by flexible polymers often have a finite contact angle for good solvents. *Langmuir* 22:1722–28

Cortet PP, Ciccotti M, Vanel L. 2007. Imaging the stick–slip peeling of an adhesive tape under a constant load. *J. Stat. Mech. Theory Exp.* 2007:P03005

Creton C. 2003. Pressure-sensitive adhesives: an introductory course. *MRS Bull.* 28:434–39

Creton C, Ciccotti M. 2016. Fracture and adhesion of soft materials: a review. *Rep. Prog. Phys.* 79:046601

Creton C, Kramer EJ, Hui CY, Brown HR. 1992. Failure mechanisms of polymer interfaces reinforced with block copolymers. *Macromolecules* 25:3075–88

Davidovitch B, Vella D. 2018. Partial wetting of thin solid sheets under tension. *Soft Matter* 14:4913–34

de Gennes PG. 1979. *Scaling Concepts in Polymer Physics*. Ithaca, NY: Cornell Univ. Press

de Gennes PG. 1980. Conformations of polymers attached to an interface. *Macromolecules* 13:1069–75

de Gennes PG. 1989. Weak adhesive junctions. *J. Phys. France* 50:2551–62

de Gennes PG. 1996. Soft adhesives. *Langmuir* 12:4497–500

de Gennes PG, Brochart-Wyart F, Quéré D. 2002. *Capillarity and Wetting Phenomena: Drops, Bubbles, Pearls, Waves*. New York: Springer-Verlag

De Zotti V, Rapina K, Cortet PP, Vanel L, Santucci S. 2019. Bending to kinetic energy transfer in adhesive peel front microinstability. *Phys. Rev. Lett.* 122:068005

Deplace F, Carelli C, Mariot S, Retsos H, Chateauminois A, et al. 2009. Fine tuning the adhesive properties of a soft nanostructured adhesive with rheological measurements. *J. Adhes.* 85:18–54

Divoux T, Fardin MA, Manneville S, Lerouge S. 2016. Shear banding of complex fluids. *Annu. Rev. Fluid Mech.* 48:81–103

Doi M. 2009. Gel dynamics. *J. Phys. Soc. Jpn.* 78:052001

Doi M, Edwards SF. 1988. *The Theory of Polymer Dynamics*. Oxford: Oxford Univ. Press

Dupas J, Verneuil E, Van Landeghem M, Bresson B, Forny L, et al. 2014. Glass transition accelerates the spreading of polar solvents on a soluble polymer. *Phys. Rev. Lett.* 112:188302

Duprat C, Protière S, Beebe AY, Stone HA. 2012. Wetting of flexible fibre arrays. *Nature* 482:510–13

Fleer G, Stuart MC, Scheutjens JM, Cosgrove T, Vincent B. 1993. *Polymers at Interfaces*. Dordrecht, Neth.: Springer

Garcia L, Barraud C, Picard C, Giraud J, Charlaix E, Cross B. 2016. A micro-nano-rheometer for the mechanics of soft matter at interfaces. *Rev. Sci. Instrum.* 87:113906

Gent AN. 1996. Adhesion and strength of viscoelastic solids: Is there a relationship between adhesion and bulk properties? *Langmuir* 12:4492–96

Ghatak A, Mahadevan L, Chung JY, Chaudhury MK, Shenoy V. 2004. Peeling from a biomimetically patterned thin elastic film. *Proc. R. Soc. Lond. A* 460:2725–35

Grandgeorge P, Krins N, Hourlier-Fargette A, Laberty-Robert C, Neukirch S, Antkowiak A. 2018. Capillarity-induced folds fuel extreme shape changes in thin wicked membranes. *Science* 360:296–99

Grzelka M, Bostwick JB, Daniels KE. 2017. Capillary fracture of ultrasoft gels: variability and delayed nucleation. *Soft Matter* 13:2962–66

Hong W, Zhao X, Zhou J, Suo Z. 2008. A theory of coupled diffusion and large deformation in polymeric gels. *J. Mech. Phys. Solids* 56:1779–93

Hourlier-Fargette A, Antkowiak A, Chateauminois A, Neukirch S. 2017. Role of uncrosslinked chains in droplets dynamics on silicone elastomers. *Soft Matter* 13:3484–91

Hourlier-Fargette A, Dervaux J, Antkowiak A, Neukirch S. 2018. Extraction of silicone uncrosslinked chains at air–water–polydimethylsiloxane triple lines. *Langmuir* 34:12244–50

Howland CJ, Antkowiak A, Castrejón-Pita JR, Howison SD, Oliver JM, et al. 2016. It's harder to splash on soft solids. *Phys. Rev. Lett.* 117:184502

Hui CY, Liu T, Salez T, Raphael E, Jagota A. 2015. Indentation of a rigid sphere into an elastic substrate with surface tension and adhesion. *Proc. R. Soc. A* 471:20140727

Humphrey JD, Dufresne ER, Schwartz MA. 2014. Mechanotransduction and extracellular matrix homeostasis. *Nat. Rev. Mol. Cell Biol.* 15:802–12

Jagota A, Hui CY. 2011. Adhesion, friction, and compliance of bio-mimetic and bio-inspired structured interfaces. *Mater. Sci. Eng. R* 72:253–92

Jensen KE, Sarfati R, Style RW, Boltyanskiy R, Chakrabarti A, et al. 2015. Wetting and phase separation in soft adhesion. *PNAS* 112:14490–94

Jerison ER, Xu Y, Wilen LA, Dufresne ER. 2011. Deformation of an elastic substrate by a three-phase contact line. *Phys. Rev. Lett.* 106:186103

Johnson DL. 1982. Elastodynamics of gels. *J. Chem. Phys.* 77:1531–39

Johnson KL. 1985. *Contact Mechanics*. Cambridge, UK: Cambridge Univ. Press

Johnson KL, Kendall K, Roberts A. 1971. Surface energy and the contact of elastic solids. *Proc. R. Soc. Lond. A* 324:301

Kajiya T, Brunet P, Royon L, Daerr A, Receveur M, Limat L. 2014. A liquid contact line receding on a soft gel surface: dip-coating geometry investigation. *Soft Matter* 10:8888–95

Kajiya T, Daerr A, Narita T, Royon L, Lequeux F, Limat L. 2011. Dynamics of the contact line in wetting and diffusing processes of water droplets on hydrogel (PAMPS–PAAM) substrates. *Soft Matter* 7:11425–32

Kajiya T, Daerr A, Narita T, Royon L, Lequeux F, Limat L. 2013. Advancing liquid contact line on viscoelastic gel substrates: stick-slip versus continuous motions. *Soft Matter* 9:454–61

Kajiya T, Schellenberger F, Papadopoulos P, Vollmer D, Butt HJ. 2016. 3D imaging of water-drop condensation on hydrophobic and hydrophilic lubricant-impregnated surfaces. *Sci. Rep.* 6:23687

Karpitschka S, Das S, van Gorcum M, Perrin H, Andreotti B, Snoeijer JH. 2015. Droplets move over viscoelastic substrates by surfing a ridge. *Nat. Commun.* 6:7891

Karpitschka S, Pandey A, Lubbers LA, Weijs JH, Botto L, et al. 2016a. Liquid drops attract or repel by the inverted Cheerios effect. *PNAS* 113:7403–7

Karpitschka S, van Wijngaarden L, Snoeijer JH. 2016b. Surface tension regularizes the crack singularity of adhesion. *Soft Matter* 12:4463–71

Keiser A, Keiser L, Clanet C, Quéré D. 2017. Drop friction on liquid-infused materials. *Soft Matter* 13:6981–87

Kirkwood J, Buff F. 1949. The statistical mechanical theory of surface tension. *J. Chem. Phys.* 17:338–43

Lafuma A, Quéré D. 2011. Slippery pre-suffused surfaces. *EPL* 96:56001

Lake G, Thomas A. 1967. The strength of highly elastic materials. *Proc. R. Soc. Lond. A* 300:108–19

Leonforte F, Mueller M. 2011. Statics of polymer droplets on deformable surfaces. *J. Chem. Phys.* 135:214703

Lhermerout R, Perrin H, Rolley E, Andreotti B, Davitt K. 2016. A moving contact line as a rheometer for nanometric interfacial layers. *Nat. Commun.* 7:12545

Liang H, Cao Z, Wang Z, Dobrynin AV. 2018a. Surface stress and surface tension in polymeric networks. *ACS Macro Lett.* 7:116–21

Liang H, Cao Z, Wang Z, Dobrynin AV. 2018b. Surface stresses and a force balance at a contact line. *Langmuir* 34:7497–502

Limat L. 2012. Straight contact lines on a soft, incompressible solid. *Eur. Phys. J. E* 35:134

Liu T, Long R, Hui CY. 2014. The energy release rate of a pressurized crack in soft elastic materials: effects of surface tension and large deformation. *Soft Matter* 10:7723–29

Long D, Ajdari A, Leibler L. 1996. How do grafted polymer layers alter the dynamics of wetting? *Langmuir* 12:1675–80

Lubbers LA, Weijs JH, Botto L, Das S, Andreotti B, Snoeijer JH. 2014. Drops on soft solids: free energy and double transition of contact angles. *J. Fluid Mech.* 747:R1

Maas J, Fleer G, Leermakers F, Cohen Stuart M. 2002. Wetting of a polymer brush by a chemically identical polymer melt: Phase diagram and film stability. *Langmuir* 18:8871–80

Manning ML, Foty RA, Steinberg MS, Schoetz EM. 2010. Coaction of intercellular adhesion and cortical tension specifies tissue surface tension. *PNAS* 107:12517–22

Marchand A, Das S, Snoeijer JH, Andreotti B. 2012a. Capillary pressure and contact line force on a soft solid. *Phys. Rev. Lett.* 108:094301

Marchand A, Das S, Snoeijer JH, Andreotti B. 2012b. Contact angles on a soft solid: from Young's law to Neumann's law. *Phys. Rev. Lett.* 109:236101

Marchand A, Weijs JH, Snoeijer JH, Andreotti B. 2011. Why is surface tension a force parallel to the interface? *Am. J. Phys.* 79:999–1008

Masurel R, Roché M, Limat L, Ionescu I, Dervaux J. 2018. Elastocapillary ridge as a noninteger disclination. *Phys. Rev. Lett.* 122:248004

Maugis D, Barquins M. 1978. Fracture mechanics and the adherence of viscoelastic bodies. *J. Phys. D* 11:1989–2023

Mensink LI, Snoeijer JH, de Beer S. 2019. Wetting of polymer brushes by polymeric nanodroplets. *Macromolecules* 52:2015–20

Milner S, Witten T, Cates M. 1988. Theory of the grafted polymer brush. *Macromolecules* 21:2610–19

Mora S, Phou T, Fromental JM, Pismen LM, Pomeau Y. 2010. Capillarity driven instability of a soft solid. *Phys. Rev. Lett.* 105:214301

Muller P, Saul A. 2004. Elastic effects on surface physics. *Surf. Sci. Rep.* 54:157–258

Nase J, Lindner A, Creton C. 2008. Pattern formation during deformation of a confined viscoelastic layer: from a viscous liquid to a soft elastic solid. *Phys. Rev. Lett.* 101:074503

Neukirch S, Antkowiak A, Marigo JJ. 2014. Soft beams: when capillarity induces axial compression. *Phys. Rev. E* 89:012401

Newby BZ, Chaudhury MK, Brown HR. 1995. Macroscopic evidence of the effect of interfacial slippage on adhesion. *Science* 269:1407–9

Nijmeijer M, Bruin C, Bakker A, Van Leeuwen J. 1990. Wetting and drying of an inert wall by a fluid in a molecular-dynamics simulation. *Phys. Rev. A* 42:6052–59

Onogi S, Masuda T, Kitagawa K. 1970. Rheological properties of anionic polystyrenes. I. Dynamic viscoelasticity of narrow-distribution polystyrenes. *Macromolecules* 3:109–16

Park SJ, Bostwick JB, De Andrade V, Je JH. 2017. Self-spreading of the wetting ridge during stick-slip on a viscoelastic surface. *Soft Matter* 13:8331–36

Park SJ, Weon B, Lee J, Lee J, Kim J, Je J. 2014. Visualization of asymmetric wetting ridges on soft solids with X-ray microscopy. *Nat. Commun.* 5:4369

Pericet-Camara R, Best A, Butt HJ, Bonaccurso E. 2008. Effect of capillary pressure and surface tension on the deformation of elastic surfaces by sessile liquid microdrops: an experimental investigation. *Langmuir* 24:10565–68

Perrin H, Eddi A, Karpitschka S, Snoeijer JH, Andreotti B. 2019. Peeling an elastic film from a soft viscoelastic adhesive: experiments and scaling laws. *Soft Matter* 15:770–78

Pham JT, Schellenberger F, Kappl M, Butt HJ. 2017. From elasticity to capillarity in soft materials indentation. *Phys. Rev. Mater.* 1:015602

Pu G, Ai J, Severtson SJ. 2010. Drop behavior on a thermally-stripped acrylic polymer: influence of surface tension induced wetting ridge formation on retention and running. *Langmuir* 26:12696–702

Pu G, Severtson SJ. 2008. Characterization of dynamic stick-and-break wetting behavior for various liquids on the surface of a highly viscoelastic polymer. *Langmuir* 24:4685–92

Raphael E, De Gennes P. 1992. Rubber-rubber adhesion with connector molecules. *J. Phys. Chem.* 96:4002–7

Restagno F, Crassous J, Charlaix E, Cottin-Bizonne C, Monchanin M. 2002. A new surface forces apparatus for nanorheology. *Rev. Sci. Instrum.* 73:2292–97

Rice J, Ruina AL. 1983. Stability of steady frictional slipping. *J. Appl. Mech.* 50:343–49

Robbe-Valloire F, Barquins M. 1998. Adhesive contact and kinetics of adherence between a rigid cylinder and an elastomeric solid. *Int. J. Adhes. Adhes.* 18:29–34

Roman B, Bico J. 2010. Elasto-capillarity: deforming an elastic structure with a liquid droplet. *J. Phys. Cond. Matt.* 22:493101

Rowlinson JS, Widom B. 1982. *Molecular Theory of Capillarity*. Oxford: Clarendon

Rubinstein M, Colby RH. 2003. *Polymer Physics*. New York: Oxford Univ. Press

Rusanov A. 1975. Theory of wetting of elastically deformed bodies. 1. Deformation with a finite contact-angle. *Colloid J. USSR* 37:614–22

Saez A, Ghibaudo M, Buguin A, Silberzan P, Ladoux B. 2007. Rigidity-driven growth and migration of epithelial cells on microstructured anisotropic substrates. *PNAS* 104:8281–86

Salez T, Benzaquen M, Raphaël É. 2013. From adhesion to wetting of a soft particle. *Soft Matter* 9:10699–704

Schapery RA. 1975. A theory of crack initiation and growth in viscoelastic media. *Int. J. Fract.* 11:141–59

Schellenberger F, Xie J, Encinas N, Hardy A, Klapper M, et al. 2015. Direct observation of drops on slippery lubricant-infused surfaces. *Soft Matter* 11:7617–26

Schroll R, Adda-Bedia M, Cerda E, Huang J, Menon N, et al. 2013. Capillary deformations of bendable films. *Phys. Rev. Lett.* 111:014301

Schulman RD, Dalnoki-Veress K. 2015. Liquid droplets on a highly deformable membrane. *Phys. Rev. Lett.* 115:206101

Schulman RD, Trejo M, Salez T, Raphaël E, Dalnoki-Veress K. 2018. Surface energy of strained amorphous solids. *Nat. Commun.* 9:982

Schwarz US, Safran SA. 2013. Physics of adherent cells. *Rev. Mod. Phys.* 85:1327

Seveno D, Blake TD, De Coninck J. 2013. Young's equation at the nanoscale. *Phys. Rev. Lett.* 111:096101

Shanahan M. 1987. The influence of solid micro-deformation on contact-angle equilibrium. *J. Phys. D* 20:945–50

Shanahan M, Carré A. 1995. Viscoelastic dissipation in wetting and adhesion phenomena. *Langmuir* 11:1396–402

Shuttleworth R. 1950. The surface tension of solids. *Proc. Phys. Soc. Lond. A* 63:444–57

Singh M, Pipkin AC. 1965. Note on Ericksen's problem. *Z. Angew. Math. Phys.* 16:706–9

Snoeijer JH, Andreotti B. 2008. A microscopic view on contact angle selection. *Phys. Fluids* 20:057101

Snoeijer JH, Andreotti B. 2013. Moving contact lines: scales, regimes, and dynamical transitions. *Annu Rev. Fluid Mech.* 45:269–92

Snoeijer JH, Rolley E, Andreotti B. 2018. Paradox of contact angle selection on stretched soft solids. *Phys. Rev. Lett.* 121:068003

Sokuler M, Auernhammer G, Roth M, Liu C, Bonacurrso E, Butt H. 2010. The softer the better: fast condensation on soft surfaces. *Langmuir* 26:1544–47

Solomon BR, Subramanyam SB, Farnham TA, Khalil KS, Anand S, Varanasi KK. 2016. Lubricant-impregnated surfaces. In *Non-Wettable Surfaces*, ed. RHA Ras, A Marmur, pp. 285–318. Cambridge, UK: R. Soc. Chem.

Style RW, Boltyanskiy R, Che Y, Wettlaufer JS, Wilen LA, Dufresne ER. 2013a. Universal deformation of soft substrates near a contact line and the direct measurement of solid surface stresses. *Phys. Rev. Lett.* 110:066103

Style RW, Che Y, Park SJ, Weon BM, Je JH, et al. 2013b. Patterning droplets with durotaxis. *PNAS* 110:12541–44

Style RW, Dufresne ER. 2012. Static wetting on deformable substrates, from liquids to soft solids. *Soft Matter* 8:7177–84

Style RW, Hyland C, Boltyanskiy R, Wettlaufer JS, Dufresne ER. 2013c. Surface tension and contact with soft elastic solids. *Nat. Commun.* 4:2728

Style RW, Jagota A, Hui CY, Dufresne ER. 2017. Elastocapillarity: surface tension and the mechanics of soft solids. *Annu. Rev. Condens. Matter Phys.* 8:99–118

Teisseire J, Nallet F, Fabre P, Gay C. 2007. Understanding cracking versus cavitation in pressure-sensitive adhesives: the role of kinetics. *J. Adhes.* 83:613–77

Trichet L, Le Digabel J, Hawkins RJ, Vedula SRK, Gupta M, et al. 2012. Evidence of a large-scale mechanosensing mechanism for cellular adaptation to substrate stiffness. *PNAS* 109:6933–38

van Brummelen EH, Roudbari MS, Şimşek G, van der Zee KG. 2017. Binary-fluid–solid interaction based on the Navier–Stokes–Cahn–Hilliard equations. In *Fluid Structure Interaction*, ed. S Frei, B Holm, T Richter, T Wick, H Yang, pp. 283–328. Berlin: De Gruyter

van Gorcum M, Andreotti B, Snoeijer JH, Karpitschka S. 2018. Dynamic solid surface tension causes droplet pinning and depinning. *Phys. Rev. Lett.* 121:208003

Villey R, Creton C, Cortet PP, Dalbe MJ, Jet T, et al. 2015. Rate-dependent elastic hysteresis during the peeling of pressure sensitive adhesives. *Soft Matter* 11:3480–91

Vilmin T, Ziebert F, Raphaël E. 2010. Simple view on fingering instability of debonding soft elastic adhesives. *Langmuir* 26:3257–60

Watanabe H. 1999. Viscoelasticity and dynamics of entangled polymers. *Prog. Polym. Sci.* 24:1253–403

Weijs JH, Andreotti B, Snoeijer JH. 2013. Elasto-capillarity at the nanoscale: on the coupling between elasticity and surface energy in soft solids. *Soft Matter* 9:8494–503

White L. 2003. The contact angle on an elastic substrate. 1. The role of disjoining pressure in the surface mechanics. *J. Colloid Interface Sci.* 258:82–96

Winter H, Chambon F. 1986. Analysis of linear viscoelasticity of a cross-linking polymer at the gel point. *J. Rheol.* 30:367–82

Wong TS, Kang SH, Tang SKY, Smythe EJ, Hatton BD, et al. 2011. Bioinspired self-repairing slippery surfaces with pressure-stable omniphobicity. *Nature* 477:443–47

Wu H, Liu Z, Jagota A, Hui CY. 2018. Effect of large deformation and surface stiffening on the transmission of a line load on a neo-Hookean half space. *Soft Matter* 14:1847–55

Xu Q, Jensen K, Boltyanskiy R, Sarfat R, Style RW, Dufresne ER. 2017. Direct measurement of strain-dependent solid surface stress. *Nat. Commun.* 8:555

Xu Q, Style RW, Dufresne ER. 2018. Surface elastic constants of a soft solid. *Soft Matter* 14:916–20

Zhao M, Dervaux J, Narita T, Lequeux F, Limat L, Roché M. 2018a. Geometrical control of dissipation during the spreading of liquids on soft solids. *PNAS* 115:1748–53

Zhao M, Lequeux F, Narita T, Roché M, Limat L, Dervaux J. 2018b. Growth and relaxation of a ridge on a soft poroelastic substrate. *Soft Matter* 14:61–72