AC Electrokinetic micropumps

Olesen, Laurits Højgaard

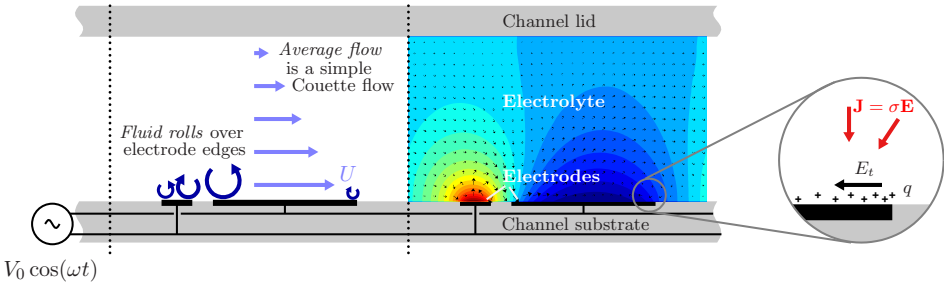
Publication date:
2006

Document Version
Publisher's PDF, also known as Version of record

PhD thesis, s971820

AC Electrokinetic micropumps

Laurits Højgaard Olesen



Supervisor: Henrik Bruus

MIC – Department of Micro and Nanotechnology
Technical University of Denmark

31 August 2006

Abstract

The goal of the present thesis has been to investigate ways of controlling and manipulating fluids and suspensions in microfluidic systems. In particular, we focus on a theoretical description of AC electroosmotic micropumps with asymmetric electrode arrays, that have recently been demonstrated to permit fast pumping (velocities $\sim$ mm/s) with low driving voltage of a few volt only.

The dynamical description of electrokinetics and electrochemical transport at driving voltages of just a few volt is a theoretically challenging subject, and therefore simplifying assumptions such as the Debye–Hückel approximation or linear response for weak applied field have often been employed in the literature.

We extend previous linear theory for AC electroosmotic flow into the "weakly nonlinear" regime by accounting for nonlinear capacitance of the Debye screening layer, and also consider the effect of Faradaic current injection from electrochemical electrode reactions. This allows us to explain why the frequency of maximum pumping is sometimes shifted down when the driving voltage is increased, but neither the linear nor weakly nonlinear models are able to account for the reversal of the pumping direction that has been observed experimentally.

Therefore we also study the "strongly nonlinear" regime where classical circuit models with uniform bulk electrolyte concentration break down. We extend recent theoretical work in this regime, by accounting for dynamics in the diffusion layer developing when an AC voltage with driving frequency around the inverse RC time is applied, and by considering fluid motion and convection of ions. Moreover, we attempt to include existing theory for double layers driven out of quasiequilibrium from problems of DC Faradaic conduction at "very large" (but experimentally relevant) voltage into our dynamical model.

We solve the coupled electrohydrodynamical problem numerically for experimental micropump geometries and display contributions to the net pumping velocity from the different flow sources in the model: Our results indicate that both bulk electroconvection and electroosmotic flow of the "second kind" may contribute a significant fraction of the overall induced flow already at driving voltages of a few volt. However, a rigorous account

for nonequilibrium double layers in a dynamical setting would be necessary to justify several *ad hoc* assumptions in our model.

Finally, we investigate different ways of breaking the symmetry of an electrode array and determine in each case the optimal device geometry to maximize the pumping velocity at a given low driving voltage as described by the simple linear theory.

Resumé

Formålet med nærværende afhandling har været at undersøge måder at kontrollere og manipulere væsker og partikler i mikrofluide systemer. Især fokuserer vi på en teoretisk beskrivelse af AC elektroosmotiske mikropumper med asymmetriske sæt af elektroder, der tillader hurtig pumpning (hastigheder $\sim$ mm/s) ved lav spænding på nogle få volt.

Dynamisk beskrivelse af elektrokinetik og elektrokemisk transport ved spændinger på bare få volt er teoretisk udfordrende, og derfor benyttes i litteraturen ofte forudsætninger såsom Debye–Hückel approksimationen eller lineært respons for svage påtrykte felter.

Vi udvider den eksisterende lineære teori for AC elektroosmose til at gælde i det "svagt ulineære" regime, idet vi tager højde for ulineær kapacitans i Debye skærmningslaget, og undersøger også effekten af faradaisk strøm fra elektrokemiske elektrode reaktioner. Dermed kan vi forklare hvorfor den frekvens hvor pumpehastigheden maksimeres i nogle tilfælde falder når spændingen øges, men hverken den lineære eller svagt ulineære teori kan forklare hvorfor pumperetningen vender, således som det ses eksperimentelt.

Derfor studerer vi også det "stærkt ulineære" regime hvor de klassiske kredsløbsmodeller bryder sammen, idet vi udvider tidligere teoretisk arbejde i dette regime ved at tage højde for dynamik i diffusionslaget der dannes når en vekselspænding med frekvens omkring den inverse RC tid påtrykkes, og ved at medregne væskestrømning og deraf følgende konvektion af ioner. Desuden forsøger vi at inkludere eksisterende teori for elektriske dobbeltlag drevet ud af kvasiligevægt fra problemer med DC faradaisk strøm ved "meget høj" spænding i vores dynamiske model.

Vi løser det koblede elektrohydrodynamiske problem numerisk for eksperimentelt relevante mikropumpe geometrier, og bestemmer bidrag til pumpehastigheden fra de forskellige strømningsled i modellen: Vores resultater indikerer at både elektrokonvektion og elektroosmose af den "anden slags" kan bidrage med en betydelig del af den samlede væskestrøm allerede ved spændinger på nogle få volt. Imidlertid er en mere grundig behandling af dobbeltlag drevet dynamisk ud af kvasiligevægt nødvendig for at retfærdiggøre flere *ad hoc* antagelser i vores model.

Endelig undersøger vi forskellige måder at bryde symmetrien af et sæt af elektroder, og bestemmer i hvert tilfælde den geometri der maksimerer

pumpehastigheden for en given lav spænding ved en beskrivelse med den simple lineære teori.

Preface

The present thesis is submitted in partial fulfilment of the requirements for the PhD degree at the Technical University of Denmark (DTU). The work has been carried out at MIC – Department of Micro and Nanotechnology in the Microfluidic Theory and Simulation Group (MIFTS) from September 2003 to August 2006. In the period January to May 2005 I worked at the Laboratoire de Physico-Chimie Théorique at Ecole Supérieure de Physique et de Chimie Industrielles (ESPCI) in Paris. The project was supervised by Prof. Henrik Bruus, and while abroad by Prof. Armand Ajdari.

I would like to thank Henrik Bruus for good spirits and dedicated effort as supervisor, and for allowing me great freedom to pursue my many wild ideas. I would also like to thank Armand Ajdari for welcoming me at ESPCI, and for providing me invaluable insight on electrokinetic phenomena; I am grateful for many thoughtful comments on my work also after I left Paris. Vincent Studer digged up old impedance measurements for me to validate my theoretical model. I also thank Patrick Tabeling and his group at ESPCI for welcoming me and for exposing me to some of their interesting work and problems.

Much of the work I did in the final part of my PhD studies depends on a theoretical model developed by Martin Bazant and Kevin Chu, with whom I had many helpful discussions. Also I would like to thank Nicolas Green for inviting me to see his lab in Southampton, and for sharing his long experience with AC electrokinetic experiments and recent observations.

My collaboration with Fridolin Okkels on topology optimization was an exciting sidetrack on the route I otherwise pursued; Misha Gregersen never lost her temper when I once and again came up with just another interesting measurement to do. I thank the people at MIC and in the MIFTS cave in particular for providing a very nice and inspiring working environment.

Finally, I thank my girlfriend Hanne for always being there and for reminding me that there is life beyond work.

Laurits Højgaard Olesen

MIC – Department of Micro and Nanotechnology
Technical University of Denmark
31 August 2006

Contents

List of Figures

List of Tables

List of Symbols

To reduce the notational overload we denote dimensionless quantities by the same symbol as their dimensional counterparts; it should be obvious from the context whether a given quantity is dimensionless or not.

Symbol	Description		Unit
c	(Molar) concentration		(mol) m^{-3}
D	Diffusivity		$\mathrm{m}^2\ \mathrm{s}^{-1}$
D_0	Average diffusivity		$\mathrm{m}^2\ \mathrm{s}^{-1}$
D_a	Ambipolar diffusion constant		$\mathrm{m}^2\ \mathrm{s}^{-1}$
ΔD	Diffusivity difference		$\mathrm{m}^2\ \mathrm{s}^{-1}$
e	Elementary charge	1.602×10^{-19}	C
$\mathbf{E}$	Electric field		$\mathrm{V}\ \mathrm{m}^{-1}$
f	Frequency		Hz $[\equiv \mathrm{s}^{-1}]$
f_a	Excess ambipolar flux		(mol) $\mathrm{s}^{-1}\ \mathrm{m}^{-2}$
$\mathbf{F}$	Flux density		(mol) $\mathrm{s}^{-1}\ \mathrm{m}^{-2}$
$\mathbf{F}_a$	Ambipolar flux		(mol) $\mathrm{s}^{-1}\ \mathrm{m}^{-2}$
j_0	Exchange current		$\mathrm{A}\ \mathrm{m}^{-2}$
j_{ext}	Faradaic current injection		$\mathrm{A}\ \mathrm{m}^{-2}$
$\mathbf{j}$	Surface excess current density		$\mathrm{A}\ \mathrm{m}^{-1}$
$\mathbf{J}$	Current density		$\mathrm{A}\ \mathrm{m}^{-2}$
k_B	Boltzmann constant	1.381×10^{-23}	$\mathrm{J}\ \mathrm{K}^{-1}$
ℓ_0	Geometrical length scale		m
$\mathbf{n}$	Surface outward normal		–
O	Oxidized species		–
p	Pressure		Pa $[\equiv \mathrm{N}\ \mathrm{m}^{-2}]$
q	Surface charge density		$\mathrm{C}\ \mathrm{m}^{-2}$
Q	Volume flow rate		$\mathrm{m}^3\ \mathrm{s}^{-1}$
$\mathbf{r}$	Position vector		m
R	Reduced species		–
R_{ct}	Charge transfer resistance		$\Omega\ \mathrm{m}^2$

Symbol	Description		Unit
R_{mt}	Warburg constant		$\Omega\ \mathrm{m}^2\ \mathrm{s}^{-1/2}$
t	Time		s
T	Period in time		s
T	Temperature		K
$\mathbf{u}$	Fluid velocity		$\mathrm{m\ s}^{-1}$
U	Pumping velocity		$\mathrm{m\ s}^{-1}$
V_0	Driving voltage		V
V_{ext}	External potential on electrode		V
w	Surface excess concentration		$(\mathrm{mol})\ \mathrm{m}^{-2}$
z	Ion valence		–
γ	Excess concentration		$(\mathrm{mol})\ \mathrm{m}^{-3}$
$\mathbf{\Gamma}_i$	Surface excess flux density		$(\mathrm{mol})\ \mathrm{s}^{-1}\ \mathrm{m}^{-1}$
δ	Surface capacitance ratio		–
ϵ	Permittivity		$\mathrm{F\ m}^{-1}$
ϵ_0	Permittivity of vacuum	8.854×10^{-12}	$\mathrm{F\ m}^{-1}$
ε	Dimensionless Debye length		–
ζ	Zeta potential		V
ζ^{eq}	Intrinsic zeta potential		V
ς	Diffusion layer potential drop		V
η	Dynamic viscosity		Pa s
ϑ	Overpotential		V
λ_D	Debye length		m
μ	Electrochemical potential		J
ν	Kinematic viscosity		$\mathrm{m}^2\ \mathrm{s}^{-1}$
ρ	Charge density		$\mathrm{C\ m}^{-3}$
ρ_m	Mass density		$\mathrm{kg\ m}^{-3}$
σ	Conductivity		$\mathrm{S\ m}^{-1}$
$\boldsymbol{\sigma}$	Stress tensor		Pa
τ_0	RC time		s
τ_D	Debye time		s
ϕ	Electrostatic potential		V
Φ	Space-charge layer potential drop		V
ψ	Excess potential		V
ω	Angular frequency		$\mathrm{rad\ s}^{-1}$
Ω	Computational domain		–

Accent	Description
a^*	Equilibrium bulk reference value
$\bar{a}$	Bulk variable in asymptotic expansion
$\tilde{a}$	Debye layer variable in asymptotic expansion
$\breve{a}$	Diffusion layer variable in asymptotic expansion
$\breve{a}$	Space-charge layer variable in asymptotic expansion
$\hat{a}$	Complex variable
$\hat{a}^*$	Complex conjugate

Chapter 1

Introduction

1.1 Microfluidics and the lab-on-a-chip

Microfluidics refers to hydrodynamics and flow manipulation in systems and devices with characteristic dimensions less than a millimeter. Of course, such systems have long been studied in biology and colloid science, but recently there has been an enormous increase of attention on the subject, for several reasons: First, the fabrication techniques originally developed by the microelectronics industry now allow accurate design of microfluidic circuits, and is being used to construct microelectromechanical systems (MEMS) and microfluidic devices, e.g., inkjet printer heads. Secondly, advances in biology and biotechnology require manipulation of particles or living cells and chemical detection of very small quantities. Also in more conventional analytical chemistry, miniaturization has the advantage of low sample and reagent consumption. The ultimate goal is development of "lab-on-a-chip" systems with integrated pumps, reagent dispensers, mixers, separators, and detection units, that would automatize the complete analysis from sample to electrical read-out [1, 2]. Fabrication from polymer materials would lower the production cost and allow for cheap and portable devices to perform simple analyses. Equally important, miniaturization opens for new types of analysis that would not be possible in a macrosystem, because mass diffusion and heat transfer is very rapid on the microscale and the relative importance of surface to volume forces increases as the system is scaled down [3, 4].

An important question when designing such microfluidic lab-on-a-chip systems is how to get the sample into the chip and manipulate it around. For a simple disposable device an elegant approach is the use of capillary forces, but for more complex systems it is necessary to use active pumping to control the flow. Depending on the application there can be very different requirements for such a pump: In a laboratory test setup it is often sufficient to use external syringe pumps, and multilayer soft lithography provides a very flexible platform for microfluidic plumbing with valves and peristaltic

pumps powered by external pressure transducers [5]. Devices for drug delivery need to be extremely fail-safe, and especially if designed for implantation they should have very low power consumption. On the other hand, for disposable devices the production costs and ease of integration should be much more important. An integrated pump design that has been subject to intensive research over the years is the (piezo-actuated) membrane pump, where a membrane is displaced to create pulsating flow that is rectified by valves at the inlet and outlet to the pumping chamber [6]. However, moving parts makes the fabrication delicate, and many valve designs are prone to failure if there are particles or bubbles in the liquid. Micropumps based on electroosmosis have no moving parts which makes them easy to fabricate, but their performance depends strongly on the pH and chemical composition of the pumping liquid. The basic design where an electric field is simply applied down a section of the microchannel can in principle pump the liquid in a plug flow with velocities of a few mm/s in free flow conditions, but only generate very small pressures so the thermodynamic (energy) efficiency is low [7]. This can be overcome by introducing a micro- or nanoporous medium, e.g., a sintered glass frit, into the channel which increases the hydraulic resistance dramatically but does not affect the electroosmotic flow; in this way both a decent flow rate and pressure can be generated at a thermodynamic efficiency of a few percent [7, 8, 9, 10, 11, 12]. Another problem is that in order to maintain an electric field down the pumping channel, a DC Faradaic current has to be drawn between electrodes submerged in the pumping liquid, which in many cases generates bubbles due to electrolysis. This can be avoided by separating the electrodes from the pumping channel by means of ion exchange membranes [11, 12] or a gel salt bridge [13] at the expense of increased fabrication complexity.

An interesting alternative to standard DC electroosmotic pumps is the use of AC electroosmosis with an asymmetric array of microelectrodes to generate pumping [14]. This requires driving the system with an AC voltage in the kHz range, but allows pumping velocities of the order of mm/s to be attained with just a few volt applied between the electrodes [15]. The fabrication is simple and only requires an array of interdigitated electrodes to be deposited on the channel substrate, which can be achieved by a single photolithography step followed by metal deposition and lift-off. Similar to the basic DC electroosmotic pump design, the AC electroosmotic micropump is very poor at generating pressure, but it allows local control of the flow and could be used, e.g., to circulate sample between two chambers on a chip. Moreover, the pumping velocity is inversely proportional to the size of the electrodes, i.e., the efficiency increases with miniaturization.

1.2 AC electrokinetics

In electrochemistry impedance spectroscopy, where a small harmonic modulation is added to a (slowly varying) DC voltage and the current response of the electrochemical cell is measured, is a standard and very powerful tool for determining reaction rate constants and also for probing the surface properties at the electrode [16]. Usually one stirs the solution to ensure uniform bulk reagent concentrations and simplify the experiment.

In colloid science the focus is most often on colloidal particles with a certain intrinsic surface charge and zeta potential, and one studies, e.g., the electrophoretic velocity of the particles in an applied DC field, or investigate their influence on the dielectric response when an AC field is applied [17, 18]. An AC field also gives rise to electrokinetic flow around the particles but no net particle motion (if the field is uniform and the particle is spherical). The flow does, however, induce interaction between particles and influence their tendency to agglomerate [19, 20]. Interestingly, Squires and Bazant predict that a nonspherical egg-shaped metal colloid will align with the applied field and "swim" due to induced-charge electroosmosis [21].

Ramos and co-workers first observed electroosmotic flow over pairs of microelectrodes when driven with an AC voltage in the kHz range, around the inverse RC time form the system [22, 23, 24]. Ajdari predicted that the same mechanism would give rise to pumping on an asymmetric electrode array [14], which was soon after demonstrated experimentally by Brown *et al.* [25] and later by several other groups [15, 26, 27, 28, 29, 30, 31, 32]. While the general symmetry arguments of Ajdari predicts pumping in *some* direction for *any* locally asymmetric geometry, Ramos *et al.* did a detailed theoretical study with realistic boundary conditions for the particular planar geometry used by Brown *et al.*, and obtained the same direction of pumping as observed experimentally and similar streamline patterns [33]. They were also able to match quantitatively the experimental pumping velocity and frequency of maximal pumping by assuming a compact Stern layer on the electrodes with roughly the same capacitance as the diffuse Debye layer. In some sense this good agreement is surprising because their theory is based on the Debye–Hückel approximation and linearization which is strictly valid only at low driving voltages $V_0 < k_BT/e \simeq 25$ mV, while experiments are typically performed with at least a few volt.

Later, Studer *et al.* observed reversal of the pumping direction when driving the system at around 50 kHz and above 5 V_{rms}, well above the frequency for maximal pumping in the "forward" direction around 1 kHz [15]. Applying the same voltage at 1 kHz, bubble formation and electrode degradation was observed because they used bare platinum electrodes that support electrochemical electrode reactions and Faradaic current. Gregersen found reversal of the pumping direction at 2.5 kHz and *below* 1.2 V_{rms}, changing to forward pumping above 1.2 V_{rms} [32]; this was found for a dilute

KCl working electrolyte on bare platinum electrodes with the same device geometry as used by Studer. Ramos and co-workers found reversal of the pumping direction on a travelling wave device at 1 kHz and above 2 V [34], and on an asymmetric electrode array at 10 kHz and above 5 V_{rms} [31]. They used titanium coated gold electrodes that do not support Faradaic current due to the formation of a thin natural titanium oxide layer. We note that while electrode reactions leading to bubble formation is a major challenge for DC electroosmosis, the same needs not be the case in AC because the reactions could be run reversibly [35]. However, our guess is that many biochemists would be uncomfortable with uncontrolled electrode reactions taking place in same the channel where a delicate sample is to be pumped through.

Most previous work on AC or induced-charge electroosmosis has been based on the Debye–Hückel approximation with linearization of the electrical problem [19, 23, 33, 36], which we denote "linear theory" even though the induced time-average electroosmotic flow scales as the square of the driving voltage. However, this theory is unable to account for the reversal of the pumping direction observed experimentally. Therefore it has been a major focus of our work to extend the existing theory to be able to describe these phenomena, and also to be able to compare the theory more quantitatively with experiments as a function of both driving voltage and frequency.

Bazant, Thornton, and Ajdari recently studied diffuse-charge dynamics in a simple 1D model system analytically by means of matched asymptotic expansions [37]. In particular they went beyond the linear theory to analyze the "weakly nonlinear" regime, where the capacitance of the Debye screening layer at the electrodes becomes nonlinear, but the bulk electrolyte concentration remains uniform and constant at leading order, with perturbations showing up only as higher order corrections. They also derived a dimensionless criterion for the zeta potential at which the weakly nonlinear asymptotics break down, and argued that for weakly nonlinear dynamics to hold, the driving voltage V_0 cannot greatly exceed the thermal voltage k_BT/ze. Further, they wrote the governing equations for the "strongly nonlinear" regime, where the assumption of uniform bulk electrolyte concentration breaks down but bulk charge neutrality remains valid; however, because that problem is not analytically tractable they did not attempt to solve it. Chu extended this work to three dimensions by studying a spherical metal particle, and investigated both the weakly nonlinear dynamical response to a suddenly applied DC field and the steady-state solution to the strongly nonlinear problem arising at larger applied DC field [38, 39].

We have taken this fruitful approach further by actually solving the strongly nonlinear dynamical model numerically in 2D, and by accounting also for electroosmotic fluid motion and convection. We agree that for weakly nonlinear dynamics to hold, the total zeta potential indeed cannot greatly exceed the thermal voltage, but because the induced zeta potential

is related to the driving voltage in a nonlinear way we also find that the latter can often be an order of magnitude larger than the thermal voltage at the point where the weakly nonlinear dynamics break down.

As the driving voltage is further increased, even the strongly nonlinear asymptotics break down due to complete salt depletion in the diffusion layers close to the electrodes. In steady-state problems of Faradaic conduction it is well known that beyond the "diffusion-limited" current, the electrical double layer changes from quasiequilibrium to a nonequilibrium structure with an extended space charge layer, as described first by Rubinstein and Shtilman [40]. Further, when tangential fields act upon this space charge layer, nonlinear electrokinetic flow is induced. Dukhin and coworkers introduced the term "electroosmosis of the second kind" to describe this phenomenon, and studied it extensively in the context of nonlinear electrophoresis of conductive particles made from ion exchanger material [41]. More recently, Rubinstein and Zaltzman theoretically predicted that second kind electroosmosis renders linearly instable the diffusion layer at a planar permselective (ion exchange) membrane and provides an efficient mixing mechanism, which can account for overlimiting conductance [42, 43].

Because even our strongly nonlinear model does not seem to be able to account for reversal of the pumping velocity, and because the driving voltages used experimentally are often large enough to make the strongly nonlinear model break down, we try to incorporate the basic features of the nonequilibrium double layer and electroosmosis of the second kind into our dynamical model. This is a challenging problem, and to be able to handle it numerically we make some fairly crude approximations. Our results indicate that electroosmosis of the second gives rise to net pumping in the same ("forward") direction as the usual electroosmosis of the "first kind", while bulk electroconvection contributes in the "reverse" direction.

1.3 Outline of the thesis

Chapter 2: Electrokinetic theory

We first give a brief overview of some of the classical electrokinetic theory, before we proceed to construct a mathematical model based on matched asymptotic expansions. The latter is a perturbation technique where one identifies the ratio of the Debye screening length to the characteristic geometrical length scale as a small parameter. Indeed the assumption of thin Debye layers underlies most classical analysis in electrochemistry, so the perturbation technique is often just a systematic way of obtaining already known results, but it also provides a firm starting point for going beyond the classical theory. In this we follow the work of Bazant, Thornton, and Ajdari [37] and Chu [38] on diffuse charge dynamics, but where they primarily focus on transient

solutions for a suddenly applied DC voltage, we study AC driving and steady-state periodic solutions.

Chapter 3: Linear and weakly nonlinear analysis

We study the pumping from an array of asymmetric pairs of microelectrodes, extending the existing linear theory for AC electroosmosis by accounting for Faradaic current injection both in a linearized scheme and using the full nonlinear Butler–Volmer reaction kinetics. We also include the nonlinear surface capacitance of the Debye layer as described by Gouy–Chapman–Stern theory. Most of this work was published in our paper *'AC electrokinetic micropumps: the effect of geometrical confinement, Faradaic current injection, and nonlinear surface capacitance'* [44], see Appendix B, but here we also take into account the effects of mass transfer on the Faradaic electrode reaction.

Chapter 4: Strongly nonlinear analysis

In this chapter we first analyse the strongly nonlinear dynamics for a simple 1D geometry with parallel-plate blocking electrodes, both to gain a better understanding of the dynamical solutions, but also to validate the asymptotic model by comparison to a full numerical solution of the Poisson–Nernst–Planck equations. Next we study a 2D model of the planar AC electroosmotic pump and determine the contributions to the net pumping from both electroosmotic flow of the first and second kind, and from electroconvection due to concentration gradients in the electroneutral diffusion layers and bulk electrolyte.

Chapter 5: Optimization

There are many ways of breaking the left-right symmetry of an electrode array to induce net pumping. In this chapter we consider both the "standard" pump with asymmetric pairs of planar microelectrodes, a design with asymmetric coating of the individual electrodes, and a 3D structured electrode design recently suggested by Bazant and Ben [45]. In all cases the performance increases with miniaturization, so we investigate the optimal electrode design for a fixed minimal linewidth or feature size in the geometry. This work is related to our work done on topology optimization earlier during the PhD studies, which has been published in the paper *'A high-level programming-language implementation of topology optimization applied to steady-state Navier–Stokes flow'* [46]. We have chosen not to present those results in this thesis but have included the paper in Appendix C.

Chapter 6: Conclusion and outlook

We present concluding remarks on our work on AC electrokinetic pumping and give some directions for future research.

Appendix A: FEMLAB code

While the standard linear model for induced-charge electroosmosis has been studied numerically by several authors, not many attempts have been made in the weakly nonlinear regime. Part of the reason may be the increased complexity involved with solving a time evolution problem with special variables defined on the boundary. However, using commercial finite element software like FEMLAB [47] this is actually straightforward to implement, and to demonstrate this we have included a code example in Appendix A. We also include the implementation of our strongly nonlinear model, and discuss how to deal with the nonlocal coupling in time introduced by our semi-analytical solution for the dynamics in the diffusion layer.

1.4 Publications during the PhD studies

Papers in peer reviewed journals

1. *Electro-hydrodynamics of binary electrolytes driven by modulated surface potentials*,
   N. A. Mortensen, L. H. Olesen, L. Belmon, and H. Bruus,
   Phys. Rev. E **71**, 056306 (2005).

2. *A high-level programming-language implementation of topology optimization applied to steady-state Navier–Stokes flow*,
   L. H. Olesen, F. Okkels, and H. Bruus,
   Int. J. Num. Meth. Eng. **65**, 975 (2006).

3. *Transport coefficients for electrolytes in arbitrarily shaped nano and micro-fluidic channels*,
   N. A. Mortensen, L. H. Olesen, and H. Bruus,
   New J. Phys. **8**, 37 (2006).

4. *AC electrokinetic micropumps: the effect of geometrical confinement, Faradaic current injection, and nonlinear surface capacitance*,
   L. H. Olesen, H. Bruus, and A. Ajdari,
   Phys. Rev. E **73**, 056313 (2006).

Conference proceedings

1. *Topology optimization of Navier–Stokes flow in microfluidics*,
   L. H. Olesen, F. Okkels, and H. Bruus,
   ECCOMAS 2004, Jyväskylä, Finland, July 2004, proc. vol. 2, p. 224.

2. *Applications of topology optimization in the design of micro- and nanofluidic devices*,
   F. Okkels, L. H. Olesen, and H. Bruus,
   Nanotech 2005, Anaheim, USA, May 2005, proc. vol. 1, p. 575–578.

3. *Applications of topology optimization in microfluidics: fluid flows, thermal transport, and chemical reactions*,
   F. Okkels, L. H. Olesen, and H. Bruus,
   6th EPS Liquid Matter Conf., Utrecth, the Netherlands, July 2005.

4. *AC electrokinetic micropumps: the effect of geometrical confinement, Faradaic currents, and nonlinear surface capacitance*,
   L. H. Olesen, H. Bruus, and A. Ajdari
   MicroTAS 2005, Boston, USA, October 2005, proc. vol. 1, p. 560–562.

5. *Nonlinear studies of AC electrokinetic micropumps*,
   H. Bruus, L. H. Olesen, and A. Ajdari
   APS MAR06 Meeting, Baltimore, USA, March 2006.

6. *Mass and charge transport in micro and nano-fluidic channels*,
   N. A. Mortensen, L. H. Olesen, and H. Bruus,
   2nd Int. Conf. Transport Phenomena in Micro and Nanodevices,
   Barga, Italy, June 2006.

7. *Universality in microfluidic phenomena inside microchannels with arbitrarily shaped cross-sections*,
   H. Bruus, N. Mortensen, F. Okkels, and L. H. Olesen,
   Euromech Fluid Mechanics Conf. 6, Stockholm, Sweden, June 2006.

Chapter 2

Electrokinetic theory

In this chapter we start by giving a brief overview of the classical theory of electrochemical transport and electrode reaction processes. This is a widely studied subject due to its importance both for industrial applications such as batteries and fuel cells, electroplating, and desalination by electrodialysis, and in natural processes such as corrosion and ion transport through ion channels in biological cells. A common feature underlying most of the classical analysis is the assumption that the system can be divided into an electroneutral bulk region and charged boundary layers that are very thin compared to the characteristic geometrical length scale in the system. This simple picture can be formalized by perturbation theory, and in Sec. 2.2 we go through a systematic derivation of a mathematical model based on matched asymptotic expansions.

2.1 Classical theory

A fundamental approximation in the classical theory of electrochemical transport is the continuum approximation where the different chemical species are represented by their concentration fields. For a concentrated solution of 1 M the mean separation between ions is around 1 nm, whereas for a dilute 1 mM solution the mean separation is around 10 nm. Therefore, in most micro- and macroscopic applications the continuum approximation is well justified, but in true nanofluidics it may not be. However, even in a macroscopic electrochemical experiment, the system spontaneously develops microscopic structure in the form of screening layers at electrodes or any other charged surfaces and interfaces in the electrolyte. The thickness of the screening layer is given by the Debye length

$$\lambda_D = \sqrt{\frac{\epsilon k_B T}{\sum_i c_i^* z_i^2 e^2}}, \tag{2.1}$$

where ϵ is the electrical permittivity of the solvent, k_B is Boltzmanns constant, T the absolute temperature, e the elementary charge, c_i^* is the bulk concentration (number density) of the ith ionic species, and z_i the valence number. For a typical system λ_D is of the order of 1–10 nm, that is, on average there are no more than a few ions across it. This means that the instantaneous "concentration profile" over a particular point on the surface will be highly nonuniform and rapidly fluctuating, but upon proper averaging in time or in the transverse direction to the screening layer one should still be able to use continuum theory on the averaged fields.

2.1.1 Electrochemical transport

Within a continuum model the transport of the individual chemical species is governed by conservation laws of the form

$$\partial_t c_i = -\boldsymbol{\nabla} \cdot \mathbf{F}_i, \tag{2.2}$$

where $\mathbf{F}_i$ is the flux of the ith species and we disregard any bulk chemical reactions producing or consuming that species. In general the flux is a function of the electrochemical potential μ_j of all the species present in solution [53, 54]

$$\mathbf{F}_i = -c_i\Big(\sum_j \mathcal{L}_{ij}\boldsymbol{\nabla}\mu_j\Big) + c_i\mathbf{u}, \tag{2.3}$$

where $\mathcal{L}_{ij}$ are generalized mobilities that account for interactions between different species, and $c_i\mathbf{u}$ represents transport by convection with the fluid velocity $\mathbf{u}$.

Nernst–Planck equation

For an ideal dilute system the interaction between different species can be neglected and the electrochemical potential can be decomposed into separate entropic and electrical terms as

$$\mu_i = k_B T \log c_i + z_i e\,\phi \tag{2.4}$$

where is ϕ the mean-field electrostatic potential. Then Eq. (2.3) for the flux reduces to the Nernst–Planck equation

$$\mathbf{F}_i = -D_i\Big(\boldsymbol{\nabla}c_i + \frac{z_i e}{k_B T}c_i\boldsymbol{\nabla}\phi\Big) + c_i\mathbf{u}, \tag{2.5}$$

where we used the Einstein relation $D_i = k_B T\mathcal{L}_{ii}$ to express the mobility in terms of the diffusivity D_i. The electrostatic potential ϕ is determined by the charge distribution ρ through the Poisson equation

$$-\epsilon\boldsymbol{\nabla}^2\phi = \rho = e\sum_i z_i c_i, \tag{2.6}$$

assuming constant permittivity ϵ of the solvent.

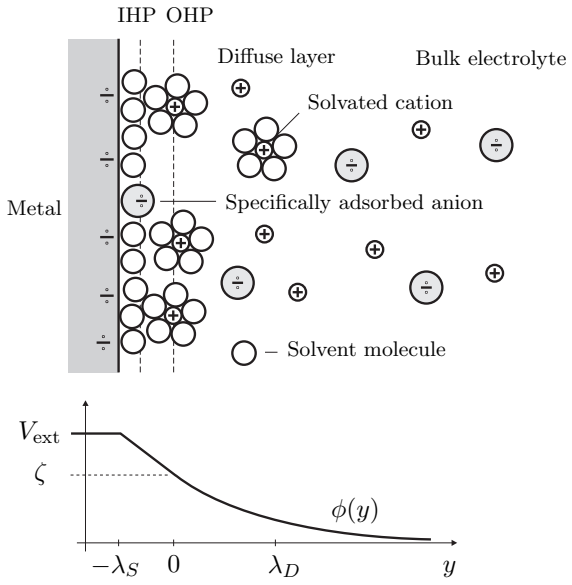


Figure 2.1: Schematic picture of the diffuse screening layer in the electrolyte next to a negatively biased metal electrode. Cations are attracted to the electrode and anions are repelled so the interface appears charge neutral when seen from the bulk. The outer Helmholtz plane (OHP) defines the closest distance that the solvated cations can move towards the electrode. Also shown is $\phi(y)$, the mean-field potential variation from the electrode to the charge neutral bulk.

2.1.2 Electrical double layer

The common continuum model view of the screening layer is based on division into a compact and a diffuse part. This is shown schematically in Fig. 2.1 for the screening layer next to a negatively biased metal electrode. Cations in the electrolyte move towards the electrode to screen the negative charge, and anions are repelled. The finite size of the ions can be taken into account by defining a smallest distance that the solvated cations can move towards the electrode, and this point of closest approach is called the outer Helmholtz plane (OHP). The region between the OHP and the electrode is called the "compact layer", and in Fig. 2.1 it is roughly two solvent molecules thick because the approach of the cations is obstructed both by their own solvation shell and by a monolayer of solvent molecules held at the surface by van der Waals forces. The ions in the screening layer outside the OHP are not static but undergo thermal motion, and this region is therefore called the "diffuse layer". Some ions in the electrolyte may interact more strongly with the electrode material and adsorb directly on the electrode, and the locus of their electrical centers is denoted the inner Helmholtz plane (IHP). Together the compact and diffuse layers are called the "double layer", and the model is known as the Gouy–Chapman–Stern model [16].

In order to avoid confusion with the diffusion layer of width $\ell_i(\omega) \sim$

$\sqrt{D_i/\omega}$, to and from which the ions have time to diffuse during one cycle of the driving voltage at angular frequency ω, we prefer to denote the diffuse part of the double layer as the "Debye layer". We also tend to denote the compact part as the "Stern layer", and mention that this can also be used to model an oxide layer or insulating coating on the electrode.

Gouy–Chapman–Stern model

In thermal equilibrium the electrochemical potential is constant thoughout the system and the ions in the Debye layer are Boltzmann distributed, i.e.,

$$c_i = c_i^* \exp\biggl[-\frac{z_i e(\phi-\phi^*)}{k_B T}\biggr], \tag{2.7}$$

where c_i^* is the reference concentration attained in the charge neutral bulk electrolyte where $\phi = \phi^*$. Substituting into Eq. (2.6) we get the Poisson–Boltzmann equation

$$-\boldsymbol{\nabla}^2\psi = \frac{e}{\epsilon}\sum_i c_i^* z_i \exp\biggl(-\frac{z_i e\psi}{k_B T}\biggr), \tag{2.8}$$

where we introduced $\psi = \phi - \phi^*$, and for a symmetric binary $z{:}z$ electrolyte this reduces to the more familiar form

$$\boldsymbol{\nabla}^2\psi = \frac{1}{\lambda_D^2}\frac{k_B T}{ze}\sinh\biggl(\frac{ze\psi}{k_B T}\biggr). \tag{2.9}$$

The Poisson–Boltzmann equation constitutes a single equation to solve for the equilibrium electrostatic potential; the ion distributions then follow directly from Eq. (2.7). For one-dimensional systems it has an analytical solution, namely the Gouy–Chapman solution[1]

$$\psi(y) = \frac{4k_B T}{ze}\tanh^{-1}\biggl[\tanh\biggl(\frac{ze\zeta}{4k_B T}\biggr)e^{-y/\lambda_D}\biggr], \tag{2.11}$$

where y is the distance perpendicular to the surface and ζ is the potential attained at the outer Helmholtz plane where $y = 0$, cf. Fig. 2.1. The total amount of charge accumulated in the Debye layer per unit area is

$$q = \int_0^\infty \rho\,\mathrm{d}y = \epsilon\partial_y\psi\big|_{y=0} \tag{2.12}$$

$$= -\mathrm{sgn}(\zeta)\sqrt{2\epsilon k_B T\textstyle\sum_i c_i^*(e^{-z_i e\zeta/k_B T}-1)} \tag{2.13}$$

$$= -\frac{\epsilon}{\lambda_D}\frac{2k_B T}{ze}\sinh\biggl(\frac{ze\zeta}{2k_B T}\biggr), \tag{2.14}$$

[1] Gouy also showed that Eq. (2.8) can be integrated for an asymmetric binary electrolyte with $z_+ = -2z_- = 2z$, where

$$\psi(y) = \frac{k_B T}{ze}\log\biggl[1 - \frac{3}{1+\cosh\bigl(\,y/\lambda_D + 2\tanh^{-1}\sqrt{(1+2e^{ze\zeta/k_B T})/3}\,\bigr)}\biggr]. \tag{2.10}$$

where Eq. (2.13) reduces to Eq. (2.14) for a symmetric z:z electrolyte only. Neglecting specifically adsorbed ions, the region $-\lambda_S \leq y \leq 0$ behind the outer Helmholtz plane is free of charge and the potential varies linearly, i.e., the Stern layer ideally behaves as a capacitor of capacitance $C_S = \epsilon_S/\lambda_S$. The overall potential drop across the double layer is then

$$V_{\text{ext}} - \phi^* = \zeta - \frac{q}{C_S}, \tag{2.15}$$

and the differential capacitance per unit area becomes

$$C_{\text{dl}} = -\frac{\mathrm{d}q}{\mathrm{d}(V_{\text{ext}} - \phi^*)} = \left[\frac{1}{C_S} + \frac{\lambda_D}{\epsilon \cosh(ze\zeta/2k_BT)}\right]^{-1}. \tag{2.16}$$

In the Debye–Hückel limit $\zeta \ll k_BT/e$ where the potential variations are small compared to the thermal voltage, the double layer capacitance reduces to

$$C_{\text{dl}} = \left[\frac{1}{C_S} + \frac{\lambda_D}{\epsilon}\right]^{-1} = \frac{\epsilon}{\lambda_D(1+\delta)}, \tag{2.17}$$

where we introduced the parameter $\delta = \epsilon\lambda_S/\epsilon_S\lambda_D$ for the capacitance ratio between the Debye and Stern layers. At larger potential the Debye layer capacitance becomes very large and the overall double layer capacitance is dominated by the Stern layer.

The relevant parameter range for δ is large: If we compare the Debye layer capacitance for a dilute 1 mM aqueous electrolyte with $\lambda_D = 10$ nm and $\epsilon = 80\epsilon_0$ to the molecular Stern layer defined by the distance closest approach $\lambda_S \sim 1$ Å of an ion to a bare metal electrode then we get a very small capacitance ratio $\delta \sim 0.01$, whereas for a titanium electrode with a natural TiO_2 oxide of thickness $\lambda_S \sim 4$ nm and $\epsilon_S = 110\epsilon_0$ we find $\delta \sim 0.3$, and for even thin polymer coatings we easily obtain $\delta \gg 1$. In real electrochemical systems the Stern layer on bare metal electrodes often does not behave exactly as an ideal capacitor but shows some voltage dependence in C_S, and the specific adsorbtion and desorption of ions give rise to an additional component in Eq. (2.16). However, because we are not concerned with any particular material system but wish to study electrokinetics in more general, we stick to the simplest form of the Gouy–Chapman–Stern model and always assume a constant Stern layer capacitance and neglect specific adsorption.

Modified Poisson–Boltzmann equation

The Gouy–Chapman–Stern model is widely used because it is simple and yet captures the generic features observed in typical electrochemical experiments. The exponential increase in ion concentration with potential implied in Eq. (2.7) effectively limits the range of zeta potentials that makes sense

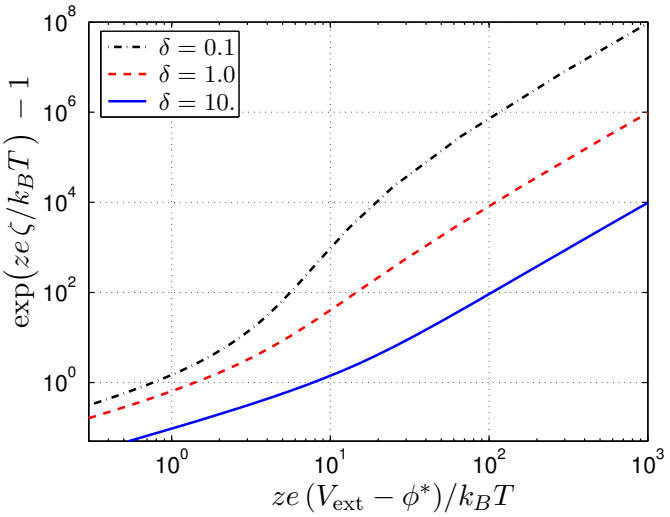


Figure 2.2: Double logarithmic plot of the relative increase in anion concentration at the outer Helmholtz plane as a function of the overall potential drop $V_{\mathrm{ext}} - \phi^*$ across the double layer for different capacitance ratios δ in the Gouy–Chapman–Stern model.

physically: Substituting $\zeta = 500$ mV for a bulk dilute 1 mM KCl solution, the concentration predicted at the outer Helmholtz plane exceeds 10^{32} ions per m^3 or 100 ions per Å^3 which is ridiculously large. However, the "purpose" of the Stern layer is exactly to supress this unphysical behaviour. At large voltage when the double layer capacitance is dominated by the Stern layer, the accumulated charge depends linearly on the applied voltage $q \simeq -C_S(V_{\mathrm{ext}} - \phi^*)$, whereas the zeta potential grows only logarithmically with q, cf. Eq. (2.13). Therefore the increase in ion concentrations in the Debye layer also becomes less dramatic, as shown in Fig. 2.2 where the relative increase in the anion concentration at the outer Helmholtz plane is plotted as a function $V_{\mathrm{ext}} - \phi^*$. The figure shows that the ion concentration grows as $(V_{\mathrm{ext}} - \phi^*)^2/\delta^2$ at large voltage, but also that this large voltage regime is not reached until $V_{\mathrm{ext}} - \phi^*$ exceeds ≈ 1 V. If we consider again a dilute 1 mM KCl solution then the bulk concentration is 6×10^{23} ions per m^3. The ions have radii of a few ångstrøm, with 1.33 (2.32) Å for a (hydrated) K^+ ion and 2.32 Å for Cl^-, so the steric limit of a close packed phase of ions corresponds to around 10^{28} ions per m^3. The voltage at which this limit is reached depends on the Stern layer capacitance: Assuming a low capacitance ratio of $\delta = 0.1$ it occurs already at 500 mV while for $\delta = 10$ the limit is reached around 10 V.

One problem with this reasoning, where almost the entire potential drop across the double layer is placed on the Stern layer, is that most insulators suffer dielectric breakdown around 1–100 MV/m, which is far exceeded if we put 1 V across a 1 Å molecular Stern layer, or even a 5 nm titanium oxide layer.

Further, the neglect of correlations between different ions inherent in the use of Eq. (2.4) for the electrochemical potential obviously becomes questionable long before the limit of steric exclusion sets in, but the violation of the steric limit is perhaps the most offending feature in the Gouy–Chapman–Stern model. Borukhov *et al.* have suggested a simple modification of the Poisson–Boltzmann equation to account for the finite size of ions [55]. They include the entropy of the solvent molecules in the expression for the free energy F, and obtain for a symmetric binary z:z electrolyte with ions of characteristic size a the following expression for the electrochemical potential

$$\mu_\pm = \frac{\delta F}{\delta c_\pm} = k_B T \log(c_\pm a^3) + z_\pm e\phi - k_B T \log(1 - c_+ a^3 - c_- a^3), \quad (2.18)$$

where a^3 is the volume occupied by a single ion. The ion concentrations $c_\pm$ cannot then exceed the steric limit $1/a^3$, essentially because the entropy associated with excluding the last solvent molecule from a condensed phase of ions is infinite. Their modified Poisson–Boltzmann equation becomes

$$\boldsymbol{\nabla}^2\psi = \frac{1}{\lambda_D^2}\frac{k_B T}{ze}\frac{\sinh\left(\frac{ze\psi}{k_B T}\right)}{1 - 2c^* a^3 + 2c^* a^3 \cosh\left(\frac{ze\psi}{k_B T}\right)}\,. \quad (2.19)$$

Kilic, Bazant, and Ajdari have recently extended this approach to the non-equilibrium case by deriving a modified Nernst–Planck equation based on Eq. (2.18) [56, 57]. This is very a interesting direction which may allow to explain, e.g., the vanishing of the AC electroosmotic pumping at electrolyte concentrations above 10 mM [15]. However, we also like stress that the ability of the double layer to absorb ions and act as a capacitor persist long after the steric limit is reached. In this regime the Gouy–Chapman–Stern model systematically underestimates the zeta potential, but may still give a good approximation to the overall capacitance provided the capacitance ratio δ is not too low and the nominal bulk concentration is not too high.

2.1.3 Electrode reactions

Electrode reactions are chemical redox reactions that involve transfer of one or more electrons from an electrode to the reactants. Familiar examples are

$$\mathrm{Zn^{2+}_{(aq)}} + 2\,\mathrm{e^-} \rightleftharpoons \mathrm{Zn_{(s)}} \qquad E^\ominus = -0.762\ \mathrm{V}, \quad (2.20)$$

$$2\,\mathrm{H^+_{(aq)}} + 2\,\mathrm{e^-} \rightleftharpoons \mathrm{H_{2(g)}} \qquad E^\ominus = 0.000\ \mathrm{V}, \quad (2.21)$$

$$\mathrm{O_{2(aq)}} + 2\,\mathrm{H_2O_{(\ell)}} + 4\,\mathrm{e^-} \rightleftharpoons 4\,\mathrm{OH^-_{(aq)}} \qquad E^\ominus = 0.444\ \mathrm{V}, \quad (2.22)$$

where $E^\ominus$ is the standard reduction potential. The reaction in Eq. (2.20) involves transfer of two electrons from the electrode to the $\mathrm{Zn^{2+}}$ ion. Such

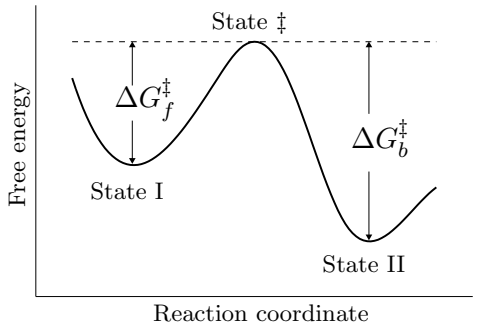


Figure 2.3: Free energy landscape along reaction path from state I (reactant) to state II (product). The activated complex (or transition state) ‡ is the configuration of maximal free energy.

a tunneling process is very unlikely, and Eq. (2.20) must therefore represent a multistep reaction where the Zn^{2+} is first reduced to Zn^{+} and then Zn. Likewise, the reaction in Eq. (2.21) must also be a multistep reaction consisting of one step where H^{+} is reduced to form a single hydrogen atom adsorbed on the electrode, followed by a second step where two hydrogen atoms join and desorb from the electrode as H_2. Often one step in such a multistep reaction is much slower than the rest. To determine the reaction kinetics it is then sufficient to focus on this rate-limiting step whereas the other steps are considered as reaction equilibria. In the present study we therefore assume a generic one-step, single electron process of the form

$$\mathrm{O} + \mathrm{e}^{-} \rightleftharpoons \mathrm{R}, \tag{2.23}$$

where $z_{\mathrm{O}} = z_{\mathrm{R}} + 1$ to ensure charge conservation.

Nernst equation

In thermal equilibrium the change in electrochemical Gibbs free energy associated with the reaction is zero

$$\begin{aligned}\Delta G = &\left[G_{\mathrm{R}}^{\ominus} + k_B T \log c_{\mathrm{R}}^{*} + z_{\mathrm{R}} e\phi^{*}\right] \\ &- \left[G_{\mathrm{O}}^{\ominus} + k_B T \log c_{\mathrm{O}}^{*} + z_{\mathrm{O}} e\phi^{*} - eV_{\mathrm{ext}}\right] = 0,\end{aligned} \tag{2.24}$$

from which we obtain the Nernst equation for the equilibrium potential drop from the electrode to the electrolyte

$$(V_{\mathrm{ext}} - \phi^{*})^{\mathrm{eq}} = E^{\ominus} + \frac{k_B T}{e} \log \frac{c_{\mathrm{O}}^{*}}{c_{\mathrm{R}}^{*}}. \tag{2.25}$$

Here $\Delta G^{\ominus} = G_{\mathrm{R}}^{\ominus} - G_{\mathrm{O}}^{\ominus} = -eE^{\ominus}$ is the increase in standard free energy associated with the reaction.

Butler–Volmer kinetics

Out of equilibrium the forward and backward reaction rates no longer balance and a net reaction takes place. According to transition state theory the rate limiting factor is associated with getting across a bump in the free energy landscape, cf. Fig. 2.3. The state ‡ of maximal free energy along the reaction path in configuration space is denoted the transition state or activated complex. The forward reaction rate ν_f is proportional to the probability of getting from the reactant state into the transition state, and this probability drops exponentially with the difference in electrochemical free energy between the two states, i.e.,

$$\nu_f \propto \exp\left(-\frac{\Delta G_f^\ddagger}{k_B T}\right), \tag{2.26}$$

and similarly for the backward reaction ν_b

$$\nu_b \propto \exp\left(-\frac{\Delta G_b^\ddagger}{k_B T}\right). \tag{2.27}$$

Further it is assumed that the electrochemical free energy at the transition state can be decomposed into a chemical term that is essentially independent of the applied potential, and an electrical term $(G^\ddagger)_e$ that is a weighted average of the electrostatic energy of the reactant and product, i.e.,

$$(G^\ddagger)_e = (1-\alpha)(z_{\mathrm{O}} e\phi^s - eV_{\mathrm{ext}}) + \alpha z_{\mathrm{R}} e\phi^s. \tag{2.28}$$

Here the transfer coefficient α is simply a measure of how close the transition state is to either the reactant or product state in terms of electrostatic energy, and ϕ^s is the potential in the electrolyte at the electrode surface (or outer Helmholtz plane). Substituting into Eqs. (2.26) and (2.27) the net reaction rate becomes

$$\nu_{\mathrm{net}} = \nu_f - \nu_b = k_f^\ominus c_{\mathrm{O}}^s e^{-\alpha e(V_{\mathrm{ext}}-\phi^s)/k_B T} - k_b^\ominus c_{\mathrm{R}}^s e^{(1-\alpha)e(V_{\mathrm{ext}}-\phi^s)/k_B T}, \tag{2.29}$$

where $k_f^\ominus$ and $k_b^\ominus$ are the forward and backward standard rate constants, and c_O^s and c_R^s are the concentrations of the oxidized and reduced species at the surface. This result and its implications is known as the Butler–Volmer formulation of reaction kinetics [16]. The Faradaic current injected from the electrode into the electrolyte by the reaction is

$$j_{\mathrm{ext}} = -e\nu_{\mathrm{net}}. \tag{2.30}$$

If we assume quasiequilibrium in the double layer we can relate the concentrations at the electrode surface to those immediately outside the Debye

layer by a Boltzmann factor, $c_i^s = c_i\, e^{-z_i e\zeta/k_BT}$. Then Eqs. (2.29) and (2.30) can be re-expressed as the current-overpotential equation

$$j_{\text{ext}} = j_0 e^{-(z_{\text{O}}-\alpha)e\zeta/k_BT}\left(\frac{c_{\text{R}}}{c_{\text{R}}^*}e^{(1-\alpha)e\vartheta/k_BT} - \frac{c_{\text{O}}}{c_{\text{O}}^*}e^{-\alpha e\vartheta/k_BT}\right), \tag{2.31}$$

where $j_0 = ek^{\ominus}(c_{\text{R}}^*)^{\alpha}(c_{\text{O}}^*)^{1-\alpha}$ is the exchange current, $k^{\ominus}$ is the standard rate constant, c_{O} and c_{R} are the concentrations immediately outside the Debye layer, c_{O}^* and c_{R}^* are the corresponding bulk values at equilibrium, and the overpotential $\vartheta = (V_{\text{ext}} - \phi) - (V_{\text{ext}} - \phi)^{\text{eq}}$ is the difference between the actual potential drop across the double layer and the thermal equilibrium for the given redox process and bulk concentration. Linearizing for small overpotentials we get

$$j_{\text{ext}} = \frac{j_0 e}{k_BT}\exp\left[-\frac{(z_{\text{O}}-\alpha)e\zeta^{\text{eq}}}{k_BT}\right]\vartheta = \frac{1}{R_{\text{ct}}}\vartheta, \tag{2.32}$$

defining the charge transfer resistance R_{ct}.

Reversibility of the electrode reactions

In our model calculations we always assume that the electrode reaction is reversible. This seems natural because our primary focus is on a micropump based on an array of electrode pairs that is driven with an AC voltage: We model this as an infinite periodic system in space and look for steady-state periodic solutions in time after all transients have died out. If the device is pumping the electrolyte in a closed loop and for a long time then this should be a good approximation. However, if the device is pumping from an inlet with "fresh" electrolyte to an outlet that is sent to "waste", then the chemical composition in the electrolyte might change upon passage through the pump, e.g., by a change in pH, by absorbtion of a small amount of electrolytic gasses, or by gradually dissolving the electrode material and flushing it out to waste. While it may or may not be feasible or desirable to operate a pump in this way we do not account for it in our model.

2.1.4 Electroosmosis

The fluid motion in the electrolyte is governed by the Navier–Stokes equation

$$\rho_m[\partial_t\mathbf{u} + (\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{u}] = -\boldsymbol{\nabla}p + \eta\boldsymbol{\nabla}^2\mathbf{u} + \rho\mathbf{E}, \tag{2.33}$$

where $\mathbf{u}$ is the fluid velocity, p is the pressure, ρ_m is the mass density, η the dynamic viscosity, and $\rho\mathbf{E}$ is the electrical body force acting the fluid. In the bulk electrolyte the charge density ρ is small, but when a tangential field acts on the charge in the Debye layers it can induce a flow which in turn drags the bulk fluid along. This phenomenon is called electroosmosis.

If we consider a uniformly charged flat surface at $y = 0$ in contact with an electrolyte and with a tangential field E_x acting on the charge in the screening layer then the tangenial component of the Navier–Stokes equation reduces to a balance between electrical and viscous forces

$$0 = \eta \partial_y^2 u_x - \epsilon \partial_y^2 \psi \, E_x. \tag{2.34}$$

Integrating twice and using that $u_x = 0$ and $\psi = \zeta$ at the surface while $\partial_y u_x = 0$ and $\partial_y \psi = 0$ in the bulk we obtain

$$u_x = -\frac{\epsilon(\zeta - \psi)}{\eta} E_x. \tag{2.35}$$

The fluid velocity increases from no-slip at the surface to a constant value

$$u_x = -\frac{\epsilon \zeta}{\eta} E_x \tag{2.36}$$

outside the Debye layer, which effectively acts as a slip velocity for the bulk fluid motion. This classical result is known as the Helmholtz–Smoluchowski formula [17, 53].

In deriving Eq. (2.36) we associated the zeta potential with the no-slip plane, whereas in Sec. 2.1.2 we defined ζ as the potential at the outer Helmholtz plane. Apriori there is no guarantee that the no-slip and the outer Helmholtz planes should coincide, but experiments for some particularly smooth, impermeable, and well-defined surfaces indicate that they do [17, 58]. On the other hand, for rough, porous, or inhomogeneous surfaces we cannot expect this, e.g., if the charged surface has a thin adsorption layer of macromolecules or polymer gel then the electroosmotic flow would be greatly suppressed due to the large hydrodynamic resistance of the polymer fragments [58].

2.2 Mathematical model

The solution of the full Poisson–Nernst–Planck equations is a challenging task; even numerically the solution is made difficult because the intrinsic length scale λ_D in the electrolyte is often several orders of magnitude smaller than the relevant geometrical length scale ℓ_0. Yet our physical intuition is that the bulk electrolyte is simply charge neutral, and that only the screening layers at the boundary display structure on the Debye length scale. This simple picture can be formalized by perturbation theory, rewriting the full Poisson–Nernst–Planck problem as a perturbation expansion in the small parameter

$$\varepsilon = \frac{\lambda_D}{\ell_0}. \tag{2.37}$$

The Poisson equation turns out to be singularly perturbed in the sense that the small parameter ε multiplies the highest (second) order derivative, and mathematically this singular perturbation can be seen as the origin of the disparate length scales appearing in the solution.

The method of matched asymptotic expansions is a generalization of Prandtl's viscous boundary layer theory, and it is a method for treating singular perturbations [59]. Essentially the method works by computing separate "outer" and "inner" approximations to the solution in the bulk and boundary regions, respectively, where the inner approximation in the boundary layer is obtained by scaling the spatial variables to remove the singular perturbation before solving the problem. The two solutions are matched by requiring that the outer limit of the inner approximation $\tilde{c}_\pm$ should be asymptotically equal to the inner limit of the outer approximation $\bar{c}_\pm$, i.e.,

$$\lim_{\tilde{y}\to\infty} \tilde{c}_\pm(x,\tilde{y},t) = \lim_{y\to 0} \bar{c}_\pm(x,y,t), \tag{2.38}$$

where $\tilde{y} = y/\varepsilon$ is the rescaled spatial variable normal to the boundary in the inner region. The standard procedure begins by seeking regular expansions in the form of power series

$$\bar{c}_\pm = \bar{c}_\pm^{(0)} + \varepsilon\,\bar{c}_\pm^{(1)} + \varepsilon^2 \bar{c}_\pm^{(2)} + \ldots, \tag{2.39}$$

substituting into the governing equations, and collecting like powers of ε. The primary interest is on the leading order term $\bar{c}_\pm^{(0)}$, but the higher order corrections $\bar{c}_\pm^{(i)}$ can also contain important information. The expansions need not always take the form Eq. (2.39) but can also contain fractional powers of ε, and logarithmic terms like $\varepsilon \log \varepsilon$ sometimes "pop up" in the process of deriving higher order terms [59]. The approximations to the solution are "asymptotic" in the sense that they converge to the true solution in the limit $\varepsilon \to 0$ with all other parameters in the problem held fixed. However, for any fixed $\varepsilon > 0$ there could be ε dependent restrictions on, e.g., the driving voltage V_0 for a truncation of the series to produce accurate results [37], and strictly there is no guarantee that the series will converge at all, even if an infinite number of terms are included [59].

We denote "weakly nonlinear" the regime where the leading order term in the standard asymptotic expansion remains a good approximation to the solution. However, experimentally there are several interesting and intriguing phenomena that show up at larger voltage, so we would like to extend the validity of the model as far up in driving voltage as possible. It turns out that for $V_0 \sim k_B T/ze\sqrt{\varepsilon}$ the weakly nonlinear solution breaks down, but by reordering the terms in the expansion to reflect the dominant balance we can develop another "strongly nonlinear" model valid in the joint limit $\varepsilon \to 0$ and $V_0 \to \infty$.

Because we focus on the periodic solution for an AC driving voltage after all transients have died out, there are actually three different charac-

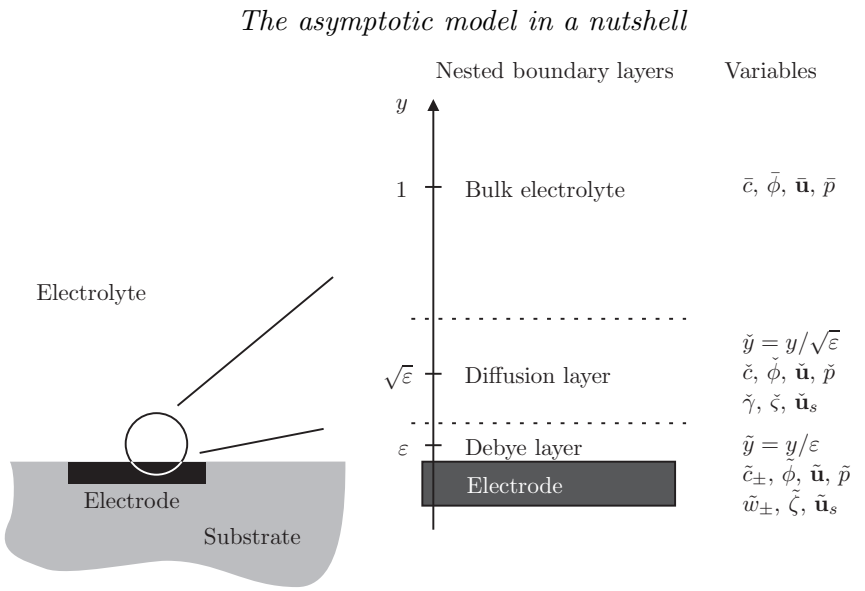


Figure 2.4: Schematic picture of the nested boundary layers in the matched asymptotic expansion: The "outer" bulk region is connected via the "middle" diffusion layer to the "inner" Debye layer. Variables associated with the different regions are denoted by bar, check, and tilde accents, respectively. The bulk is electroneutral with a salt concentration $\bar{c}$ that is constant in time but not necessarily in space. The diffusion layer is also electroneutral but has a dynamically varying salt concentration; the excess salt concentration $\check{\gamma} = \check{c} - \bar{c}$ is determined by a simple 1D diffusion problem. The excess potential drop $\check{\zeta}$ across the diffusion layer induces an effective slip velocity $\check{\mathbf{u}}_s$. The Debye layer is in quasiequilibrium at the RC time scale with a structure that is completely determined by the excess amount of each ion $\tilde{w}_\pm = \int_0^\infty (\tilde{c}_\pm - \check{c})\,\mathrm{d}\tilde{y}$ accumulated in the layer together with the salt concentration $\check{c}$ seen from the Debye layer. The potential drop $\tilde{\zeta}$ across the Debye layer determines the electroosmotic contribution to the slip velocity $\tilde{\mathbf{u}}_s$.

teristic length scales in the system that we can identify, namely, the Debye screening length λ_D, the geometrical length scale ℓ_0, and the diffusion length $\ell(\omega) = \sqrt{D_0/\omega}$ which is the typical distance that an ion can diffuse during one period of the driving signal. Likewise we can identify three different characteristic time scales, namely, the charge relaxation (Debye) time $\tau_D = \lambda_D^2/D_0$, the bulk diffusion time $\tau_\ell = \ell_0^2/D_0$, and the cell relaxation (RC) time $\tau_0 = \lambda_D \ell_0/D_0$. AC electroosmosis occurs primarily for driving frequencies around the inverse RC time $\omega \sim 1/\tau_0$. Here the diffusion length becomes $\ell(\omega = 1/\tau_0) = \sqrt{\lambda_D \ell_0}$, which means that in the asymptotic analysis we need to account for three nested regions: The "inner" screening layer of (dimensionless) width ε, the "middle" diffusion layer of width $\sqrt{\varepsilon}$, and the "outer" bulk region at the unit characteristic length scale [38].

Applying the method of matched asymptotic expansions to study problems in electrochemistry and electrokinetics is not a new idea: Much work has been done on steady-state solutions for electrochemical cells operated at

DC conditions, e.g., [40, 60], and the low voltage regime, where the Debye–Hückel approximation $\zeta \ll k_B T/ze$ can be employed or the driving voltage or applied field is weak enough that the problem can be linearized, has been studied extensively, e.g., [23, 58, 61, 62]. Dynamical solution have been looked at recently by Bazant and co-workers [37, 38, 39]. However, they primarily focus on transient solutions for a suddenly applied DC voltage. We extend their work by taking into account dynamics in the diffusion layers when the system is driven into the strongly nonlinear regime by a large AC voltage around the inverse RC time. Also we include electroosmotic fluid motion into our model and account for bulk ion transport by convection.

Below we first cast the problem into dimensionless form and identify the relevant dimensionless groups. Then we proceed determine the leading order solution in the bulk, diffusion layer, and Debye layer, respectively, and state the surface conservation laws governing the dynamics of the solution in the boundary layers. The structure is sketched in Fig. 2.4; we focus exclusively on the leading order solution and therefore drop the superscript $^{(0)}$ to simplify the notation. In the derivation we retain all terms that are important in the strongly nonlinear regime, which makes it easy in subsequent sections to extract the linear (low voltage), weakly nonlinear, and strongly nonlinear dynamical models. Finally we discuss the conditions for validity or breakdown of both the weakly and strongly nonlinear models, and propose an *ad hoc* modification to the strongly nonlinear model when the double layer is driven out of quasiequilibrium at "very large" voltage.

2.2.1 Dimensionless form

For simplicity we assume a symmetric binary electrolyte with $z_+ = -z_- = z$, but we allow for asymmetry in the diffusivity $D_+ \neq D_-$. The problem is cast into dimensionless form, denoting dimensionless variables by an acute accent,

$$\acute{\mathbf{r}} = \frac{\mathbf{r}}{\ell_0}, \qquad \acute{t} = \frac{t}{\tau_0}, \qquad \acute{c}_\pm = \frac{c_\pm}{c_0}, \qquad \acute{\phi} = \frac{\phi}{\phi_0}, \qquad \acute{\mathbf{u}} = \frac{\mathbf{u}}{u_0}, \tag{2.40}$$

where ℓ_0 is the characteristic geometrical length scale,

$$\tau_0 = \frac{\lambda_D \ell_0}{D_0} \tag{2.41}$$

is the RC time for the system,

$$D_0 = \frac{D_+ + D_-}{2} \tag{2.42}$$

is the average diffusivity, c_0 is the nominal bulk salt concentration,

$$\phi_0 = \frac{k_B T}{ze} \tag{2.43}$$

is the thermal voltage, and

$$u_0 = \frac{\epsilon\phi_0^2}{\eta\ell_0} \tag{2.44}$$

is the characteristic electroosmotic velocity. In dimensionless form the governing equations become

$$\frac{1}{\varepsilon}\partial_{\acute{t}}\acute{c}_\pm = -\acute{\boldsymbol{\nabla}}\cdot\acute{\mathbf{F}}_\pm, \tag{2.45}$$

$$\acute{\mathbf{F}}_\pm = -\acute{D}_\pm\big(\acute{\boldsymbol{\nabla}}\acute{c}_\pm \pm \acute{c}_\pm\acute{\boldsymbol{\nabla}}\acute{\phi}\big) + Pe\,\acute{\mathbf{u}}\acute{c}_\pm, \tag{2.46}$$

$$-\varepsilon^2\acute{\boldsymbol{\nabla}}^2\acute{\phi} = \acute{\rho} = \frac{1}{2}(\acute{c}_+ - \acute{c}_-), \tag{2.47}$$

$$\frac{1}{Sc}\left[\frac{1}{\varepsilon}\partial_{\acute{t}}\acute{\mathbf{u}} + Pe(\acute{\mathbf{u}}\cdot\acute{\boldsymbol{\nabla}})\acute{\mathbf{u}}\right] = -\acute{\boldsymbol{\nabla}}\acute{p} + \acute{\boldsymbol{\nabla}}^2\acute{\mathbf{u}} - \frac{1}{\varepsilon^2}\acute{\rho}\acute{\boldsymbol{\nabla}}\acute{\phi} \tag{2.48}$$

$$0 = \acute{\boldsymbol{\nabla}}\cdot\acute{\mathbf{u}}. \tag{2.49}$$

Here $\varepsilon = \lambda_D/\ell_0$ is the dimensionless thickness of the Debye layer,

$$Pe = \frac{u_0\ell_0}{D_0} = \frac{\epsilon\phi_0^2}{\eta D_0} \tag{2.50}$$

is the Péclet number defined as the ratio of the convective to the diffusive ion flux, and

$$Sc = \frac{\eta}{\rho_m D_0} = \frac{\nu}{D_0} \tag{2.51}$$

is the Schmidt number defined as the ratio of momentum diffusion to ionic diffusion in the electrolyte. The factors of $1/\varepsilon$ in front of the time derivatives in Eqs. (2.45) and (2.48) arise because we have chosen to focus on the dynamics at the RC time $\acute{t}$ and not the bulk diffusion time $\bar{t} = \varepsilon\acute{t}$. We also introduced dimensionless diffusion constants

$$\acute{D}_+ = \frac{D_+}{D_0} = \frac{2D_r}{1+D_r} \qquad \text{and} \qquad \acute{D}_- = \frac{D_-}{D_0} = \frac{2}{1+D_r}, \tag{2.52}$$

where $D_r = D_+/D_-$ is the ratio of the cation to the anion diffusivity. By construction they add up to $\acute{D}_+ + \acute{D}_- = 2$, whereas (half) their difference becomes

$$\Delta\acute{D} = \frac{1}{2}(\acute{D}_+ - \acute{D}_-) = \frac{D_r - 1}{D_r + 1}, \tag{2.53}$$

and is always in the range $-1 < \Delta\acute{D} < 1$. It is convenient to operate also with the average ion concentration

$$\acute{c} = \tfrac{1}{2}(\acute{c}_+ + \acute{c}_-), \tag{2.54}$$

which we often denote as the "salt" concentration. The characteristic diffusion constant for this quantity turns out to be the harmonic mean of the individual diffusivities

$$\acute{D}_a = \acute{D}_+\acute{D}_- = \frac{4D_r}{(1+D_r)^2}. \tag{2.55}$$

Symbol	Expression	Value	Unit	Scaling
ν		1×10^{-6}	$\mathrm{m^2\,s^{-1}}$	
D_0		2×10^{-9}	$\mathrm{m^2\,s^{-1}}$	
ϕ_0	k_BT/ze	25	mV	
c_0		6×10^{23}	$\mathrm{m^{-3}}$	
ℓ_0		10	μm	
λ_D	$\sqrt{\epsilon k_BT/2c_0z^2e^2}$	10	nm	$\propto 1/\sqrt{c_0}$
τ_0	$\lambda_D\ell_0/D_0$	50	μs	$\propto \ell_0/\sqrt{c_0}$
u_0	$\epsilon\phi_0^2/\eta\ell_0$	50	$\mu\mathrm{m\,s^{-1}}$	$\propto 1/\ell_0$
ε	λ_D/ℓ_0	0.001		$\propto 1/\ell_0\sqrt{c_0}$
δ	$\epsilon\lambda_S/\epsilon_S\lambda_D$	1		$\propto \sqrt{c_0}$
Pe	$\epsilon\phi_0^2/\eta D_0$	0.25		
Sc	ν/D_0	5×10^2		
Re	$\epsilon\phi_0^2/\eta\nu$	5×10^{-4}		$\equiv Pe/Sc$

Table 2.1: Typical numbers with and without dimensions, corresponding to a dilute 1 mM KCl solution at room temperature in a device with 10 μm characteristic length.

This is denoted the *ambipolar diffusion constant* and it is always in the range $0 < \acute{D}_a \leq 1$.

Table 2.1 shows typical values for the different numbers with or without dimension in the problem, corresponding to a dilute 1 mM KCl solution at room temperature in a device with characteristic geometrical length scale $\ell_0 = 10$ μm. Notice that the Reynolds number $Re = Pe/Sc = u_0\ell_0/\nu$ is small, and also independent of ℓ_0 because the characteristic velocity scales as $1/\ell_0$. The nonlinear convection term in the Navier–Stokes equation can be neglected if $|Re\,\acute{\mathbf{u}}| \ll 1$ which is usually the case, even at large driving voltage; in the numerical example above this corresponds to $|\mathbf{u}| \ll 0.1$ m/s. The relative diffusivity is typically in the range $0.1 < D_r < 10$, but the common working electrolyte KCl is actually very close to symmetric with $D_r \approx 0.97$.

For the remainder of this thesis we drop the acute accent on dimensionless variables to reduce the notational overload; it should be obvious from the context whether a given quantity is dimensionless or not.

2.2.2 Electroneutral bulk transport

The standard procedure of matched asymptotic expansions begins by seeking regular expansions (denoted here by bar accents) in the bulk "outer" region, cf. Eq. (2.39) and similarly for $\bar{\phi}$, $\bar{\mathbf{u}}$, and $\bar{p}$. Substituting first into the Poisson equation Eq. (2.47) and equating like powers in ε we find that the bulk

charge density $\bar{\rho}$ vanishes to both zeroth and first order in ε. For a binary electrolyte this implies that the leading order ion concentrations are equal, i.e.,

$$\bar{c}_+ = \bar{c}_- = \bar{c}. \tag{2.56}$$

From the Nernst–Planck equation Eq. (2.45) we find at leading order that

$$\partial_t \bar{c} = 0, \tag{2.57}$$

i.e., the leading order bulk concentration $\bar{c} = \bar{c}(\mathbf{r})$ is stationary in time, with the dynamics showing up only as higher order corrections. Bulk charge neutrality together with the Nernst–Planck equation also imply that the leading order electric current density $\bar{\mathbf{J}} = \frac{1}{2}(\bar{\mathbf{F}}_+ - \bar{\mathbf{F}}_-)$ is divergence free

$$\boldsymbol{\nabla} \cdot \bar{\mathbf{J}} = 0, \tag{2.58}$$

where

$$\bar{\mathbf{J}} = -\bar{c}\boldsymbol{\nabla}\bar{\phi} - \Delta D \boldsymbol{\nabla}\bar{c}. \tag{2.59}$$

In the case of equal diffusivities where $\Delta D = 0$ this reduces to an Ohmic current with spatially varying conductivity $\bar{c}$, but more generally the second term related to the diffusion potential must also be taken into account. With this, the leading order charge density (that is to second order in ε) can be determined as

$$\bar{\rho} = -\varepsilon^2 \boldsymbol{\nabla}^2 \bar{\phi} = \varepsilon^2 \frac{\boldsymbol{\nabla}\bar{c} \cdot \boldsymbol{\nabla}\bar{\phi} + \Delta D \boldsymbol{\nabla}^2 \bar{c}}{\bar{c}}, \tag{2.60}$$

using Eqs. (2.58) and (2.59) to eliminate $\boldsymbol{\nabla}^2\bar{\phi}$.

We focus on the periodic solution reached after all transients in the system have died out. In order for a periodic first order correction to exist, it is necessary that the leading order time average fluxes are divergence free, i.e., that $\boldsymbol{\nabla} \cdot \langle \bar{\mathbf{F}}_\pm \rangle = 0$. Writing the mean ion flux as

$$\bar{\mathbf{F}} = \tfrac{1}{2}(\bar{\mathbf{F}}_+ + \bar{\mathbf{F}}_-) = \bar{\mathbf{F}}_a + \Delta D \bar{\mathbf{J}}, \tag{2.61}$$

where

$$\bar{\mathbf{F}}_a = \tfrac{1}{2}(D_- \bar{\mathbf{F}}_+ + D_+ \bar{\mathbf{F}}_-) = -D_a \boldsymbol{\nabla}\bar{c} + Pe\, \bar{\mathbf{u}}\, \bar{c}, \tag{2.62}$$

we then obtain[2]

$$-\boldsymbol{\nabla} \cdot \langle \bar{\mathbf{F}}_a \rangle = \boldsymbol{\nabla} \cdot \big(D_a \boldsymbol{\nabla}\bar{c} - Pe\, \langle \bar{\mathbf{u}} \rangle \bar{c} \big) = 0. \tag{2.63}$$

At leading order in the Navier–Stokes equation for the fluid motion we obtain

$$\partial_t \bar{\mathbf{u}} = 0, \tag{2.64}$$

[2]It is an interesting feature of binary electrolytes that the evolution of the leading order bulk concentration profile – at the *bulk* diffusion time scale $\bar{t} = \varepsilon t$ – is determined by a simple convection-diffusion equation with no migration term, and with a diffusivity that is simply the harmonic mean of the individual ion diffusivities, namely, the ambipolar diffusion constant D_a [54].

i.e., the leading order fluid velocity $\bar{\mathbf{u}} = \bar{\mathbf{u}}(\mathbf{r})$ is stationary and determined by a time average force balance

$$0 = -\boldsymbol{\nabla}\bar{p} + \boldsymbol{\nabla}^2\bar{\mathbf{u}} + \langle \boldsymbol{\nabla}^2\bar{\phi}\boldsymbol{\nabla}\bar{\phi}\rangle, \tag{2.65}$$

with the pressure $\bar{p}$ determined by the incompressibility constraint

$$\boldsymbol{\nabla}\cdot\bar{\mathbf{u}} = 0. \tag{2.66}$$

We denote as "bulk electroconvection" that component of the overall fluid motion that is induced by the bulk electrical body force $\bar{\mathbf{f}} = \langle \boldsymbol{\nabla}^2\bar{\phi}\,\boldsymbol{\nabla}\bar{\phi}\rangle$ to distinguish it from, e.g., electroosmotic flow induced in the Debye layer. Because momentum diffusion is typically much faster than ionic diffusion corresponding to a large Schmidt number $Sc \gg 1$, cf. Table 2.1, the first order correction of $\mathcal{O}(\varepsilon Sc)$ may be not-so-small for the ε of interest to a microsystem. It may therefore be more appropriate to state the leading order problem as

$$\frac{1}{\varepsilon Sc}\partial_t\bar{\mathbf{u}} = -\boldsymbol{\nabla}\bar{p} + \boldsymbol{\nabla}^2\bar{\mathbf{u}} + \boldsymbol{\nabla}^2\bar{\phi}\boldsymbol{\nabla}\bar{\phi}, \tag{2.67}$$

for $\varepsilon Sc \sim \mathcal{O}(1)$. In any case the time average velocity entering Eq. (2.63) remains the same.

2.2.3 Diffusion layer

The Debye layer at the electrodes periodically absorbs and expels an excess amount of ions. However, because the bulk transport of neutral salt is effectively a diffusion process with diffusion constant D_a and we consider driving frequencies around the inverse RC time, the disturbances do not have time to propagate into the bulk region, and the leading order dynamics only occur in a diffusion layer of width $\mathcal{O}(\sqrt{\varepsilon})$ around the electrode. Assuming the electrode is smooth on the diffusion length scale we model it locally as a flat surface and define x as the tangential and y as the normal direction. Further, we introduce a scaled spatial variable $\check{y} = y/\sqrt{\varepsilon}$ and seek regular asymptotic expansions (denoted by check accents) in the "middle" diffusion layer.

Substituting into the Poisson equation (2.47) we find immediately that the charge density vanishes to zeroth order in ε, so that again the leading order ion concentrations are equal

$$\check{c}_+ = \check{c}_- = \check{c}. \tag{2.68}$$

To determine the dynamical solution we introduce the excess concentration $\check{\gamma} = \check{c} - \bar{c}$ in the diffusion layer relative to the bulk, and form weighted

averages of the ionic fluxes similar to Eq. (2.62) to get

$$\begin{aligned}\partial_t \check{\gamma} &= -\boldsymbol{\nabla}\cdot(\check{\mathbf{F}}_a - \bar{\mathbf{F}}_a)\\ &= D_a\partial_{\check{y}}^2\check{\gamma} + \varepsilon D_a\partial_x^2\check{\gamma} - \varepsilon\, Pe\Big[\check{u}_x\partial_x\check{c} + \frac{1}{\sqrt{\varepsilon}}\check{u}_y\partial_{\check{y}}\check{c} - \bar{u}_x\partial_x\bar{c} - \bar{u}_y\partial_y\bar{c}\Big]\\ &= D_a\partial_{\check{y}}^2\check{\gamma} + \mathcal{O}(\varepsilon).\end{aligned} \tag{2.69}$$

At leading order the excess concentration in the diffusion layer is therefore determined by a simple 1D diffusion problem for $\check{y} \in [0,\infty[$. We argue below that the normal velocity component u_y vanishes at leading order, which means that convection is indeed negligible provided $|Pe\, u_x\partial_x c| \ll \varepsilon^{-1}$. Matching to the bulk solution requires that $\check{\gamma} = 0$ at $\check{y} = \infty$ whereas the boundary condition at $\check{y} = 0$ is determined by the excess flux fed into the diffusion layer from the Debye layer

$$-\frac{1}{\sqrt{\varepsilon}}D_a\partial_{\check{y}}\check{\gamma} = \big[\check{F}_{y,a} - \bar{F}_{y,a}\big]_{\check{y}=0} \equiv \check{f}_a. \tag{2.70}$$

The solution can be expressed as a Fourier series

$$\check{\gamma}(\check{y},t) = \sqrt{\frac{\varepsilon}{\omega D_a}}\sum_{n=1}^{\infty}\frac{\hat{f}_{a,n}}{\sqrt{in}}\, e^{in\omega t - \sqrt{in\omega/D_a}\check{y}} + \text{c.c.} \tag{2.71}$$

where $\hat{f}_{a,n}$ is the nth Fourier component of $\check{f}_a(t)$

$$\hat{f}_{a,n} = \int_0^T \check{f}_a(t)\, e^{-in\omega t}\,\mathrm{d}t. \tag{2.72}$$

Equivalently, the solution can be written as a convolution integral

$$\check{\gamma} = \sqrt{\frac{\varepsilon}{\omega D_a}}\frac{1}{T}\int_0^T \mathcal{G}(\check{y},t-t')\check{f}_a(t')\,\mathrm{d}t' \tag{2.73}$$

where $\mathcal{G}$ the periodic 1D diffusion kernel (Green's function)

$$\mathcal{G}(\check{y},t-t') = \sum_{n=1}^{\infty}\frac{1}{\sqrt{in}}\, e^{in\omega(t-t') - \sqrt{in\omega/D_a}\check{y}} + \text{c.c.} \tag{2.74}$$

In the weakly nonlinear regime where the excess flux $\check{f}_a$ is $\mathcal{O}(1)$ the excess concentration $\check{\gamma}$ only exist as an $\mathcal{O}(\sqrt{\varepsilon})$ small perturbation to the bulk concentration; the strongly nonlinear regime at driving voltages $V_0 \sim 1/\sqrt{\varepsilon}$ is then essentially defined as the point where the perturbation reaches $\mathcal{O}(1)$.

Since the diffusion layer is charge neutral, the current must be constant across it at leading order, i.e.,

$$\check{J}_y = -\frac{1}{\sqrt{\varepsilon}}\big(\check{c}\partial_{\check{y}}\check{\phi} + \Delta D\partial_{\check{y}}\check{c}\big) = \bar{J}_y = -\big(\bar{c}\partial_y\bar{\phi} + \Delta D\partial_y\bar{c}\big). \tag{2.75}$$

Introducing the excess potential $\check{\psi} = \check{\phi} - \bar{\phi}$ and rearranging a bit we obtain

$$\partial_{\check{y}}\psi = -\sqrt{\varepsilon}\,\bar{J}_y\left(\frac{1}{\check{c}} - \frac{1}{\bar{c}}\right) - \Delta D\partial_{\check{y}}\log\bigl(\check{c}/\bar{c}\bigr), \tag{2.76}$$

where we used $\partial_{\check{y}}\bar{\phi} = \sqrt{\varepsilon}\partial_y\bar{\phi}$ and similarly for $\bar{c}$. Integrating once we get

$$\check{\psi}(\check{y}) = \sqrt{\varepsilon}\,\bar{J}_y\int_{\check{y}}^{\infty}\left(\frac{1}{\check{c}(\check{y}')} - \frac{1}{\bar{c}}\right)\mathrm{d}\check{y}' - \Delta D\log\bigl[\check{c}(\check{y})/\bar{c}\bigr]. \tag{2.77}$$

The interpretation is straightforward: The last term is the diffusion potential due to the difference in concentration across the layer, whereas the first term is just the excess potential drop required to drive the bulk current through the diffusion layer when the local resistivity $1/\check{c}$ differs from the bulk value $1/\bar{c}$. In the weakly nonlinear regime where $\check{c}$ is only a small perturbation to the bulk concentration $\bar{c}$, the magnitude of the diffusion potential $-\Delta D\log(\check{c}/\bar{c})$ is only $\mathcal{O}(\sqrt{\varepsilon})$, whereas the excess Ohmic potential drop is $\mathcal{O}(\varepsilon)$. In the strongly nonlinear regime both terms reach $\mathcal{O}(1)$ because $\log(\check{c}/\bar{c}) \sim \mathcal{O}(1)$ and $\bar{J}_y \sim 1/\sqrt{\varepsilon}$. [3]

Concerning the fluid motion induced in the diffusion layer, the incompressibility constraint immediately yields $\check{u}_y = 0$. In the normal direcion the Navier–Stokes equation reduces to a force balance between the excess electrical body force and the pressure at leading order

$$0 = -\partial_{\check{y}}\check{p} + \frac{1}{\varepsilon}\bigl(\partial_{\check{y}}^2\check{\psi}\bigr)\partial_{\check{y}}\check{\psi} = \partial_{\check{y}}\Bigl[-\check{p} + \frac{1}{2\varepsilon}\bigl(\partial_{\check{y}}\check{\psi}\bigr)^2\Bigr], \tag{2.79}$$

from which

$$\check{p} = \frac{1}{2\varepsilon}\bigl(\partial_{\check{y}}\check{\psi}\bigr)^2. \tag{2.80}$$

In the tangential direction we obtain a balance between the viscous and electrical forces and the tangential pressure gradient

$$0 = -\varepsilon\partial_x\check{p} + \partial_{\check{y}}^2\check{u}_x + \partial_{\check{y}}^2\check{\psi}\partial_x\check{\phi}. \tag{2.81}$$

Neglecting terms proportional to $\partial_x\check{\psi}$ because these are much smaller than $\partial_x\bar{\phi}$ we arrive at a Helmholtz–Smoluchowski type tangential velocity component

$$\check{u}_x = (\check{\varsigma} - \check{\psi})\partial_x\bar{\phi}, \tag{2.82}$$

where $\check{\varsigma} = \check{\psi}(0)$ is the total excess potential drop across the diffusion layer.

[3] To make a rough estimate of the integral factor we can approximate $\check{c}$ by a piecewise linear profile $\check{c} \approx \check{c}_o + (\bar{c} - \check{c}_o)\sqrt{\omega/D_a}\,\check{y}$ for $\check{y} \le \sqrt{D_a/\omega}$ and $\check{c} \approx \bar{c}$ for $\check{y} \ge \sqrt{D_a/\omega}$ to get

$$\int_0^{\infty}\left(\frac{1}{\check{c}} - \frac{1}{\bar{c}}\right)\mathrm{d}\check{y} \approx -\sqrt{\frac{D_a}{\omega}}\left[\frac{\log(\check{c}_o/\bar{c})}{\bar{c} - \check{c}_o} + \frac{1}{\bar{c}}\right] \sim \mathcal{O}(1). \tag{2.78}$$

2.2.4 Quasiequilibrium double layer

The singular perturbation in the Poisson equation gives rise to a boundary layer of width ε where the charge density is nonzero to zeroth order in ε. Introducing a scaled spatial variable $\tilde{y} = y/\varepsilon$ to remove the singular perturbation, we can seek regular asymptotic expansions (denoted by tilde accents) in the Debye "inner" layer. Substituting into the Nernst–Planck equations Eqs. (2.45) and (2.46) and using $\partial_{\tilde{y}} = \varepsilon\partial_y$ we find at leading order

$$0 = \partial_{\tilde{y}}\big(\partial_{\tilde{y}}\tilde{c}_\pm \pm \tilde{c}_\pm\,\partial_{\tilde{y}}\tilde{\phi}\big), \tag{2.83}$$

that is, the Debye layer is in quasiequilibrium and the ions are Boltzmann distributed

$$\tilde{c}_\pm = \check{c}e^{\mp\tilde{\psi}}, \tag{2.84}$$

where $\tilde{\psi} = \tilde{\phi} - \check{\phi}$ is the excess potential in the Debye layer relative to the bulk, and $\check{c}$ is the limiting value of the concentration in the diffusion layer seen from the Debye layer at $\check{y} = 0$. Likewise for the charge density and average ion concentration in the Debye layer we have

$$\tilde{\rho} = -\check{c}\sinh(\tilde{\psi}) \qquad \text{and} \qquad \tilde{c} = \check{c}\cosh(\tilde{\psi}). \tag{2.85}$$

The quasiequilibrium arises because the Debye layer reaches local equilibrium on the Debye time scale $\tilde{t} = t/\varepsilon$ which is much faster than the RC time of the system. Strictly, Eq. (2.83) does not imply that the leading order fluxes

$$\tilde{F}_{y,\pm} = -\frac{1}{\varepsilon}D_\pm\big(\partial_{\tilde{y}}\tilde{c}_\pm \pm \tilde{c}_\pm\,\partial_{\tilde{y}}\tilde{\phi}\big) \tag{2.86}$$

vanish but only that they are constant; it is by matching with bulk fluxes $\ll \mathcal{O}(1/\varepsilon)$ that we arrive at the Boltzmann profiles. Substituting into the Poisson equation, the leading order problem is the dimensionless Poisson–Boltzmann equation

$$\partial_{\tilde{y}}^2\tilde{\psi} = \check{c}\sinh(\tilde{\psi}), \tag{2.87}$$

which has the well-known Gouy–Chapman solution

$$\tilde{\psi} = 4\tanh^{-1}\Big[\tanh\big(\tilde{\zeta}/4\big)e^{-\sqrt{\check{c}}\,\tilde{y}}\Big]. \tag{2.88}$$

The integration constant $\tilde{\zeta}$ is simply the leading order zeta potential, and $1/\sqrt{\check{c}}$ is the local effective Debye length.

As in the diffusion layer, the incompressibility constraint immediately yields $\tilde{u}_y = 0$ at leading order, and the balance between the electrical body force and the pressure in the normal direction that gives rise to an (excess) osmotic pressure of

$$\tilde{p} = \frac{1}{2\varepsilon^2}\big(\partial_{\tilde{y}}\tilde{\psi}\big)^2. \tag{2.89}$$

Substituting into the Navier–Stokes equation in the tangential direction we obtain at leading order

$$0 = -\varepsilon^2 \partial_x \tilde{p} + \partial_{\tilde{y}}^2 \tilde{u}_x - \tilde{\rho}\, \partial_x \tilde{\phi}. \tag{2.90}$$

Writing $\partial_x \tilde{\phi} = \partial_x \check{\phi} + \partial_x \tilde{\psi}$ and using again $\tilde{\rho} = -\partial_{\tilde{y}}^2 \tilde{\psi}$ we arrive at the general expression for electroosmotic flow of the first kind [4]

$$\tilde{u}_x = (\tilde{\zeta} - \tilde{\psi}) \partial_x \check{\phi} + 4 \log \left[\frac{\cosh(\tilde{\psi}/4)}{\cosh(\tilde{\zeta}/4)} \right] \partial_x \log \check{c}. \tag{2.91}$$

The first term is the familiar Helmholtz–Smoluchowski velocity due to the action of the tangential field from the diffusion layer, and the second term is the diffusioosmotic flow due to tangential gradients in the osmotic pressure in the Debye layer.

2.2.5 Surface conservation laws

The redistribution of ions across the Debye layer in the normal direction is instantaneous on the (RC) time scale that we consider here, but the total amount of ions accumulated locally can change only by flux in or out of the layer or by tangenial fluxes within the layer. Following Bazant and Chu [37, 39] we quantify this by defining $\varepsilon \tilde{w}_\pm(x,t)$ as the excess amount of each ionic species accumulated locally in the Debye layer, such that

$$\tilde{w}_\pm = \frac{1}{\varepsilon} \int_{\text{Debye layer}} \left(\tilde{c}_\pm - \check{c}_\pm \right) \mathrm{d}y = \int_0^\infty (\tilde{c}_\pm - \check{c}_\pm)\, \mathrm{d}\tilde{y}. \tag{2.92}$$

Likewise we define the surface charge and average excess concentration by

$$\tilde{q} = \int_0^\infty (\tilde{\rho} - \check{\rho})\, \mathrm{d}\tilde{y} \qquad \text{and} \qquad \tilde{w} = \int_0^\infty (\tilde{c} - \check{c})\, \mathrm{d}\tilde{y}, \tag{2.93}$$

respectively, where by construction $\tilde{w}_\pm = \tilde{w} \pm \tilde{q}$. At leading order the integrals can be evaluated using Eqs. (2.84) and (2.88) to get

$$\tilde{w}_\pm = 2\sqrt{\check{c}} \left(e^{\mp \tilde{\zeta}/2} - 1 \right), \tag{2.94}$$

from which also

$$\tilde{q} = -2\sqrt{\check{c}} \sinh(\tilde{\zeta}/2) \quad \text{and} \quad \tilde{w} = 4\sqrt{\check{c}} \sinh^2(\tilde{\zeta}/4). \tag{2.95}$$

[4] The notion "electroosmosis of the first kind" was introduced by Dukhin and co-workers to distinguish it from their "electroosmosis of the second kind" in double layers driven out of quasiequilibrium [41, 42].

The time evolution of $\tilde{w}_\pm$ is determined by

$$\partial_t \tilde{w}_\pm = \int_0^\infty \partial_t \big(\tilde{c}_\pm - \check{c}_\pm\big)\,\mathrm{d}\tilde{y} \tag{2.96}$$

$$= -\int_0^\infty \Big[\varepsilon\partial_x\big(\tilde{F}_{x,\pm} - \check{F}_{x,\pm}\big) + \partial_{\tilde{y}}\big(\tilde{F}_{y,\pm} - \check{F}_{y,\pm}\big)\Big]\,\mathrm{d}\tilde{y} \tag{2.97}$$

$$= -\varepsilon\partial_x \int_0^\infty \big(\tilde{F}_{x,\pm} - \check{F}_{x,\pm}\big)\,\mathrm{d}\tilde{y} - \Big[\tilde{F}_{y,\pm} - \check{F}_{y,\pm}\Big]_{\tilde{y}=0}^{\infty} \tag{2.98}$$

$$= -\varepsilon\partial_x \tilde{\Gamma}_{x,\pm} - \check{F}_{y,\pm} + F_{\pm,\mathrm{ext}}. \tag{2.99}$$

Here we used the continuity equation Eq. (2.45) to eliminate the time derivative. Next we partially integrated the normal flux, used the flux matching condition

$$\lim_{\tilde{y}\to\infty} \tilde{F}_{y,\pm} \sim \lim_{\check{y}\to 0} \check{F}_{y,\pm}, \tag{2.100}$$

and wrote $\tilde{F}_{y,\pm}|_{\tilde{y}=0} = F_{\pm,\mathrm{ext}}$ for the flux injected into the Debye layer at the electrode surface, e.g., by electrochemical reaction. Finally, we introduced the surface excess flux $\tilde{\Gamma}_{x,\pm}$ given by

$$\tilde{\Gamma}_{x,\pm} = \int_0^\infty \big(\tilde{F}_{\pm,x} - \check{F}_{\pm,x}\big)\,\mathrm{d}\tilde{y}. \tag{2.101}$$

At leading order we can evaluate $\Gamma_{x,\pm}$ by rewriting the flux in terms of the electrochemical potential $\mu_\pm = \log c_\pm \pm \phi$ to get $\mathbf{F}_\pm = -D_\pm c_\pm \boldsymbol{\nabla}\mu_\pm + Pe\,\mathbf{u}\,c_\pm$. Then

$$\tilde{\Gamma}_{\pm,x} = -D_\pm \int_0^\infty (\tilde{c}_\pm \partial_x \tilde{\mu}_\pm - \check{c}_\pm \partial_x \check{\mu}_\pm)\,\mathrm{d}\tilde{y} + Pe \int_0^\infty (\tilde{c}_\pm \tilde{u}_x - \check{c}_\pm \check{u}_x)\,\mathrm{d}\tilde{y}$$

$$= -D_\pm \tilde{w}_\pm \partial_x \check{\mu}_\pm + Pe \int_0^\infty (\tilde{c}_\pm \tilde{u}_x - \check{c}_\pm \check{u}_x)\,\mathrm{d}\tilde{y}, \tag{2.102}$$

where we used that $\mu_\pm$ is constant across the Debye layer because it is in quasiequilibrium at leading order. This is a powerful result because it is independent of the particular model for the double layer [39], i.e., it remains valid when the dilute approximation breaks down at large voltage, and may even be generalized to account for interactions between the different species by replacing the diffusivities $D_\pm$ by generalized mobilities $\mathcal{L}_{ij}$, cf. Eq. (2.3). However, in order to handle the convection term we do need a model, and using Eq. (2.91) we get

$$\int_0^\infty (\tilde{c}_\pm \tilde{u}_x - \check{c}\,\check{u}_x)\,\mathrm{d}\tilde{y} = -2\Big\{\pm\tilde{w}_\pm + \sqrt{\check{c}}\,\big[\mathcal{W}^\ddagger(\tilde{\zeta}) + \tilde{\zeta}\big]\Big\}\partial_x\check{\phi} \tag{2.103}$$

$$-2\Big\{\tilde{w}_\pm + \sqrt{\check{c}}\,\big[\mathcal{W}^\dagger(\tilde{\zeta}) \pm \tilde{\zeta}\big]\Big\}\partial_x \log\check{c}.$$

Here $\mathcal{W}^\dagger$ and $\mathcal{W}^\ddagger$ are shorthands for the expressions

$$\mathcal{W}^\dagger(\zeta) = -8\log[\cosh(\zeta/4)] - \frac{1}{2}\mathrm{Li}_2[\tanh^2(\zeta/4)] \approx -\frac{9}{32}\zeta^2 + \frac{7}{2048}\zeta^4 \tag{2.104}$$

and

$$\mathcal{W}^{\ddagger}(\zeta) = \mathrm{Li}_2[\tanh(\zeta/4)] - \mathrm{Li}_2[-\tanh(\zeta/4)] \approx \frac{1}{2}\zeta - \frac{1}{144}\zeta^3, \qquad (2.105)$$

and $\mathrm{Li}_2(z) = \sum_{n=1}^{\infty} z^n/n^2$ is the dilogarithm.

Surface conservation laws of this form were used already by Dukhin, Deryagin, and Shilov in their studies of surface conductance and the polarization of thin but highly charged double layers around spherical particles in weak applied fields [18, 63, 64, 65]. Recently Chu and Bazant revisited the problem using matched asymptotic expansions which allowed them to lift the requirements for weak applied fields and small deviations from equilibrium in the bulk [38, 39]. What we have contributed is to extend their work to account also for surface convection and evaluate the integral in Eq. (2.103). Dukhin and co-workers also accounted for electroosmotic convection and our result is fully equivalent to their work except for two features: First, in the English translation it appears that there is a sign error on their term for flux by electroosmotic convection proportional to $\partial_x \log \check{c}$, which in our notation reads $2\big\{ \pm \tilde{w}_{\pm} + \tilde{\zeta} - 8\log[\cosh(\tilde{\zeta}/4)]\big\}\partial_x \log \check{c}$. For a symmetric electrolyte a change of sign of $\check{\phi}$ and $\tilde{\phi}$ while retaining $\check{c}$ unchanged must result in a symmetric interchange of the cation and anion fluxes. Our Eq. (2.103) satisfies this property, but the results in Refs. [18, 65] do not. Secondly, the dilogarithm terms are absent in their work.

The presence of the dilogarithm terms qualitatively changes the form of the convective flux at low zeta potential: The slope of the coefficient to $\partial_x \check{\phi}$ becomes nonzero at $\tilde{\zeta} = 0$, and coefficient to $\partial_x \log \check{c}$ becomes a non-monotonic function of $\tilde{\zeta}$. However, these subtle changes will be very hard to detect experimentally due to the factor of ε multiplying the divergence of the surface fluxes in Eq. (2.99). Only at high zeta potential is the excess amount of ions accumulated in the Debye layer large enough to make the surface flux term significant, which can occur either for colloidal particles with a large intrinsic zeta potential determined by their surface chemistry, or for electrodes or polarizable objects subjected to large external voltages or applied field. In this regime the surface excess conductance is dominated by the factors $\tilde{w}_{\pm}$ that grow exponentially with $\tilde{\zeta}$ and the expression for the surface excess flux simplifies to

$$\tilde{\Gamma}_{\pm,x} = -(D_{\pm} + 2Pe)\tilde{w}_{\pm}\partial_x \check{\mu}_{\pm}. \qquad (2.106)$$

Using this we can write the surface excess current $\tilde{\jmath}_x = \frac{1}{2}(\tilde{\Gamma}_{+,x} - \tilde{\Gamma}_{-,x})$ and neutral salt flux $\tilde{\Gamma} = \frac{1}{2}(\tilde{\Gamma}_{+,x} + \tilde{\Gamma}_{-,x})$ as

$$\tilde{\jmath}_x = -\big[(1+2Pe)\tilde{w} + \Delta D\tilde{q}\big]\partial_x \check{\phi} - \big[(1+2Pe)\tilde{q} + \Delta D\tilde{w}\big]\partial_x \log \check{c}, \qquad (2.107)$$

and

$$\tilde{\Gamma}_x = -\big[(1+2Pe)\tilde{q} + \Delta D\tilde{w}\big]\partial_x \check{\phi} - \big[(1+2Pe)\tilde{w} + \Delta D\tilde{q}\big]\partial_x \log \check{c}, \qquad (2.108)$$

where we used that $\tilde{w}_+ + \tilde{w}_- = 2\tilde{w}$, $\tilde{w}_+ - \tilde{w}_- = 2\tilde{q}$, $D_+ + D_- = 2$, and $D_+ - D_- = 2\Delta D$.

To close the coupled electrical problem for $\bar{\phi}$ and $\bar{c}$ in the bulk, $\tilde{q}$ and $\tilde{w}$ in the Debye layer, and $\check{c}$ in the diffusion layer, we need another two boundary conditions. The first is the potential drop across the boundary layers obtained from the Gouy–Chapman–Stern model

$$V_{\text{ext}} - \bar{\phi} = \check{\varsigma} + \tilde{\zeta} - \tilde{q}\,\delta, \tag{2.109}$$

where $\check{\varsigma}$, $\tilde{\zeta}$, and $-\tilde{q}\,\delta$ are the potential drops over the diffusion, Debye, and Stern layers, respectively, and δ is the surface capacitance ratio. This acts as a Dirichlet condition for the bulk potential $\bar{\phi}$. The second condition is a requirement of consistency in Eq. (2.95) as

$$\tilde{w} = 4\sqrt{\check{c}}\sinh^2(\tilde{\zeta}/4) = \sqrt{\tilde{q}^2 + 4\check{c}} - \sqrt{4\check{c}}, \tag{2.110}$$

where we used $\tilde{\zeta} = -2\sinh^{-1}\big(\tilde{q}/2\sqrt{\check{c}}\big)$ to express $\tilde{w}$ directly in terms of the dynamical variables $\tilde{q}$ and $\check{c}$.

With this we are ready to formulate in the following sections the leading order dynamical models corresponding to the linear, weakly nonlinear, and strongly nonlinear regimes.

2.2.6 Linear model

In the Debye–Hückel limit $\tilde{\zeta} \ll 1$ where the potential is small compared to the thermal voltage the electrokinetic problem can be linearized. This simplifies the analysis significantly and this regime has therefore been studied extensively in the literature. Since we drive the system with a harmonic signal, at leading order all variables can be expressed as an equilibrum value with a small deviation that varies harmonically in time

$$\phi(\mathbf{r}, t) = \mathrm{Re}\big[\hat{\phi}(\mathbf{r})e^{i\omega t}\big] = \tfrac{1}{2}\hat{\phi}(\mathbf{r})e^{i\omega t} + \text{c.c.} \tag{2.111}$$

and likewise for the ionic concentrations

$$c_i(\mathbf{r}, t) = c_i^* + \mathrm{Re}\big[\hat{c}_i(\mathbf{r})e^{i\omega t}\big]. \tag{2.112}$$

Because we focus on driving frequencies around the inverse RC time, the perturbations do not have time to propagate into the bulk, so the bulk concentration is uniform and Eq. (2.58) reduces to a Laplace problem for $\hat{\phi}$

$$-\boldsymbol{\nabla}^2\hat{\phi} = 0. \tag{2.113}$$

The surface conservation law for charge accumulation in the Debye layer becomes

$$i\omega\hat{q} = \hat{\jmath}_{\text{ext}} + \mathbf{n}\cdot\boldsymbol{\nabla}\hat{\phi}, \tag{2.114}$$

where $\hat{\jmath}_{\text{ext}}$ is the Faradaic current injection at the electrode and $\mathbf{n}\cdot\boldsymbol{\nabla}\hat{\phi}$ is the Ohmic current running into the Debye layer from the bulk. The capacitance of the double layer at low voltage is $C_{\text{dl}} = 1/(1+\delta)$ so the accumulated charge is related to the potential drop over the double layer by

$$\hat{q} = -\frac{\hat{V}_{\text{ext}} - \hat{\phi}}{1+\delta}. \tag{2.115}$$

In the absense of Faradaic electrode reactions the boundary condition for the potential therefore becomes [23]

$$-\mathbf{n}\cdot\boldsymbol{\nabla}\hat{\phi} = \frac{i\omega}{1+\delta}\big(\hat{V}_{\text{ext}} - \hat{\phi}\big). \tag{2.116}$$

The slip velocity condition for the time average bulk fluid motion reduces to [33]

$$\begin{aligned}\tilde{\mathbf{u}}_s &= \big\langle \tilde{\zeta}\boldsymbol{\nabla}_s\bar{\phi}\big\rangle = \frac{1}{2}\mathrm{Re}\big[\hat{\zeta}^*\boldsymbol{\nabla}_s\hat{\phi}\,\big] = \frac{1}{2(1+\delta)}\mathrm{Re}\big[(\hat{V}_{\text{ext}} - \hat{\phi})^*\boldsymbol{\nabla}_s\hat{\phi}\,\big] \\ &= -\frac{1}{4(1+\delta)}\boldsymbol{\nabla}_s\big|\hat{V}_{\text{ext}} - \hat{\phi}\big|^2,\end{aligned} \tag{2.117}$$

where we used that $\tilde{\zeta} = (V_{\text{ext}} - \bar{\phi})/(1+\delta)$ at low voltage and that $\boldsymbol{\nabla}_s V_{\text{ext}} = 0$ because the metal electrode is an equipotential surface. Since the bulk concentration is uniform there is no body force at leading order and the Stokes flow problem for the bulk fluid motion is driven entirely by the electroosmotic slip.

Faradaic current injection

The Faradaic current is determined from Eq. (2.32) as

$$\hat{\jmath}_{\text{ext}} = \frac{1}{R_{\text{ct}}}\Big[(\hat{V}_{\text{ext}} - \hat{\phi}) + \frac{\hat{c}_{\text{R}}}{c^*_{\text{R}}} - \frac{\hat{c}_{\text{O}}}{c^*_{\text{O}}}\Big], \tag{2.118}$$

where $R_{\text{ct}} = \sigma k_B T/j_0 e\ell_0$ is the dimensionless charge transfer resistance, and $\hat{c}_{\text{O}}$ and $\hat{c}_{\text{R}}$ are the concentrations of the oxidized and reduced species immediately outside the Debye layer. Assuming we can neglect migration of O and R in the diffusion layer we employ the result from Sec. 2.2.3 to get

$$\hat{c}_{\text{O}}|_{\check{y}=0} = \sqrt{\frac{\varepsilon}{i\omega D_{\text{O}}}}\,\hat{f}_{\text{O}} \tag{2.119}$$

and similarly for $\hat{c}_{\text{R}}$. Approximating the flux $\hat{f}_{\text{O}}$ into the diffusion layer by the external current $\hat{\jmath}_{\text{ext}}$, which is equivalent to neglecting accumulation of O in the Debye layer, and similarly approximating $\hat{f}_{\text{R}}$ by $-\hat{\jmath}_{\text{ext}}$ we then arrive at

$$\hat{\jmath}_{\text{ext}} = \frac{1}{R_{\text{ct}}}\Big[(\hat{V}_{\text{ext}} - \hat{\phi}) - \frac{1}{c^*_{\text{R}}}\sqrt{\frac{\varepsilon}{i\omega D_{\text{R}}}}\,\hat{\jmath}_{\text{ext}} - \frac{1}{c^*_{\text{O}}}\sqrt{\frac{\varepsilon}{i\omega D_{\text{O}}}}\,\hat{\jmath}_{\text{ext}}\Big]. \tag{2.120}$$

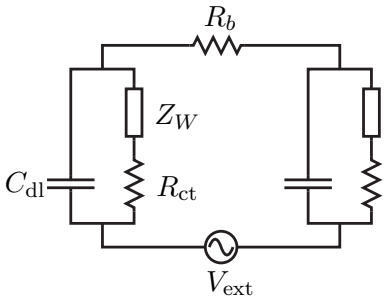


Figure 2.5: Equivalent circuit diagram for the linear model: The double layer capacitance C_{dl} is connected in parallel with the charge transfer resistance R_{ct} for the Faradaic electrode reaction and the Warburg impedance $\hat{Z}_W(\omega)$ describing the effects of mass transfer. All are connected in series to the bulk Ohmic resistance R_b.

This can be rearranged to

$$\hat{j}_{\mathrm{ext}} = \frac{\hat{V}_{\mathrm{ext}} - \hat{\phi}}{R_{\mathrm{ct}} + \hat{Z}_W(\omega)}, \tag{2.121}$$

where

$$\hat{Z}_W(\omega) = \left(\frac{\sqrt{\varepsilon}}{c_{\mathrm{O}}^* \sqrt{D_{\mathrm{O}}}} + \frac{\sqrt{\varepsilon}}{c_{\mathrm{R}}^* \sqrt{D_{\mathrm{R}}}} \right) \frac{1}{\sqrt{i\omega}} \tag{2.122}$$

is known as the Warburg impedance [16]. The expression in the parenthesis is called the Warburg constant and we denote it by R_{mt}[5]

$$R_{\mathrm{mt}} = \left(\frac{\sqrt{\varepsilon}}{c_{\mathrm{O}}^* \sqrt{D_{\mathrm{O}}}} + \frac{\sqrt{\varepsilon}}{c_{\mathrm{R}}^* \sqrt{D_{\mathrm{R}}}} \right). \tag{2.123}$$

Finally, substituting into the surface charge conservation law we obtain a closed relation between the complex bulk current running into the double layer and the potential drop across it

$$-\mathbf{n} \cdot \boldsymbol{\nabla} \hat{\phi} = \left[\frac{i\omega}{1+\delta} + \frac{1}{R_{\mathrm{ct}} + \hat{Z}_W(\omega)} \right] \left(\hat{V}_{\mathrm{ext}} - \hat{\phi} \right). \tag{2.124}$$

The equivalent circuit diagram for the entire system is shown in Fig. 2.5 with the double layer capacitance $C_{\mathrm{dl}} = 1/(1+\delta)$ connected in parallel to the charge transfer resistance R_{ct} that is in series with $\hat{Z}_W$, while all are connected in series to the bulk Ohmic resistance $R_b = 1$. This way of representing an electrochemical cell is a standard tool in electrochemical impedance spectroscopy which allows experimental parameters such as R_{ct} to be determined with high accuracy [16]. In such experiments it is common practice to add an excess amount of supporting electrolyte that is

[5]The Warburg constant is often denoted σ in electrochemistry but we prefer the non-standard symbol R_{mt} to avoid confusion with the electrolyte conductivity.

electrochemically inert but serves to minimize the bulk Ohmic resistance. This also justifies the neglect of migration of O and R in the diffusion layer and of accumulation of those species in the Debye layer. In our formulation this corresponds to the limit $c_{\mathrm{O}}^* \ll 1$ and $c_{\mathrm{R}}^* \ll 1$. However, because AC electroosmosis is strongly supressed at high electrolyte concentration it is relevant to consider the case when, e.g., the species O also plays the role of cation in the supporting electrolyte, i.e., $z_{\mathrm{O}} = z_{\mathrm{R}} + 1 = 1$ and $c_{\mathrm{O}}^* = 1$. Performing exactly the same analysis as above but taking into account migration of O in the diffusion layer we get

$$\hat{c}_{\mathrm{O}} = \sqrt{\frac{\varepsilon}{i\omega D_a}}\, \hat{f}_a, \tag{2.125}$$

where D_a is the ambipolar diffusion constant and the flux $\hat{f}_a$ is given by

$$\hat{f}_a = \tfrac{1}{2}(D_- \hat{F}_+ + D_+ \hat{F}_-) = \tfrac{1}{2} D_- \hat{\jmath}_{\mathrm{ext}} - i\omega \Delta D\, \hat{q}. \tag{2.126}$$

Using this we find a relation similar to Eq. (2.124) between the bulk current and potential drop across the double layer

$$-\mathbf{n}\cdot\boldsymbol{\nabla}\hat{\phi} = \left[\frac{i\omega}{1+\delta} + \frac{1 - \sqrt{i\omega\varepsilon}\frac{\Delta D}{(1+\delta)\sqrt{D_a}}}{R_{\mathrm{ct}} + \hat{Z}_W(\omega)}\right](\hat{V}_{\mathrm{ext}} - \hat{\phi}), \tag{2.127}$$

where now $\hat{Z}_W$ is given by

$$\hat{Z}_W(\omega) = \left(\frac{\sqrt{\varepsilon}}{c_{\mathrm{R}}^*\sqrt{D_{\mathrm{R}}}} + \frac{\sqrt{\varepsilon}}{2\sqrt{D_r}}\right)\frac{1}{\sqrt{i\omega}} \tag{2.128}$$

and $D_r = D_+/D_-$. The term proportional to ΔD in the nominator in Eq. (2.127) is related to the accumulation of O in the Debye layer, and it is negligible at leading order due to the factor of $\sqrt{\varepsilon}$. Therefore the case when O plays the role of cation in the supporing electrolyte (or R plays the role of anion) is formally equivalent to the that of excess supporting electrolyte, with the only difference being the definition of R_{mt}.

The Warburg impedance reaches $\mathcal{O}(1)$ only at driving frequencies below

$$\omega \sim R_{\mathrm{mt}}^2 \sim \varepsilon/\min\{c_{\mathrm{O}}^*, c_{\mathrm{R}}^*\}^2. \tag{2.129}$$

However, we must keep in mind that our model with thin diffusion layers is valid only when ω is well above the inverse bulk diffusion time, i.e., for $\omega \gg \varepsilon$. The effect of mass transfer is therefore significant in our model only if one of the electrochemically active species is present at a concentration much lower than the supporting electrolyte, $\min\{c_{\mathrm{O}}^*, c_{\mathrm{R}}^*\} \ll 1$. On the other hand, at very low concentration $\min\{c_{\mathrm{O}}^*, c_{\mathrm{R}}^*\} \ll \sqrt{\varepsilon}$ the mass transfer resistance is so large that the Faradaic electrode reaction is negligible alltogether.

2.2.7 Weakly nonlinear model

Our notion of the "weakly nonlinear" regime is characterized by two features in the solution, namely,

- The bulk electrolyte concentration is uniform and given by $\bar{c} = 1$.
- The diffusion layer around the electrodes exists only as a small perturbation to the bulk solution.

When these conditions are satisfied the solution again simplifies significantly. The leading order bulk problem is linear and given by

$$\bar{c} = 1, \tag{2.130}$$
$$0 = -\boldsymbol{\nabla}^2 \bar{\phi}, \tag{2.131}$$
$$0 = -\boldsymbol{\nabla} \bar{p} + \boldsymbol{\nabla}^2 \bar{\mathbf{u}}, \tag{2.132}$$
$$0 = \boldsymbol{\nabla} \cdot \bar{\mathbf{u}}, \tag{2.133}$$

which is matched with the following set of dynamical boundary conditions in the Debye layer

$$\partial_t \tilde{q} = \mathbf{n} \cdot \boldsymbol{\nabla} \bar{\phi} + j_{\text{ext}}, \tag{2.134}$$
$$V_{\text{ext}} - \bar{\phi} = \tilde{\zeta} - \tilde{q}\delta, \tag{2.135}$$
$$\bar{\mathbf{u}} = \tilde{\mathbf{u}}_s = \left\langle \tilde{\zeta} \boldsymbol{\nabla}_s \bar{\phi} \right\rangle, \tag{2.136}$$

where the zeta potential is related to the charge accumulated in the Debye layer by

$$\tilde{\zeta} = -2 \sinh^{-1}(\tilde{q}/2). \tag{2.137}$$

As we discuss below in Sec. 2.2.9 both the surface fluxes in the Debye layer and the perturbations in the local salt concentration in the diffusion layer reach $\mathcal{O}(1)$ at sufficiently large driving voltage $V_0 \sim \sqrt{\delta/\varepsilon}$. Our weakly nonlinear model is therefore only valid at relatively low voltage

$$V_0 \ll \sqrt{\delta/\varepsilon}. \tag{2.138}$$

For blocking electrodes where the Faradaic current j_{ext} vanishes, the weakly nonlinear model is described in terms of four dimensionless parameters, namely the driving voltage V_0 and frequency ω, the capacitance ratio δ, and the equilibrium zeta potential $\tilde{\zeta}^{\text{eq}}$ on the electrodes. The electrical part of the problem is nonlinear, but local in time so it is straightforward to integrate using, e.g., FEMLAB's standard time-stepping algorithm. The flow problem is linear and can be computed from the time average slip velocity once the electrical problem has been solved.

Faradaic current injection

When electrode reactions do occur the leading order Faradaic current j_{ext} is determined from the current-overpotential equation Eq. (2.31)

$$j_{\text{ext}} = \frac{1}{R_{\text{ct}}} e^{-\tilde{\zeta}/2} \left[\frac{\check{c}_{\text{R}}}{c^*_{\text{R}}} e^{\tilde{\vartheta}/2} - \frac{\check{c}_{\text{O}}}{c^*_{\text{O}}} e^{-\tilde{\vartheta}/2} \right], \tag{2.139}$$

where $R_{\text{ct}} = \sigma k_B T / j_0 e \ell_0$ is the dimensionless charge transfer resistance, $\vartheta = (V_{\text{ext}} - \bar{\phi}) - (V_{\text{ext}} - \bar{\phi})^{\text{eq}}$ is the overpotential, $\check{c}_{\text{O}}$ and $\check{c}_{\text{R}}$ are the concentrations of the oxidized and reduced species immediately outside the Debye layer, and we assume a symmetric transfer coefficient $\alpha = \frac{1}{2}$ and take $z_{\text{O}} = z_{\text{R}} + 1 = 1$. To further simplify the solution we also assume that there is excess supporting electrolyte such that $c^*_{\text{O}} \ll 1$. If this is not satisfied the Faradaic reaction would invalidate our assumption of small perturbations to the bulk concentration in the diffusion layer and make the problem "strongly nonlinear" already at relatively low voltage $V_0 \sim \mathcal{O}(\log \varepsilon)$ due to the exponential voltage dependence in Eq. (2.139). For $c^*_{\text{O}} \ll 1$ we can neglect migration of O in the diffusion layer and employ the 1D solution from Sec. 2.2.3 to get

$$\check{c}_{\text{O}} = c^*_{\text{O}} + \sqrt{\frac{\varepsilon}{\omega D_{\text{O}}}} \frac{1}{T} \int_0^T \check{\mathcal{G}}(t - t') \check{f}_{\text{O}}(t') \, \mathrm{d}t', \tag{2.140}$$

assuming the bulk concentration $\bar{c}_{\text{O}}$ is at the equilibrium value c^*_{O}, and similarly for $\check{c}_{\text{R}}$. Using $\check{f}_{\text{O}} = j_{\text{ext}}$ and $\check{f}_{\text{R}} = -j_{\text{ext}}$ and substituting into Eq. (2.139) we obtain a single integral equation for $j_{\text{ext}}(t)$, and after some rearranging we arrive at

$$\frac{1}{T} \int_0^T \check{\mathcal{G}}(t - t') j_{\text{ext}}(t') \, \mathrm{d}t' = \sqrt{\omega} \, \frac{2 \sinh\bigl(\tilde{\vartheta}/2\bigr) - R_{\text{ct}} e^{\tilde{\zeta}/2} j_{\text{ext}}(t)}{R^{\text{R}}_{\text{mt}} e^{\tilde{\vartheta}/2} + R^{\text{O}}_{\text{mt}} e^{-\tilde{\vartheta}/2}}. \tag{2.141}$$

Here the constants R^{O}_{mt} and R^{R}_{mt} are the contributions from O and R to the Warburg constant, cf. the linear analysis in Sec. 2.2.6

$$R^{\text{O}}_{\text{mt}} = \frac{1}{c^*_{\text{O}}} \sqrt{\frac{\varepsilon}{D_{\text{O}}}} \qquad \text{and} \qquad R^{\text{R}}_{\text{mt}} = \frac{1}{c^*_{\text{R}}} \sqrt{\frac{\varepsilon}{D_{\text{R}}}}. \tag{2.142}$$

At large positive overpotential the Faradaic current j_{ext} saturates at the limiting value $j^{\text{R}}_{\text{lim}}(\omega) = D_{\text{R}} c^*_{\text{R}} / \ell_{\text{R}}(\omega) = 1 / R^{\text{R}}_{\text{mt}} \sqrt{\omega}$, where $\ell_{\text{R}}(\omega) = \sqrt{\varepsilon D_{\text{R}} / \omega}$ is the diffusion length for species R, and similarly for O at large negatively overpotential. The numerical solution of the problem is more complicated than for the simple case with blocking electrodes because the solution is nonlocal in time. In Appendix A.3 we discuss how it can be handled by replacing the standard time evolution with a relaxation method to determine the steady-state periodic solution in time while employing the commercial finite element software FEMLAB to solve the spatial problem [47].

2.2.8 Strongly nonlinear model

Our notion of the "strongly nonlinear" regime is characterized by the following features in the solution

- The surface excess flux $\tilde{\mathbf{\Gamma}}$ in the Debye layer induces $\mathcal{O}(1)$ bulk electrolyte concentration gradients, although the leading order bulk concentration is constant in time $\bar{c} = \bar{c}(\mathbf{r})$.
- In the diffusion layer the perturbations to the local electrolyte concentration caused by the periodic massive uptake and release of ions in the Debye layer reach $\mathcal{O}(1)$.
- The Debye layer remains in quasiequilibrium but the solution depends on the local electrolyte concentration in the diffusion layer immediately outside the Debye layer.
- The fluid motion has components from both electroosmotic slip from the Debye layer, an additional effective slip from space charge in the diffusion layer, and bulk electroconvection.
- We disregard Faradaic electrode reactions because the problem is complicated enough without.

In the bulk we have four variables, namely, the potential $\bar{\phi}(\mathbf{r}, t)$, concentration $\bar{c}(\mathbf{r})$, and fluid velocity and pressure $\bar{\mathbf{u}}(\mathbf{r})$ and $\bar{p}(\mathbf{r})$, that are governed by

$$0 = \boldsymbol{\nabla} \cdot \bar{\mathbf{J}}, \tag{2.143}$$
$$0 = \boldsymbol{\nabla} \cdot \bar{\mathbf{F}}_a, \tag{2.144}$$
$$0 = -\boldsymbol{\nabla}\bar{p} + \boldsymbol{\nabla}^2\bar{\mathbf{u}} + \langle \boldsymbol{\nabla}^2\bar{\phi}\boldsymbol{\nabla}\bar{\phi}\rangle, \tag{2.145}$$
$$0 = \boldsymbol{\nabla} \cdot \bar{\mathbf{u}}, \tag{2.146}$$

where the electric current $\bar{\mathbf{J}}$ and ambipolar flux $\bar{\mathbf{F}}_a$ are given by

$$\bar{\mathbf{J}} = -\bar{c}\boldsymbol{\nabla}\bar{\phi} - \Delta D\boldsymbol{\nabla}\bar{c}, \tag{2.147}$$
$$\bar{\mathbf{F}}_a = -D_a\boldsymbol{\nabla}\bar{c} + Pe\,\bar{\mathbf{u}}\,\bar{c}. \tag{2.148}$$

In the Debye layer we also have four dynamical variables, namely, the accumulated charge $\tilde{q}$ and excess neutral salt $\tilde{w}$, together with the normal current $\bar{J}_n$ and salt flux $\check{F}_n$ that act as Lagrange multipliers in the model

$$\partial_t\tilde{w} = -\check{F}_n - \varepsilon\boldsymbol{\nabla}_s \cdot \tilde{\mathbf{\Gamma}}, \tag{2.149}$$
$$\partial_t\tilde{q} = -\bar{J}_n - \varepsilon\boldsymbol{\nabla}_s \cdot \tilde{\mathbf{j}}, \tag{2.150}$$
$$V_{\text{ext}} - \bar{\phi} = \tilde{\zeta} - \tilde{q}\,\delta, \tag{2.151}$$
$$\tilde{w} = \sqrt{\tilde{q}^2 + 4\check{c}} - \sqrt{4\check{c}}, \tag{2.152}$$

where

$$\tilde{\zeta} = -2\sinh^{-1}\big(\tilde{q}/2\sqrt{\check{c}}\big), \tag{2.153}$$

and we neglected the potential drop $\check{\zeta}$ across the diffusion layer in the condition Eq. (2.151) for the overall potential drop over the boundary layers because $\check{\zeta}$ is small compared to $-\tilde{q}\,\delta$. In the expressions for the surface excess current $\tilde{\mathbf{j}}$ and neutral salt flux $\tilde{\mathbf{\Gamma}}$ we drop the diffusion term proportional to $\boldsymbol{\nabla}_s \log \check{c}$ because this is negligible compared to $\boldsymbol{\nabla}_s\check{\phi}$ at large voltage, and moreover approximate the latter by $\boldsymbol{\nabla}_s\bar{\phi}$ because $\check{\zeta}$ is small; hence

$$\tilde{\mathbf{j}} = -\big[(1+2Pe)\tilde{w} + \Delta D\tilde{q}\big]\boldsymbol{\nabla}_s\bar{\phi}, \tag{2.154}$$

$$\tilde{\mathbf{\Gamma}} = -\big[(1+2Pe)\tilde{q} + \Delta D\tilde{w}\big]\boldsymbol{\nabla}_s\bar{\phi}. \tag{2.155}$$

The Lagrange multipliers $\bar{J}_n$ and $\check{F}_n$ determine the boundary conditions for the bulk current and ambipolar flux running into the Debye layer

$$\mathbf{n}\cdot\bar{\mathbf{J}} = \bar{J}_n, \tag{2.156}$$

$$\mathbf{n}\cdot\bar{\mathbf{F}}_a = \langle\check{F}_n\rangle - \Delta D\langle\bar{J}_n\rangle, \tag{2.157}$$

and the Debye layer also gives rise to an electroosmotic slip velocity of

$$\tilde{\mathbf{u}}_s = \big\langle\tilde{\zeta}\boldsymbol{\nabla}_s\bar{\phi}\big\rangle, \tag{2.158}$$

where again we dropped the term proportional to $\boldsymbol{\nabla}_s \log \check{c}$ and approximated the tangential field by $-\boldsymbol{\nabla}_s\bar{\phi}$.

In the diffusion layer we have just one dynamical variable, namely, the concentration $\check{c}$, that is determined by

$$\check{c}(\check{y},t) = \bar{c} + \sqrt{\frac{\varepsilon}{\omega D_a}}\frac{1}{T}\int_0^T \mathcal{G}(\check{y},t-t')\check{f}_a(t')\,\mathrm{d}t', \tag{2.159}$$

where the ambipolar flux entering $\check{f}_a$ is expressed in terms of $\bar{J}_n$ and $\check{F}_n$ as

$$\check{f}_a = \check{F}_n - \Delta D\bar{J}_n - \langle\check{F}_n - \Delta D\bar{J}_n\rangle. \tag{2.160}$$

With this, the excess potential drop across the diffusion layer can be evaluated

$$\check{\zeta} = \sqrt{\varepsilon}\bar{J}_n\int_0^\infty\left[\frac{1}{\check{c}} - \frac{1}{\bar{c}}\right]\mathrm{d}\check{y} - \Delta D\log(\check{c}/\bar{c}), \tag{2.161}$$

in terms of which the effective slip velocity induced in the diffusion layer becomes

$$\check{\mathbf{u}}_s = \big\langle\check{\zeta}\boldsymbol{\nabla}_s\bar{\phi}\big\rangle. \tag{2.162}$$

Now, we already argued that $\check{\zeta}$ is small so one might ask if $\check{\mathbf{u}}_s$ is not negligible compared to the electroosmotic slip $\tilde{\mathbf{u}}_s$ from the Debye layer. The

point is that $\tilde{\zeta}$ scales only as $\tilde{\zeta} \sim 2\log\tilde{q} \sim 2\log(V_0/\delta)$ at large voltage, cf. Eqs. (2.151) and (2.153), which is also a relatively weak dependence. Moreover, when the Gouy–Chapman solution in the Debye layer breaks down due to steric exclusion there may be viscoelectric effects suppressing the electroosmotic flow. While a description of that is beyond the scope of this thesis, it is interesting to take into account other sources of electrokinetic flow that could start to dominate in that situation.

In order to simplify the numerical solution we do, however, assume that the component of the bulk flow induced by $\check{\mathbf{u}}_s$ can be neglected in the convection of neutral salt in Eq. (2.148). Upon this approximation, only the concentration $\check{c}$ at $\check{y} = 0$ enters the coupled dynamical problem for $\{\bar{\phi}, \bar{c}, \bar{\mathbf{u}}, \bar{p}, \tilde{q}, \tilde{w}, \bar{J}_n, \check{F}_n, \check{c}\}$, which is convenient. Once this problem is solved, we can compute the full profile for $\check{c}$ as a function of $\check{y}$ and evaluate $\check{\zeta}$ along the electrode surface, evaluating the integral over $\check{y}$ by numerical quadrature, and determine the slip velocity $\check{\mathbf{u}}_s$ and the corresponding bulk flow component as a postprocessing step.

Finally, because the bulk concentration profile is controlled only by flux (Neumann) boundary conditions we need an additional constraint to fix the total number of ions in the system, namely,

$$\int_\Omega \bar{c}\,\mathrm{d}\mathbf{r} + \varepsilon \int_{\partial\Omega} \langle\tilde{w}\rangle\,\mathrm{d}s = \int_\Omega c^*\,\mathrm{d}\mathbf{r} = |\Omega|, \tag{2.163}$$

because the nominal bulk concentration is exactly $c^* = 1$; the excess concentration from the diffusion layer does not enter since it has no time average by construction. Such a constraint would be redundant in an ordinary time evolution problem, where the total amount of salt is fixed by the initial condition, but it is necessary in our case because we compute the steady-state periodic solution directly using a relaxation method. Alternatively, if the system is not closed but open and connected to a large reservoir of electrolyte at unit concentration, the total amount of ions is of course *not* fixed, but the salt concentration level is determined by a Dirichlet condition $\bar{c} = 1$ at the reservoir boundary. Likewise we need to impose a constraint on the total amount of charge accumulated in the Debye layers

$$\int_{\partial\Omega} \tilde{q}\,\mathrm{d}s = \int_{\partial\Omega} \tilde{q}^{\mathrm{eq}}\,\mathrm{d}s, \tag{2.164}$$

where $\tilde{q}^{\mathrm{eq}} = -2\sinh(\tilde{\zeta}^{\mathrm{eq}}/2)$ is the equilibrium intrinsic surface charge. In the special case where the ionic diffusivities are equal, $\Delta D = 0$, and the intrinsic zeta potential vanishes, $\tilde{\zeta}^{\mathrm{eq}} = 0$, the problem has additional symmetry in time: The bulk potential $\bar{\phi}$ and $\tilde{q}$ and $\bar{J}_n$ in the Debye layer have odd symmetry across one half period in time, e.g., $\bar{\phi}(\mathbf{r}, t + T/2) = -\bar{\phi}(\mathbf{r}, t)$, whereas the rest of the solution components have even symmetry, e.g., $\bar{c}(\mathbf{r}, t + T/2) = \bar{c}(\mathbf{r}, t)$.

2.2.9 Beyond quasiequilibrium

In this section we first develop an estimate for the driving voltage at which the weakly nonlinear model breaks down, namely, $V_0 \sim \sqrt{\delta/\varepsilon}$. Around this voltage both the bulk concentration gradients and the perturbations to the bulk concentration in the diffusion layer reach $\mathcal{O}(1)$. As the driving voltage is increased the electrolyte concentration in the diffusion layer approaches zero, which makes the assumption of local charge neutrality break down and with that also the leading order strongly nonlinear solution. The voltage where this occurs is *also* around $V_0 \sim \sqrt{\delta/\varepsilon}$, so it would seem that the strongly nonlinear model applies only to a relatively narrow voltage range, although certainly an experimentally relevant one. At even larger voltage the double layer is driven out of quasiequilibrium and expands to form an extended space charge layer: We give a brief summary of the existing theory for the structure of this layer in the case of DC Faradaic conduction, and discuss how the key features in this solution could be incorporated into our strongly nonlinear model to extend its validity even beyond quasieuilibrium.

The common way of expressing the condition for validity of the weakly nonlinear description of an electrokinetic system is in terms of the Dukhin number $Du = \sigma_s/\sigma\ell_0 \sim \left|\varepsilon\tilde{\mathbf{j}}\right| / \left|\bar{\mathbf{J}}\right|$, defined as the ratio of the double layer surface excess conductance σ_s to the bulk conductivity σ per geometrical length ℓ_0 [58]: Surface conductance and distortion of the double layer can be neglected only when $Du \ll 1$. For a symmetric binary electrolyte it can be expressed as, cf. Eq. (2.154)

$$Du = \varepsilon(1+2Pe)\tilde{w} = \varepsilon(1+2Pe)4\sinh^2\bigl(\tilde{\zeta}/4\bigr). \tag{2.165}$$

Bazant *et al.* showed that the same dimensionless group governs local salt depletion in the diffusion layer in the dynamical problem of a suddenly applied voltage, and they used the low voltage estimate $\tilde{\zeta} \simeq V_0/(1+\delta)$ to argue that for weakly nonlinear dynamics to hold the applied voltage cannot greatly exceed the thermal voltage [37].

In colloid science it is natural to express the Dukhin number in terms of the zeta potential because for a colloidal particle the electrokinetic properties are often determined by the intrinsic zeta potential due to bound surface charges when the particle reacts chemically with the suspending electrolyte. However, for electrokinetic problems involving microelectrodes that are biased by an external voltage, and also for initially uncharged polarizable particles subjected to an external DC or AC electric field, the induced zeta potential depends on the driving voltage in a nonlinear way and it is preferable to express the conditions for validity directly in terms of V_0 rather than $\tilde{\zeta}$. To this end we observe that at large voltage the surface capacitance is dominated by the Stern layer so the accumulated charge can be estimated as

$$\tilde{q} \sim \frac{V_0}{\delta}, \tag{2.166}$$

provided that the driving frequency is at or below the inverse RC time such the screening is not negligible, and that the intrinsic surface charge $\tilde{q}^{\mathrm{eq}} = -2\sinh\bigl(\tilde{\zeta}^{\mathrm{eq}}/2\bigr)$ can be disregarded. Eq. (2.166) is an upper bound on $\tilde{q}$ where a more accutate estimate would be

$$\tilde{q} \sim \frac{V_0}{\delta + \operatorname{sech}(\tilde{\zeta}/2)}, \tag{2.167}$$

with $\tilde{\zeta}$ understood to be a function of $\tilde{q}$. If steric effects are taken into account the Debye layer capacitance becomes a non-monotonic function of the accumulated charge, but $\tilde{q} \lesssim V_0/\delta$ remains valid also in that case. The excess accumulated salt is bounded by $\tilde{w} \leq |\tilde{q}|$ by construction and at large voltage where the absorption is dominated by uptake of coions rather than expulsion of counterions we have $\tilde{w} \simeq |\tilde{q}|$. The Dukhin number can therefore be estimated as

$$Du \sim \varepsilon(1+2Pe)\tilde{w} \sim \frac{\varepsilon V_0}{\delta}, \tag{2.168}$$

assuming the Péclet number is $\mathcal{O}(1)$. This shows that the surface excess current should become significant only at very large voltage $V_0 \sim \delta/\varepsilon$. However, because the bulk salt flux $\bar{\mathbf{F}}$ is only $\mathcal{O}(1)$, the surface excess salt flux $\tilde{\mathbf{\Gamma}}$ becomes signficant when the ratio

$$\frac{\bigl|\varepsilon\tilde{\mathbf{\Gamma}}\bigr|}{\bigl|\,\bar{\mathbf{F}}\,\bigr|} \sim \bigl|\varepsilon(1+2Pe)\tilde{q}\boldsymbol{\nabla}_s\bar{\phi}\,\bigr| \sim \frac{\varepsilon V_0^2}{\delta} \tag{2.169}$$

reaches $\mathcal{O}(1)$, which occurs already for $V_0 \sim \sqrt{\delta/\varepsilon}$. Likewise, if an amount of ions $\varepsilon\tilde{w}$ (per unit area) is taken up by the Debye layer from a diffusion layer of width $\sqrt{\varepsilon/\omega}$, the perturbation to the concentration in the diffusion layer can be estimated as

$$\check{\gamma} \sim \frac{\varepsilon\tilde{w}}{\sqrt{\varepsilon/\omega}} \sim \frac{\sqrt{\varepsilon\omega}\,V_0}{\delta}, \tag{2.170}$$

reaching $\mathcal{O}(1)$ for $V_0 \sim \delta/\sqrt{\varepsilon\omega}$, or, for $V_0 \sim \sqrt{\delta/\varepsilon}$ assuming the system is driven around the inverse RC time with $\omega \sim \delta$. If ε is small it is clear that the weakly nonlinear dynamics do not break down before the V_0 significantly exceeds unity, i.e., not before the driving voltage is well above the thermal voltage. This argument is consistent with our numerical results for pairs and arrays of microelectrodes, and also with those of Chu *et al.* for an uncharged metal particle subjected to a DC electric field where $O(1)$ perturbations to the concentration were reached at a dimensionless field strength of $\bar{E} \simeq 5$ for dimensionless Debye layer thickness $\varepsilon = 0.01$ and unit surface capacitance ratio $\delta = 1$ [38, 39].

The strongly nonlinear model both in the bulk and the diffusion layer is based fundamentally on local electroneutrality. The leading order charge

density in the bulk is nominally to second order in ε and for a symmetric electrolyte it can be expressed simply as, cf. Eq. (2.60)

$$\bar{\rho} = -\varepsilon^2 \boldsymbol{\nabla}^2 \bar{\phi} = -\varepsilon^2 \frac{\boldsymbol{\nabla}\bar{c} \cdot \bar{\mathbf{J}}}{\bar{c}^2}. \tag{2.171}$$

Obviously, as $\bar{c}$ approaches zero the charge density diverges which is inconsistent with the underlying assumption of electroneutrality; that assumption breaks down when the charge density becomes comparable to the average concentration, i.e., for $\bar{\rho} \sim \bar{c}$, which occurs when

$$\bar{c} \sim \varepsilon^2 \frac{\boldsymbol{\nabla}\bar{c} \cdot \bar{\mathbf{J}}}{\bar{c}^2} \quad \text{or} \quad \bar{c} \sim \varepsilon^{2/3} \left| \boldsymbol{\nabla}\bar{c} \cdot \bar{\mathbf{J}} \right|^{1/3}. \tag{2.172}$$

In an electrochemical cell the "diffusion-limited" DC Faradaic current is $\mathcal{O}(1)$ which leads to the conclusion that the electroneutral bulk solution breaks down for $\bar{c} \sim \varepsilon^{2/3}$ [60, 54]. However, in our case of strong AC forcing applied to blocking electrodes the Ohmic currents involved and the average ion flux fed into the diffusion layer at the electrode surface are much larger: Taking $\bar{\mathbf{J}} \sim V_0$ and using $-\partial_y \check{c} = \check{F}_y \sim \omega\tilde{w}$ we obtain

$$|\partial_y \check{c} \bar{J}_y| \sim \frac{\omega V_0^2}{\delta}. \tag{2.173}$$

In effect, the leading order strongly nonlinear solution based on electroneutrality breaks down already for

$$\check{c} \sim (\omega\varepsilon)^{1/3}, \tag{2.174}$$

taking $V_0 \sim \sqrt{\delta/\varepsilon}$.

Nonequilibrium double layer structure

When the electroneutral solution breaks down, the structure of the double layer changes from quasiequilibrium to a nonequilibrium structure and expands in width from $\mathcal{O}(\varepsilon)$ to $\mathcal{O}(\varepsilon^{2/3})$; this was first described by Smyrl and Newman [60]. Rubinstein and Shtilman later showed that a region completely depleted of counterions can develop at an electrode or ion exchange membrane when driven above the diffusion-limited current [40]. In this regime one can identify three sublayers within the nonequilibrium double layer, namely [42]

- A quasiequilibrium "Debye" layer of width $\mathcal{O}(\varepsilon)$ at the electrode surface.
- An extended "space charge" layer of width $y_o > \mathcal{O}(\varepsilon^{2/3})$ that is completely depleted of counterions.

- A "Smyrl–Newman" transition layer of width $\mathcal{O}(\varepsilon^{2/3})$ around $y = y_o$ connecting the space charge layer to the quasi-electroneutral diffusion layer.

To our knowledge this fundamental structure of the nonequilibrium double layer has only been studied for electrochemical systems in steady state with a DC Faradaic current, although it was argued by Bazant *et al.* that a similar solution with a transient space charge layer should form also in systems driven dynamically at very large voltage [37]. Our numerical solution to the full Poisson–nernst–Planck equations in Chap. 4 show that such a layer does indeed form.

In the existing theory for DC Faradaic conduction, the Poisson–Nernst–Planck equations can be reduced to a single master equation for the electric field whose solution can be expressed in terms of Painlevé trancendents [42, 60, 66]. In the space charge layer migration dominates and the solution for the electric field reduces to (denoting variables by a breve accent)

$$\partial_y \breve{\phi} = \pm \frac{1}{\varepsilon} \sqrt{\mp 2 \bar{J}_n (y_o - y)/D_{\pm}}, \tag{2.175}$$

where the positive sign applies to a space charge layer completely depleted of anions ($\breve{c}_+ \neq 0$ and $\breve{c}_- = 0$) and the negative sign to depletion of cations, and the thichkess y_o of the layer arises simply as an integration constant. The potential drop across the space charge layer is found by integrating Eq. (2.175) to get

$$\breve{\Phi} = \breve{\phi}(0) - \breve{\phi}(y_o) = \mp \frac{2}{3\varepsilon} \sqrt{\frac{2|\bar{J}_n|}{D_{\pm}}} y_o^{2/3}, \tag{2.176}$$

and the (small) charge density from the coions by differentiation

$$\breve{\rho} = -\varepsilon^2 \partial_y^2 \breve{\phi} = \pm \frac{\varepsilon}{2} \sqrt{\frac{2|\bar{J}_n|}{D_{\pm}(y_o - y)}}. \tag{2.177}$$

The ions in the inner $\mathcal{O}(\varepsilon)$ quasiequilibrium Debye layer are Boltzmann distributed and the solution for the potential takes the form [66]

$$\tilde{\phi} = \pm \log \left[\sinh^2(a\tilde{y} + b) \right] + \text{const.} \tag{2.178}$$

where a is determined by matching to the field in the space charge layer

$$a = \sqrt{|\bar{J}_n|/2D_{\pm}}, \tag{2.179}$$

and b is related to the total amount of charge $\tilde{q}$ accumulated in *both* the Debye and space charge layers by

$$\tilde{q} = \pm 2a \coth(b). \tag{2.180}$$

The excess potential drop across the $\mathcal{O}(\varepsilon)$ Debye layer can be determined as

$$\tilde{\zeta} = \mp 2 \log\left(\frac{1}{2} + \frac{|\tilde{q}|}{4a}\right). \tag{2.181}$$

Finally, in the Smyrl–Newman transition layer $|y - y_o| \leq \mathcal{O}(\varepsilon^{2/3})$ all terms in the master equation play in, but at least the fundamental form of the solution is an invariant function of the rescaled spatial variable [42]

$$\grave{y} = (y - y_o)\frac{|\bar{J}_n/D_\pm|^{1/3}}{\varepsilon^{2/3}}. \tag{2.182}$$

When we try to apply this solution structure in a dynamical rather that steady-state setting there are several things we need to worry about, e.g., in the space charge layer the migration term must dominate over $\partial_t \check{c}_\pm$ in the Nernst–Planck equation, and the detailed structure of the Smyrl–Newman transition layer might change due to the dynamic nature of $y_o(t)$. However, the most important complication seems to be that the solution structure is only valid during the *charging* of the Debye layer, but not during *decharging*. This is indicated in Eq. (2.175) where the expression under the square root is positive only when $\bar{J}_n$ has the opposite sign of the coions, i.e., when $\bar{J}_n < 0$ for a space charge layer completely depleted of anions with $\check{c}_+ \neq 0$ and $\check{c}_- = 0$, developing during charging of the negatively biased electrode, or vice versa at the positively biased electrode. During decharging the interpretation of the integration constant y_o as the thickness of the space charge layer does not apply and matching to the solution in the diffusion layer seems impossible. Moreover, at the end of the charging period when $\bar{J}_n$ approaches zero before it changes sign, the thickness of the $\mathcal{O}(\varepsilon^{2/3}/|\bar{J}_n|^{1/3})$ Smyrl–Newman transition layer diverges. It is not clear to us how the dynamics at this point should be properly described, but from our numerical solutions in Chap. 4 it appears that the space charge layer collapses rapidly when the current is reversed.

Ad hoc modification to strongly nonlinear model

In order to incorporate the key features from the nonequilibrium double layer into our strongly nonlinear model we attempt the following modifications:

- The breakdown of the quasiequilibrium solution is identified with the point where the concentration $\check{c}$ in the diffusion layer drops to zero at the electrode surface. Beyond this point we *continue* to solve a simple diffusion problem for the concentration profile in the diffusion layer, i.e., we maintain Eq. (2.159) even if that results in unphysical negative concentrations close to the electrode surface.

- Based on the solution for $\check{c}$ we determine $y_o(t)$ as the width of the region where $\check{c}$ is negative during charging of the Debye layer. However,

we always set $y_o = 0$ during decharging to model a rapid collapse of the space charge layer when the current is reversed.

- The nonequilibrium zeta potential is modelled as

$$\tilde{\zeta} = -2\sinh^{-1}\left[\frac{\tilde{q}}{2\sqrt{\underline{\check{c}} + 8|\bar{J}_n/D_\pm|y_o}}\right], \tag{2.183}$$

where $\underline{\check{c}}$ is a strictly positive model for the concentration at the surface

$$\underline{\check{c}} = \max\{\check{c}(0,t), \varepsilon^{1/3}\}. \tag{2.184}$$

Here $\varepsilon^{1/3}$ is a small number introduced to avoid division by zero at the point where $\check{c}(0,t)$ and $y_o(t)$ vanish simultaneously, and the exponent is chosen in accordance with Eq. (2.174). This approximation for $\tilde{\zeta}$ is identical to the quasiequilibrium solution for $\check{c}(0,t) \geq \varepsilon^{1/3}$ and asymptotically equal to Eq. (2.181) for $|\bar{J}_n/D_\pm|y_o \gg \varepsilon^{1/3}$.

- When computing the potential drop $\check{\varsigma}$ across the diffusion layer in Eq. (2.161) we only consider the interval where $\check{c} \geq \underline{\check{c}}$.

- We include the potential drop $\breve{\Phi}$ across the space charge layer as given by Eq. (2.176) into the boundary condition for the bulk potential

$$V_{\text{ext}} - \bar{\phi} = \breve{\Phi} + \tilde{\zeta} - \tilde{q}\,\delta, \tag{2.185}$$

but continue to exclude $\check{\varsigma}$.

- The relation between the surface excess accumulated charge and neutral salt remains

$$\tilde{w} = \sqrt{\tilde{q}^2 + 4\max\{\check{c}, 0\}} - \sqrt{4\max\{\check{c}, 0\}}, \tag{2.186}$$

i.e., we simply impose $\tilde{w} = |\tilde{q}|$ in the nonequilibrium double layer completely depleted of counterions.

- The surface conservation laws can still be applied to the nonequilibrium double layer provided the thickness is small enough, $y_o \ll 1$. The (excess) surface flux is dominated by the $\mathcal{O}(\varepsilon)$ Debye layer because the concentration in the space charge layer is small in comparison. Since the Debye layer is in quasiequilibrium the electrochemical potential of the coions is constant, which allows the diffusion and migration part of the surface flux to be evaluated exactly as in Sec. 2.2.5 to get $-\tilde{w}_\pm \boldsymbol{\nabla}_s \breve{\mu}_\pm$. However, the electrochemical potential is *not* constant across the space charge layer, so $\breve{\mu}_\pm \neq \bar{\mu}_\pm$. What we do assume is that $\breve{\mu}_\pm$ is dominated by the electrical term so that we can write $\breve{\mu}_\pm \approx \pm\breve{\phi}$. The convection part of the surface flux is more difficult to evaluate, but for simplicity we continue to use $\tilde{\boldsymbol{\Gamma}}_\pm = \mp\tilde{w}_\pm(1 + 2Pe)\boldsymbol{\nabla}_s\breve{\phi}$, which is consistent with our assumption that the Helmholtz–Smoluchowski term dominates the electroosmotic flow of the first kind.

- For the electrosmotic slip induced in the nonequilibrium double layer we have both the usual component of the "first kind" from the $\mathcal{O}(\varepsilon)$ quasiequilibrium Debye layer

$$\tilde{\mathbf{u}}_s = \tilde{\zeta}\boldsymbol{\nabla}_s\breve{\phi}, \tag{2.187}$$

  but also an additional component of the "second kind" from the space charge layer, for which we use the result of Rubinstein *et al.* [43]

$$\breve{\mathbf{u}}_s = -\frac{y_o^3}{9\varepsilon^3 D_\pm}\boldsymbol{\nabla}_s|\bar{J}_n| + \breve{\Phi}\boldsymbol{\nabla}_s\bar{\phi}. \tag{2.188}$$

  Here the first term is expected to dominate at very large voltage, but the second Helmholtz–Smoluchowski type term is significant at more realistic finite voltages.

Clearly, there are several crude approximations in this approach. Especially our model for the transport in the diffusion layer is offending because of the region of negative concentration, which corresponds to ions absorbed by the double layer that were really not there, but should have been taken up from the region $y \geq y_o$. This effectively means that we systematically underestimate y_o, and therefore also $\breve{\Phi}$ and $\breve{\mathbf{u}}_s$. A more correct approach would be to solve a diffusion problem for $\check{\gamma}$ in the dynamically changing domain $\check{y}_o(t) \leq \check{y} < \infty$ with the flux boundary condition $-D_a\partial_{\check{y}}\check{\gamma} = \sqrt{\varepsilon}\check{f}_a$ applied at $\check{y}_o = \sqrt{\varepsilon}y_o$. However, this problem is significantly more challenging to implement numerically than the simple form Eq. (2.159) so we have chosen not to pursue it. On the contrary, to avoid the complexity involved with determining y_o from the spatial profile of $\check{c}$ dynamically and as a function of the position along the electrode surface, we make a simple estimate for it based on a quadratic extrapolation of $\check{c}$ from the value at the surface, i.e.,

$$\check{c}(\check{y},t) \approx \check{c}(0,t) + \partial_{\check{y}}\check{c}\,\check{y} + \tfrac{1}{2}\partial_{\check{y}}^2\check{c}\,\check{y}^2, \tag{2.189}$$

to get an algorithm of the form

$$\begin{aligned}
&\textbf{if } \check{c} \geq 0 \textbf{ or } \operatorname{sign}\big(\tilde{q}\,\bar{J}_n\big) \geq 0\\
&\qquad \check{y}_o \approx 0,\\
&\textbf{elseif } \partial_{\check{y}}^2\check{c} < 0\\
&\qquad \check{y}_o \approx -\check{c}/\partial_{\check{y}}\check{c},\\
&\textbf{else}\\
&\qquad \check{y}_0 \approx -\frac{\partial_{\check{y}}\check{c}}{\partial_{\check{y}}^2\check{c}} + \sqrt{\left(\frac{\partial_{\check{y}}\check{c}}{\partial_{\check{y}}^2\check{c}}\right)^2 - 2\frac{\check{c}}{\partial_{\check{y}}^2\check{c}}}\,.\\
&\textbf{end}
\end{aligned} \tag{2.190}$$

Because both $\check{c}(0,t)$, $\partial_{\check{y}}\check{c} = -\sqrt{\varepsilon}\check{f}_a/D_a$, and $\partial^2_{\check{y}}\check{c} = \partial_t\check{c}/D_a$ are readily available as solution variables at the surface, this estimate for $\check{y}_o$ is straightforward and cheap to implement numerically. In Chap. 4 we compare our simple model to a full solution of the Poisson–Nernst–Planck equations in a simple 1D geometry, and find that they agree remarkably well: Even at $V_0 = 200$ for $\varepsilon = 10^{-3}$ where the space charge layer expands a significant depth $y_o \approx 0.025$ into the diffusion layer of characteristic length $\sqrt{\varepsilon} \approx 0.03$, the rough estimate from Eq. (2.190) comes within 10% of the true thickness of the space charge layer as observed from the full numerical solution, except just around the collapse of the space charge layer.

Another issue is that the bulk concentration variations due to surface excess salt flux in the Debye layer can become so large that $\bar{c}$ approaches zero. This occurs first close to the electrode edges or other sharp features where the electric field is particularly strong. We take this as an indication that space charge layers should develop that are steady, i.e., not transient, in space but with a charge density that changes with the polarity of the driving voltage. However, it is not clear to us how the dynamics in such a situation should be described. For the driving voltages that we consider in Chap. 4 the region where $\bar{c}$ approaches zero is small, so we simply choose neglect this problem except for another *ad hoc* modification to our model:

- We replace $\bar{c}$ with the strictly positive $\underline{\bar{c}} = \max\{\bar{c}, \varepsilon^{1/3}\}$ in Eqs. (2.59) and (2.60) to avoid unphysically large predictions of bulk electroconvection and low conductivity.

2.3 Summary

In this chapter we first gave a brief overview of some of the classical theory for electrochemical transport, electrode reactions, and the standard Gouy–Chapman–Stern model for the electrical double layer. Next we constructed a mathematical model based on matched asymptotic expansions in the three nested regions, namely, the "inner" Debye layer, "middle" diffusion layer, and "outer" bulk region. In this we build on recent work by Bazant *et al.* and Chu on dynamical solutions in electrochemical systems, where our contribution has been to write an explicit semi-analytical solution (in terms of a Fourier series) for the salt concentration in the diffusion layer developing when the system is driven with an AC voltage around the inverse RC time, and also to account for electroosmotic fluid motion and convection of salt in the bulk region. We formulated separate model equations corresponding to the linear, weakly nonlinear, and strongly nonlinear regimes, which form the basis for our numerical results in Chap. 3 and 4. Finally, we derived an estimate for the driving voltage where the weakly and strongly nonlinear models break down, and also discussed how existing theory for nonequilibrium double layers in electrochemical cells driven above the diffusion-limited

current could be incorporated into our dynamical model. However, this attempt to extend the validity of our strongly nonlinear model beyond the breakdown of local electroneutrality in the diffusion layers has character of an *ad hoc* modification and lacks the mathematical rigor otherwise pursued in this chapter.

Chapter 3

Linear and weakly nonlinear analysis

In this chapter we study the standard planar AC electroosmotic pump design with an asymmetric array of interdigitated thin microelectrodes on a planar substrate, as sketched in Fig. 3.1. This design was first used by Brown *et al.* [25] and later by several other groups [15, 26, 27, 29, 30, 31, 32].

It was on a device like this that Studer *et al.* observed pumping in the "reverse" direction [15], where the "forward" direction is defined by the experimental observations of Brown *et al.* and the predictions from the linear theory of Ramos *et al.* [33]. This theory predicts a single maximum in the pumping velocity at a driving frequency around the inverse RC time, with rapid and monotonical decrease in pumping at both lower and higher frequencies. The only free adjustable parameter is the surface capacitance ratio between the Debye and Stern layers, which allows the inverse RC time and to some extent also the magnitude of the pumping velocity to be fitted to experiments. However, the theory is completely unable to account for reversal of the pumping velocity at higher frequency.

This discrepancy between experimental observations and predictions from existing theory has motivated us to extend the latter by including nonlinear capacitance in the Debye layer, and Faradaic electrode reactions in both a linearized scheme and using the full nonlinear Butler–Volmer reaction kinetics. Most of our results have been published in the paper Ref. [44], see Appendix B, but in the present chapter we take into account also the effect of mass transfer limitations on the Faradaic electrode reaction. Several groups have observed bubble formation and electrode degradation at driving frequencies below 1 kHz and voltages of a few volt [26, 28, 29, 15, 32, 31], which is experimental evidence that electrode reactions do actually occur. Another motivation for including electrode reactions into the theory was that the original results of Ajdari indicated reverse pumping in this case, although in the low, not high frequency limit [14].

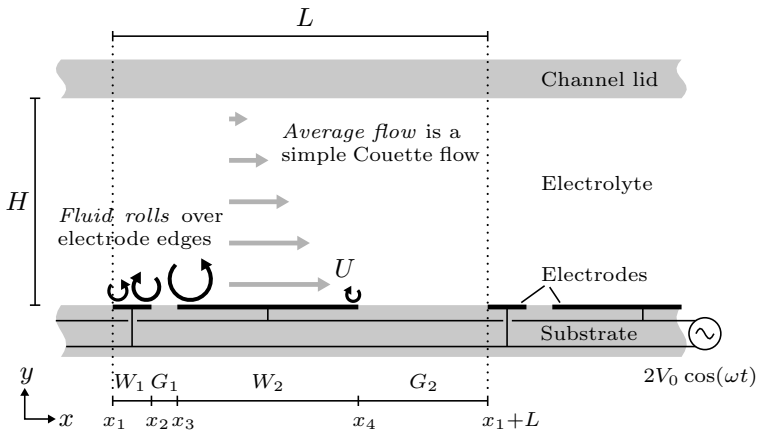


Figure 3.1: Sketch of the device geometry for the planar pump design with asymmetric pairs of interdigitated electrodes. Each pair has a narrow electrode of width W_1 and a wide electrode of width W_2 separated by a narrow and wide gap of width G_1 and G_2. The fluid flow pattern is complex with vortices above the electrode edges, but the spatial average in the x direction is a simple Couette flow profile.

Idealized geometry model

As sketched in Fig. 3.1 the device consist of an array of interdigitated electrodes, where each pair has a narrow electrode of width W_1 and a wide electrode of width W_2 separated by a narrow and a wide gap of width G_1 and G_2, respectively, and the full period is $L = W_1 + G_1 + W_2 + G_2$. Notice that in order to break the left-right symmetry it is necessary that both $W_1 \neq W_2$ and $G_1 \neq G_2$. On top of the array a microfluidic channel of height H is placed, and we assume that the transverse width of the channel is large enough that a 2D description in the xy plane is appropriate. In order to reduce the number of geometrical parameters we assume an idealized geometry with infinitely thin and flat electrodes, and a channel height $H \gg L$ large enough to not influence the solution. Brown *et al.* originally used $W_1 = 4.2$ μm, $G_1 = 4.5$ μm, $W_2 = 25.7$ μm, $G_2 = 15.6$ μm, and $L = 50$ μm, but in this chapter we study a design with (dimensionless) electrode widths and gaps $W_1 = 1.5$, $G_1 = 1$, $W_2 = 7$, $G_2 = 5$, and $L = 14.5$. As discussed in Chap. 5 this choice is close to optimal in the sense of maximizing the pumping velocity for a given (low) driving voltage, but the qualitative features in the solution are not very sensitive to the particular choice of electrode geometry.

In order to determine the net pumping we observe that because the system is periodic in the x direction, the fluid velocity $\bar{\mathbf{u}}$ can be expressed as a Fourier series [33]

$$\bar{\mathbf{u}}(x, y) = \sum_{n=0}^{\infty} \hat{\mathbf{u}}_n(y)\, e^{i2\pi n x/L} + \text{c.c.} \tag{3.1}$$

where $\hat{\mathbf{u}}_n$ is the nth Fourier component. The net flow rate is given by

$$Q = \int_0^H \bar{u}_x \, \mathrm{d}y, \tag{3.2}$$

and due to incompressibility Q must be independent of x, i.e., only the zeroth order Fourier component $\hat{\mathbf{u}}_0$ can contribute to the net flow rate. The solution for $\hat{\mathbf{u}}_0$ is simply a Couette profile, $\hat{\mathbf{u}}_0 = y\,U/H\,\mathbf{e}_x$, where U is the spatial average slip velocity

$$U = \frac{1}{L} \int_{\text{electrodes}} \tilde{u}_s \, \mathrm{d}x, \tag{3.3}$$

and the net flow rate becomes $Q = HU/2$. When the channel height is sufficiently large $H \gg L$ to not affect the electric field line distribution around the electrodes, the slip velocity and hence U becomes independent of H, whereas Q grows linearly with H. To quantify the pump performance it is therefore convenient to report the "pumping velocity" U instead of the net flow rate Q.

In Ref. [33] Ramos *et al.* carried out a complete Fourier analysis in the limit $H \to \infty$ and determined all components $\hat{\mathbf{u}}_n(y)$ in terms of the slip velocity distribution $\tilde{u}_s(x)$. They also expressed the streamfunction Ψ for the flow in terms of the Green's function $\mathcal{H}(x - x', y)$ for an array of fluid velocity line sources

$$\Psi(x, y) = \frac{1}{L} \int_0^L \mathcal{H}(x - x', y)\, \tilde{u}_s(x') \, \mathrm{d}x', \tag{3.4}$$

where the Green's function is given by

$$\mathcal{H}(x - x', y) = \frac{y \sinh\bigl(2\pi y/L\bigr)}{\cosh\bigl(2\pi y/L\bigr) - \cos\bigl(2\pi(x - x')/L\bigr)}. \tag{3.5}$$

We adopt this for calculating the streamfunction and visualizing the flow pattern around the electrodes also at finite $H \gg L$.

Importantly, this Fourier analysis applies not only to the low voltage limit, but also in the weakly nonlinear regime, or more generally to any slip velocity driven flow above a planar surface. Moreover, the linearity of the Stokes flow problem allows us to include (back) pressure driven flow simply by superposition of a parabolic Poiseuille flow profile [25, 31].

3.1 Linear analysis at low voltage

In the Debye–Hückel limit $\zeta \ll k_B T/ze$ the electrical problem can be linearized, and the dimensionless boundary condition for the Ohmic current running into the Debye layer becomes, cf. Eq. (2.124)

$$-\mathbf{n} \cdot \boldsymbol{\nabla} \hat{\phi} = \left(\frac{i\omega}{1 + \delta} + \frac{1}{R_{\text{ct}} + R_{\text{mt}}/\sqrt{i\omega}} \right) (\hat{V}_{\text{ext}} - \hat{\phi}), \tag{3.6}$$

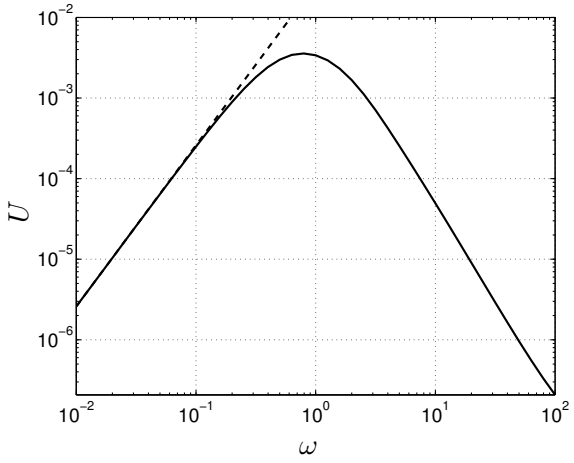


Figure 3.2: Pumping velocity U as a function of frequency ω. The pumping is maximized around the inverse RC time $\omega \sim 1$ when the screening of the electrodes is partial. Dashed line: Analytical result at low frequency where $U \propto \omega^2$.

whereas the time average electroosmotic slip velocity reduces to

$$\tilde{\mathbf{u}}_s = \langle \tilde{\zeta} \boldsymbol{\nabla}_s \bar{\phi} \rangle = -\frac{1}{1+\delta} \boldsymbol{\nabla}_s \left| \hat{V}_{\text{ext}} - \hat{\phi} \right|^2. \tag{3.7}$$

Here $\hat{V}_{\text{ext}} - \hat{\phi}$ is the potential drop across the double layer; for a given electrode geometry there are only five dimensionless parameters describing the problem, namely, the driving frequency ω and voltage V_0, the surface capacitance ratio δ, charge transfer resistance R_{ct}, and Warburg constant R_{mt}. The dependence on the driving voltage in the model can be scaled out by rescaling the potential with V_0 instead of the thermal voltage $k_B T/ze$, and the fluid velocity by V_0^2. Also the influence of the capacitance ratio δ is almost trivial as it simply shifts the inverse RC time from $\omega \sim 1$ to $\omega \sim 1 + \delta$, and reduces the electroosmotic slip by a factor of $(1+\delta)^{-1}$. In our paper Ref. [44] we included this into the rescaling of the problem, but here we simply set $V_0 = 1$ and $\delta = 0$ throughout the linear analysis, with the understanding that the effect of nonzero δ and other $V_0 \lesssim 1$ is trivial.

3.1.1 Capacitive charging

In the absense of Faradaic electrode reactions the only dimensionless parameter left in the problem is the driving frequency ω. At low frequency $\omega \ll 1$ the Debye layers are fully charged and the screening is complete at all times, so there is no tangential electric field to drive the electroosmotic flow. Conversely, at high frequency $\omega \gg 1$ the screening is negligible, so there is no charge in the Debye layers and again no flow. Only around the inverse RC time $\omega \sim 1$ when there is partial screening do we have both charge and

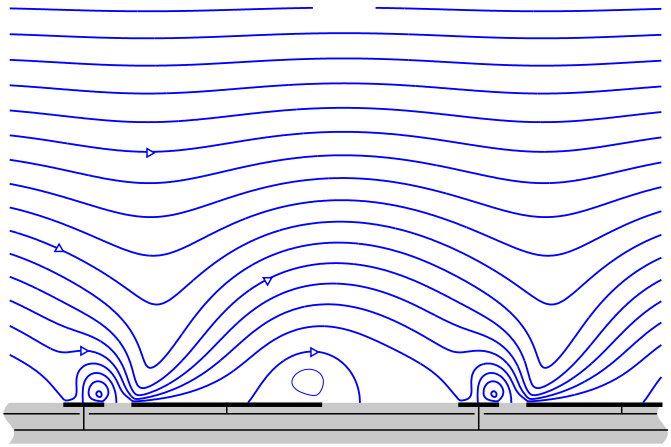


Figure 3.3: Streamlines for the time average fluid flow at $\omega = 0.8$ where the pumping velocity is maximized. The streamlines are drawn as equidistant contours of the stream-function Ψ, cf. Eq. (3.4), following Ramos *et al.* [33].

tangential field in the Debye layers. Fig. 3.2 shows the pumping velocity U as a function of ω. There is a unique maximum around $\omega \simeq 0.8$ where $U = 0.0035$, and the pumping falls off rapidly as ω^2 for $\omega \ll 1$ or as ω^{-2} for $\omega \gg 1$.[1] This scaling of the pumping velocity at low and high frequency was predicted already in the original paper by Ajdari [14] for a simple generic model of an electrode array as a harmonic potential modulation at the bottom of the channel, whereas the particular direction and magnitude was obtained by Ramos *et al.* using more realistic boundary conditions [33].

The streamline pattern for the flow at $\omega = 0.8$ is shown in Fig. 3.3: There is a uniform flow from the left to the right far above the electrodes and more tortuous streamlines close to the electrodes, including two recirculating vortices at the right edge of both electrodes. Notice that the stagnation point where the slip velocity changes sign is close to the center of the narrow electrode but is shifted more to the right on the wide electrode.

Electrical impedance

The impedance of the system can be determined as $\hat{Z} = \Delta\hat{V}/\hat{I}$, where $\Delta V = 2$ is the magnitude of the voltage difference applied between the electrodes, and $\hat{I}$ is the total current running onto the electrodes through the power supply, i.e.,

$$I(t) = \int_{\text{electrode } \#1} \left(j_{\text{ext}} - \partial_t(\tilde{q} + \varepsilon\, \mathbf{n} \cdot \boldsymbol{\nabla}\bar{\phi}\,) \right) \mathrm{d}x. \tag{3.8}$$

[1] In Fig. 3.2 it appears that U falls off faster than ω^{-2} at $\omega \gg 1$, but this is an artefact due to our idealized geometry model with infinitely thin electrodes. For a more realistic geometry with electrodes of finite thickness and radius of curvature R at the edges, the pumping U matches ω^{-2} for $\omega \gg R^{-1}$.

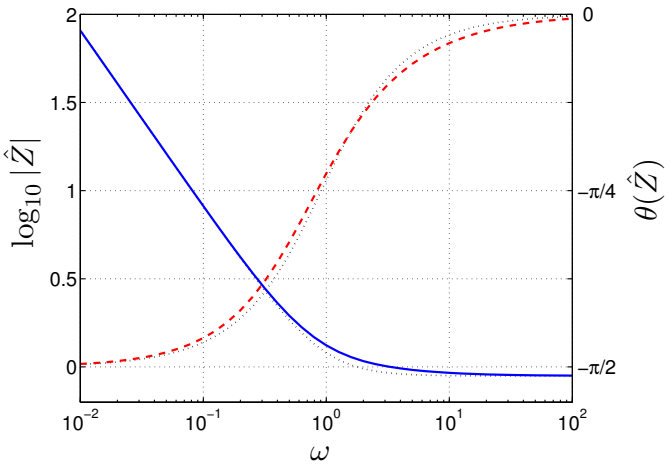


Figure 3.4: Bode plot of the dimensionless impedance $\hat{Z}$ as a function of frequency ω at zero capacitance ratio $\delta = 0$: Magnitude (solid line) and phase angle (dashed line). For comparison also the impedance $\hat{Z}_0 = R_\infty + \frac{1}{i\omega C_0}$ for a simple RC circuit with linear circuit elements is shown (dotted lines). The phase angle passes through $-\pi/4$ when the frequency is at the inverse RC time at $\omega \approx 0.8$.

Here j_{ext} is the Faradaic current passed from the electrode into the electrolyte, $-\tilde{q}$ is the charge on the electrode required to screen the accumulated charge in the Debye layer, and $-\varepsilon\,\mathbf{n}\cdot\boldsymbol{\nabla}\bar{\phi}$ is the dielectric polarization charge. For $\omega \ll \varepsilon^{-1}$ we can neglect the last term and obtain

$$\hat{I} = \left(\frac{i\omega}{1+\delta} + \frac{1}{R_{\text{ct}} + R_{\text{mt}}/\sqrt{i\omega}}\right) \int_{\text{electrode }\#1} \left(\hat{V}_{\text{ext}} - \hat{\phi}\right) \mathrm{d}x. \tag{3.9}$$

Fig. 3.4 shows a Bode plot of the dimensionless impedance of the system in the absense of Faradaic electrode reactions and for zero capacitance ratio δ. At low frequency the impedance is dominated by the capacitance C_0 of the Debye layers on the two electrodes that are connected in series, i.e.,

$$\frac{1}{C_0} = \frac{1}{W_1} + \frac{1}{W_2} = \frac{1}{1.24}, \tag{3.10}$$

whereas at high frequency the bulk Ohmic resistance $R_\infty = 0.88$ dominates.[2] For $\omega \sim 1$ we observe a slight frequency dispersion in $\hat{Z}$ as compared to a constant circuit element model $\hat{Z}_0 = R_\infty + 1/i\omega C_0$, which is due to a change

[2]At high frequency the resistance R_∞ can be computed with high accuracy from the power dissipation $P = \Delta V^2/R_\infty$ in the electrolyte

$$\frac{1}{R_\infty} = \frac{1}{\Delta V^2} \int_\Omega \boldsymbol{\nabla}\phi \cdot \boldsymbol{\nabla}\phi \,\mathrm{d}\mathbf{r} = 1.125, \tag{3.11}$$

where ϕ is a solution to the Laplace equation with a Dirichlet boundary condition $\phi = V_{\text{ext}}$ on the electrodes. Likewise, the capacitance due to dielectric polarization at high frequency

in the electric field line pattern between low and high frequency. The inverse RC time can be read off as the point when the phase angle of $\hat{Z}$ passes through $-\pi/4$, which coincides almost exactly with the frequency where the pumping velocity is maximized, namely at $\omega \approx 0.8$. This "coincidense" is not particular to our specific electrode geometry from Fig. 3.1, but holds for other choices of the electrode widths and gaps as well. Which is perhaps not surprising, given that both the impedance $\hat{Z}$ and the pumping velocity U are integral quantities over the electrodes.

3.1.2 Faradaic current injection

When we include electrode reactions and Faradaic current into our model, one of the most important new features is that the complete screening at frequencies below the inverse RC time is unbalanced, that is, provided the electrode reaction is not limited by mass transfer. In general, we expect AC electroosmotic flow whenever there is partial screening of the electrodes. Inspecting the boundary condition Eq. (3.6) there are three regimes in parameter space where the surface impedance is $\mathcal{O}(1)$, namely, (i) when capacitive charging dominates for $\omega \sim 1$ and either $R_{\mathrm{ct}} \gtrsim 1$ or $R_{\mathrm{mt}} \gtrsim 1$, (ii) when the electrode reaction kinetics dominate for $R_{\mathrm{ct}} \sim 1$ and $R_{\mathrm{mt}}^2 \leq \omega \leq 1$, and (iii) when mass transfer dominates for $\omega \sim R_{\mathrm{mt}}^2$ and $R_{\mathrm{ct}} \lesssim 1$. These different regimes are illustrated in Fig. 3.5, where the pumping velocity U is shown as a function of ω and R_{ct} for different values of R_{mt}. In Fig. 3.5(a) we have $R_{\mathrm{mt}} = 1$, so at the baseline $R_{\mathrm{ct}} \gg 1$ we recover the previous results from Fig. 3.2 for purely capacitive charging, whereas for $R_{\mathrm{ct}} \leq 1$ we see a minor downshift in the optimal frequency for pumping and a factor of 3 reduction in the magnitude of the pumping velocity. Figs. 3.5(b) to (d) show results for $R_{\mathrm{ct}} = 0.46$, 0.1, and 0.01, respectively, where we see the peak around $\omega \sim R_{\mathrm{ct}}^2$ shifting down and even outside the figure window in Fig. 3.5(d). What we do *not* see, however, is the peak in the pumping velocity for $R_{\mathrm{ct}} \sim 1$ and $R_{\mathrm{mt}}^2 \lesssim \omega \lesssim 1$. This is puzzling, in particular because there is actually electroosmotic fluid motion in this parameter range, but only no *net* pumping. It turns out that the vanishing of the pumping is particular to the case when the charge transfer resistance takes the same value on both electrodes. If R_{ct} is larger on the wide electrode than on the narrow electrode we do get pumping in the usual "forward" direction, whereas if R_{ct} is larger on the narrow electrode the pumping is in the *"reverse"* direction; see our paper in Appendix B for more details on this.

Although we have not gained an intuitive understanding of why the

can be determined from the electric field energy $\mathcal{U} = \frac{1}{2} C_\infty \Delta V^2$ as

$$C_\infty = \frac{2\varepsilon}{\Delta V^2} \int_\Omega \frac{1}{2} \boldsymbol{\nabla}\phi \cdot \boldsymbol{\nabla}\phi \, \mathrm{d}\mathbf{r} = 1.125\varepsilon. \tag{3.12}$$

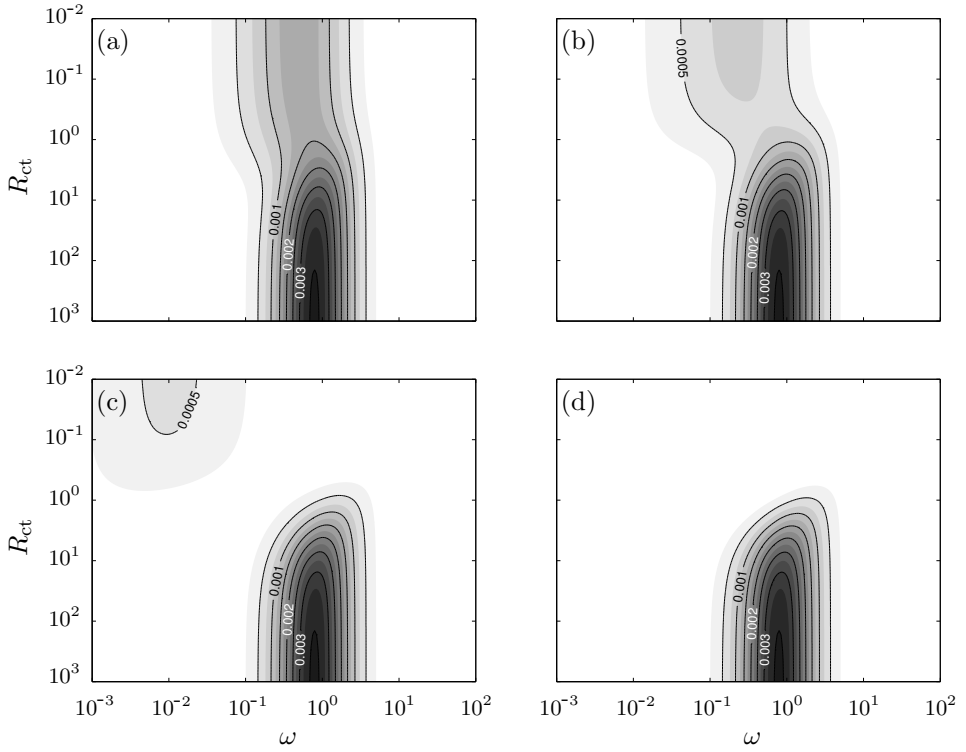


Figure 3.5: Contour plots of the pumping velocity U as a function of frequency ω and charge transfer resistance R_{ct} for different values of the Warburg constant R_{mt}: (a) $R_{mt} = 1$, (b) $R_{mt} = 0.46$, (c) $R_{mt} = 0.1$, (d) $R_{mt} = 0.01$.

Faradaic charging does not give rise to any pumping when R_{ct} is uniform, we do have a qualitative explanation for the direction of the pumping when R_{ct} takes different values on the two electrodes: The net pumping contribution from the narrow electrode tends to be in the reverse direction, whereas the wide electrode contributes in the forward direction. In the simple case with purely capacitive charging, the wide electrode dominates and the net pumping is in the forward direction. At lower frequency the Debye layer on the electrode where R_{ct} is smaller tends to be "short-circuited" by the electrode reaction, which inhibits the electroosmotic flow and leaves the other electrode dominating and determining the pumping direction. This idea can be useful in other designs as well where, e.g., one could inhibit the flow from the reverse slipping electrode by modifying the intrinsic zeta potential on the electrodes as in Sec. 3.1.3 below, by depositing a porous layer permable to ions but with a large hydraulic resistance, or by geometrically lowering the reverse slipping sections from the channel surface as suggested by Bazant and Ben [45] and investigated in detail in Sec. 5.3.

3.1.3 Linear response at nonzero intrinsic zeta potential

Many electrode-electrolyte systems spontaneously form a Debye layer and possess an intrinsic zeta potential typically of the order $\zeta^{\mathrm{eq}} \sim 100$ mV, which is well beyond the Debye–Hückel limit $\zeta \ll k_B T/ze \simeq 25$ mV. Our previous linear analysis does not immediately apply to such systems, but we can still attempt a linear response analysis if the driving voltage is low enough $V_0 \ll (1+\delta)k_B T/ze$. In the absense of Faradaic electrode reactions we obtain essentially the same results as in Sec. 3.1.1, except that the RC time for the system is determined by the differential capacitance, and the dimensionless frequency where the pumping occurs is therefore shifted down from $\omega \sim 1+\delta$ to

$$\omega \sim \left[-\frac{d\tilde{q}}{d(V_{\mathrm{ext}} - \bar{\phi})} \right]^{-1} = \operatorname{sech}(\tilde{\zeta}^{\mathrm{eq}}/2) + \delta. \tag{3.13}$$

Also the characteristic fluid velocity is proportional, not to $(1+\delta)^{-1}$ but to

$$\frac{d\tilde{\zeta}}{d(V_{\mathrm{ext}} - \bar{\phi})} = \frac{1}{1 + \delta \cosh(\tilde{\zeta}^{\mathrm{eq}}/2)} \tag{3.14}$$

which can be come very small if $\tilde{\zeta}^{\mathrm{eq}}$ is large. Of course, at very large $\tilde{\zeta}^{\mathrm{eq}}$ where the Dukhin number reaches $\mathcal{O}(1)$ we would also have to include tangential surface fluxes into the model. When Faradaic electrode reactions do occur we need to replace R_{ct} by the differential charge transfer resistance

$$R_{\mathrm{ct}}(\tilde{\zeta}^{\mathrm{eq}}) = \left[\frac{dj_{\mathrm{ext}}}{d(V_{\mathrm{ext}} - \bar{\phi})} \right]^{-1} = R_{\mathrm{ct}}(0) e^{(z_{\mathrm{O}} - \alpha)\tilde{\zeta}^{\mathrm{eq}}}. \tag{3.15}$$

However, we also need to take into account the possibility that a time average current to second order in V_0 could run between the electrodes. This would give rise to an additional component to the flow when the corresponding time average electric field interacts with the constant intrinsic zeta potential, as we discuss at the end of Sec. 3.2.2.

3.2 Weakly nonlinear analysis

The weakly nonlinear model is a first step towards better understanding of the dynamics at larger voltage, and yet there have been few theoretical investigations that go beyond the Debye–Hückel limit (or linear response) in the context of AC electrokinetics. One reason for this may be that the numerical solution is slightly more complicated: Instead of a single linear problem for the complex potential $\hat{\phi}(\mathbf{r})$ one has to solve a time evolution problem for the bulk potential $\bar{\phi}(\mathbf{r}, t)$ with a special variable $\tilde{q}(x, t)$ defined only on the electrode boundary. However, using commercial finite element software like FEMLAB [47] this is straightforward to implement, and to demonstrate this we have included a code example in Appendix A.1.

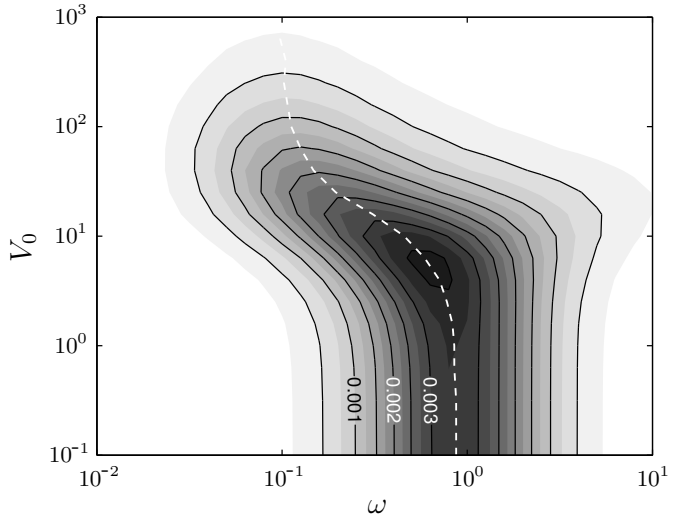


Figure 3.6: Contour plot of the rescaled dimensionless pumping velocity U/V_0^2 as a function of driving frequency ω and voltage V_0 for a capacitance ratio of $\delta = 0.1$ and no Faradaic current. The dashed white line marks the frequency where the pumping velocity is maximized as a function of V_0.

3.2.1 Nonlinear surface capacitance

In the absense of Faradaic electrode reactions there are four dimensionless parameters in the model, namely, the driving voltage V_0 and frequency ω, the capacitance ratio δ, and the intrinsic zeta potential $\tilde{\zeta}^{\mathrm{eq}}$ on the electrodes. Fig. 3.6 shows the pumping velocity U as a function of V_0 and ω for $\delta = 0.1$ and $\tilde{\zeta}^{\mathrm{eq}} = 0$, with the data rescaled to U/V_0^2 to remove the inherent quadratic voltage dependence and ease comparison to the linear analysis. At the baseline $V_0 \ll 1$ we recover the result from Fig. 3.2 with a peak in the pumping velocity for $\omega \sim 1 + \delta$. However, at higher voltage the double layer capacitance is dominated by the Stern layer, and the inverse RC time where the pumping is maximized drops to $\omega \sim \delta$. It also appears from Fig. 3.6 that the pumping velocity falls off at large voltage, but this is simply because the electroosmotic slip $\tilde{\mathbf{u}}_s = \tilde{\zeta}\boldsymbol{\nabla}_s\bar{\phi}$ no longer scales as V_0^2 at large voltage: There $\tilde{\zeta} \sim 2\log(V_0/\delta)$ rather than $\tilde{\zeta} \sim V_0/(1+\delta)$, cf. the discussion in Sec. 2.2.9, and the slip velocity grows only as $V_0 \log V_0$.

When the intrinsic zeta potential is nonzero, the pumping velocity follows the linear response analysis from Sec. 3.1.3 up to the point where the induced perturbations in $\tilde{\zeta}$ becomes comparable to $\tilde{\zeta}^{\mathrm{eq}}$, and well above that point the results are largely unaffected by $\tilde{\zeta}^{\mathrm{eq}}$.

We must emphasize that the upper voltage limit in Fig. 3.6 where V_0 extends to 10^3 is far outside the range of validity of weakly nonlinear model for a typical microsystem. Nevertheless, the conclusion that according to the weakly nonlinear model the *only* effect of having a nonlinear surface

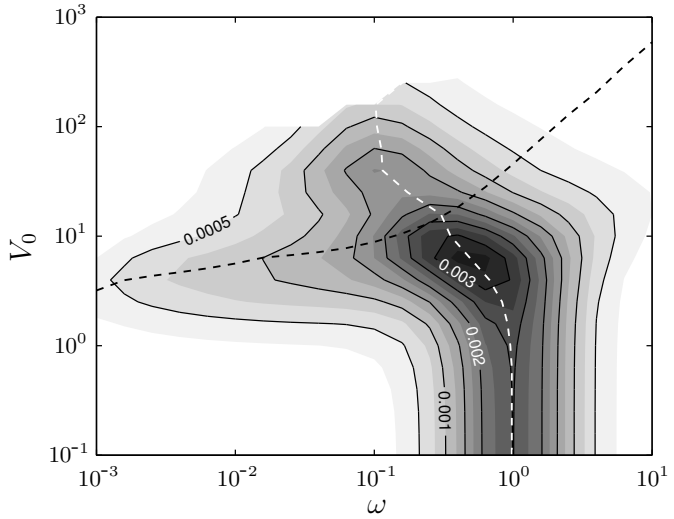


Figure 3.7: Contour plot of the rescaled dimensionless pumping velocity U/V_0^2 as a function of driving frequency ω and voltage V_0, for capacitance ratio $\delta = 0.1$ and Faradaic current with charge transfer resistance $R_{\mathrm{ct}} = 10$ and $R_{\mathrm{mt}}^{\mathrm{O}} = R_{\mathrm{mt}}^{\mathrm{R}} = 0.1$. The dashed white line marks the frequency where the pumping velocity is maximized as a function of V_0, and the dashed black line roughly marks the voltage where the limiting current is reached. Some data points at large voltage $V_0 \geq 10^2$ are missing due to numerical difficulties with obtaining a self-consistent solution for $j_{\mathrm{ext}}(t)$ in Eq. (2.141).

capacitance is to downshift the inverse characteristic RC time and to change the scaling of U with driving voltage from V_0^2 to $V_0 \log V_0$ is important as a starting point for interpreting the results from the strongly nonlinear model in Chap. 4.

3.2.2 Nonlinear Faradaic current injection

With electrode reactions included into our model, the number of dimensionless parameters increases by the charge transfer resistance R_{ct}, mass transfer coefficients $R_{\mathrm{mt}}^{\mathrm{O}}$ and $R_{\mathrm{mt}}^{\mathrm{R}}$, and the intrinsic zeta potential $\tilde{\zeta}^{\mathrm{eq}}$ on the electrodes and channel substrate. From the analysis in Sec. 3.1.2 we know that if $R_{\mathrm{mt}} \gtrsim 1$, the effect of the Faradaic electrode reaction is always relatively small, so in order to actually see a nonvanishing effect we consider $R_{\mathrm{mt}}^{\mathrm{O}} = R_{\mathrm{mt}}^{\mathrm{R}} = 0.1$. Further we take $R_{\mathrm{ct}} = 10$ and $\zeta^{\mathrm{eq}} = 0$ on the electrodes. Fig. 3.7 shows the pumping velocity U as a function of the driving voltage V_0 and frequency ω, where again we rescaled to U/V_0^2 to remove the inherent quadratic voltage dependence. At the baseline for $V_0 \ll 1$ we recover the results from Fig. 3.5(c), except that the scaling differs by a factor of $1 + \delta$. For $V_0 \sim 5$ there is a peak in the pumping velocity at low frequency which is due to Faradaic current injection. Although R_{ct} is relatively large, the exponential increase in Faradaic current with overpotential

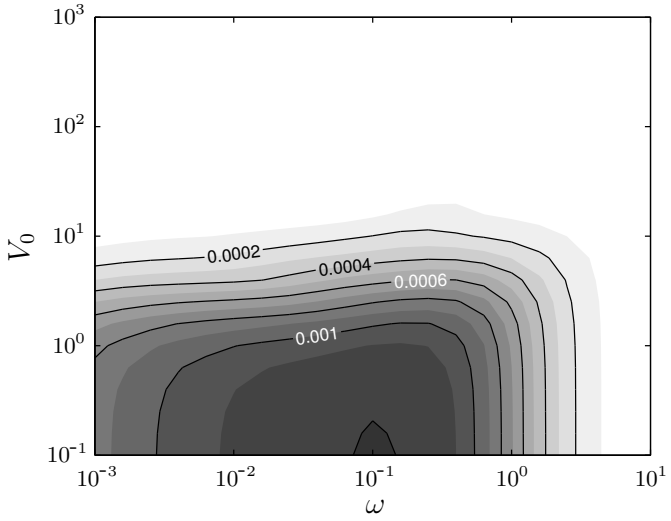


Figure 3.8: Contour plot of the rescaled average tangential field E_s/V_0^2 along the insulating substrate, cf. Eq. (3.17), for the same parameters as in Fig. 3.7. Depending on the intrinsic zeta potential $\tilde{\zeta}^{\mathrm{eq}}$ this can induce pumping in either the "forward" or "reverse" direction, but above the diffusion-limited current E_s saturates and the effect becomes negligible compared to the "ordinary" induced-charge electroosmosis.

always makes it significant for $V_0 \gg 1$. Moreover, because the potential drop across the double layer is generally larger on the narrow electrode than on the wide, the effective value of R_{ct} is smaller, i.e., the nonlinearity induces an asymmetry between R_{ct} on the two electrodes that leads to pumping in the forward direction, cf. the discussion in Sec. 3.1.2 and in more detail in our paper [44] in Appendix B. The peak fades out for $\omega \ll 10^{-2}$ due to the mass-transfer limitations on the Faradaic current, a feature which was missing in our earlier results in that paper. Specifically, the dashed black line in Fig. 3.7 roughly marks the voltage where the deviations in $\check{c}_{\mathrm{O}}$ and $\check{c}_{\mathrm{R}}$ relative to the bulk reach $\mathcal{O}(1)$; beyond this voltage the diffusion-limited current is reached, i.e., the Faradaic current saturates, and for $V_0 \gg 1$ the system behaviour is again dominated by the capacitive charging.

The nonlinearity in the reaction kinetics also implies that a DC Faradaic current can be running between the electrodes even when a harmonic AC voltage difference is applied. The corresponding DC tangential electric field $\langle -\partial_x \bar{\phi} \rangle$ will interact with any intrinsic zeta potential on the electrodes or channel substrate and induce electroosmotic flow. In the special case when $\tilde{\zeta}^{\mathrm{eq}}$ takes the same value on both electrodes and substrate, this flow does not give rise to any net pumping on a planar device because there is no global potential gradient along the electrode array, i.e.,

$$\int_0^L \langle \tilde{\zeta}^{\mathrm{eq}} \partial_x \bar{\phi} \rangle \, \mathrm{d}x = \tilde{\zeta}^{\mathrm{eq}} \int_0^L \langle \partial_x \bar{\phi} \rangle \, \mathrm{d}x = \tilde{\zeta}^{\mathrm{eq}} \big[\langle \bar{\phi} \rangle \big]_{x=0}^{L} = 0. \tag{3.16}$$

However, if $\tilde{\zeta}^{\mathrm{eq}}$ is nonuniform there can be pumping, and Fig. 3.8 shows the average tangential field along the channel substrate

$$E_s = \frac{1}{L}\int_{\text{substrate}} -\langle \partial_x \bar{\phi} \rangle \, \mathrm{d}x, \tag{3.17}$$

which gives rise to an additional component to the pumping velocity of $U' = \left(\tilde{\zeta}_e^{\mathrm{eq}} - \tilde{\zeta}_s^{\mathrm{eq}}\right) E_s$, where $\tilde{\zeta}_e^{\mathrm{eq}}$ is the constant intrinsic zeta potential on the electrodes and $\tilde{\zeta}_s^{\mathrm{eq}}$ is the corresponding value on the channel substrate. As before the data are rescaled to E_s/V_0^2 to remove the inherent voltage dependence. The figure shows that at low voltage and for $\omega \sim 0.1$ there is a maximum in the average tangential field along the substrate with $E_s = 0.0012$. Therefore, if the (difference in) dimensionless intrinsic zeta potential is larger that 2, the pumping induced by the time average field becomes comparable in magnitude to the peak pumping velocity from the "ordinary" induced-charge electroosmotic flow, while the direction of this flow component depends on the sign of $\Delta\tilde{\zeta}^{\mathrm{eq}} = \tilde{\zeta}_e^{\mathrm{eq}} - \tilde{\zeta}_s^{\mathrm{eq}}$. Beyond the voltage where the diffusion-limited current is reached, E_s saturates and therefore appears to fall off rapidly in Fig. 3.8 due to the rescaling with V_0^2. Since the induced-charge electroosmotic flow continues to grow as $V_0 \log V_0$, this means that any net pumping from E_s is negligible at large voltage.

Strictly, a DC Faradaic current implies a time average flux of O and R between the two electrodes. In our model we assumed uniform bulk concentrations $\bar{c}_{\mathrm{O}} = c_{\mathrm{O}}^*$ and $\bar{c}_{\mathrm{R}} = c_{\mathrm{R}}^*$ of the electrochemically active species, which means that we neglected any bulk mass transfer limitations. Taking such into account would reduce the value of the DC current and hence E_s, but we do not expect it to have much impact on the results in Fig. 3.7.

3.3 Summary

In this chapter we have studied the planar AC electroosmotic pump design used experimentally by several groups, extending the existing linear theory to account for Faradaic electrode reactions. When the electrode reaction is fast enough to effectively "short-circuit" the Debye layer, i.e., when the charge transfer resistance is small $R_{\mathrm{ct}} \ll 1$, the electroosmotic flow is suppressed because it depends fundamentally on partial screening of the electrodes. However, the electrode reaction can run no faster than the electrochemically active species O and R are transported to and from the electrode by diffusion, and depending on whether this mass-transfer limitation is weak or strong, the pumping velocity is either suppressed and shifted down in frequency to $\omega \sim R_{\mathrm{mt}}^2$ or virtually unaffected by the electrode reaction.

We also extended the theory into the weakly nonlinear regime, where the nonlinear capacitance of the Debye layer causes the characteristic inverse RC

time to shift down from $\omega \sim 1 + \delta$ to $\omega \sim \delta$, and the electroosmotic slip to scale as $V_0 \log V_0$ with the driving voltage rather than V_0^2. González *et al.* have reported a nonlinear theoretical study of a travelling wave device, and their results are fully equivalent to ours [67]. The nonlinearity in the Butler–Volmer electrode reaction kinetics gives rise to more dramatic results: A DC Faradaic current scaling as V_0^2 at low voltage should interact with any intrinsic zeta potential on the electrodes or channel substrate, and the direction of the induced pumping velocity depends on the sign of the (difference in) intrinsic zeta potential. At larger voltage there is pumping at low frequency due to an asymmetry in the effective charge transfer resistance between the narrow and wide electrode. Ultimately, when the diffusion-limited current is reached and the concentration of O and R (dynamically) approaches zero at the electrodes, the Faradaic current saturates and our weakly nonlinear model becomes dominated by capacitive charging. We should stress that this last conclusion is valid only because we assume an excess of inert supporting electrolyte, while otherwise the Faradaic current would induce bulk concentration gradients and strong concentration polarization which we consider "strongly nonlinear" phenomena.

In many aspects the weakly nonlinear model does not differ qualitatively from the linear theory, and could therefore be classified as a "circuit model" with nonlinear circuit elements. Neither the linear or weakly nonlinear models are able to account for the reversal observed experimentally at large voltage and frequencies above the inverse RC time. Another feature is that the bulk concentration almost completely scales out of the problem, entering only through λ_D in the surface capacitance ratio $\delta = \epsilon\lambda_S/\epsilon_S\lambda_D$. Therefore these models are also unable to account for the experimental fact that AC electroosmosis is suppressed at concentrations above 10 mM [15].

Chapter 4

Strongly nonlinear analysis

In this chapter we present numerical results obtained using the strongly nonlinear model derived in Chap. 2. We first apply the model to a simple 1D geometry in Sec. 4.1. The purpose of studying this system is both that it gives good understanding of the dynamical properties of the model, which is necessary when we later interpret the results in higher dimensions, but also that it allows us to test the validity by comparing with a dynamical solution of the full Poisson–Nernst–Planck equations.

In Sec. 4.2 we study the electrokinetic flow induced on both a symmetric and asymmetric electrode array. Of course, the symmetric device does not pump, but it allows us to monitor the strength of the different flow components, namely, electroosmotic flow of the first and second kind, and electroconvection induced when a current is passed through concentration gradients in the electroneutral diffusion layer and bulk. Such a distinction between different flow components is artificial in the sense that experimentally only the overall flow is observed, but it helps us gain understanding of the system behaviour. Finally, we investigate the net pumping induced in the asymmetric device.

4.1 1D model problem

In this section we study the dynamics of the strongly nonlinear model in a one-dimensional geometry as shown in Fig. 4.1 with an electrolyte confined between parallel-plate blocking electrodes, i.e., we neglect Faradaic electrode reactions, and the (dimensionless) width of the electrode gap is $L = 2$. The system is identical to that studied by Bazant *et al.* in Ref. [37], except that we apply an AC voltage rather than a step voltage, and study the periodic response after all transients have died out.

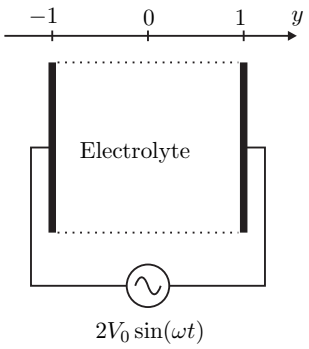


Figure 4.1: Sketch of 1D model problem. The electrolyte is confined between parallel-plate blocking electrodes separated by a gap of width $L = 2$, and a harmonic potential of $\pm V_0 \sin(\omega t)$ is applied to the left and right electrode, respectively, so the overall voltage drop across the system is $2V_0 \sin(\omega t)$.

In one dimension the solution to the leading order bulk problem simplifies a lot. By symmetry there can be no fluid motion,[1] so the only dynamical variables in the electroneutral bulk are the potential $\bar{\phi}(y,t)$ and the average concentration $\bar{c}(y,t)$. Moreover, there can be no time average flux from the boundary layers into the bulk, so the leading order bulk concentration is simply constant $\bar{c}(y,t) = \bar{c}_o$, and using Eq. (2.163) to impose mass conservation across the entire system we obtain

$$\bar{c}_o = 1 - \epsilon\langle\tilde{w}\rangle. \tag{4.1}$$

The bulk electric current $\bar{J}(t)$ is constant in space and the potential is given by

$$\bar{\phi} = \bar{\phi}_o(t) - \frac{\bar{J}(t)y}{\bar{c}_o}, \tag{4.2}$$

where $\bar{\phi}_o$ is the potential at the centre of the cell. The bulk fields are matched to the solution in the boundary layers at $y = \pm 1$ by requiring

$$\partial_t \tilde{q} = \pm\bar{J}, \tag{4.3}$$

$$\partial_t \tilde{w} = \pm\check{F}, \tag{4.4}$$

$$V_{\text{ext}} - \bar{\phi}_o \pm \frac{\bar{J}}{\bar{c}_o} = \tilde{\zeta} - \tilde{q}\,\delta, \tag{4.5}$$

$$\tilde{\zeta} = -2\sinh^{-1}\!\big(\tilde{q}/2\sqrt{\check{c}}\big), \tag{4.6}$$

$$\tilde{w} = \sqrt{\tilde{q}^2 + 4\check{c}} - \sqrt{4\check{c}}, \tag{4.7}$$

in the Debye layers, where $\check{c}$ is the concentration in the diffusion layer seen

[1] Unless, of course, some kind of electroconvective instability occurs, but we do not consider that here.

from the electrode surface

$$\check{c} = \bar{c}_o \mp \sqrt{\frac{\varepsilon}{D_a \omega}} \frac{1}{T} \int_0^T \mathcal{G}(t-t') \big[\check{F}(t') - \Delta D \bar{J}(t') \big] \, \mathrm{d}t'. \tag{4.8}$$

The bulk potential must have odd symmetry upon reflection in space and translation by one half period in time

$$\bar{\phi}(-y, t + T/2) = -\bar{\phi}(y, t), \tag{4.9}$$

which implies that also $\bar{J}$ and $\bar{\phi}_o$ must have odd symmetry in time. In the special case where $\Delta D = 0$, the central potential $\bar{\phi}_o$ vanishes and both $\tilde{q}$ and $\tilde{\zeta}$ have odd symmetry, whereas $\tilde{w}$, $\check{F}$, and $\check{c}$ have even symmetry, e.g., $\tilde{w}(t + T/2) = \tilde{w}(t)$. We do not include the excess potential drop $\check{\varsigma}$ across the diffusion layer into Eq. (4.5) because this vanishes at low voltage and typically remains small compared to Ohmic potential drop across the bulk at larger voltage.

The coupled differential, algebraic, and integral equations are solved numerically by substituting a truncated Fourier series and using a collocation scheme, e.g.,

$$\bar{J}(t) = \sum_{m=-N/2}^{N/2\,\prime} \hat{J}_m e^{im\omega t} = \sum_{n=1}^{N} = \bar{J}(t_n) S_N(t - t_n), \tag{4.10}$$

where S_N is the periodic sinc function

$$S_N(t - t') = \frac{1}{N} \sum_{m=-N/2}^{N/2\,\prime} e^{im\omega(t-t')}, \tag{4.11}$$

see Appendix A.3 or Ref. [68] for details.

4.1.1 Dynamical solutions

We solve the problem for $\varepsilon = 0.001$, $\omega = 1$, $\delta = 1$, and $\Delta D = 0$, at a relatively large voltage $V_0 = 50$, and using $N = 128$ grid points in time. The solution for the different components in the overall potential drop across (half) the system, namely, the Ohmic potential drop $\bar{J}/\bar{c}_o$ across the bulk electrolyte, the Stern layer potential drop $-\tilde{q}\,\delta$, and the zeta potential $\tilde{\zeta}$ across the Debye layer, is shown in Fig. 4.2(a). At this driving voltage the Debye layer capacitance is large so $\tilde{\zeta}$ is small and the potential drop across the double layer is primarily on the Stern layer. The nonlinearity is not very pronounced and the system behaves almost as an ideal linear impedance $\hat{Z} = 1 - i$, except when $\tilde{\zeta}$ changes sign and the Debye layer capacitance shortly drops to 1, which results in a small decrease in the current. As usual for an RC circuit driven at a frequency around the inverse

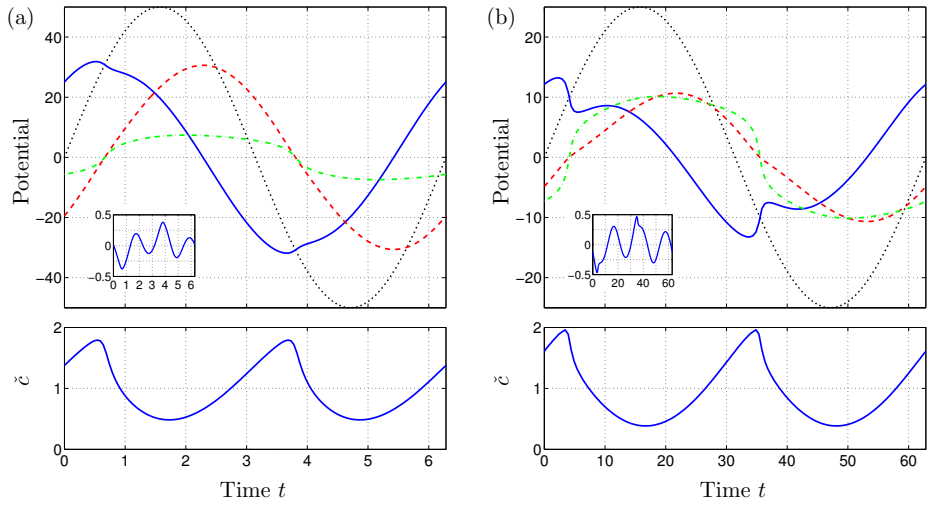


Figure 4.2: Dynamical solution to the 1D model problem as a function of time for two sets of parameter values: (a) $V_0 = 50$, $\omega = 1$, and $\delta = 1$; (b) $V_0 = 25$, $\omega = 0.1$, and $\delta = 0.1$. The top panel shows the different components to the overall potential drop across (half) the system, namely, the driving voltage $V_0 \sin(\omega t)$ (*dotted*), bulk Ohmic potential $\bar{J}/\bar{c}_o$ (*solid*), Stern layer potential $-\tilde{q}\,\delta$ (*dashed*), and zeta potential $\tilde{\zeta}$ (*dash-dot*). The inset shows the potential drop $\check{\varsigma}$ across the diffusion layer which is $\mathcal{O}(\sqrt{\varepsilon})$ smaller than $\bar{J}/\bar{c}_o$. The bottom panel shows the concentration $\check{c}$ in the diffusion layer as seen from the surface.

RC time, the phase angle of the current is around $\pi/4$ ahead of the driving voltage, whereas the charge accumulated in the Debye layer is $\pi/2$ ahead the current. Fig. 4.2(a) also shows the surface concentration $\check{c}$ that oscillates at the double frequency 2ω. Although the variations in $\check{c}$ are $\mathcal{O}(1)$, the diffusion layer potential drop is small, $\check{\varsigma} \lesssim 0.4$, as shown in the inset.

For comparison Fig. 4.2(b) shows the solution for a different set of parameter values where now $\omega = 0.1$, $\delta = 0.1$, and $V_0 = 25$. Again, at this voltage the Debye layer capacitance is large, but because the intrinsic capacitance ratio δ is small, the potential drop across the Debye and Stern layers are comparable. Therefore the frequency $\omega = 0.1$ is actually also a little below the inverse RC time. The current has a significant bump at the point where $\tilde{\zeta}$ changes sign and the double layer capacitance drops from $\sim 1/\delta$ to $1/(1+\delta)$. The maximal amount of charge $\tilde{q}$ and salt $\tilde{w}$ accumulated in the Debye layer is about a factor of 3 larger than in Fig. 4.2(a), but the variations in the surface concentration $\check{c}$ are similar because the diffusion length $\ell(\omega) \propto 1/\sqrt{\omega}$ is also increased by roughly a factor of 3.

The concentration variations in the diffusion layer are directly linked to the amount of salt accumulated in the Debye layer. At low voltage $\tilde{w} = 4\sqrt{\tilde{c}}\sinh^2\!\big(\tilde{\zeta}/4\big)$ scales as V_0^2, while at higher voltage our estimate $\tilde{w} \sim V_0/\delta$ from Sec. 2.2.9 applies. Therefore we also expect the perturbations to the concentration in the diffusion layer to grow linearly with V_0

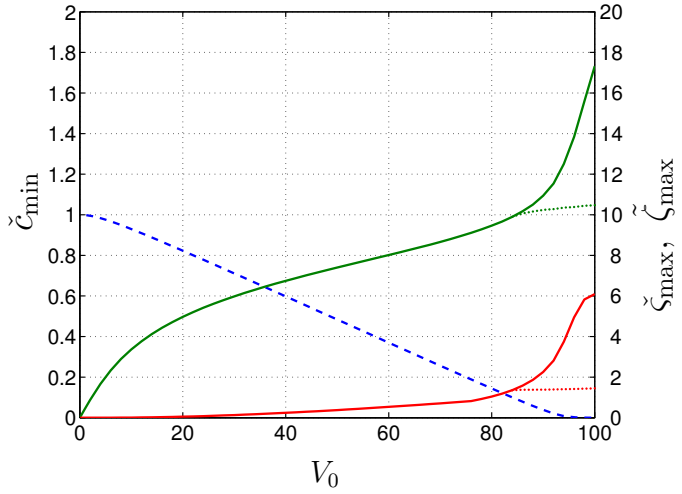


Figure 4.3: Minimal salt concentration $\min_t \check{c}$ in the diffusion layer (dashed line), and maximal values of the zeta potential $\max_t \tilde{\check{\zeta}}$ (upper solid line) and diffusion layer potential drop $\max_t \check{\varsigma}$ (lower solid line), as a function of the driving voltage V_0 for $\omega = 1$ and $\delta = 1$. Around $V_0 \approx 90$ the concentration $\check{c}$ approaches zero which results in unphysically large values of both $\tilde{\check{\zeta}}$ and $\check{\varsigma}$. The dotted lines show the result when this is suppressed by replacing $\check{c}$ with $\underline{\check{c}} = \max\{\check{c}, \varepsilon^{1/3}\}$ in Eq. (4.6) and similarly when computing $\check{\varsigma}$, cf. the *ad hoc* modifications suggested in Sec. 2.2.9.

for $V_0 \gg 1$. Fig. 4.3 shows the minimal concentration $\min_t \check{c}$ in the diffusion layer as a function of the driving voltage V_0 at $\omega = 1$ and $\delta = 1$, and indeed for $V_0 \gtrsim 10$ there is a clear linear dependence. Around $V_0 \approx 90$ the minimal concentration approaches zero, which marks the breakdown of local electroneutrality in the diffusion layer close to the electrodes. The figure also shows the maximal values of $\tilde{\check{\zeta}}$ and $\check{\varsigma}$. At low voltage the zeta potential $\tilde{\check{\zeta}} \sim V_0/(1+\delta)$ grows linearly, while at higher voltage the scaling is only logarithmic with $\tilde{\check{\zeta}} \sim 2\log(V_0/\delta)$. The excess potential drop across the diffusion layer depends on both the current passing though the layer and the perturbations to the concentration, and therefore scales as V_0^3 at low voltage changing to V_0^2 at higher voltage. When $\check{c}$ approaches zero both $\tilde{\check{\zeta}}$ and $\check{\varsigma}$ diverge logarithmically, which in turns prevents $\check{c}$ from going negative in Fig. 4.3. However, this divergent behaviour of $\tilde{\check{\zeta}}$ and ς is unphysical: As discussed in Sec. 2.2.9 the leading order strongly nonlinear solution based on local electroneutrality in the diffusion layer breaks down already when $\check{c} \sim \varepsilon^{1/3}$. The divergence is suppressed by replacing $\check{c}$ with $\underline{\check{c}} = \max\{\check{c}, \varepsilon^{1/3}\}$ in Eq. (4.6) and similarly when computing $\check{\varsigma}$, as suggested also in Sec. 2.2.9, and the result is shown with dotted lines in Fig. 4.3.

4.1.2 Comparison to a full numerical solution of the Poisson–Nernst–Planck equations

Based on the matched asymptotic expansions we can construct a uniformly valid approximation in space by adding the solutions from the inner, middle, and outer regions, and subtracting the overlaps, i.e.,

$$\begin{aligned} c(y,t) \;=\; \bar{c}_o + \left[\check{c}\left(\frac{1+y}{\sqrt{\varepsilon}},t\right) - \bar{c}_o\right] + \left[\tilde{c}\left(\frac{1+y}{\varepsilon},t\right) - \check{c}(0,t)\right] \\ + \left[\check{c}\left(\frac{1-y}{\sqrt{\varepsilon}},t\right) - \bar{c}_o\right] + \left[\tilde{c}\left(\frac{1-y}{\varepsilon},t\right) - \check{c}(0,t)\right], \end{aligned} \tag{4.12}$$

where $\tilde{y} = (1 \pm y)/\epsilon$ and $\check{y} = (1 \pm y)/\sqrt{\epsilon}$ are the rescaled spatial variables in the inner and middle regions. In order to validate the asymptotic model we can compare this uniformly valid approximation to a full numerical solution of the Poisson–Nernst–Planck equations in time and space. For a symmetric electrolyte in one dimension Eqs. (2.45) to (2.47) reduce to

$$-\varepsilon^2 \partial_y^2 \phi = \rho, \tag{4.13}$$

$$\partial_t \rho = \varepsilon \partial_y (\partial_y \rho + c \partial_y \phi), \tag{4.14}$$

$$\partial_t c = \varepsilon \partial_y (\partial_y c + \rho \partial_y \phi), \tag{4.15}$$

with boundary conditions at $y = \pm 1$

$$V_{\text{ext}} - \phi = \mp \varepsilon \delta \partial_y \phi, \tag{4.16}$$

$$0 = -\partial_y \rho - c \partial_y \phi, \tag{4.17}$$

$$0 = -\partial_y c - \rho \partial_y \phi. \tag{4.18}$$

When solving this problem for $\varepsilon = 0.001$ it is necessary to use a very fine finite element mesh of $h_{\min} \lesssim 10^{-4}$ close to the electrodes in order to resolve the highly compressed Debye layer structure, or even finer at larger voltage. In 1D this is straightforward to achieve with a graded mesh, but in 2D or 3D it can become a serious problem.

We compare the two models for the same parameter values as in Fig. 4.2(a), i.e., $\varepsilon = 0.001$, $\delta = 1$, $\omega = 1$, and $V_0 = 50$. The results for the average concentration $c(y,t)$ is shown in Fig. 4.4 as a number of snapshots over one half period in time. The first frame at $t = \frac{1}{16}T$ is just about when the charge density in the Debye layer changes sign, cf. Fig. 4.2(a), so for $t = \frac{2}{16} - \frac{4}{16}T$ we see the Debye layer charging up and accumulating a lot of ions, holding them at $t = \frac{5}{16}T$, before it decharges and expels them back into the diffusion layer for $t = \frac{6}{16} - \frac{8}{16}T$. The agreement between the full numerical solution and the asymptotic expansion is very good, with a maximal relative error less than 0.5% in the diffusion layer and around around 2.5% in the Debye

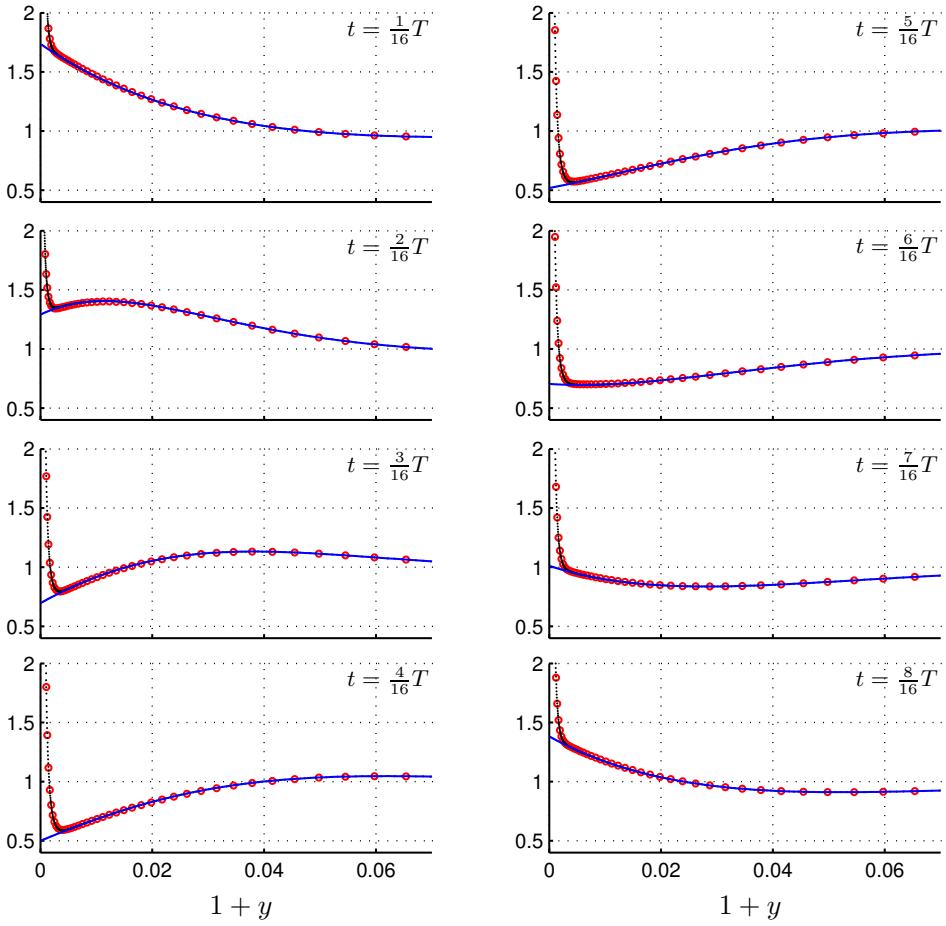


Figure 4.4: Snapshots of the average concentration c in the diffusion layer over one half period in time for $\varepsilon = 0.001$, $\omega = 1$, $\delta = 1$, and $V_0 = 50$. The figure compares a full numerical solution of the Poisson–Nernst–Planck equations (*circles*) to the uniformly valid approximation from the asymptotic model (*dotted line*), and the agreement between the two is very good. Also the leading order approximation $\check{c}$ in the diffusion layer alone is shown (*solid line*). Notice that the vertical axis does not go to zero.

layer. The error is of the order $\mathcal{O}(\sqrt{\varepsilon})$ because we neglected the excess potential drop across the diffusion layer $\check{\varsigma}$, which is only $\mathcal{O}(\sqrt{\varepsilon})$ smaller that the bulk Ohmic potential drop.

If we increase ε the quality of the leading order asymptotic approximation of course decreases, and, e.g., for $\varepsilon = 0.05$ and $\delta = 0.1$, which are the parameter values used throughout Ref. [37], the error for $\omega \sim 0.1$ at the inverse RC time is large because the diffusion layer of thickness $\ell(\omega) \sim \sqrt{\varepsilon/\omega} \simeq 0.7$ extends across the entire system. However, even at very small ε the leading order strongly nonlinear solution breaks down when the driving voltage is

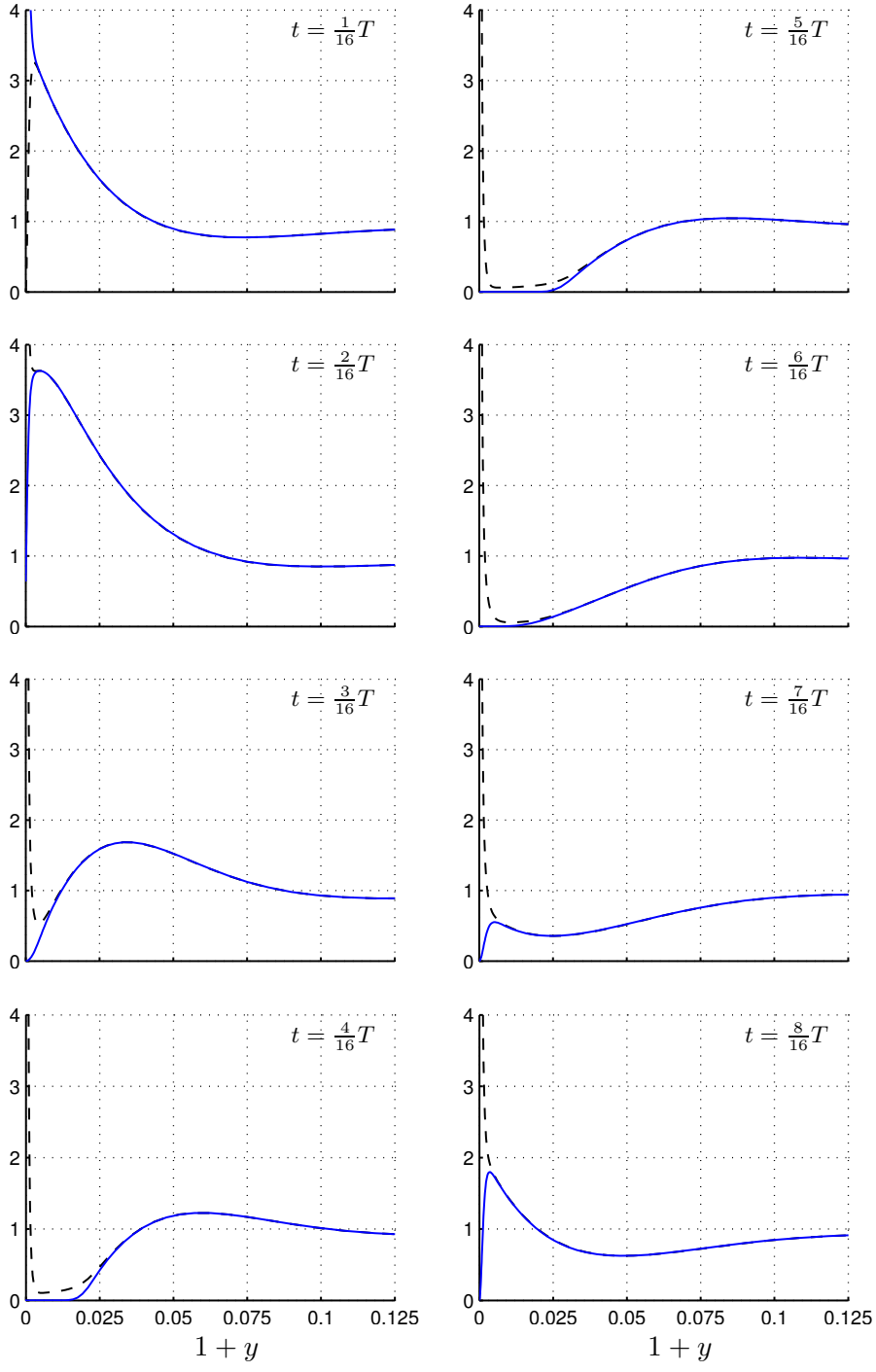


Figure 4.5: Individual ion concentrations for $\varepsilon = 0.001$, $\omega = 1$, and $\delta = 1$, at a driving voltage $V_0 = 200$ which is large enough to drive the double layer out of quasiequilibrium. The figure shows the full numerical solution for the concentration of the cations c_+ (*solid line*) and anions c_- (*dashed line*). Between $t = \frac{3}{16}T$ and $\frac{6}{16}T$ a layer completely depleted of cations appear between the electrode and the electroneutral diffusion layer.

increased enough that the solution in the diffusion layer approaches zero. For $\varepsilon = 0.001$ and $\omega = \delta = 1$ this occurs around $V_0 \approx 90$, cf. Fig. 4.3, and for other values of ε and δ the breakdown occurs approximately at $V_0 \approx 3\sqrt{\delta/\varepsilon}$; the prefactor 3 is particular to the simple 1D geometry that we study here, but the scaling with $\sqrt{\delta/\varepsilon}$ should hold generally.

When the driving voltage is further increased beyond the point where the concentration at the surface approaches zero, the Debye layer does not stop charging but continues to take up ions from the diffusion layer, and a layer almost completely depleted of ions develops between the Debye layer and the electroneutral diffusion layer. This is shown in Fig. 4.5, again as a number of snapshots over one half period for $\varepsilon = 0.001$, $\delta = 1$, $\omega = 1$, and $V_0 = 200$. The solid and dashed lines show the cation and anion concentrations, respectively, from a full solution to the Poisson–Nernst–Planck equations. Between $t = \frac{1}{16}T$ and $\frac{2}{16}T$ we see the charge density in the Debye layer change sign, followed by a rapid uptake of ions for $t = \frac{3}{16}T$ to $\frac{5}{16}T$. At $t = \frac{4}{16}T$ and $\frac{5}{16}$ we can clearly distinguish a region completely depleted of cations, while the anion concentration $c_- \sim 0.1$ is also low but finite. This is the extended space charge layer discussed in Sec. 2.2.9, which was first described by Rubinstein and Shtilman for a permselective membrane driven above the DC "diffusion-limited" current [40]. Here we see the same basic structure of the nonequilibrium double layer developing and dynamically propagating into the diffusion layer. At $t = \frac{6}{16}T$, which is immediately before the current reverses, we see the space charge layer retreating, and for $t = \frac{7}{16}T$ and $\frac{8}{16}T$ the diffusion layer is quickly refilled by an equal amount of anions expelled from the Debye layer and cations transported by migration from the opposite electrode.

Fig. 4.6 shows the same data as Fig. 4.5 but on double logarithmic axis which better allow the different boundary layers to be distinguished. The figure also shows the results from our strongly nonlinear asymptotic model when the *ad hoc* modifications suggested in Sec. 2.2.9 are employed, namely, we continue to solve a simple 1D diffusion problem for $\check{c}$ even beyond the breakdown of local electroneutrality, and identify the unphysical region of negative "salt" concentration with the width y_o of the space charge layer. In the electroneutral diffusion layer for $y \geq y_o(t)$ we have $c_+ = c_- = \check{c}$ at leading order, while in the space charge layer $c_+ = 0$ and $c_- = |2\check{\rho}|$, where the charge density $\check{\rho}$ is given by Eq. (2.177), and vice versa during the second half period when a space charge layer completely depleted of anions develops. Finally, in the $\mathcal{O}(\varepsilon)$ quasiequilibrium "Debye" layer the leading order anion concentration is

$$c_- = |2\tilde{\rho}| = 2|\partial_{\tilde{y}}^2 \tilde{\phi}| = \frac{4a^2}{\sinh^2(a\tilde{y} + b)}, \tag{4.19}$$

cf. Eq. (2.178), where $a = \sqrt{|\bar{J}_n/2D_-}$ and $b = \coth^{-1}(|\tilde{q}|/2a)$. Fig. 4.6 shows that the asymptotic model captures the overall dynamical behaviour

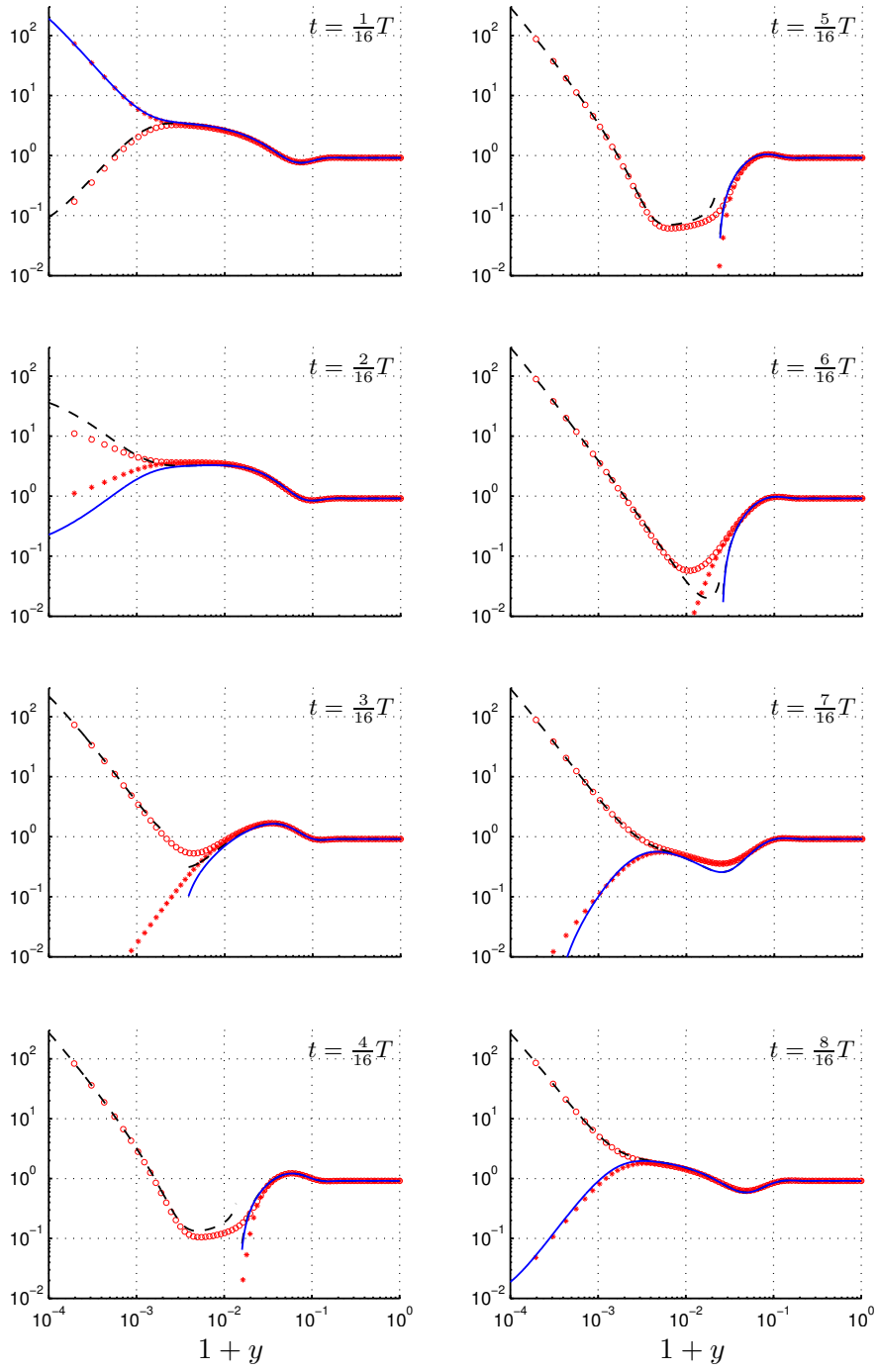


Figure 4.6: Individual ion concentrations: Same data as in Fig. 4.5 but plotted on double logarithmic axis. Comparison of the full numerical solution for the concentration of the cations c_+ (*stars*) and anions c_- (*circles*) to the asymptotic model for the cations (*solid line*) and anions (*dashed line*).

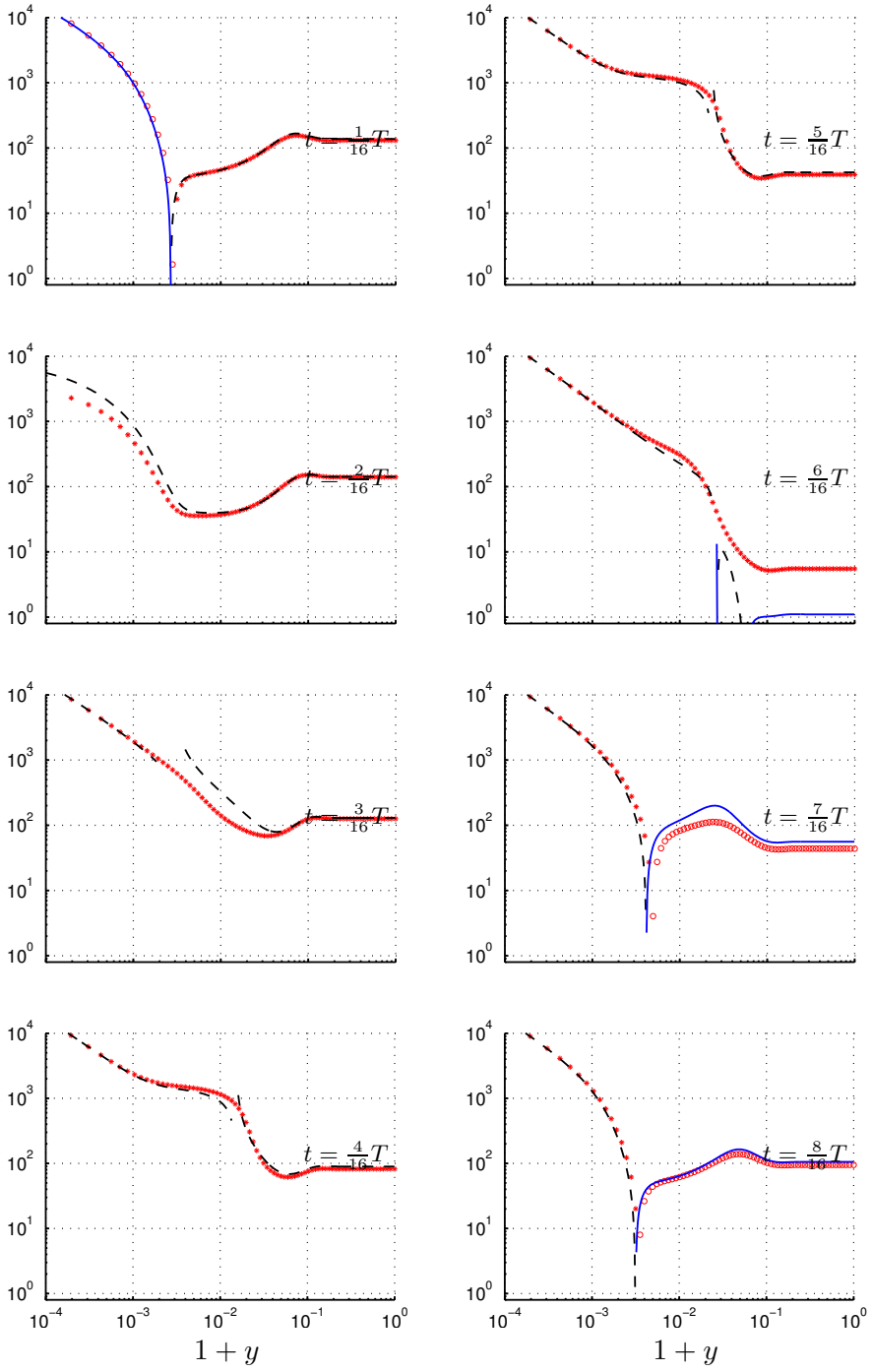


Figure 4.7: Magnitude of the electric field $E = -\partial_y \phi$ for the same parameters as in Fig. 4.5. Comparison of the full numerical solution for the field (*stars* or *circles* depending in the sign of E) to the asymptotic model (*solid* or *dashed* lines). The agreement is good except for $t = \frac{6}{16}T$ and $\frac{7}{16}T$ just before and after the space charge layer collapses.

of the full numerical solution well, both in the electroneutral bulk and diffusion layer, and in the quasiequilibrium or nonequilibrium double layer. At $t = \frac{2}{16}T$, just after the charge density in the Debye layer has changed sign, there is a visible difference between the solutions inside the $\mathcal{O}(\varepsilon)$ Debye layer, but once the space charge layer has formed the asymptotic model is spot on the full numerical solution in the Debye layer. For $t = \frac{6}{16}T$ and $\frac{7}{16}T$, immediately before and after the collapse of the space charge layer, there is a significant discrepancy in the diffusion layer around $1 + y \sim \sqrt{\varepsilon} \simeq 0.03$. And the quality of the asymptotic model within the space charge layer is generally not as good as in the other regions, part of the reason for which is simply that we have not included the $\mathcal{O}(\varepsilon^{2/3}/|\bar{J}_n|^{1/3})$ Smyrl–Newman transition layer, and therefore the divergent tail of $\check{\rho}$ close to y_o is visible in the plots. We already argued in Sec. 2.2.9 that the nonequilibrium solution breaks down close to the point where the current reverses, because migration then no longer dominates in the space charge layer, and because the thickness of the Smyrl–Newman transition layer diverges. Also, much of the discrepancy is due to our crude model for salt diffusion, allowing for negative values of $\check{c}$. A solution for $\check{c}$ with the ambipolar flux boundary condition applied on a dynamically changing domain $\check{y}_o \leq \check{y} < \infty$ would certainly improve the quality of the model. However, having said this, we still find that our simple model captures the qualitative features of the full numerical solution well. To further support this, Fig. 4.7 compares the magnitude of the electric field $E = -\partial_y \phi$ from the asymptotic model to the full numerical solution. Again, the agreement is good, except in an $\mathcal{O}(\varepsilon^{2/3})$ region around y_o where the field from the diffusion layer $-\partial_y \check{\phi} = \bar{J}_n/\check{c}$ diverges, and except for $t = \frac{6}{16}T$ and $\frac{7}{16}T$ immediately before and after the current reverses and the space charge layer collapses.

Finally, Fig. 4.8 compares the actual thickness of the space charge layer from the full numerical solution, defined (arbitrarily) as the width of the region where the cation or anion concentration drops below $\varepsilon^{2/3}$, to the width of the region of negative concentration from the asymptotic model, and to the approximate expression for y_o based on Eq. (2.190). The figure shows that the width of negative concentration tends to overestimate y_o, whereas the approximate expression tends to underestimate. Both are within roughly 10 % of the full numerical solution, except close to the point where the space charge layer collapses. This accuracy is good enough that we dare employ the simple *ad hoc* modifications to the strongly nonlinear model and use it below to predict the direction and magnitude of electroosmotic flow of the second kind.

Returning to Fig. 4.6 we notice that the Debye layer is highly compressed at this large voltage. The concentration exceeds 10^2 at $1 + y = 10^{-4}$ and increases to almost 10^4 for $1 + y \sim 10^{-5}$. For a dilute electrolyte with nominal bulk concentration of 1 mM this corresponds to a concentration of about 10^{28} ions per m^3 or 10 ions per nm^3 at the surface which is approx-

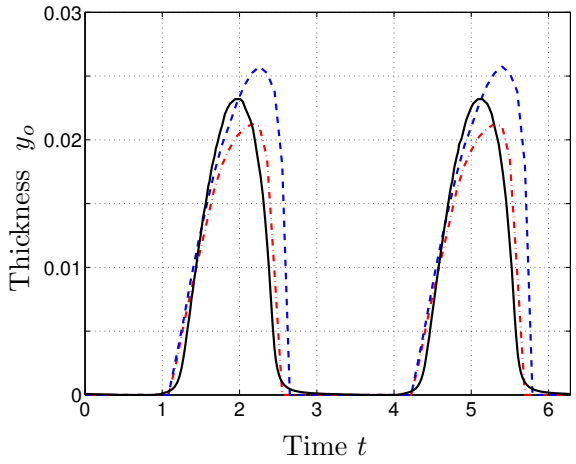


Figure 4.8: Thickness y_o of the space charge layer as a function of time: Comparison of the space charge layer thickness from the full numerical solution (*solid line*) to the width of the region of negative concentration in the asymptotic model (*dashed line*) and to the approximation for y_o based on Eq. (2.190) (*dash-dot*).

imately at the steric limit, cf. the discussion in Sec. 2.1.2, and for more concentrated solutions this limit is reached already at lower voltages. At the same time, even though the Debye length is nominally $\lambda_D = 10$ nm for a 1 mM electrolyte, the highly compressed Debye layer is effectively no more than about an ångstrøm wide. This means that the Gouy–Chapman–Stern model or more generally the use of dilute solution theory to describe the structure of the $\mathcal{O}(\varepsilon)$ quasiequilibrium Debye layer at the electrode becomes questionable. In particular, our model for the zeta potential $\tilde{\zeta}$ breaks down, but as we argued also in Sec. 2.1.2 this does not necessarily imply that our entire model becomes invalid: Even beyond the steric limit the Debye layer continues to accumulate charge, so the effects of salt depletion in the diffusion layer, including the development of an extended space charge layer, are still relevant. The surface conductance could also be affected by strong interactions between the ions. While this is an interesting topic, it is also beyond the scope of the present thesis, so we continue to investigate the dynamical behaviour in our model as derived from dilute solution theory.

4.2 2D model problems

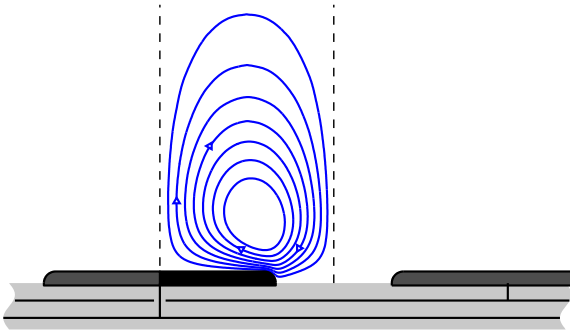


Figure 4.9: Sketch of the geometry of the symmetric electrode array. The interdigitated electrodes all have width $W = 2$ and finite thickness and radius of curvature $R = 0.1$ at the edges, and are separated by gaps of $G = 1$. Dashed lines mark the computational domain or unit cell, and the streamlines show the usual flow pattern with a vortex above the electrode edge, driven by electroosmotic slip from the electrode edges towards the center.

4.2.1 Symmetric electrode array

We consider a symmetric array of interdigitated electrodes as sketched in Fig. 4.9. The electrodes have (dimensionless) width $W = 2$ and are separated by a gap $G = 1$, and to avoid introducing unphysically strong salt depletion due to the field singularity at the edges we do not model the electrodes as being infinitely thin but allow a finite thickness and radius of curvature $R = 0.1$ at the edges. Further, we take $\varepsilon = 0.001$, $\delta = 1$, $\omega = 1$, and assume a symmetric electrolyte with $\Delta D = 0$. Fig. 4.9 also shows the fluid flow pattern for $V_0 = 50$ and no convective transport, i.e., $Pe = 0$. The dominant component in the flow is electroosmosis, so we observe the "usual" pattern with slip from the electrode edges towards the center, driving pairs of fluid vortices above the electrodes. The slip pattern is shown in detail in Fig. 4.10, where the solid line shows electroosmotic slip and the dashed line is electroconvection induced in the diffusion layer. The two components have opposite sign, but the electroosmotic slip dominates because the zeta potential $\tilde{\zeta}$ is much larger than the excess Ohmic potential drop $\check{\zeta}$ across the diffusion layer.

Fig. 4.11(a) shows the bulk concentration profile as a contour plot: At the electrode edges the concentration is lower than the equilibrium bulk value $\bar{c} = 1$, whereas at the electrode center the concentration is higher. This nonuniform concentration profile is maintained by a continuous surface flux in the Debye layer running from the electrode edges to the center, i.e., on average in time the Debye layer sucks in salt at the electrode edges, and transports it towards the center where it is expelled and diffuses back; this

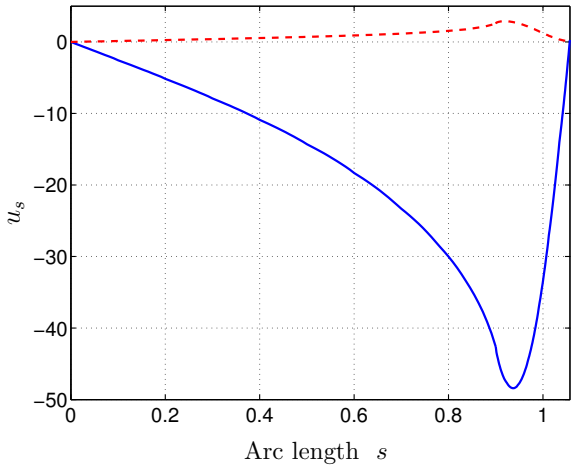


Figure 4.10: Slip velocity u_s as a function of the distance s from the electrode center. The electroosmotic slip (solid line) and electroconvection from the diffusion layer (dashed) have opposite directions but electroosmosis dominates.

diffusive flux is shown in Fig. 4.11(b). The feature that the surface excess flux $\tilde{\mathbf{\Gamma}} = -(1+2Pe)\tilde{q}\boldsymbol{\nabla}_s\bar{\phi}$ is in the same direction as the electroosmotic slip $\tilde{\mathbf{u}}_s = \tilde{\zeta}\boldsymbol{\nabla}_s\bar{\phi}$ holds more generally for a symmetric electrolyte because $\tilde{q}$ and $\tilde{\zeta}$ have opposite sign. Fig. 4.11(c) shows the bulk electroconvection component of the overall fluid motion due to the volume electrical body force $\bar{\mathbf{f}} = \boldsymbol{\nabla}^2\bar{\phi}\boldsymbol{\nabla}\bar{\phi}$, arising when a current is passed through a region of nonuniform concentration. This flow takes the form of a fluid vortex with the same direction as that from electroosmotic slip, but the magnitude is much smaller than the electroosmotic flow at $V_0 = 50$.

Now, it turns out that neglecting bulk salt transport by convection is a poor approximation. Although for a typical aqueous electrolyte the nominal

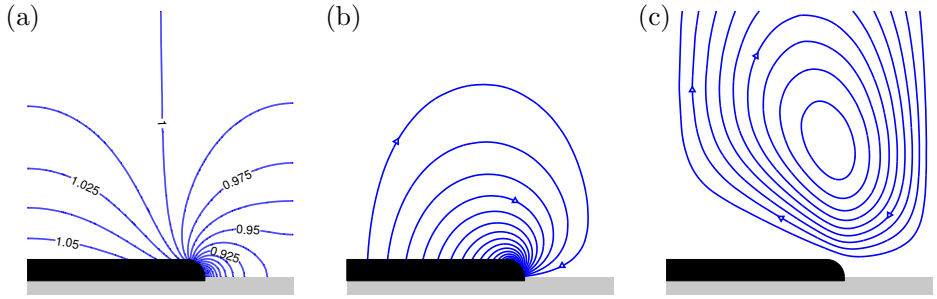


Figure 4.11: Solution at $V_0 = 50$ and $Pe = 0$. (a) Contour plot of the bulk concentration profile. The concentration at the electrode edges is lower than the equilibrium concentration $\bar{c} = 1$ and higher at the electrode center. (b) Purely diffusive bulk salt flux $\bar{\mathbf{F}} = -\boldsymbol{\nabla}\bar{c}$. (c) Bulk electroconvection component of fluid flow. The flow takes the form of a fluid vortex with the same direction as that from electroosmotic slip.

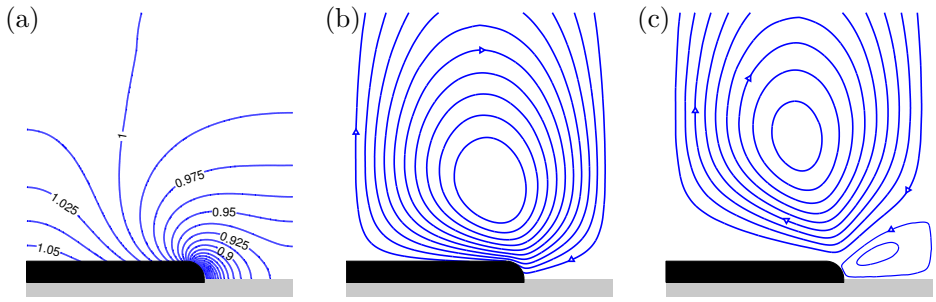


Figure 4.12: Solution for $V_0 = 50$ and $Pe = 0.25$. (a) The salt concentration profile is distorted and (b) the bulk salt flux is dominated by convection. (c) The bulk electroconvection induced with the distorted concentration profile features a second counterrotating vortex in the electrode gap.

Peclet number $Pe = 0.25$ is relatively small, cf. Table 2.1, the true ratio of convective to diffusive ion transport

$$Pe^* = \frac{|Pe\bar{\mathbf{u}}\bar{c}|}{|\boldsymbol{\nabla}\bar{c}|} \sim Pe|\bar{\mathbf{u}}| \tag{4.20}$$

is large because the dimensionless fluid velocity is much larger than unity at $V_0 = 50$, cf. Fig. 4.10. Therefore in Fig. 4.12 we show the solution at $Pe = 0.25$: The salt flux in Fig. 4.12(b) is dominated by convection, which distorts the concentration profile in Fig. 4.12(a) as compared to Fig. 4.11(a). Interestingly, the bulk electroconvective flow in Fig. 4.12(c) induced with the distorted concentration profile has different topology than for $Pe = 0$: There is again a large fluid vortex rotating in the same direction as the electroosmotic flow, but a second smaller vortex appears in the gap between the electrodes. However, the overall flow pattern is still dominated by electroosmosis so this second small vortex does not show up in the overall picture.

The minimal concentration in the bulk region close to the electrode edge is around $\min_{\mathbf{r}} \bar{c} \approx 0.8$ and $\min_{\mathbf{r}} \bar{c} \approx 0.7$ for $Pe = 0$ and $Pe = 0.25$, respectively, but the concentration in the diffusion layer is very close to dynamically reaching zero at $V_0 = 50$. Upon a further increase in the driving voltage, a transient space charge layer forms and gives rise to electroosmotic flow of the second kind. Fig. 4.13 shows the slip velocity distribution along the electrode for $V_0 = 100$ and $Pe = 0.25$, and in particular we notice a significant contribution from electroosmotic flow of the second kind close to the electrode edge, that drives the fluid in the same direction as electroosmosis of the first kind: The potential drop $\breve{\Phi}$ across the space charge layer has the same sign as $\tilde{\zeta}$, i.e., opposite of $\tilde{q}$, so in general the Helmholtz–Smoluchowski component $\breve{\Phi}\boldsymbol{\nabla}_s\bar{\phi}$ of the second kind electroosmotic slip should have the same direction as that of the first kind.

In order to quantify the relative strength of the different components in

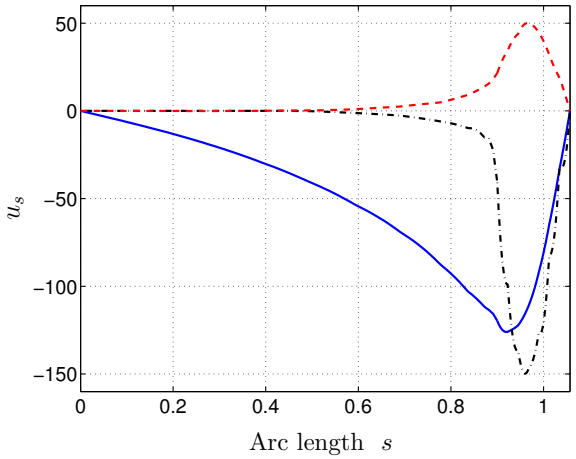


Figure 4.13: Slip velocity distribution at $V_0 = 100$. The electroosmotic flow of the first kind (solid line) and second kind (dash-dot) are in the same direction, while electroconvection from the diffusion layer (dashed) is opposite.

the overall fluid flow we compute the total angular momentum

$$L_{\text{tot}} = \int_{\Omega} \mathbf{r} \times \bar{\mathbf{u}} \, \mathrm{d}\mathbf{r}, \tag{4.21}$$

where Ω is the computational domain, cf. Fig. 4.9. Because of symmetry the linear momentum of the fluid inside Ω vanishes, which makes L_{tot} independent of the choice of origin for $\mathbf{r}$. The contribution to L_{tot} from the different flow components is shown in Fig. 4.14(a) for $Pe = 0$ and in Fig. 4.14(b) for $Pe = 0.25$. In both cases electroosmosis of the first and second kind and bulk electroconvection all contribute negatively to L_{tot}, i.e., they give rise to fluid motion in the clockwise direction, while electroconvection induced in the diffusion layer contributes positively. In general it appears that bulk convection of salt for $Pe \neq 0$ enhances all components in the flow, except bulk electroconvection. The mechanism behind this is as follows: The surface excess salt flux from the electrode edges towards the center tends to produce a region of low salt concentration and therefore low conductivity in the gap between the electrodes. This slows down the charging close to the electrode edges, which means that the accumulated charge becomes more uniformly distributed along the electrode. That in turns suppresses the tangential electric field because that is fundamentally dependent on nonuniform charging of the Debye layer, and both electroosmotic slip of the first and second kind and electroconvection from the diffusion layer are driven by the tangential field. However, bulk convection of salt tends to counteract the development of a low conductivity region between the electrodes by stirring the solution and bringing down unit concentration electrolyte into the gap. This explains why electroosmosis of the first kind saturates already for

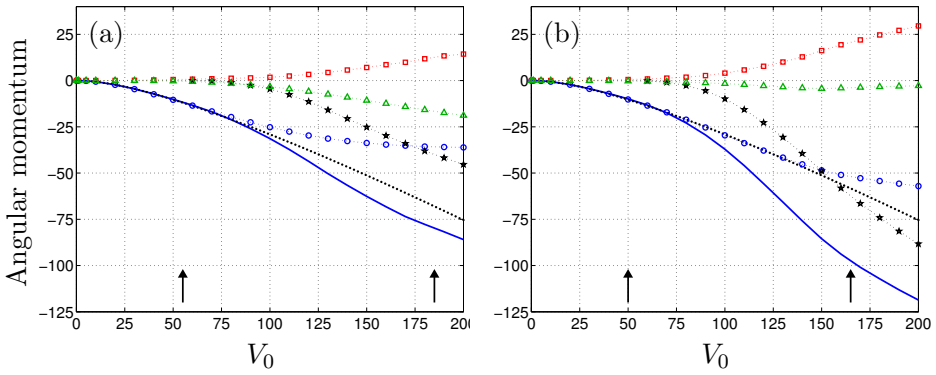


Figure 4.14: Total angular momentum in fluid vortex as a function of driving voltage V_0 for: (a) $Pe = 0$ and (b) $Pe = 0.25$. The solid line shows the overall L_{tot}, and the symbols show contributions from electroosmosis of the first kind (circles) and the second kind (stars), and from electroconvection from the diffusion layer (squares) and bulk (triangles). For comparison the dotted line shows the corresponding result from the weakly nonlinear model. The two arrows mark the voltage where local electroneutrality first breaks down in the diffusion layer and bulk region, respectively.

$V_0 \gtrsim 100$ in Fig. 4.14(a) but not until above $V_0 \gtrsim 175$ in Fig. 4.14(b), and why both electroosmosis of the second kind and electroconvection from the diffusion layer is about twice as strong in Fig. 4.14(b) as in Fig. 4.14(a). The development of the low conductivity zone between the electrodes is shown in Fig. 4.15 for $V_0 = 50$, 100, 150, and 200, and $Pe = 0$ and 0.25, respectively.

The bulk electroconvection only contributes very little to L_{tot} at $Pe = 0.25$, but this is partly because the secondary vortex from Fig. 4.12(c) grows so large that the overall contribution from the two vortices almost integrates to zero. The arrows in Fig. 4.14 marks the point around $V_0 \approx 50$ where local electroneutrality first breaks down in the diffusion layer close to the electrode edge, as well as the point around $V_0 \approx 175$ where the bulk concentration drops below $\varepsilon^{1/3} = 0.1$ and our *ad hoc* modification, replacing $\bar{c}$ by $\underline{\bar{c}} = \max\{\bar{c}, \varepsilon^{1/3}\}$ in the bulk conductivity and fluid body force, becomes active.

4.2.2 Asymmetric electrode array

We now break the left-rigth symmetry of the electrode array and consider a planar device as that sketched in Fig. 3.1, with pairs of electrodes of width $W_1 = 1.5$ and $W_2 = 7$, separated by gaps of $G_1 = 1$ and $G_2 = 5$, respectively. To avoid problems associated with the field singularity at sharp edges we allow for electrodes of finite thickness and curvature $R = 0.2$ at the edges. Further, we take $\varepsilon = 0.001$ and $\delta = 1$, and consider first a symmetric electrolyte with $\Delta D = 0$. Because the device is no longer completely flat, and because bulk electroconvection is not induced at the surface, we cannot apply the simple Fourier analysis from Chap. 3 to compute the pumping

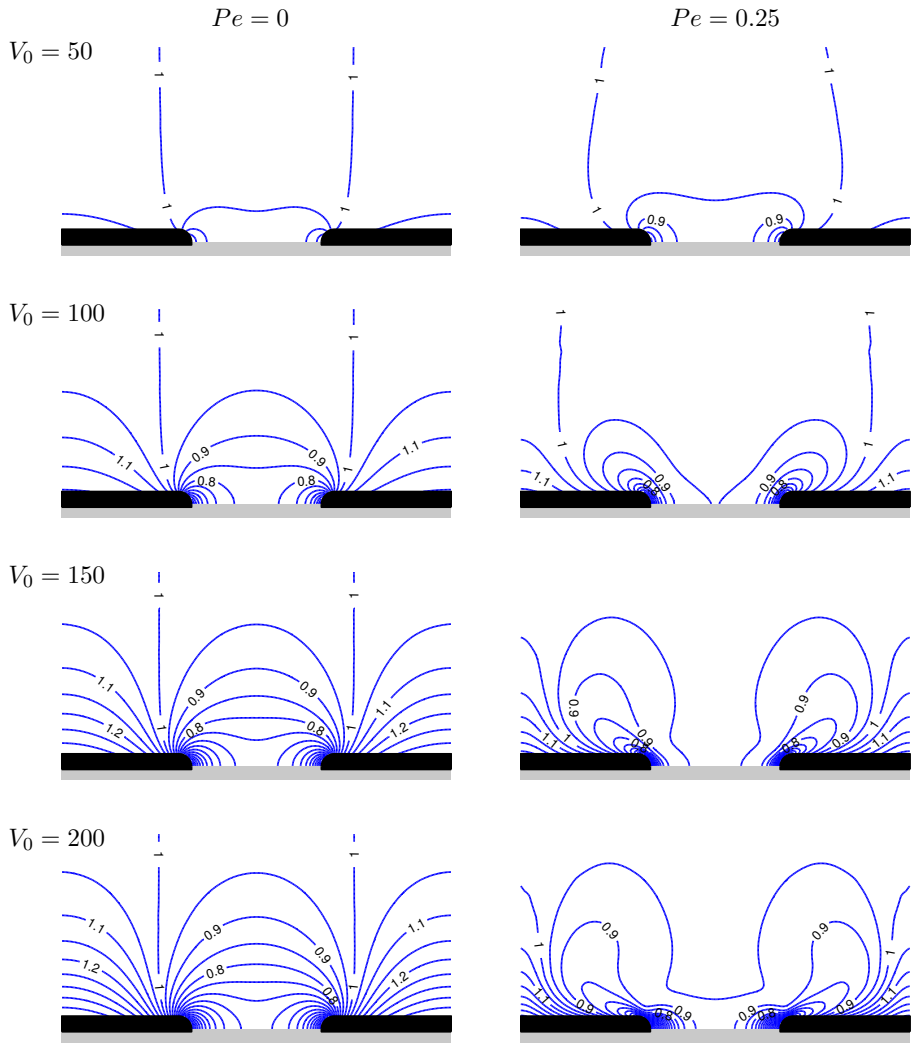


Figure 4.15: Salt concentration profile for $Pe = 0$ (left) and 0.25 (right) at $V_0 = 50$, 100, 150, and 200. The low conductivity region between the electrodes locally reduces the current into the Debye layer at the electrode edges, which makes the charging of the Debye layer more uniform and reduces the tangential field. Bulk convection of salt when $Pe \neq 0$ counteracts this by flushing "fresh" electrolyte down into the gap region.

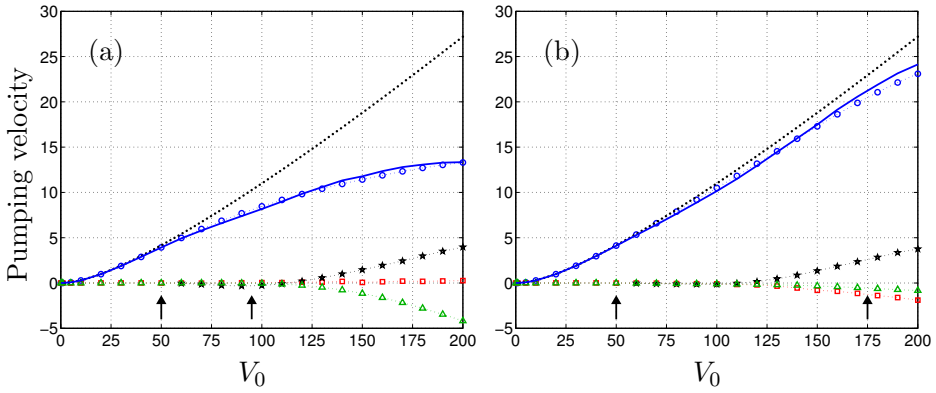


Figure 4.16: Pumping velocity as a function of driving voltage V_0 at $\omega = 1$ for: (a) $Pe = 0$ and (b) $Pe = 0.25$. The solid line shows the net pumping U, and the symbols show contributions from electroosmosis of the first kind (circles) and the second kind (stars), and from electroconvection from the diffusion layer (squares) and bulk (triangles). For comparison the dotted line shows the corresponding result from the weakly nonlinear model. The two arrows mark the voltage where local electroneutrality first breaks down in the diffusion layer and bulk region, respectively.

velocity from the spatial average slip. Instead we determine the full bulk flow profile, compute the net volume flow rate Q by integrating over a cross section of the channel, cf. Eq. (3.2), and define the effective pumping velocity as $U = 2Q/H$, where H is the height of the channel.

Fig. 4.16 shows the pumping velocity as a function of driving voltage V_0 for $\omega = 1$ at the inverse RC time, both with and without convective transport of salt. In both cases the net pumping is dominated by electroosmotic flow of the first kind, so the model does not predict any reversal of the pumping direction for $V_0 \lesssim 200$. The contribution from electroconvection falls off as V_0^4 at low voltage, where the flow induced in the diffusion layer contributes in the "forward" direction, while bulk electroconvection pumps in the "reverse" direction with much smaller strength. Note that we might actually have expected the opposite, since we learned in Sec. 4.2.1 that bulk electroconvection induces vortices in the same direction as electroosmosis, while electroconvection induced in the diffusion layer slips in the opposite direction. However, for a planar asymmetric electrode array it is not only the slip direction, i.e., towards the center or the edges of the electrodes, that determines the net pumping, but also the balance between different sections with opposite slip direction at different strength.

Around $V_0 \approx 50$ the concentration in the diffusion layer dynamically approaches zero close to the right edge of the narrow electrode, and above this voltage the electroosmotic flow of the second kind contributes to pumping in the reverse direction. When the driving voltage is further increased, the narrow electrode experiences breakdown of local electroneutrality and formation of transient space charge layers over the entire surface of the elec-

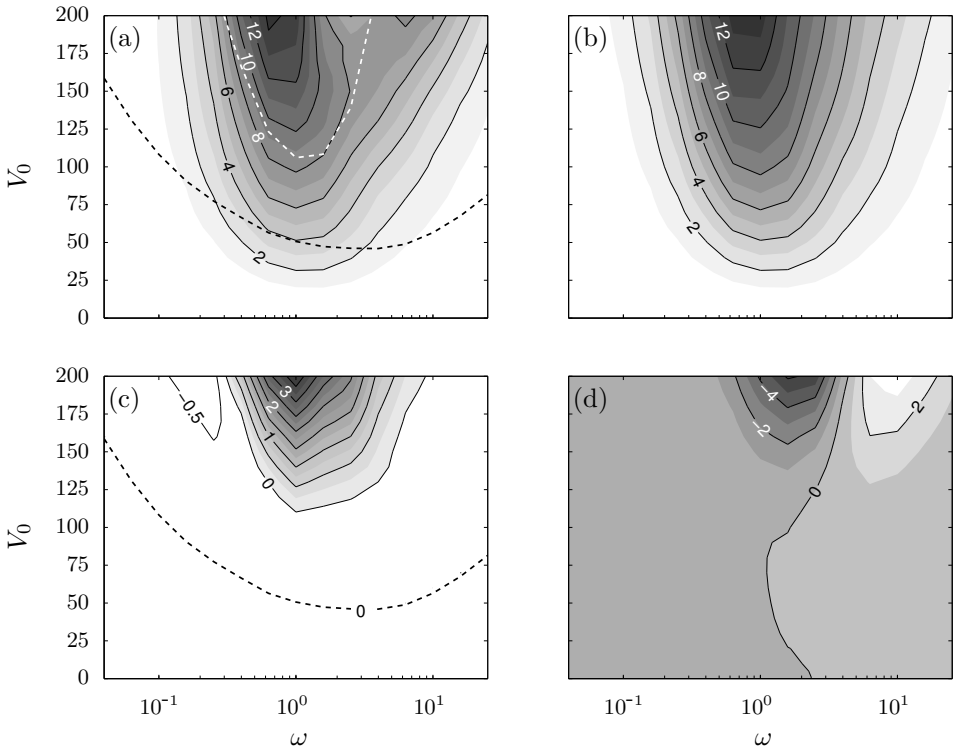


Figure 4.17: Contour plot of pumping velocity as a function of driving frequency ω and voltage V_0 for $Pe = 0$. (a) Net pumping velocity U and contributions from: (b) Electroosmotic flow of the first kind, (c) electroosmotic flow of the second kind, and (d) bulk electroconvection. The dashed black and white lines mark the voltage where local electroneutrality breaks down in the diffusion layer and bulk region, respectively.

trode, while the wide electrode experiences this only close to the left edge. Therefore for $V_0 \gtrsim 110$ the overall contribution from electroosmosis of the second kind to the net pumping changes to be in the forward direction. The electroosmotic flow of the first kind saturates for $V_0 \gtrsim 125$ in Fig. 4.16(a) with $Pe = 0$, but not in Fig. 4.16(b) where $Pe = 0.25$. As for the symmetric device this occurs because the surface excess flux gives rise to a region of low conductivity in the gap between the electrodes which reduces the tangential field, and again bulk convection of salt counteracts the formation of such a region for $Pe \neq 0$.

Fig. 4.17 shows the pumping velocity on a contour plot as a function of the driving frequency ω and voltage V_0 for $Pe = 0$. In Fig. 4.17(a) the total pumping velocity U is maximized around $\omega \sim 1$ at the inverse RC time, but at large voltage there is a more complex frequency dependence due to the contribution from bulk electroconvection. The dashed lines mark the breakdown of local electroneutrality in the diffusion layer and bulk region, respectively, cf. the arrows in Fig. 4.16. Fig. 4.17(b) shows the contribution

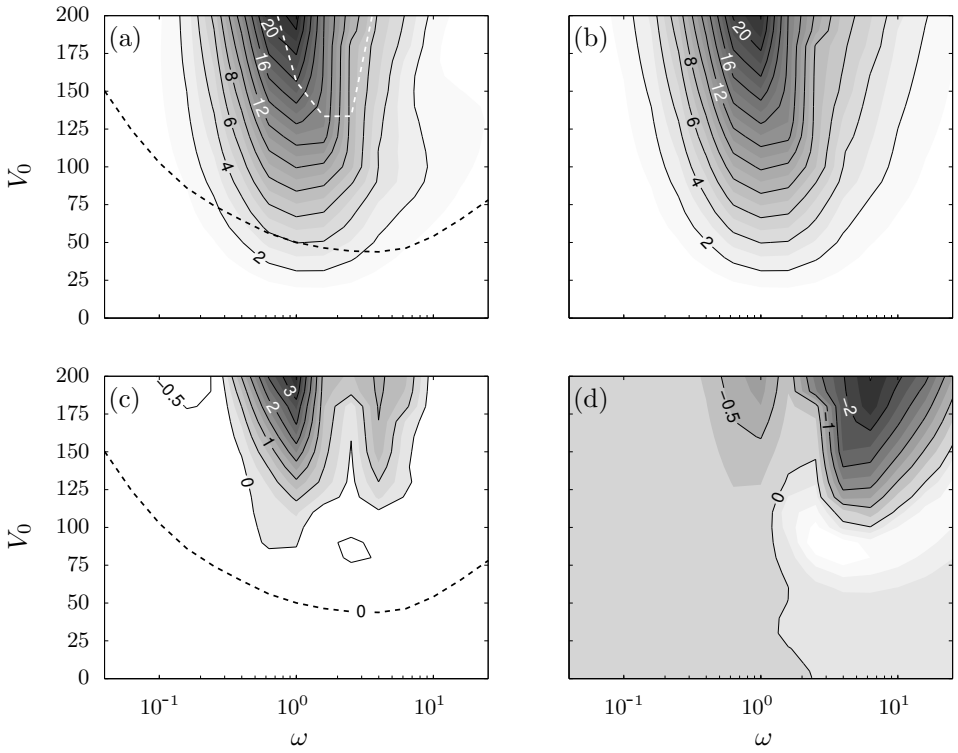


Figure 4.18: Contour plot of pumping velocity U as a function of driving frequency ω and voltage V_0 for $Pe = 0.25$. (a) Net pumping velocity U and contributions from: (b) Electroosmotic flow of the first kind, (c) electroosmotic flow of the second kind, and (d) bulk electroconvection. Again, dashed black and white lines mark the voltage where local electroneutrality breaks down in the diffusion layer and bulk region, respectively.

from electroosmosis of the first kind, featuring a single peak around $\omega \sim 1+\delta$ at low voltage that shifts down to $\omega \sim \delta$ at larger voltage. The contribution from electroosmosis of the second in Fig. 4.17(c) starts out negative, i.e., in the reverse direction, but changes to positive at larger voltage. Finally, bulk electroconvection in Fig. 4.17(d) contributes in the reverse direction around $\omega \sim 1$ but in the forward direction at larger frequency. We do not show the electroconvection induced in the diffusion layer because its contribution to the net pumping is small.

Fig. 4.18 shows the results at $Pe = 0.25$, which are qualitatively very similar to those at $Pe = 0$, but quantitatively the electroosmotic flow of the first kind is about twice as strong because convection of salt prevents formation of a low conductivity region in the electrode gap. The electroosmotic flow of the second kind displays a twin peak at large voltage, and the same is true for bulk electroconvection in the reverse direction. A similar feature with a secondary peak in the pumping velocity has been observed experimentally by Urbanski *et al.*, although for a different electrode geometry [69].

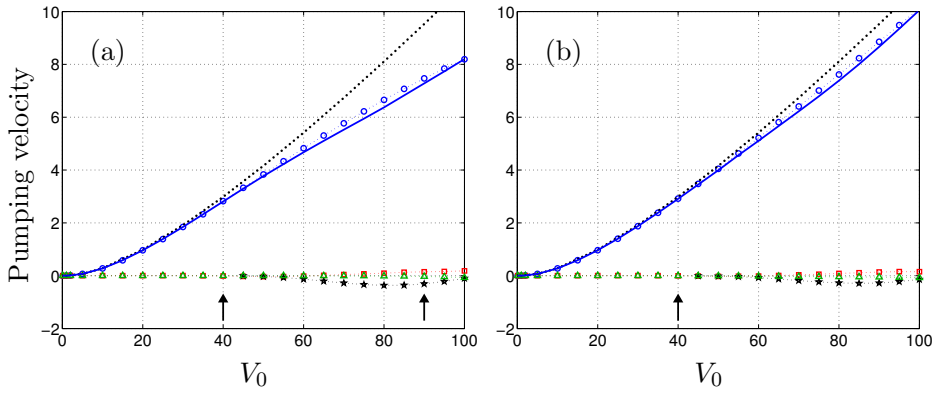


Figure 4.19: Pumping velocity as a function of driving voltage V_0 at $\omega = 1$ for an asymmetric electrolyte with $\Delta D = 0.33$: (a) $Pe = 0$ and (b) $Pe = 0.25$. The solid line shows the net pumping U, and the symbols show contributions from electroosmosis of the first kind (circles) and the second kind (stars), and from electroconvection from the diffusion layer (squares) and bulk (triangles). For comparison the dotted line shows the corresponding result from the weakly nonlinear model. Arrows mark the breakdown of local electroneutrality in the diffusion layer and bulk region.

However, the convergence of the solution for large voltage and frequencies $\omega \gtrsim 3$ was relatively slow, so we are not fully confident with our numerical results here.

Until this point we have focused on the symmetric electrolyte case with $\Delta D = 0$ in the strongly nonlinear model which is already farily complex, but as a final twist let us consider an asymmetric electrolyte with cations twice as mobile as the anions, i.e., $D_+/D_- = 2$, such that $\Delta D = 0.33$ and $D_a = 0.88$. The results are shown in Fig. 4.19 as a function of driving voltage at $\omega = 1$, and they are very similar to those for a symmetric electrolyte from Fig. 4.16. The contribution from electroconvection induced in the diffusion layer scales as V_0^2 at low voltage because the diffusion potential $-\Delta D \log(1 + \check{\gamma}/\bar{c})$ is linear in the excess concentration $\check{\gamma}$, which in turns is linear in V_0 due to the electric current component in the excess ambipolar flux, cf. Eqs. (2.159) and (2.160). Moreover, salt depletion in the diffusion layer is more pronounced due to the lower effective diffusivity D_a, and the concentration at the electrode approaches zero already around $V_0 \approx 40$. Therefore electroosmotic flow of the second kind also shows up earlier, but with magnitude similar to the symmetric electrolyte solution.

Limitations of the numerical scheme

The temporal resolution in Fig. 4.16 is $N = 32$ points in time over one period, while in Figs. 4.17, 4.18, and 4.19 it is only $N = 16$. The difference between the results obtained with $N = 16$ and $N = 32$ is small, around 5% or less, but also not completely negligible. We compute the periodic solution

using a relaxation method as described in Appendix A.3, which allows us to account for the nonlocal time dependence of the surface concentration on the flux injected from the Debye layer. However, because the computer memory required to solve the problem grows as $\mathcal{O}(N^2)$, this effectively limits the time resolution that we can obtain; for the spatial resolution in the finite element mesh that we use, we run out of computer memory already at $N = 64$. Obviously, it would be a great advantage to develop a timestepping algorithm to compute the solution because the memory requirements for a single timestep would be $\mathcal{O}(1)$ only, allowing for a much better resolution in time and space than we have been able to achieve here. However, we have not had the time to pursue this.

4.3 Summary

In this chapter we have studied numerical solutions to our strongly nonlinear dynamical model, both for a simple 1D geometry and for 2D models of a symmetric and an asymmetric electrode array.

We compared the strongly nonlinear model to a full numerical solution of the Poisson–Nernst–Planck equations in one dimension, and found excellent agreement at driving voltages below the point where local electroneutrality breaks down in the diffusion layer. At even larger voltage we found that the simple modifications to the model to account for formation of transient space charge layers captures the qualitative features of the full numerical solution well, except close to the point where the current reverses and the space charge layer collapses, but also admit that there is room for general improvement of the model.

For the symmetric electrode array we found that both electroosmotic flow of the first and second kind and bulk electroconvection gives rise to fluid motion in the same direction, whereas the electroconvection induced in the electroneutral diffusion layer slips in the opposite direction of electroosmosis. We also found that electroosmotic flow is limited at large voltage by the formation of a region of low conductivity in the gap between the electrodes, but bulk convection of salt tends to counteract this.

On the asymmetric array our model predicts that electroosmotic flow of the first and second kind contributes to pumping in the same "forward" direction at large voltage, whereas bulk electroconvection contributes in the "reverse" direction, and electroconvection induced in the diffusion layer contributes only little to the net pumping. However, the electroosmotic flow dominates over the entire voltage range that we have considered, so we are not able to account for the reversal of the net pumping direction that has been observed experimentally [15, 31, 34].

The particular parameter values of $\varepsilon = 0.001$ and $\delta = 1$ used throughout this chapter corresponds to, e.g., a dilute 1 mM KCl solution with $\lambda_D =$

10 nm in a device with 10 μm characteristic feature size, cf. Table 2.1, with an electrode material such as titanium that forms a natural oxide layer and therefore has a moderate Stern layer capacitance. For electrolytes of even lower concentration or devices of smaller feature size, the relevant value of ε would be larger, and the strongly nonlinear phenomena would become significant at lower driving voltage. Likewise, for electrodes made from gold or platinum that can have very large Stern layer capacitance, the amount of ions accumulated in the Debye layer is larger, which again makes the strongly nonlinear phenomena show up at lower voltage.

For electrolytes of higher concentration the use of dilute solution theory to describe the Debye screening layer also breaks down at lower driving voltage. The effect of steric exclusion due to finite size of the ions is first to decrease the capacitance of the Debye layer, i.e., to increase the zeta potential for a given amount of accumulated charge [56], which would in turn enhance the electroosmotic flow of the first kind. However, one might also envisage viscoelectric effects from strong interactions in a dense phase of ions that would tend to inhibit fluid motion. In that case the other sources of electrokinetic flow, i.e., electroosmosis of the second kind and electroconvection from the quasielectroneutral diffusion layers and bulk as studied in this chapter, would become more significant to the overall fluid motion.

Chapter 5

Optimization

Device optimization is an important discipline in engineering. It is also a complex task with many objectives to be pursued simultaneously, including device performance, price, robustness, ease of manufacturing, etc. There are many trade-off's to be made and it is often not easy to measure what is the "optimal" design.

In mathematics an optimization problem is formulated in a much more clean and abstract way: The goal of the optimization is to minimize or maximize some quantity $f(\mathbf{x})$, the objective function, which is a function of all the parameters $\mathbf{x}$ that can be adjusted in the optimization process. In the present chapter we focus on optimizing the electrode geometry to maximize the pumping velocity U as a function of driving frequency and of the different parameters describing the electrode shape. E.g., for a simple planar array of asymmetric electrode pairs as studied in detail Chap. 3 we have $\mathbf{x} = [\omega, W_1, W_2, G_1, G_2]^T$. There may be constraints on the values that the parameters can take, e.g., a geometrical length scale cannot be negative, or more complex nonlinear constraints that depends on several of the parameters together which are typically written as $g_i(\mathbf{x}) \leq 0$. Hence a large class of optimization problems can be formulated as

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & g_i(\mathbf{x}) \leq 0 \qquad \text{for} \quad i = 1, \ldots, m, \\ & x_j^{\min} \leq x_j \leq x_j^{\max} \qquad \text{for} \quad j = 1, \ldots, n. \end{aligned} \tag{5.1}$$

Depending on the nature of the problem there are several different solution algorithms to choose from: For constrained linear problems the method of choice is the so-called *simplex algorithm*, whereas for unconstrained nonlinear problems one can apply either a *conjugate gradient* or a *quasi-Newton* method, provided the gradient of the objective function with respect to $\mathbf{x}$ can be computed [70]. For low-dimensional nonlinear constrained optimization a popular method is *sequential quadratic programming* where in each iteration the objective and constraints are approximated by quadratic functions

of $\mathbf{x}$; implementations can be found in many software libraries, including the MATLAB Optimization Toolbox. The algorithm requires the gradient and Hessian (second derivative) of the objective and constraints to be evaluated, either exactly or by finite differencing. However, for large-dimensional problems typical of structural optimization, in particular topology optimization, it becomes very expensive to compute the Hessian. In such cases it is preferable to use a simple gradient based algorithm, like *steepest descent*, to determine the optimal solution. We have used the so-called *method of moving asymptotes* (MMA) where the approximation for the objective and constraints is chosen to be both convex and separable [71, 72, 73].

Topology optimization is an approach to designing mechanical structures where one does not a priori specify or even know the topology of the optimal geometrical structure [74, 75, 76]. It was originally developed in mechanical engineering to optimize beams for maximal stiffness at a given amount of material. More recently the method has been extended to optimize acoustical systems [77], photonic crystals [78], and fluidic structures [79, 80]. We have been working with the method in the fluidic domain, extending the original results of Borrvall and Petterson for Stokes flow [79] to account for inertial forces, and using it to study a small fluidic network whose optimal topology changes as a function of the Reynolds number [46]; see our paper in Appendix 5. The basic idea is to replace the discrete transition between solid structure and void with a smooth one by introducing an (artifical) porous medium whose permeability to the fluid depends on the local material density. The optimization algorithm is then allowed to distribute a fixed amount of material in a continuous way, which allows changes in topology to occur smoothly rather than discretely. Often this improves the convergence towards the globally optimal topology, but for fluidic systems we observed problems when there are two strong compeeting local optima [46]. F. Okkels continues our work by optimizing flow in chemical enzyme reactors and cell metabolism [51, 81].

While topology optimization is a powerfull tool, there are two major obstacles that makes it impractical for optimizing the electrode geometry for AC electroosmotic pumping: First, the electroosmotic slip is an inherent surface phenomenon which is incompatible with a fuzzy transition between solid electrode and fluid electrolyte. Second, the electrode topology is not really up for discussion: A device with free-hanging metal islands above the base electrodes would certainly be an interesting sort of combination of AC- and induced-charge electroosmosis but it would also be extremely difficult to fabricate. Therefore, in the present chapter we use more classical shape optimization of the electrode structures, both with low- and high-dimensional design spaces.

We always stick to the low voltage regime $V_0 \ll k_B T/e$. The primary reason is that optimization is an expensive task numerically and therefore we prefer to work with the linear model which is in itself inexpensive to solve.

Our experience from Chap. 3 is that the weakly nonlinear regime is not very different from the low voltage regime, so it seems reasonable that a device geometry optimized at low voltage will also perform well at larger voltage. In the strongly nonlinear regime we still need a better understanding of the system behaviour before we can start thinking of systematic optimization.

5.1 Planar asymmetric electrode pairs

Most of the AC electroosmotic pump designs so far tested experimentally [25, 26, 28, 29, 15, 30] have been planar devices with asymmetric electrode pairs as studied in Chap. 3 and sketched in Fig. 3.1. Often the relative size of the electrodes has been chosen as in the original paper by Brown *et al.* [25], with narrow and wide electrodes of width $W_1 = 4.2$ μm and $W_2 = 25.7$ μm, respectively, separated by gaps of $G_1 = 4.5$ μm and $G_2 = 15.6$ μm. Mpholo *et al.* tested devices with three different overall periods $L = 50$ μm, 30 μm, and 20 μm, but used the same relative electrode sizes in each case; this allowed them to confirm that the pumping velocity increases with downscaling of the system.

In practice the minimal size of the microelectrodes and gaps is limited by the lithography used to fabricate the device. However, for a given minimal feature size one can still improve on the pumping velocity U by choosing the relative electrode sizes optimally. This problem can be expressed as a constrained optimization of the form of Eq. (5.1) writing

$$\begin{aligned} \text{maximize} \quad & U(\log\omega, W_1, W_2, G_1, G_2) \\ \text{subject to} \quad & 1 \le W_1 \le W_2, \\ & 1 \le G_1 \le G_2, \end{aligned} \tag{5.2}$$

where ω is the driving frequency.[1] Because the pumping velocity increases with downscaling we can be sure that in the optimal design at least one of W_1 and G_1 will attain their lower limit, whereas both W_2 and G_2 must be strictly larger than W_1 and G_1, respectively, because the net pumping vanishes for $W_1 = W_2$ or $G_1 = G_2$ due to symmetry. It turns out that in the optimal design $G_1 = 1$ and $W_1 > G_1$, and the exact location of the optimum can therefore be determined using a simple unconstrained optimization algorithm such as the MATLAB function `fminunc` to solve the problem[2]

$$\text{maximize} \quad U(\log\omega, W_1, W_2, G_1 = 1, G_2), \tag{5.3}$$

[1] Optimizing for $\log\omega$ effectively removes any problems associated with zero or negative frequencies; indeed we would argue this is the natural way of reporting *any* frequency.

[2] In order to use the `fminunc` function from the MATLAB Optimization Toolbox, all the user has to do is write a function that takes $(\log\omega, W_1, W_2, G_2)$ as arguments, sets up and solves the electrokinetic problem, e.g., using FEMLAB, and returns (minus) the pumping velocity U. However, due to the truncation error and noise introduced by remeshing the geometry in each function call, it is important choose appropriate not-so-small values for the termination tolerance and for the step size used when `fminunc` approximates the

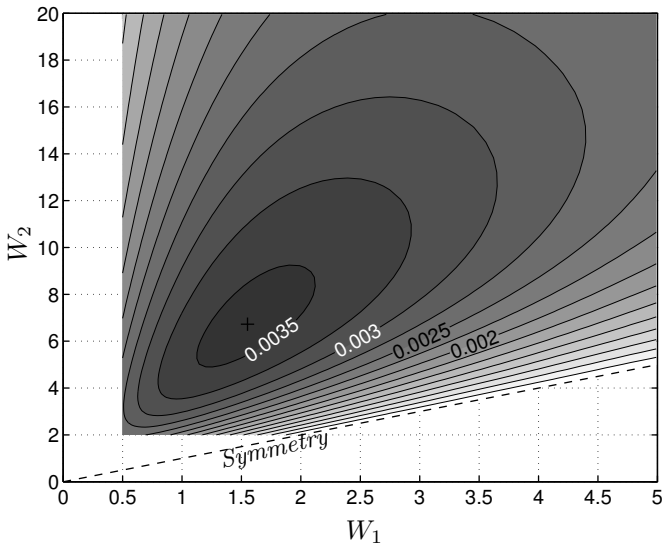


Figure 5.1: Contour plot of the maximum pumping velocity $\max_\omega U$ as a function of W_1 and W_2 for $G_1 = 1$ and $G_2 = 5$. When $W_1 = W_2$ the net pumping vanishes due to symmetry. There is a unique optimum close to $W_1 = 1.5$ and $W_2 = 7$ with $U = 0.00358$ at $\omega = 0.8$.

provided we start the algorithm somewhere in the interior of the feasible region. In this way we find that the pumping velocity is maximized for $W_1 = 1.51$, $W_2 = 6.55$, and $G_2 = 4.74$, where $U = 0.00359$ at $\omega = 0.80$.

The pumping velocity is actually not very sensitive to the particular choice of planar electrode geometry. This is demonstrated in Fig. 5.1 where the maximum pumping velocity $\max_\omega U$ is shown as a function of W_1 and W_2 for $G_1 = 1$ and $G_2 = 5$. The maximum is close to $W_1 = 1.5$ and $W_2 = 7$ where $U = 0.00358$ at $\omega = 0.8$, but the figure also shows that the range of designs that come within the 0.0035 contour is fairly broad. It is for this reason that we have chosen in Chap. 3 to work with a design with rounded numbers, namely $W_1 = 1.5$, $G_1 = 1$, $W_2 = 7$, and $G_2 = 5$. This does not correspond to any device tested experimentally (yet), but it is close to the global optimum according to our present linear analysis.

5.2 Planar electrodes with nonuniform coating

There are many ways that one can break the left-right symmetry of an electrode array. From a fabrication point of view the asymmetric pairs of planar electrodes are particularly simple because the device can be processed with just a single lithography step to define metal electrodes on the substrate. Another simple approach is to use pairs of planar electrodes that are sym-

gradient and Hessian by finite differencing. A robust but also less efficient alternative is direct search algorithms like the MATLAB `fminsearch` Nelder–Mead simplex method.

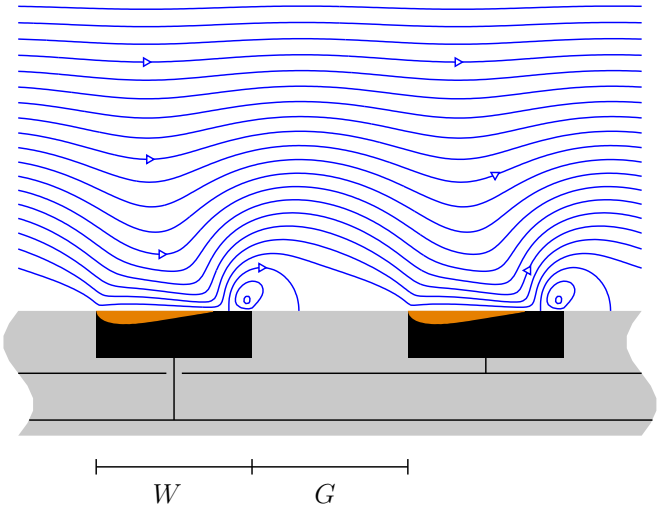


Figure 5.2: Planar electrodes (black) with a spatially nonuniform thin oxide layer or coating (orange). The coating thickness is not drawn to scale; streamlines correspond to the optimal solution at driving frequency $\omega = 5.1$.

metric in size, but with an oxide layer or coating that makes the Stern layer capacitance left-right asymmetric on the individual electrodes, see Fig. 5.2. This was actually one of the suggestions for breaking the symmetry in the original paper by Ajdari [14].

We consider the general problem of an array of blocking electrode where the Stern layer capacitance varies along the electrodes, i.e., the capacitance ratio $\delta(x)$ is a function on the position with $0 \leq \delta(x) \leq \infty$. For our present purpose it turns out to be more convenient to parametrize the problem in terms of $\Lambda = (1+\delta)^{-1}$, which is simply the ratio of the overall capacitance of the double layer to the Debye layer capacitance [23, 33], because this stays in the unit interval $0 \leq \Lambda(x) \leq 1$. In the low voltage regime the dimensionless problem then takes the form

$$-\boldsymbol{\nabla}^2 \hat{\phi} = 0 \tag{5.4}$$

in the bulk with

$$-\mathbf{n} \cdot \boldsymbol{\nabla} \hat{\phi} = i\omega\Lambda(x)(\hat{\phi} - \hat{V}_{\text{ext}}) \tag{5.5}$$

on the electrodes and $\mathbf{n} \cdot \boldsymbol{\nabla} \hat{\phi} = 0$ on the channel walls. The time average slip velocity can be expressed as

$$u_s = \Lambda(x)\frac{1}{2}\text{Re}\Big[(\hat{V}_{\text{ext}} - \hat{\phi})\partial_x \hat{\phi}^*\Big] = -\Lambda(x)\frac{1}{4}\partial_x \big|\hat{V}_{\text{ext}} - \hat{\phi}\big|^2. \tag{5.6}$$

Discretization and sensitivity analysis

We approximate both the complex potential $\hat{\phi}$ and the design variable Λ by expanding on a finite basis set

$$\hat{\phi}(\mathbf{r}) = \sum_m \varphi_m(\mathbf{r})\hat{\phi}_m \quad \text{and} \quad \Lambda(x) = \sum_k \chi_k(x)\Lambda_k. \tag{5.7}$$

Multiplying Eq. (5.4) by a test function φ_n and integrating over the computational domain Ω, we obtain the weak form of the problem

$$\int_\Omega \boldsymbol{\nabla}\varphi_n \cdot \boldsymbol{\nabla}\hat{\phi}\,\mathrm{d}\mathbf{r} + i\omega \int_{\text{electrodes}} \varphi_n \Lambda(x)(\hat{\phi} - \hat{V}_{\text{ext}})\,\mathrm{d}x = 0 \qquad \text{for all } n, \tag{5.8}$$

where we used Gauss' law with Eq. (5.5) to eliminate $\mathbf{n}\cdot\boldsymbol{\nabla}\hat{\phi}$. Inserting the expansion for $\hat{\phi}$ we obtain a linear problem for the coefficients $\hat{\phi}_m$

$$\sum_m \left[\int_\Omega \boldsymbol{\nabla}\varphi_n \cdot \boldsymbol{\nabla}\varphi_m\,\mathrm{d}\mathbf{r} + i\omega \int_{\text{electrodes}} \varphi_n \Lambda \varphi_m\,\mathrm{d}x\right]\hat{\phi}_m = i\omega \int_{\text{electrodes}} \varphi_n \Lambda \hat{V}_{\text{ext}}\,\mathrm{d}x. \tag{5.9}$$

This can be written more compactly in matrix form

$$\hat{\mathbf{K}}\hat{\boldsymbol{\Phi}} = \hat{\mathbf{F}}, \tag{5.10}$$

where

$$\hat{K}_{nm} = \int_\Omega \boldsymbol{\nabla}\varphi_n \cdot \boldsymbol{\nabla}\varphi_m\,\mathrm{d}\mathbf{r} + i\omega \int_{\text{electrodes}} \varphi_n \Lambda(x) \varphi_m\,\mathrm{d}x \tag{5.11}$$

and

$$\hat{F}_n = i\omega \int_{\text{electrodes}} \varphi_n \Lambda(x) \hat{V}_{\text{ext}}\,\mathrm{d}x. \tag{5.12}$$

Notice that $\hat{\mathbf{K}}$ is symmetric but not Hermitian. We also define the residual $\hat{\mathbf{L}} = \hat{\mathbf{F}} - \hat{\mathbf{K}}\hat{\boldsymbol{\Phi}}$ in terms of which the electrical problem becomes

$$\hat{\mathbf{L}}(\hat{\boldsymbol{\Phi}}, \boldsymbol{\Lambda}, \omega) = \mathbf{0}. \tag{5.13}$$

We want to maximize the pumping velocity U, and for an array of planar electrodes this is simply defined as as the spatial average slip velocity

$$U = \frac{1}{L}\int_{\text{electrodes}} u_s\,\mathrm{d}x = -\frac{1}{4L}\int_{\text{electrodes}} \Lambda\,(\hat{V}^*_{\text{ext}} - \hat{\phi}^*)\partial_x(\hat{V}_{\text{ext}} - \hat{\phi})\,\mathrm{d}x + \text{c.c.} \tag{5.14}$$

Since the number of parameters Λ_k describing the distribution of the coating material is large, we use the gradient based algorithm MMA [71, 72] to determine the optimum. Hence we need to compute the derivatives

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\Lambda}}\Big[U\big(\hat{\boldsymbol{\Phi}}(\boldsymbol{\Lambda},\omega),\boldsymbol{\Lambda}\big)\Big] \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}\omega}\Big[U\big(\hat{\boldsymbol{\Phi}}(\boldsymbol{\Lambda},\omega),\boldsymbol{\Lambda}\big)\Big], \tag{5.15}$$

where $\boldsymbol{\Phi}(\boldsymbol{\Lambda},\omega)$ is considered an implicit function of $\boldsymbol{\Lambda}$ and ω, defined by the constraint Eq. (5.13). However, because $\hat{\boldsymbol{\Phi}}(\boldsymbol{\Lambda},\omega)$ is implicit and $\partial\hat{\boldsymbol{\Phi}}/\partial\boldsymbol{\Lambda}$ and $\partial\hat{\boldsymbol{\Phi}}/\partial\omega$ are unknown, we cannot simply use the chain rule. Instead we use the adjoint method [82] and introduce the Lagrangian $\mathcal{L}$,

$$\mathcal{L}(\hat{\boldsymbol{\Phi}},\hat{\boldsymbol{\Phi}}^*,\hat{\mathbf{W}},\hat{\mathbf{W}}^*,\boldsymbol{\Lambda},\omega) = U(\hat{\boldsymbol{\Phi}},\hat{\boldsymbol{\Phi}}^*,\boldsymbol{\Lambda}) + \frac{1}{2}\hat{\mathbf{W}}^T\hat{\mathbf{L}} + \frac{1}{2}\hat{\mathbf{W}}^{*T}\hat{\mathbf{L}}^*, \tag{5.16}$$

where both the potential $\hat{\boldsymbol{\Phi}}$ and its complex conjugate $\hat{\boldsymbol{\Phi}}^*$ are considered as independent variables, and $\hat{\mathbf{W}}$ is a Lagrange multiplier for the constraint $\hat{\mathbf{L}}(\hat{\boldsymbol{\Phi}},\boldsymbol{\Lambda},\omega) = \mathbf{0}$. The derivatives in Eq. (5.15) can be computed as the gradient of $\mathcal{L}$ with respect to $\boldsymbol{\Lambda}$ and ω at a point in $(\hat{\boldsymbol{\Phi}},\hat{\boldsymbol{\Phi}}^*,\hat{\mathbf{W}},\hat{\mathbf{W}}^*,\boldsymbol{\Lambda},\omega)$ space where $\mathcal{L}$ is stationary with respect to both $\hat{\boldsymbol{\Phi}}$ and $\hat{\mathbf{W}}$ and their complex conjugates. That is, for a given $\boldsymbol{\Lambda}$ and ω we first compute the potential $\hat{\boldsymbol{\Phi}}$ from

$$\frac{\partial\mathcal{L}}{\partial\hat{\mathbf{W}}} = \hat{\mathbf{F}} - \hat{\mathbf{K}}\hat{\boldsymbol{\Phi}} = \mathbf{0}, \tag{5.17}$$

and the Lagrange multiplier $\hat{\mathbf{W}}$ from

$$\frac{\partial\mathcal{L}}{\partial\hat{\boldsymbol{\Phi}}} = \frac{\partial U}{\partial\hat{\boldsymbol{\Phi}}} - \frac{1}{2}\hat{\mathbf{K}}^T\hat{\mathbf{W}} = \mathbf{0}. \tag{5.18}$$

Then the desired derivatives are given by

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\Lambda}}\Big[U\big(\hat{\boldsymbol{\Phi}}(\boldsymbol{\Lambda},\omega),\boldsymbol{\Lambda}\big)\Big] = \frac{\partial\mathcal{L}}{\partial\boldsymbol{\Lambda}} = \frac{\partial U}{\partial\boldsymbol{\Lambda}} + \mathrm{Re}\Big[\hat{\mathbf{W}}^T\frac{\partial\hat{\mathbf{L}}}{\partial\boldsymbol{\Lambda}}\Big], \tag{5.19}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}\omega}\Big[U\big(\hat{\boldsymbol{\Phi}}(\boldsymbol{\Lambda},\omega),\boldsymbol{\Lambda}\big)\Big] = \frac{\partial\mathcal{L}}{\partial\omega} = \mathrm{Re}\Big[\hat{\mathbf{W}}^T\frac{\partial\hat{\mathbf{L}}}{\partial\omega}\Big]. \tag{5.20}$$

Numerical implementation

We build the physical model and solve the problem Eq. (5.17) using FEMLAB. Therefore the both system matrix $\hat{\mathbf{K}} = -\partial\hat{\mathbf{L}}/\partial\hat{\boldsymbol{\Phi}}$ and the matrix $\partial\hat{\mathbf{L}}/\partial\boldsymbol{\Lambda}$ are computed automatically as part of the solution process and can be obtained directly as MATLAB sparse matrices. Further we need to compute the partial derivatives of U with respect to $\hat{\boldsymbol{\Phi}}$ and $\boldsymbol{\Lambda}$

$$\frac{\partial U}{\partial\hat{\phi}_n} = \frac{1}{4L}\int_{\text{electrodes}}\Lambda\big[(\hat{V}^*_{\text{ext}}-\hat{\phi}^*)\partial_x\varphi_n + \varphi_n\partial_x(\hat{V}^*_{\text{ext}}-\hat{\phi}^*)\big]\,\mathrm{d}x \tag{5.21}$$

$$\frac{\partial U}{\partial\Lambda_k} = -\frac{1}{4L}\int_{\text{electrodes}}\chi_k\,\partial_x\big|\hat{V}_{\text{ext}}-\hat{\phi}\big|^2\mathrm{d}x. \tag{5.22}$$

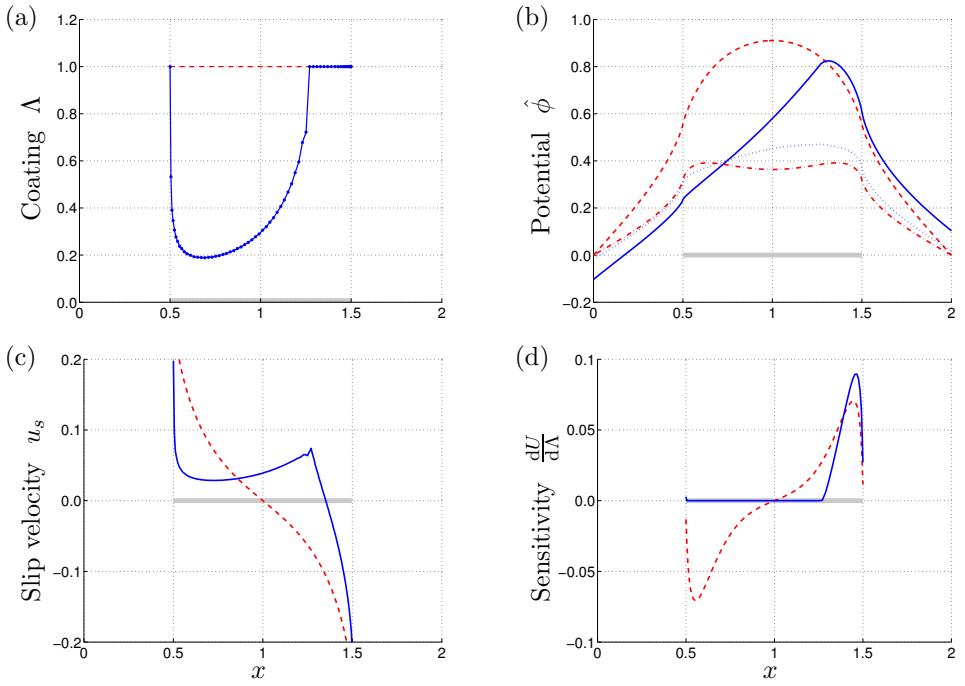


Figure 5.3: Comparison of the solution for optimal coating (solid) at $\omega = 5.1$ where $U = 0.011$ to the solution for uncoated electrodes (dashed) at $\omega = 4.0$ where $U = 0$: (a) Design variable $\Lambda(x)$. (b) Potential $\hat{\phi}(x, y = 0)$ on electrodes; optimal $\mathrm{Re}[\hat{\phi}]$ (solid) and $\mathrm{Im}[\hat{\phi}]$ (dotted) vs. uncoated $\mathrm{Re}[\hat{\phi}]$ (dashed) and $\mathrm{Im}[\hat{\phi}]$ (dash-dot). (c) Time-average slip velocity $u_s(x)$. (d) Sensitivity $\mathrm{d}U/\mathrm{d}\Lambda(x)$. The position of the electrode on the x-axis is marked in gray.

This can also be done with FEMLAB although on the way we need to make an auxiliary copy of the original model; see our paper in Appendix C for details. Finally we need to evaluate

$$\hat{\mathbf{W}}^T \frac{\partial \hat{\mathbf{L}}}{\partial \omega} = i \int_{\text{electrodes}} \Lambda \hat{W} \, (\hat{V}_{\text{ext}} - \hat{\phi}) \, \mathrm{d}x \tag{5.23}$$

$$= \hat{\mathbf{W}}^T \big[\hat{\mathbf{L}}(\hat{\mathbf{\Phi}}, \mathbf{\Lambda}, 1) - \hat{\mathbf{L}}(\hat{\mathbf{\Phi}}, \mathbf{\Lambda}, 0) \big], \tag{5.24}$$

where the last equality follows because $\hat{\mathbf{L}}(\hat{\mathbf{\Phi}}, \mathbf{\Lambda}, \omega)$ is a linear function of ω.

Optimal solution

We consider a symmetric device with electrodes of a fixed width $W = 1$ and gap $G = 1$, cf. Fig. 5.2. The design variable Λ is approximated with first order Lagrange elements, i.e., it is piecewise linear, and we use second order elements for the potential $\hat{\phi}$. The optimization algorithm is initialized with a uniform distribution $\Lambda(x) = 1$ corresponding to uncoated electrodes, and with a frequency of $\omega = 4$ which is about twice the relaxation frequency

for the uncoated device. The initial solution is shown with dashed lines in Fig. 5.3: The potential $\hat{\phi}$ has left-right symmetry so the induced electroosmotic slip velocity does not give rise to any net pumping. The sensitivity $\mathrm{d}U/\mathrm{d}\Lambda$ of the pumping velocity U to local changes in Λ shows that U can be increased by decreasing Λ on the left half of the electrode or by increasing Λ on the right half; however, the latter is not an option due to the bound $\Lambda \leq 1$. At first this result may seem counterintuitive: From Eq. (5.6) we know that at fixed tangential field and potential drop across the double layer, the slip velocity decreases locally with decreasing Λ. However, a decrease in Λ on the left half of the electrode also induces changes in the potential in such way that the overall result is an increase in the net pumping U.

The optimization algorithm adjusts $\Lambda(x)$ according the sensitivity and quickly reaches a solution that is close to optimal, except for some small wiggles at the resolution of the finite element mesh, but these slowly fade away as the design iterations proceed. After 71 iterations the relative increase in pumping velocity is less than 10^{-5} for two successive iterations and we terminate the algorithm. The maximum pumping velocity is $U = 0.011$ at $\omega = 5.1$, and the solution is shown with solid lines in Fig. 5.3: Close to the right edge of the electrode there is an uncoated section of width ~ 0.25 where $\Lambda = 1$, but for the remaining part it turns out to be optimal to add coating. The potential varies almost linearly and the time-average slip velocity u_s is positive throughout the coated section. Notice that in the optimal solution, $\mathrm{d}U/\mathrm{d}\Lambda \neq 0$ only where Λ is limited by the upper bound; in optimization jargon this is known as the complementary slackness condition. Finally, the thickness of the coating in Fig. 5.2 corresponds to the optimal capacitance ratio $\delta = \Lambda^{-1} - 1$, and the streamlines correspond to the optimal solution at $\omega = 5.1$.

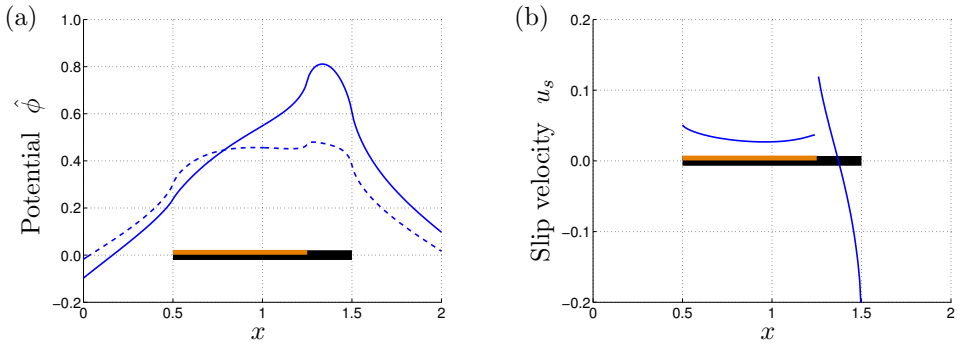


Figure 5.4: Approximating the optimal solution from Fig. 5.3 by a design with electrodes with a uniformly coated section of width $W_1 = 0.75$ where $\Lambda = 0.25$ and an uncoated section of width $W_2 = 0.25$. The maximal pumping velocity is $U = 0.0105$ at $\omega = 5.4$. (a) Potential on electrodes; $\mathrm{Re}[\hat{\phi}]$ (solid) and $\mathrm{Im}[\hat{\phi}]$ (dashed). (b) Slip velocity u_s.

From an engineering point of view it may be possible to fabricate devices

with a nonuniformly graded coating thickness, but with standard lithography techniques it is much easier to obtain piecewise constant coating thickness. Therefore, inspired by the optimal solution we consider a device where the electrodes are divided into two sections of width $W_1 = 0.75$ and $W_2 = 0.25$ where the design variable is $\Lambda_1 = 0.25$ and $\Lambda_2 = 1.0$, respectively. The maximum pumping velocity for this simple design is $U = 0.0105$ at $\omega = 5.4$, which is merely 5% lower than the optimum from Fig. 5.3. The solution is shown in Fig. 5.4; notice that the slip velocity is discontinuous at the point where Λ changes abruptly from Λ_1 to Λ_2.

One may object to the above solutions because they have features below the (unit) characteristic length scale. If the same lithography technique is used to define both the electrodes and the coating on them it is more appropriate to formulate the problem as

$$\begin{aligned} \text{maximize} \quad & U(\log\omega, W_1, W_2, G, \Lambda_1, \Lambda_2 = 1) \\ \text{subject to} \quad & 1 \le W_1, \\ & 1 \le W_2, \\ & 1 \le G, \\ & 0 \le \Lambda_1 \le 1. \end{aligned} \tag{5.25}$$

Here the optimal design is $W_1 = 1.4$, $W_2 = 1$, $G = 1$, $\Lambda_1 = 0.31$, and $\Lambda_2 = 1$, where $U = 0.0096$ at $\omega = 1.9$. This is about 15% lower than the optimum from Fig. 5.3, but better suited for fabrication at the lower limit on resolution set by standard lithography.

Compared to the device with asymmetric pairs of electrodes, asymmetric coating allows the pumping velocity to be increased by a almost a factor of 3, at the expense of an additional lithography step to define the coating. A disadvantage of the use an asymmetric coating to break the left-right symmetry is that the design variable Λ depends on the electrolyte concentration through the Debye layer capacitance. Therefore the coating thickness will have to be designed for a particular concentration, while off that concentration the performance is suboptimal.

5.3 3D structured electrodes with regular shape

While most work so far have focused on planar designs with thin flat electrodes on a substrate, one may obtain a significant gain in pumping efficiency by stepping into the third dimension: Bazant and Ben recently suggested the use electrodes like those shown in Fig. 5.5 [45], for which they predict pumping velocities an order of magnitude larger than the planar design of Brown *et al.* They also tested the 3D design experimentally and obtained already a factor of 7 larger pumping velocity than with planar devices of the same minimal feature size [83, 69].

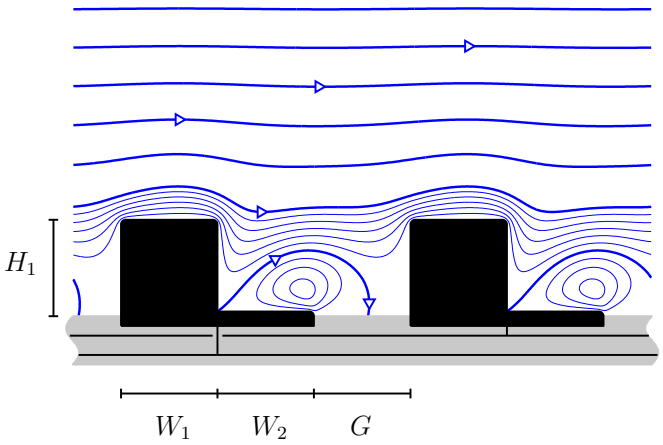


Figure 5.5: Stepped 3D electrode geometry: Each electrode in the array has a raised section of width W_1 and height H_1 and a base section of width W_2 and the gap between adjacent electrodes is G. Streamlines correspond to the solution with $W_1 = W_2 = H_1 = G = 1$ at driving frequency $\omega = 0.8$; close to the electrode extra streamlines are plotted to show the details of the flow pattern.

The basic design principle that they propose is to raise those sections of the electrodes where the slip is in the "forward" direction and to lower the sections with "reverse" slip. This enhances the forward flow because the reverse vortex is living in a trench, where it is confined by the vertical solid electrode surface rather than struggling up against the forward flow. On the contrary: At the level of the forward slipping surface the recirculation from the vortex enhances the forward pumping instead of inhibiting it. Effectively, the electrode array now acts as a fluid "conveyor belt" where all sections help pumping the fluid in the forward direction [45].

We consider first a geometry as shown in Fig. 5.5 with electrodes that have both a raised section of width W_1 and height H_1, a base section of width W_2, and with a gap G between adjacent electrodes. In the simple case when all dimensions are set to unity, $W_1 = W_2 = H_1 = G = 1$, the net pumping is maximized around $\omega = 0.8$. By analogy with the planar pumps we calculate the effective pumping velocity as $U \approx 2\,Q/H_{\mathrm{tot}} \approx 0.05$, where $H_{\mathrm{tot}} \gg 1$ is the overall channel height measured from the substrate to the channel top wall. This is an order of magnitude larger than the optimal result from Sec. 5.1, which clearly demonstrates the potential of 3D designs.

For planar designs the pumping velocity U, defined as the spatial average slip velocity, comes out of a simple Fourier analysis as the natural way of quantifying the pump performance. For 3D structured electrodes it is not immediately obvious how the pumping velocity should actually be defined, or if it is appropriate for quantifying the pump performance at all. To clarify this we compute the net flow rate Q as a function of the channel height H_{tot}. The result is shown in Fig. 5.6 where we clearly see a linear dependence for

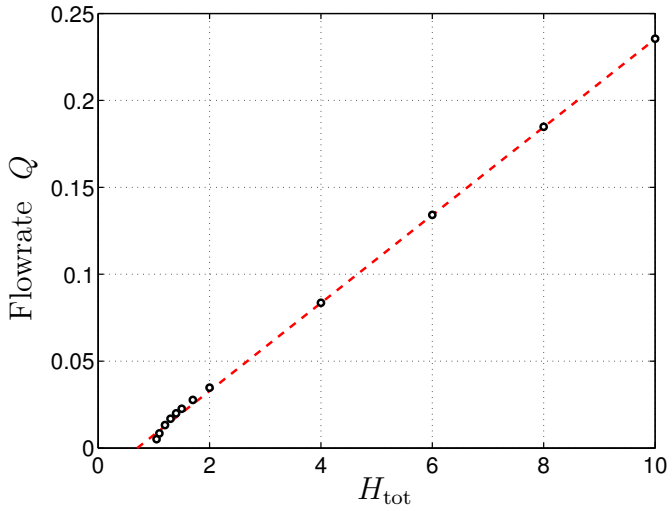


Figure 5.6: Maximal flowrate $\max_\omega Q$ vs. total channel height H_{tot} measured from the substrate to the channel top wall in a geometry with $W_1 = W_2 = H_1 = G = 1$. For $H_{\mathrm{tot}} \gg 1$ there is a linear dependence: Data points (circles) and linear fit (dashed line). For $H_{\mathrm{tot}} = 1$ the top wall touches the raised electrode section and the flow is blocked.

$H_{\mathrm{tot}} \gg 1$. For a given electrode design we can therefore uniquely define the pumping velocity as

$$U = \lim_{H_{\mathrm{tot}} \to \infty} \frac{2\,Q(H_{\mathrm{tot}})}{H_{\mathrm{tot}}}. \tag{5.26}$$

Extrapolating $Q(H_{\mathrm{tot}})$ to zero we get a nonzero intercept at around $H^* = 0.7$, which is perhaps not so surprising, given the ambiguity in the definition of the channel height, i.e., should it be measured from the substrate or from the top of the electrodes? For $H_{\mathrm{tot}} = H_1 = 1$ the channel top wall touches the raised electrode section and the flow is blocked. In order to estimate U from data at finite H_{tot} we could either use $U \approx 2\,Q/(H_{\mathrm{tot}} - H^*)$ or $U \approx 2\,\partial Q/\partial H_{\mathrm{tot}}$. However, because H^* depends on H_1, and $\partial Q/\partial H_{\mathrm{tot}}$ requires finite differencing, we simply choose a "large" $H_{\mathrm{tot}} = 40$ and use $U \approx 2\,Q/H_{\mathrm{tot}}$ throughout this section.

In order to compute the net flow rate Q we need to solve the Stokes flow problem with a slip condition on the electrodes. To avoid numerical difficulties associated with infinite slip velocity from the electric field singularity at sharp corners on the electrode we use a geometry model with rounded corners of radius R. Smaller R gives stronger field at the protruding corners which actually tends to increase the pumping velocity U, but the limit of sharp corners is finite. We use $R = 0.02$ which is not very sharp, and therefore the pumping velocity that we report here can be up to a few percent lower than the sharp limit.

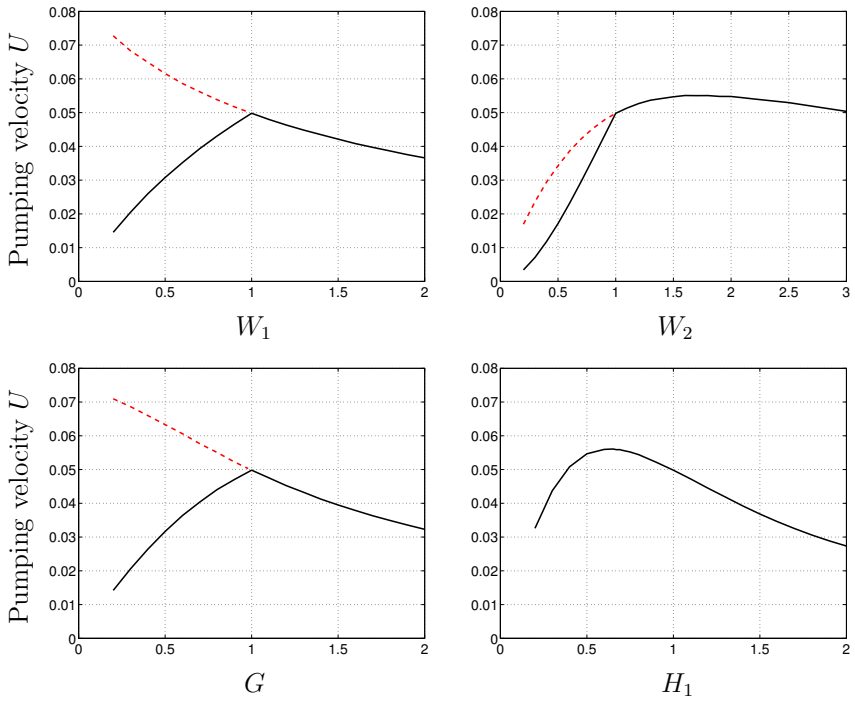


Figure 5.7: Maximum pumping velocity $\max_\omega U$ as a function of each one of the individual geometrical parameters W_1, W_2, G, and H_1, with the other parameters fixed at unity. Dashed and solid lines show the results using 1 and $\min\{1, W_1, W_2, G\}$ as the characteristic length scale, respectively.

Optimal solution

Using standard planar lithography to process the device, it is much easier to control the thickness of the raised section on the electrodes than the transverse dimensions. Therefore we formulate the optimal design problem as

$$\begin{aligned} \text{maximize} \quad & U(\log\omega, W_1, W_2, H_1, G) \\ \text{subject to} \quad & 1 \le W_1, \\ & 1 \le W_2, \\ & 1 \le G. \end{aligned} \tag{5.27}$$

In Fig. 5.7 we investigate the dependence of the maximum pumping velocity $\max_\omega U$ on each the individual geometrical parameters W_1, W_2, G, and H_1, when the other parameters are fixed at unity. The figure shows that the optimum for both W_1 and G is at the lower limit $W_1 = G = 1$ whereas for W_2 and H_1 it is around $W_2 \approx 1.7$ and $H_1 \approx 0.6$, respectively. The global optimum can be determined from a contour plot as in Fig. 5.8 showing the maximum pumping velocity as a function of both W_2 and H_1: It is close to

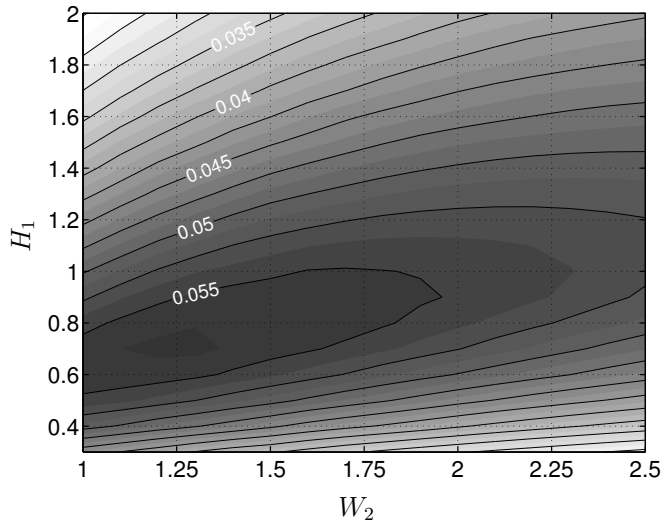


Figure 5.8: Contour plot of the maximum pumping velocity $\max_\omega U$ as a function of W_2 and H_1 for $W_1 = G = 1$. There is a unique optimum around $W_2 = 1.25$ and $H_1 = 0.7$ where $U = 0.0566$ at $\omega = 0.8$.

$W_2 = 1.25$ and $H_1 = 0.7$, where $U = 0.0566$ at $\omega = 0.8$. However, the figure also shows that the optimum is broad. In particular, the simple design with $H_1 = 0.7$ and $W_1 = W_2 = G = 1$ has $U = 0.0545$ at $\omega = 0.9$ which is within 2% of the global optimum, and our starting point $W_1 = W_2 = H_1 = G = 1$ is within 12%.

A virtue of the simple 3D structured rectangular design in Fig. 5.5 is that it can be fabricated with just two lithography steps to define the base and raised sections of the metal electrodes on top of a flat substrate. However, the basic design principle in [45] speaks of both raising sections of "forward" slip and recessing sections of "reverse" slip on the electrode. Therefore we can generalize the simple design by allowing the base section of the electrode to be lowered by an amount H_2 relative to the substrate level as shown in Fig. 5.9(a). The optimal solution to this problem turns out to be $W_1 = 1$, $W_2 = 1$, $G = 1$, $H_1 = 0.4$, and $H_2 = 1$, where $U = 0.078$ at $\omega = 0.65$. Yet another possibility is to add thin metal electrodes on top of a 3D structured substrate as shown in Fig. 5.9(b); D. Burch reported that such a design would increase the pumping velocity over the solid metal electrode design [84]. The optimal solution is, roughly, $W_1 = 1$, $W_2 = 1.2$, $G = 1.0$, $H_1 = 0.9$, and $H_2 = 0.3$, where $U = 0.089$ at $\omega = 2.1$. Whether or not these more complex designs are actually superior to the simple design in Fig. 5.5 depends on how difficult it is to pattern both the metal electrodes and the substrate: If it turns out that the minimal linewidth that one can (routinely) produce in the more complex designs is, say, twice as large as for the simple design, then the simple design is actually superior due to the

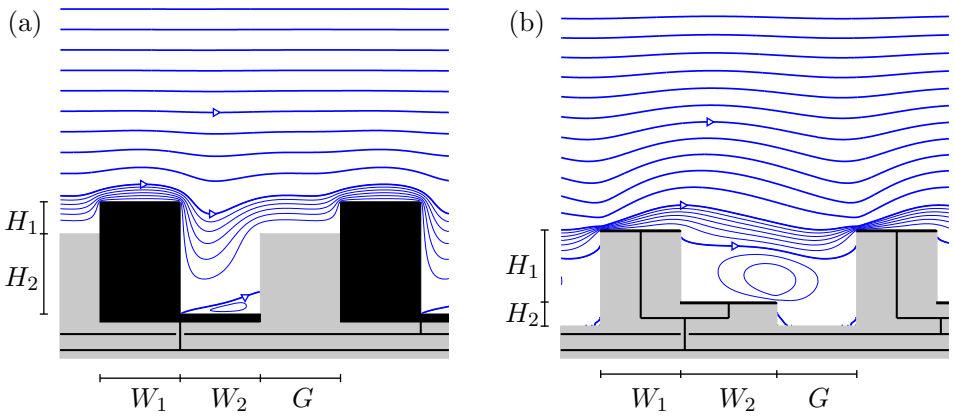


Figure 5.9: Stepped and grooved 3D electrode geometries: (a) Design with the "reverse" slipping section of the electrode put into a groove. The optimal solution is $W_1 = 1$, $W_2 = 1$, $G = 1$, $H_1 = 0.4$, $H_2 = 1$, where $U = 0.078$ at $\omega = 0.65$. (b) Design with thin metal electrodes deposited on a 3D structured substrate. The optimal solution is $W_1 = 1$, $W_2 = 1.2$, $G = 1$, $H_1 = 0.9$, $H_2 = 0.3$, where $U = 0.089$ at $\omega = 2.1$.

fundamental scaling of the pumping velocity with one over the characteristic length scale.

5.4 3D structured electrodes with arbitrary shape

A natural question to ask is whether the rectangular electrode structures we investigated in the previous section are (close to) optimal in general, or if some other shape with sloped side walls could improve the performance even further. In order to answer this question we now consider electrodes with a general shape described by some continuous function $\bar{\mathbf{r}}(s)$. Because the number of parameters to describe $\bar{\mathbf{r}}(s)$ is large (infinite) we will need to compute the derivative $\mathrm{d}U/\mathrm{d}\bar{\mathbf{r}}$ of the pumping velocity U with respect to local changes in the electrode shape exactly rather than using finite difference approximation. One way to obtain this is by an *arbitrary Lagrangian Eulerian* (ALE) type formulation, introducing a mapping from the physical frame (x, y) to a fixed reference frame (X, Y). The idea is to continuously deform the physical domain Ω to match the electrode shape, but keep a fixed finite element mesh in the reference domain Ω^* where we solve the problem numerically. When we transform the weak form of the governing equations from the physical frame to the reference frame we make use of the transformation rules

$$\int_{\Omega} f(\mathbf{r})\,\mathrm{d}\mathbf{r} = \int_{\Omega^*} f\big(\mathbf{r}(\mathbf{R})\big)\,\mathrm{det}\mathbf{J}\,\mathrm{d}\mathbf{R} \tag{5.28}$$

and

$$\frac{\partial}{\partial x_i} = \frac{\partial X_j}{\partial x_i}\frac{\partial}{\partial X_j}, \tag{5.29}$$

where det$\mathbf{J}$ is the determinant of the Jacobian matrix $J_{ij} = \partial x_i/\partial X_j$ for the mapping from the reference domain Ω^* to the physical domain Ω, and $\partial X_j/\partial x_i$ is an element of the inverse Jacobian inv$\mathbf{J}$. The gradient $\mathrm{d}U/\mathrm{d}\bar{\mathbf{r}}$ is computed using the adjoint method as described in Sec. 5.2. It is straightforward in principle, but more cumbersome because there are more dependent variables in the present problem, including the mesh deformation $\mathbf{r}(\mathbf{R})$, complex potential $\hat{\phi}(\mathbf{r})$, fluid velocity $\mathbf{u}(\mathbf{r})$ and pressure $p(\mathbf{r})$, and a number of Lagrange multipliers that we account for explicitly to gain full control over which constraint forces act where in the governing equations.

Discretization

Our bold statement that we consider a general continuous curve $\bar{\mathbf{r}}(s)$ for the electrode shape is subject to a number of constraints, e.g., in order to be a valid shape $\bar{\mathbf{r}}(s)$ cannot intersect itself. Moreover we want it to be fairly "nice" with not too many or too sharp wiggles.

We discretize the problem by introducing a piecewise linear approximation to $\bar{\mathbf{r}}(s)$ with N line segments defined by $N+1$ nodes $\bar{\mathbf{r}}_n$ where the endpoints $\bar{\mathbf{r}}_0$ and $\bar{\mathbf{r}}_N$ are fixed. In order to gain more control over the shape of $\bar{\mathbf{r}}$ and avoid clustering of the nodes we force all line segments to have the same length a. Then the nodal coordinates can be expressed as

$$\bar{x}_n = \bar{x}_0 + a\sum_{m=1}^{n-1}\cos\theta_m \quad \text{and} \quad \bar{y}_n = \bar{y}_0 + a\sum_{m=1}^{n-1}\sin\theta_m, \tag{5.30}$$

where θ_m is the angle between the mth segment and the horizontal axis, i.e., $\tan\theta_m = \Delta\bar{y}_m/\Delta\bar{x}_m$. The shape of $\bar{\mathbf{r}}$ is uniquely determined by the $N-1$ first angles whereas the final angle θ_N and the segment length a are chosen to match the final node with the fixed endpoint $\bar{\mathbf{r}}_N$. Alternatively we can parametrize in terms of the $N-1$ relative angles $\Theta_m = N\left(\theta_{m+1} - \theta_m\right)$ between adjacent line segments; Θ is also a measure of the local curvature $\kappa = \partial_s\theta$ of $\bar{\mathbf{r}}$, where s is the arc length in the physical frame, since

$$\Theta_m = N\big(\theta_{m+1} - \theta_m\big) = Na\frac{\Delta\theta_m}{a} \approx L_{\mathrm{tot}}\kappa_m, \tag{5.31}$$

writing $L_{\mathrm{tot}} = Na$ for the overall arc length of $\bar{\mathbf{r}}$. In the examples below we control the shape of $\bar{\mathbf{r}}$ by limiting Θ_m to a finite interval $\Theta_{\min} \le \Theta_m \le \Theta_{\max}$. This effectively limits the curvature of $\bar{\mathbf{r}}$ and determines how sharp features can be on the electrode, but also sets a limit on the overall arc length, cf. Eq. (5.31). In the design process the optimization algorithm will therefore have to make a trade-off between curvature and arc length. This makes the

interpretation of the results more complex, but since we like the electrodes to have neither too sharp features nor too large a surface area we are still content with our approach.

Optimal solutions

As a first design problem, inspired by the results for rectangular electrodes in the previous section, we consider a device with a fixed width $W = 2$ of the electrode and a fixed gap of $G = 1$. To avoid difficulties with fabrication we do not allow undercut or backward sloping structures, i.e., we require $-\pi/2 \leq \theta \leq \pi/2$, and formulate the optimization problem as

$$\begin{aligned} &\text{maximize} && U(\log\omega, \boldsymbol{\Theta}) && && (5.32)\\ &\text{subject to} && -\Theta_{\max} \leq \Theta_n \leq \Theta_{\max} && \text{for } n = 1, \ldots, N-1,\\ &&& -\pi/2 \leq \theta_m \leq \pi/2 && \text{for } m = 1, \ldots, N. \end{aligned}$$

The optimal solution to this problem for different values of $\Theta_{\max}$ with $N = 100$ is shown in Fig. 5.10. In panel (a) we set $\Theta_{\max} = 6.66$ which is just enough to allow for two half circles of radius $R \approx 0.5$, i.e., curvature $\kappa = 1/R \approx 2$, and a total arc length around π. The maximal pumping velocity is $U = 0.078$, which is almost as good as the corresponding optimal rectangular grooved structure from Sec. 5.3, even though the present electrode shape completely lacks sharp corners to enhance the electric field strength. In panels (b)-(f) we gradually increase $\Theta_{\max}$: We see first the width of the raised section decrease, which is in accordance with the trend in Fig. 5.7(a) for $W_1 < 1$. Secondly, we see the lowered section develop into a double groove, where the "purpose" of the extra "pillar" dividing the groove into two seems to be to weaken the counter-rotating vortex. In the first four panels (a)-(d) the optimal solution is everywhere limited either by $|\Theta_n| \leq \Theta_{\max}$ or by $|\theta_m| \leq \pi/2$, but in panels (e) and (f) this is no longer the case: In particular in panel (f) neither of the two "pillars" are capped with half circles but have a width and shape that start to be "optimal" in itself rather than limited by the particular constraints that we impose.

In our second design we relax the constraint on backward sloping structures, but maintain that all nodes should conform with the overall electrode width W. Thus we formulate the problem as

$$\begin{aligned} &\text{maximize} && U(\log\omega, \boldsymbol{\Theta}) && && (5.33)\\ &\text{subject to} && -\Theta_{\max} \leq \Theta_n \leq \Theta_{\max} && \text{for } n = 1, \ldots, N-1,\\ &&& \bar{x}_0 \leq \bar{x}_m \leq \bar{x}_N && \text{for } m = 1, \ldots, N-1. \end{aligned}$$

The result is shown in Fig. 5.11, again with $N = 100$ line segments to describe the electrode shape. The overall features in the solutions are the same as in Fig. 5.10: The width of the raised section decreases and a double

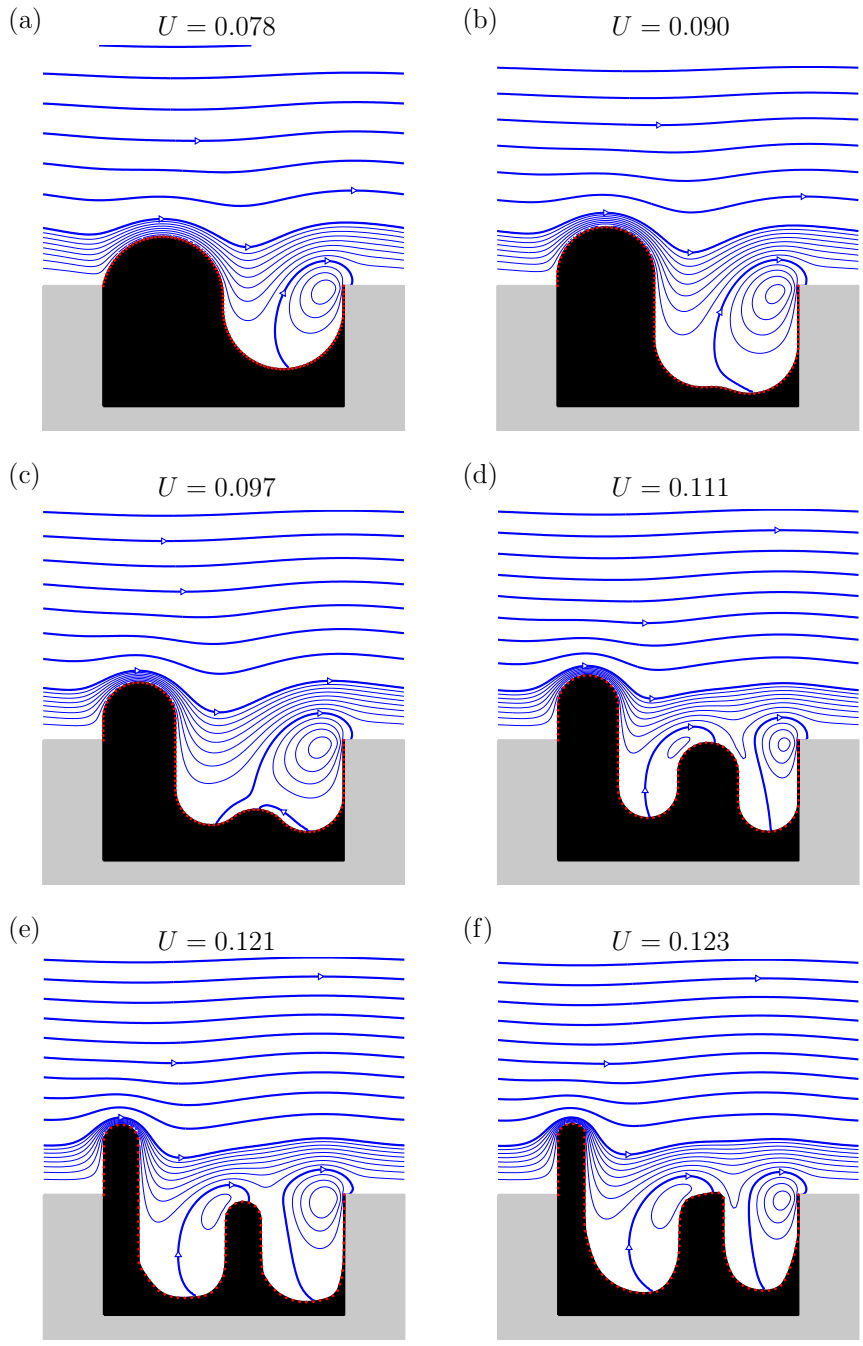


Figure 5.10: Electrode shapes (black) computed as optimal solutions to the problem Eq. (5.32) for different values of the maximum curvature $\Theta_{\max}$: (a) $\Theta_{\max} = 6.66$. (b) $\Theta_{\max} = 10$. (c) $\Theta_{\max} = 13.3$. (d) $\Theta_{\max} = 20$. (e) $\Theta_{\max} = 40$. (f) $\Theta_{\max} = 80$. The electrode width is fixed at $W = 2$ and the gap at $G = 1$ and the electrode shape is described by $N = 100$ line segments (nodes in red).

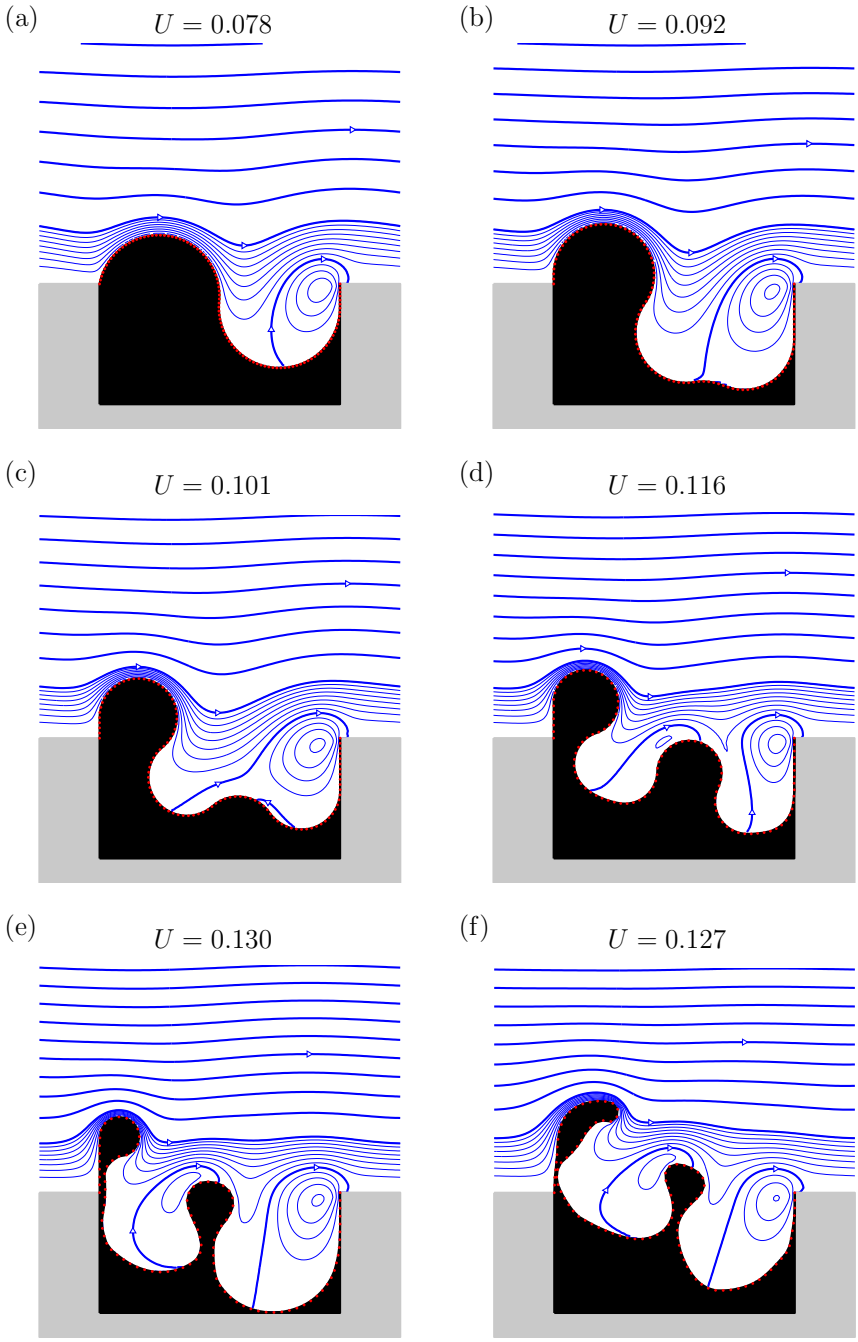


Figure 5.11: Electrode shapes (black) computed as optimal solutions to the problem Eq. (5.33) for different values of the maximum curvature $\Theta_{\max}$: (a) $\Theta_{\max} = 6.66$. (b) $\Theta_{\max} = 10$. (c) $\Theta_{\max} = 13.3$. (d) $\Theta_{\max} = 20$. (e) $\Theta_{\max} = 40$. (f) $\Theta_{\max} = 80$. The shapes in (e) and (f) are not fully converged solutions but were terminated after 200 design iterations.

groove develops when $\Theta_{\max}$ is increased. At the same time the extent of the backward sloping becomes more and more pronounced, and the two last solutions in panels (e) and (f) look pathological in the sense that they contain very narrow parts in the electrode shape. Actually, panels (e) and (f) do not correspond to fully converged optimal solutions because in both cases the design process was terminated after a total of 200 iterations.

It is difficult to compare the performance of our optimized electrode shapes from Figs. 5.10 and 5.11 with the rectangular geometries from Sec. 5.3 because strictly they are not optimal solutions to the same problem: In Sec. 5.3 we optimized the geometry subject to a constraint on the minimal transverse linewidth allowed, whereas in the present section we fix the total width of the electrode at $W = 2$ and optimize subject to a constraint on the minimal feature size as represented by the maximal curvature. However, looking at the results in Fig. 5.10 and 5.11 we find that, largely, the optimization algorithm seems to follow the basic design principle of raising the forward slipping and recessing the reverse slipping sections on the electrodes, even though it does not know of any such principle but only has information on the local gradient $\mathrm{d}U/\mathrm{d}\bar{\mathbf{r}}$. It therefore seems likely that the simple rectangular geometries from Sec. 5.3 are indeed (close to) optimal because we constructed their low-dimensional design spaces exactly based on this basic principle together with our knowledge of what electrode structures can most easily be manufactured in practice.

5.5 Thermodynamic efficiency

Throughout this chapter we have focused on the maximum pumping velocity that can be obtained for a given driving voltage. Equally important might be the maximum pumping velocity for a given electrical power consumption used to drive the electrode array, or the maximum amount of "useful work" that the pump can deliver at a given power consumption, i.e., the thermodynamic efficiency [7, 9]. Provided convection can be neglected in the electrical part of the problem, which is true for our linear and weakly nonlinear models but not in the strongly nonlinear regime, the net flow rate Q can be decomposed into separate and independent electroosmotic and pressure-driven components

$$Q = Q_{\mathrm{eo}} + Q_p. \tag{5.34}$$

The electroosmotic contribution to the net flow rate per unit width of the channel is

$$Q_{\mathrm{eo}} = UH/2 \tag{5.35}$$

where U is the pumping velocity and H is the height of the pumping channel. The pressure driven flow is given by

$$Q_p = -\Delta p/R_{\mathrm{hyd}} \tag{5.36}$$

where Δp is the backpressure that the pump is pumping up against, and $R_{\rm hyd}$ is the (dimensionless) hydraulic resistance per unit width

$$R_{\rm hyd} = \frac{12 L_{\rm tot}}{H^3} \tag{5.37}$$

where $L_{\rm tot}$ is the total length of the channel. The Q-Δp diagram for the pump is linear, with maximal flow rate $Q_{\rm max} = Q_{\rm eo}$ obtained at zero backpressure, whereas the maximal backpressure $\Delta p_{\rm max} = Q_{\rm eo}/R_{\rm hyd}$ is defined as the point where the pump stalls and Q vanishes. The work done when pumping up against an external backpressure is $P_{\rm out} = Q\Delta p$ which is maximized at $\Delta p = \Delta p_{\rm max}/2$ where

$$P_{\rm out} = \frac{1}{4} R_{\rm hyd} Q_{\rm eo}^2. \tag{5.38}$$

The electrical power consumption is

$$P_{\rm in} = \frac{1}{2}{\rm Re}[\hat{I}^* \Delta \hat{V}] = 2N{\rm Re}\left[\frac{1}{\hat{Z}}\right] \tag{5.39}$$

where $\Delta\hat{V} = 2$ is the (dimensionless) voltage applied between the electrodes when $V_{\rm ext} = \pm{\rm Re}[e^{i\omega t}]$, I is the total current passed though the device, N is the number of electrode pairs, and $\hat{Z}$ is the impedance for a single pair, cf. Sec. 3.1.1. To optimize for thermodynamic efficiency one should therefore maximize the ratio

$$\frac{P_{\rm out}}{P_{\rm in}} = \beta \frac{LU^2}{{\rm Re}[1/\hat{Z}]}, \tag{5.40}$$

where $L = W_1 + G_1 + W_2 + G_2$ is the array period and β is a dimensionless parameter given by

$$\beta = \frac{3}{8}\frac{\epsilon}{\sigma H}\frac{\epsilon V_0^2}{\eta \ell_0 (1+\delta)^2} = \frac{3\tau_D u_0}{8H(1+\delta)^2}. \tag{5.41}$$

The dimensionless function $LU^2/{\rm Re}[1/\hat{Z}]$ should be optimized geometrically, but otherwise the scaling of β shows that it pays to pump an electrolyte of low conductivity σ because this lowers the power consumption at a given driving voltage, but does not (in theory) affect the pumping velocity. It pays to squeeze down the channel lid and lower H because this increases the internal hydraulic resistance and enables to pump up against a more decent backpressure. It pays scale down the device and decrease ℓ_0 as much as possible. And finally, according to the low-voltage linear theory, it pays to increase the driving voltage V_0 and decrease the capacitance ratio δ, which should hold also at larger voltage although the scaling is different. Substituting typical numbers for a 1 mM KCl solution in a device with $\ell_0 = 5$ μm, $H = 50$ μm, $\delta = 0$, and $V_0 = 1$ V, and using $U \simeq 0.0035$

and $L \simeq 15$ as corresponding to the optimized planar device, we obtain a thermodynamic efficiency of

$$\frac{P_{\text{out}}}{P_{\text{in}}} \approx 10^{-8}. \tag{5.42}$$

What this tells us is that an AC electroosmotic pump may be employed to move liquid rapidly through a microchannel, but not against any significant backpressure.

This should actually be no surprise to us because the same limitation applies to ordinary inline DC electroosmotic pumps. To overcome this, several authors have introduced a porous material or sintered glass frit into the pumping channel, with effective pore radius a from a few μm down to 100 nm [8, 9, 10, 11, 12]. The thermodynamic efficiency is maximized at $a \approx 2.5\lambda_D$, and experimentally efficiencies of around 1 % have been obtained pumping deionized water [9] an up to 5.6 % for a buffer solution with the strongly dipolar additive tris(hydroxymethyl)aminomethane hydrocloride [10]. However, introducing porous material in the pumping channel does not enhance the efficiency of the AC electroosmotic pump. Our results from Ref. [44] show that strong geometrical confinement of the pumping channel down to $H \ll L$ does not, as one might have worried, mask the asymmetry of the electrodes and destroy the ability to pump. Hence geometrical confinement could be a route for increasing the thermodynamic efficiency, but there is a long way to go from 10^{-8}. We believe it is more fruitful to think of AC electroosmosis as a means for local flow control in a microfluidic system, and to focus on ease integration rather than (lack of) energy efficiency.

5.6 Summary

In this chapter we have investigated different strategies for breaking the left-right asymmetry of an electrode array to create an AC electroosmotic pump, and determined in each case the design that maximizes the pumping velocity for a given (low) driving voltage. From a fabrication point of view the easiest device to produce is that based on asymmetric pairs of electrodes, but the 3D structured design suggested by Bazant and Ben [45] could increase the pumping velocity by an order of magnitude at the expense of an additional lithography step. The thermodynamic efficiency is low, so one should focus on the use of AC electroosmotic pumps for local generation of flow but not for pumping up against large external backpressure or hydraulic resistance.

Chapter 6

Conclusion and outlook

Conclusion

In the present thesis we have studied AC electrokinetics with focus on micropumps based on asymmetric electrode arrays.

We have extended the existing linear theory for AC electroosmosis into the weakly nonlinear regime by accounting for nonlinear capacitance of the Debye layer and Faradaic current injection from electrochemical reactions at the electrodes. The primary effect of the nonlinear capacitance is to change the scaling of the pumping velocity with driving voltage, and to shift down the inverse RC time to lower frequency at large driving voltage. Depending on the application it may be feasible to operate a micropump with Faradaic current injection, provided that the electrode reaction runs reversibly such that electrolysis and bubble formation is avoided. We conclude that fast electrode reactions act to "short-circuit" the Debye layer and thereby suppress electroosmotic flow, unless the reaction itself is limited by mass-transfer of the reactants. If the electrochemically active species are only present in small concentrations with an excess of inert supporting electrolyte, the Faradaic current from the reaction levels off when the diffusion-limited current is reached, and the dynamical behaviour at larger voltage is dominated by capacitive charging. On the other hand, if there is no inert supporting electrolyte the AC Faradaic current should give rise to strong concentration polarization above the diffusion-limited current with transient space charge layers and electroosmosis of the second kind. However, that is a strongly nonlinear phenomenon that we have have not investigated.

Neither the linear nor weakly nonlinear models are able to account for the reversal of the pumping velocity observed experimentally, and therefore we have also studied the strongly nonlinear regime where the classical assumption of uniform bulk electrolyte concentration breaks down. Building on recent theoretical work for diffuse-charge dynamics in electrochemical systems, we have constructed a strongly nonlinear model based on bound-

ary layer analysis and matched asymptotic expansions. In particular, we have contributed by coupling an explicit dynamical solution in the diffusion layer to the evolution of the quasiequilibrium Debye layer and to the bulk concentration profile that is nonuniform but stationary in time. We also account for fluid motion and convection of salt in the bulk.

We have obtained numerical solutions for this model in both a simple 1D geometry and in 2D for a symmetric and an asymmetric electrode array. Moreover, because even our strongly nonlinear model based on local electroneutrality breaks down at "very large" but still experimentally relevant voltage, we have attempted to include existing theory for nonequilibrium double layers and electroosmosis of the second kind into our model. We predict that electroosmotic flow of the first and second kind contributes to pumping in the forward direction, while bulk electroconvection contributes to pumping in the reverse direction. However, for the voltage range that we have considered, electroosmosis dominates, so we are not able to account for the experimentally observed reversal of the pumping direction with our strongly nonlinear model, either.

Finally, we have looked at various ways of breaking the symmetry of an electrode array and determined in each case the design that maximizes the pumping velocity at a given low driving voltage. The solutions are expected to be close to optimal also at larger voltage as long as electroosmosis of the first kind dominates the electrokinetic flow. It is important to realize that AC electrokinetic micropumps can provide fast pumping of the order of $\sim$ mm/s, but not against a significant backpressure. Thermodynamic efficiency will therefore always be low, but the technology is still attractive due to ease of integration in a microsystem or lab-on-a-chip, and because it offers local control over the fluid motion in the device.

Outlook

There remains many open problems and possible directions for future research on strongly nonlinear electrokinetics. What makes this subject so challenging, and the experimental results so difficult to interpret, is that there is a whole zoo of strongly nonlinear effects showing up at roughly the same time, i.e., salt depletion in the diffusion layer followed by breakdown of local electroneutrality and development of an extended space charge layer, large surface excess flux in the Debye layer leading to nonuniform bulk concentration with convection dominated bulk salt flux, and even breakdown of dilute solution theory in the highly charged Debye layer. A full solution of the Poisson–Nernst–Planck equations provides only little insight into the solution structure, but for simple geometries it can be useful for validating the strongly nonlinear theory.

Because the different effects interact in a nontrivial way we can envisage

complex dynamical behaviour in geometries such as the AC electroosmotic pump with 3D structured electrodes. However, before this can fully be exploited to optimize the pump performance or to engineer robust "forward" operation at large voltage, it is necessary to gain a more detailed understanding of the dynamical behaviour, e.g., by studying simple geometries such as metal spheres or cylinders.

The theoretical description for DC rather than AC driving is simplified because the diffusion layer and bulk region merge and the number of nested boundary layers drop by one. It would therefore be interesting to extend the work of Chu on a metal sphere in an externally applied DC field to even larger fields, where local electroneutrality breaks down in the bulk region close to the poles of the sphere due to large surface fluxes towards the equator. The space charge layers developing in this case are immediately describable by the existing theory for DC Faradaic conduction.

Alternatively, the existing theory for nonequilibrium double layers could deserve a rigorous account for the solution structure when the width of the space charge layer changes in time with the applied current, and in particular the dynamics of the collapse of the layer when the current reverses is a challenging problem. In relation to the simple *ad hoc* modifications that we suggested, the quality of the model would certainly improve by solving a 1D semi-infinite diffusion problem on the interval $\check{y}_o(t) \leq \check{y} < \infty$ for the concentration in the diffusion layer, and it would be a great advantage to device a timestepping algorithm to allow for better resolution in time when computing the periodic solution.

Finally, the dependence on the model parameters ε, δ, ΔD, and possibly $\tilde{\zeta}^{\mathrm{eq}}$, deserves more attention and systematic variation. In particular, lower values of the capacitance ratio δ are relevant to many experiments. To this end it would also be interesting to actually determine the experimental values of δ and $\tilde{\zeta}^{\mathrm{eq}}$ by fitting the linear response theory to impedance spectra recorded as a function of frequency and with a DC voltage applied relative to a distant reference electrode.

Appendix A

FEMLAB code

A.1 Weakly nonlinear finite element model

In this section we give an example of our implementation of the weakly nonlinear model from Sec. 2.2.7 in the absense of Faradaic electrode reaction. The leading order electrical problem reduces to a Laplace problem in the bulk

$$0 = -\boldsymbol{\nabla}^2 \bar{\phi}, \tag{A.1}$$

with

$$0 = \mathbf{n} \cdot \boldsymbol{\nabla} \bar{\phi}, \tag{A.2}$$

on the insulating channel walls and a set of dynamical boundary conditions

$$\partial_t \tilde{q} = \mathbf{n} \cdot \boldsymbol{\nabla} \bar{\phi}, \tag{A.3}$$

$$V_{\text{ext}} - \bar{\phi} = \tilde{\zeta} - \tilde{q}\,\delta, \tag{A.4}$$

on the electrodes, where the zeta potential is given by

$$\tilde{\zeta} = -2 \sinh^{-1}(\tilde{q}/2). \tag{A.5}$$

The solution to the electrical problem determines the electroosmotic slip velocity that is boundary condition for the bulk time average Stokes flow problem for the fluid velocity $\bar{\mathbf{u}}$ and pressure $\bar{p}$

$$\mathbf{0} = \boldsymbol{\nabla} \cdot \bar{\boldsymbol{\sigma}}, \tag{A.6}$$

$$0 = \boldsymbol{\nabla} \cdot \bar{\mathbf{u}}, \tag{A.7}$$

where $\bar{\boldsymbol{\sigma}}$ is the stress tensor defined by

$$\bar{\sigma}_{ij} = -\bar{p}\,\delta_{ij} + \partial_i \bar{u}_j + \partial_j \bar{u}_i. \tag{A.8}$$

The slip condition on the electrodes is

$$\bar{\mathbf{u}} = \langle \tilde{\zeta} \boldsymbol{\nabla}_s \bar{\phi} \rangle, \tag{A.9}$$

and on the channel walls we impose the no-slip condition $\bar{\mathbf{u}} = 0$.

The electrical problem is converted to weak form by multiplying Eq. (A.1) by a test function φ for the potential and integrating over the volume Ω of the computational domain, employing Gauss' law, and using Eq. (A.3) to eliminate $\mathbf{n} \cdot \boldsymbol{\nabla} \bar{\phi}$, i.e.,

$$\begin{aligned} 0 = -\int_\Omega \varphi \, \boldsymbol{\nabla}^2 \bar{\phi} \, \mathrm{d}\mathbf{r} \; &= \int_\Omega \boldsymbol{\nabla}\varphi \cdot \boldsymbol{\nabla}\bar{\phi} \, \mathrm{d}\mathbf{r} + \int_{\partial\Omega} \varphi \, \mathbf{n} \cdot \boldsymbol{\nabla}\bar{\phi} \, \mathrm{d}s \\ &= \int_\Omega \boldsymbol{\nabla}\varphi \cdot \boldsymbol{\nabla}\bar{\phi} \, \mathrm{d}\mathbf{r} + \int_{\text{electrodes}} \varphi \, \partial_t \tilde{q} \, \mathrm{d}s, \end{aligned} \tag{A.10}$$

which is supposed to be satisfied for all test functions φ.[1] The Gouy–Chapman–Stern boundary condition for the potential drop across the double layer is imposed by multiplying a test function ϱ for the accumulated charge and integrating over the electrodes

$$0 = \int_{\text{electrodes}} \varrho \left(V_{\text{ext}} - \bar{\phi} - \tilde{\zeta} + \tilde{q} \, \delta \right) \mathrm{d}s, \tag{A.11}$$

which is to be satisfied for all ϱ. The flow problem is converted to weak form by multiplying the force balance (A.6) by a test function $\boldsymbol{u}$ for the velocity, and the incompressibility constraint by a test function ξ for the pressure, and integrating over Ω, i.e.,

$$\begin{aligned} 0 = -\int_\Omega \boldsymbol{u} \cdot (\boldsymbol{\nabla} \cdot \bar{\boldsymbol{\sigma}}) + \xi \, \boldsymbol{\nabla} \cdot \bar{\mathbf{u}} \, \mathrm{d}\mathbf{r} \\ = \int_\Omega \boldsymbol{\nabla}\boldsymbol{u} : \bar{\boldsymbol{\sigma}} - \xi \, \boldsymbol{\nabla} \cdot \bar{\mathbf{u}} \, \mathrm{d}\mathbf{r} + \int_{\partial\Omega} \boldsymbol{u} \cdot (\mathbf{n} \cdot \bar{\boldsymbol{\sigma}}) \, \mathrm{d}s. \end{aligned} \tag{A.12}$$

In FEMLAB we use so-called *strong* or *ideal* constraints to impose the no-slip condition on the channel walls, but for the slip condition on the electrodes we need *weak* constraints. The reason is that with an ideal constraint FEMLAB would consider Eq. (A.9) equally much as a constraint on $\bar{\phi}$ as on $\bar{\mathbf{u}}$, whereas we want it to be a constraint on $\bar{\mathbf{u}}$ only, and also want the reaction forces or Lagrange multipliers from that constraint to act in the flow problem only and not in the electrical problem. Therefore we introduce the reaction force $\bar{\mathbf{f}} = \mathbf{n} \cdot \bar{\boldsymbol{\sigma}}$ from the electrode surface as a separate variable, and use this to eliminate $\mathbf{n} \cdot \bar{\boldsymbol{\sigma}}$ in the boundary integral to get

$$0 = \int_\Omega \boldsymbol{\nabla}\boldsymbol{u} : \bar{\boldsymbol{\sigma}} - \xi \, \boldsymbol{\nabla} \cdot \bar{\mathbf{u}} \, \mathrm{d}\mathbf{r} + \int_{\text{electrodes}} \boldsymbol{u} \cdot \bar{\mathbf{f}} + \boldsymbol{f} \cdot \left(\bar{\mathbf{u}} - \langle \tilde{\zeta} \boldsymbol{\nabla}_s \bar{\phi} \rangle \right) \mathrm{d}s, \tag{A.13}$$

[1] Unlike the usual notation with Gauss' law our boundary normal $\mathbf{n}$ points *out* of the electrodes and channel walls and *into* the electrolyte and Ω.

for all test functions $\boldsymbol{u}$, ξ, $\boldsymbol{f}$, and subject to the no-slip condition $\bar{\mathbf{u}} = 0$ on the channel walls. In this form the problem can be typed directly into FEMLAB [47] as shown in the code example below:

Code also available on `http://www.mic.dtu.dk/mifts`

```
% AC ELECTOOSMOTIC PUMPING BY WEAKLY NONLINEAR MODEL

clear fem

% GEOMETRY
G1 = 1;           % gap between electrodes = reference dimension
W1 = 1.5;         % width of narrow electrode
W2 = 7;           % width of wide electrode
G2 = 5;           % gap between adjacent electrode pairs
H = 25;           % height of channel
L = W1+W2+G1+G2;  % period of electrode array
R = 0.05;         % electrode corner radius

fem.geom = rect2(0,L,0,H) - fillet(rect2(G2/2,G2/2+W1,-1,R),'Radii',R) ...
   - fillet(rect2(G2/2+W1+G1,G2/2+W1+G1+W2,-1,R),'Radii',R);

% MESH
fem.mesh = meshinit(fem,'Hmax',2,'Hmaxedg',[4 6; 0.05 0.2]);

% PARAMETERS
fem.const.w = 1;        % driving frequency
fem.const.V0 = 10;      % driving voltage
fem.const.delta = 1;    % surface capacitance ratio

% VARIABLES
fem.sdim = {'x' 'y'};
fem.dim = {'phi' 'u' 'v' 'p' 'q' 'fx' 'fy'};
fem.shape = [2 2 2 1 2 2 2];

% GOVERNING EQUATIONS
fem.form = 'weak';
fem.equ.shape = {[1:4]};
fem.equ.weak = {{'phix_test*phix+phiy_test*phiy' '-p_test*(ux+vy)' ...
   'ux_test*(-p+2*ux)+vy_test*(-p+2*vy)+(uy_test+vx_test)*(uy+vx)'}};

% BOUNDARY CONDITIONS
% group [1 2] narrow and wide electrode
% group [3]   insulating walls
% group [4]   master periodic boundary
% group [5]   slave periodic boundary
fem.bnd.ind = {[4 9 10] [6 11 12] [2 3 5 7] [1] [8]};
fem.bnd.shape = {[1:7] [1:7] [1:4] [1:4] [1:4]};
fem.bnd.expr = {'Vext' {'V0*sin(w*t)' '-V0*sin(w*t)' {} {} {}} ...
   'zeta' '-2*asinh(q/2)' 'us' 'zeta*phiTx' 'vs' 'zeta*phiTy'};
fem.bnd.weak = {{'phi_test*qt' 'q_test*(Vext-phi-zeta+q*delta)' ...
    'fx_test*(u-us)+fy_test*(v-vs)' 'u_test*fx+v_test*fy'} ...
```

```
   {'phi_test*qt' 'q_test*(Vext-phi-zeta+q*delta)' ...
    'fx_test*(u-us)+fy_test*(v-vs)' 'u_test*fx+v_test*fy'} {} {} {}};

% implement no-slip and periodic boundary conditions as ideal constraints
fem.bnd.constr = {{} {} {'u' 'v'} {} ...
   {'phi-phi_cpl' 'u-u_cpl' 'v-v_cpl' 'p-p_cpl'}};

% POINT CONDITIONS
% clamp pressure at a single point to fix the pressure level
% clamp constraint force [fx fy] where the electrodes and substrate meet
% and [fx fy] are overruled by the ideal constraint Lagrange multipliers
fem.pnt.ind = {[1] [3 6 7 10]};
fem.pnt.shape = {[4] [6 7]};
fem.pnt.constr = {'p' {'fx' 'fy'}};

% COUPLING ELEMENT FOR PERIODIC B.C.
% This is black magic. For an explanation see: elcplextr_example.m
% from www.comsol.com/support/knowledgebase/942.php
clear cpl src dst map
% extrusion coupling variables acting in geometry one
cpl.elem = 'elcplextr';
cpl.g = {'1'};
cpl.var = {'phi_cpl' 'u_cpl' 'v_cpl' 'p_cpl'};
% source expressions and trivial map {0} *from* source *to* intermediate mesh
src.expr = {{'phi'} {'u'} {'v'} {'p'}};
src.map = {'0' '0' '0' '0'};
src.ind = {{'1'}};
cpl.src = {{{} src {}}};
% linear transformation *from* destination *to* intermediate mesh
map.type = 'linear';
map.sg = '1';
map.sv = {'11','12'};
map.dg = '1';
map.dv = {'1','2'};
cpl.map = {map};
% declare {1} as map *from* destination *to* intermediate mesh
dst.map = {{'1'} {'1'} {'1'} {'1'}};
dst.ind = {{'8'}};
cpl.geomdim = {{{} dst {}}};
% finally add cpl to fem.elem and clean up
fem.elem = {cpl};
clear cpl src dst map

% BUILD EXTENDED MESH
fem.xmesh = meshextend(fem);
fem.sol = asseminit(fem);


% TRANSIENT SOLUTION OVER 2 PERIODS USING FEMTIME
% solve nonlinear electrical problem
fem.sol = femtime(fem,'Tlist',[0 4*pi],'Solcomp',{'phi' 'q'});
% solve linear flow problem with constant system matrix
t = fem.sol.tlist;
fem.sol = femlin(fem,'Solcomp',{'u' 'v' 'p' 'fx' 'fy'}, ...
```

```
   'Pname','t','Plist',t,'U',fem.sol,'Keep',{'K' 'N'});
% compute volume flow rate Q(t)  [large Schmidt number limit]
Q = postint(fem,'u','Edim',1,'Dl',1,'Solnum',1:length(t));

% COMPUTE PERIODIC SOLUTION USING CUSTOM WRITTEN ALGORITHM FEMFREQ
fem0 = fem;
nt = 16; tt = 2*pi*(0:nt-1)/nt;
% solve nonlinear electrical problem
fem.sol = femfreq(fem,'Tlist',tt,'Solcomp',{'phi' 'q'},'Symmetry','odd');
% solve linear flow problem
fem.sol = femlin(fem,'Solcomp',{'u' 'v' 'p' 'fx' 'fy'}, ...
   'Pname','t','Plist',tt,'U',fem.sol,'Keep',{'K' 'N'});
% compute time average flow rate <Q>
Qavg = postint(fem,'u','Edim',1,'Dl',1,'U',mean(fem.sol.u,2));

plot(t,Q,'b.-',[0 4*pi],[Qavg Qavg],'r:')
```

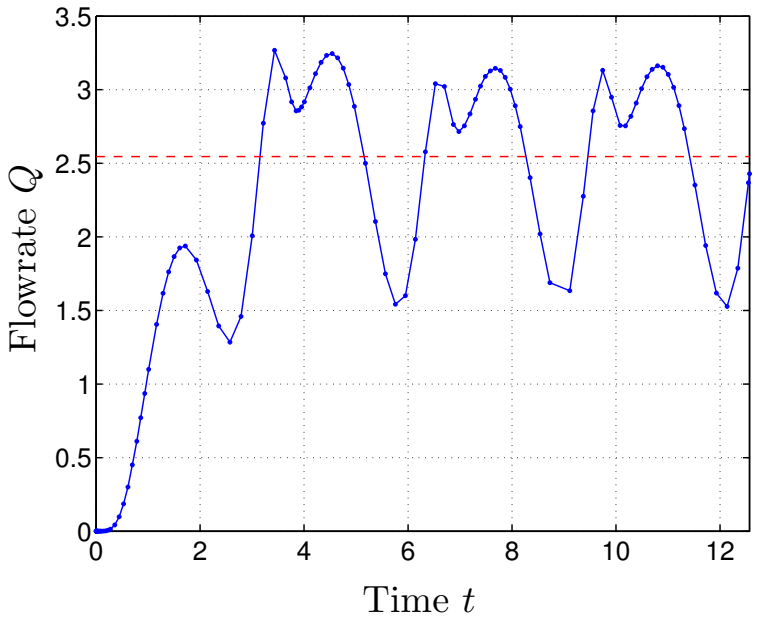


Figure A.1: Flowrate Q as a function of time t from weakly nonlinear FEMLAB code example. Result from FEMLAB's built-in timestepping algorithm `femtime` (solid line) and time average from custom written pseudospectral algorithm `femfreq` (dashed). Strictly, for $\varepsilon Sc \ll 1$ we ought to solve a stationary Stokes flow problem driven by the time average slip $\langle \tilde{\zeta} \boldsymbol{\nabla}_s \bar{\phi} \rangle$, while the code example computes a stationary solution at each point in time corresponding to the limit $\varepsilon Sc \gg 1$. However, because the problem is linear we get the same result for the time average flow rate $\langle Q \rangle$ in any case.

A.2 Strongly nonlinear finite element model

In this section we give an example of our implementation of the strongly nonlinear model from Sec. 2.2.8. The weak form of the bulk electrical problem $\boldsymbol{\nabla} \cdot \bar{\mathbf{J}}$ becomes

$$\begin{aligned} 0 &= -\int_\Omega \varphi \boldsymbol{\nabla} \cdot \bar{\mathbf{J}} \, \mathrm{d}\mathbf{r} = \int_\Omega \boldsymbol{\nabla}\varphi \cdot \bar{\mathbf{J}} \, \mathrm{d}\mathbf{r} + \int_{\partial\Omega} \varphi \, \mathbf{n} \cdot \bar{\mathbf{J}} \, \mathrm{d}s \\ &= \int_\Omega \boldsymbol{\nabla}\varphi \cdot \bar{\mathbf{J}} \, \mathrm{d}\mathbf{r} + \int_{\text{electrodes}} \varphi \, \bar{J}_n \, \mathrm{d}s, \end{aligned} \tag{A.14}$$

for all test functions φ for the potential $\bar{\phi}$ at all times, where we used the Lagrange multiplier $\bar{J}_n$ to eliminate the normal current $\mathbf{n} \cdot \bar{\mathbf{J}}$ in the boundary integral, and the boundary condition for the potential drop across the double layer is imposed by requiring

$$0 = \int_{\text{electrodes}} \mathcal{J}_n \left(V_{\text{ext}} - \bar{\phi} - \tilde{\zeta} + \tilde{q} \, \delta \right) \mathrm{d}s, \tag{A.15}$$

for all test functions $\mathcal{J}_n$ for $\bar{J}_n$. The conservation law for the accumulated charge $\tilde{q}$ in the Debye layer takes the form

$$0 = \int_{\text{electrodes}} \varrho \left(\partial_t \tilde{q} + \bar{J}_n + \varepsilon \boldsymbol{\nabla}_s \cdot \tilde{\mathbf{j}} \right) \mathrm{d}s = \int_{\text{electrodes}} \varrho \left(\partial_t \tilde{q} + \bar{J}_n \right) - \varepsilon \boldsymbol{\nabla}_s \varrho \cdot \tilde{\mathbf{j}} \, \mathrm{d}s, \tag{A.16}$$

for all test functions ϱ for $\tilde{q}$, where we used partial integration to move the derivative $\boldsymbol{\nabla}_s$ onto ϱ and assumed $\tilde{\mathbf{j}} = 0$ at the electrode edges. For the bulk salt concentration $\bar{c}$ we write

$$\begin{aligned} 0 &= -\int_\Omega c \, \boldsymbol{\nabla} \cdot \bar{\mathbf{F}}_a \, \mathrm{d}\mathbf{r} = \int_\Omega \boldsymbol{\nabla} c \cdot \bar{\mathbf{F}}_a \, \mathrm{d}\mathbf{r} + \int_{\partial\Omega} c \, \mathbf{n} \cdot \bar{\mathbf{F}}_a \, \mathrm{d}s \\ &= \int_\Omega \boldsymbol{\nabla} c \cdot \bar{\mathbf{F}}_a \, \mathrm{d}\mathbf{r} + \int_{\text{electrodes}} c \left\langle \check{F}_n - \Delta D \bar{J}_n \right\rangle \mathrm{d}s, \end{aligned} \tag{A.17}$$

for all test functions c for $\bar{c}$. Alternatively, it sometimes gives more stable numerical results to rewrite $\boldsymbol{\nabla} \cdot \bar{\mathbf{F}}_a$ in non-conservative form

$$\boldsymbol{\nabla} \cdot \bar{\mathbf{F}}_a = \boldsymbol{\nabla} \cdot \left(-D_a \boldsymbol{\nabla} \bar{c} + Pe \, \bar{\mathbf{u}} \, \bar{c} \right) = -D_a \boldsymbol{\nabla}^2 \bar{c} + Pe \, \bar{\mathbf{u}} \cdot \boldsymbol{\nabla} \bar{c}, \tag{A.18}$$

using $\boldsymbol{\nabla} \cdot \bar{\mathbf{u}} = 0$, and performing partial integration only on the $\boldsymbol{\nabla}^2 \bar{c}$ term. The conservation law for the surface excess accumulated salt $\tilde{w}$ becomes

$$0 = \int_{\text{electrodes}} \varpi \left(\partial_t \tilde{w} + \check{F}_n + \varepsilon \boldsymbol{\nabla}_s \cdot \tilde{\boldsymbol{\Gamma}} \right) \mathrm{d}s = \int_{\text{electrodes}} \varpi \left(\partial_t \tilde{w} + \check{F}_n \right) - \varepsilon \boldsymbol{\nabla}_s \varpi \cdot \tilde{\boldsymbol{\Gamma}} \, \mathrm{d}s, \tag{A.19}$$

for all test functions ϖ for $\tilde{w}$, which effectively determines the normal flux $\check{F}_n$, whereas the excess amount of salt is determined by

$$0 = \int_{\text{electrodes}} \mathcal{F}_n \left(\tilde{w} - \sqrt{\tilde{q}^2 + 4\check{c}} + \sqrt{4\check{c}} \right) \mathrm{d}s, \tag{A.20}$$

for all test functions $\mathcal{F}_n$ for $\check{F}_n$. Finally, the surface excess concentration $\check{\gamma} = \check{c} - \bar{c}$ at $\check{y} = 0$ is determined by the (excess) ambipolar flux in and out of the diffusion layer

$$0 = \int_{\text{electrodes}} \mathcal{Y}\Big[\check{\gamma} - \sqrt{\frac{\varepsilon}{\omega D_a}}\frac{1}{T}\int_0^T \mathcal{G}(t-t')\left(\check{F}_n - \Delta D \bar{J}_n\right)\mathrm{d}t'\Big]\,\mathrm{d}s, \tag{A.21}$$

for all test functions $\mathcal{Y}$ for $\check{\gamma}$; it is not necessary to subtract $\langle \check{F}_n - \Delta D \bar{J}_n \rangle$ because the kernel $\mathcal{G}(t-t')$ is blind to any time average component by construction.

In the strongly nonlinear model we need to include the bulk electrical body force $\bar{\mathbf{f}}$ into the Stokes flow problem, where

$$\bar{\mathbf{f}} = \boldsymbol{\nabla}^2\bar{\phi}\,\boldsymbol{\nabla}\bar{\phi} = -\frac{\boldsymbol{\nabla}\bar{c}\cdot\boldsymbol{\nabla}\bar{\phi} + \Delta D\boldsymbol{\nabla}^2\bar{c}}{\bar{c}}\,\boldsymbol{\nabla}\bar{\phi}, \tag{A.22}$$

cf. Eq. (2.60). The dependence on $\boldsymbol{\nabla}^2\bar{\phi}$ or $\boldsymbol{\nabla}^2\bar{c}$ is problematic because with standard finite element approximations the solution is only continuous and piecewise differentiable, with the second derivative undefined on element boundaries. In the special case of equal diffusivities where $\Delta D = 0$, Eq. (A.22) expresses $\bar{\mathbf{f}}$ in terms of the gradients $\boldsymbol{\nabla}\bar{c}$ and $\boldsymbol{\nabla}\bar{\phi}$ only, and the problem does not arise. Otherwise it is necessary to rewrite the body force as

$$\bar{\mathbf{f}} = \boldsymbol{\nabla}^2\bar{\phi}\,\boldsymbol{\nabla}\bar{\phi} = \boldsymbol{\nabla}\cdot\bar{\boldsymbol{T}}, \tag{A.23}$$

where $\bar{\boldsymbol{T}}$ is the Maxwell stress tensor

$$\bar{T}_{ij} = \partial_i\bar{\phi}\,\partial_j\bar{\phi} - \delta_{ij}\,\partial_k\bar{\phi}\,\partial_k\bar{\phi}/2. \tag{A.24}$$

With this formulation we can use Gauss' law to move the derivative from $\bar{\boldsymbol{T}}$ onto the test function $\boldsymbol{u}$ for the fluid velocity, and the weak form of the Stokes flow problem becomes

$$0 = \int_\Omega \boldsymbol{\nabla}\boldsymbol{u} : (\bar{\boldsymbol{\sigma}} + \bar{\boldsymbol{T}}) - \xi\,\boldsymbol{\nabla}\cdot\bar{\mathbf{u}}\,\mathrm{d}\mathbf{r} + \int_{\text{electrodes}} \boldsymbol{u}\cdot\bar{\boldsymbol{\mu}} + \boldsymbol{f}\cdot(\bar{\mathbf{u}} - \tilde{\mathbf{u}}_s)\,\mathrm{d}s, \tag{A.25}$$

for all test functions $\boldsymbol{u}$, ξ, $\boldsymbol{f}$, and subject to the no-slip condition on the channel walls. Here $\bar{\boldsymbol{\mu}} = \mathbf{n}\cdot(\bar{\boldsymbol{\sigma}} + \bar{\boldsymbol{T}})$ is the total reaction force from the electrode surface, and $\tilde{\mathbf{u}}_s = \langle(\check{\zeta}+\check{\varsigma})\boldsymbol{\nabla}_s\bar{\phi}\rangle$ is the effective slip velocity induced in the Debye and diffusion layers.

In the code example below we consider the simple equal diffusivity case with $\Delta D = 0$ for the symmetric geometry from Sec. 4.2.1. For simplicity we assume an open system and impose unit concentration $\bar{c} = 1$ and no shear force $\mathbf{n}\cdot\bar{\boldsymbol{\sigma}} = \mathbf{0}$ boundary conditions far above the electrode array. For $\Delta D = 0$ both $\bar{\phi}$, $\tilde{q}$, and $\bar{J}_n$ have odd symmetry in time, $\tilde{w}$, $\check{F}_n$, and $\check{\gamma}$ have even symmetry, whereas $\bar{c}$, $\bar{\mathbf{u}}$, $\bar{p}$, and $\bar{\boldsymbol{\mu}}$ are constant in time. This allows the computational cost of solving the full coupled problem to be reduced

by a factor of two as compared to the case $\Delta D \neq 0$. The convection from fluid motion induced in the diffusion layer is neglected when solving the coupled electrohydrodynamical problem, but in a postprocessing step we single out the components of the net fluid motion due to electroosmotic slip, bulk electroconvection, and flow induced in the diffusion layer. The auxiliary function `fkernel` evaluates the kernel $\mathcal{G}(t-t')$ at $\check{y}=0$, and `excessr` computes the integral $\sqrt{\omega/D_a}\int_0^\infty \frac{1}{\check{c}} - \frac{1}{\bar{c}}\,\mathrm{d}\check{y}$ given $\check{\gamma}(t)$ and $\bar{c}$ by numerical quadrature.

Code also available on `http://www.mic.dtu.dk/mifts`

```
% AC ELECTOKINETIC FLOW BY STRONGLY NONLINEAR MODEL

clear fem

% GEOMETRY
G = 1;      % gap between electrodes
W = 2;      % width of electrodes
H = 5;      % height of channel
L = W+G;    % period of electrode array
R = 0.1;    % electrode corner radius

fem.geom = rect2(0,L/2,0,H) - fillet(rect2(-W/2,W/2,-1,R),'Radii',R);

% MESH
fem.mesh = meshinit(fem,'Hmax',0.75,'Hmaxedg',[2; 0.1]);

% PARAMETERS
fem.const.omega = 1;         % driving frequency
fem.const.V0 = 50;           % driving voltage
fem.const.delta = 1;         % surface capacitance ratio
fem.const.epsilon = 0.001;   % Debye layer thickness
fem.const.Pe = 0.25;         % Peclet number

% VARIABLES
fem.sdim = {'x' 'y'};
fem.dim = {'phi' 'c' 'u' 'v' 'p' 'q' 'w' 'Jn' 'Fn' 'gamma' 'Zeta' 'fx' 'fy'};
fem.shape = [2 2 2 2 1 2 2 2 2 2 2 2 2];

% GOVERNING EQUATIONS
fem.form = 'weak';
fem.equ.shape = {[1:5]};
fem.equ.expr = {'rho' '(cx*phix+cy*phiy)/c' ...
   'Jx' '-c*phix' 'Jy' '-c*phiy' 'Fx' '-cx+Pe*u*c' 'Fy' '-cy+Pe*v*c'};
fem.equ.weak = {{'phix_test*Jx+phiy_test*Jy' ...
   '-cx_test*cx-cy_test*cy-c_test*Pe*(u*cx+v*cy)' ...
   'ux_test*(-p+2*ux)+vy_test*(-p+2*vy)+(uy_test+vx_test)*(uy+vx)' ...
   '-p_test*(ux+vy)' 'u_text*rho*phix+v_test*rho*phiy'}};
fem.equ.init = {{0 1 0 0 0 0 0 0 0 0 0 0 0}};

% BOUNDARY CONDITIONS
```

```
% group [1] electrode
% group [2] substrate
% group [3] even symmetry
% group [4] odd symmetry
% group [5] "infinity"
fem.bnd.ind = {[2 6] [4] [1] [5] [3]};
fem.bnd.shape = {[1:13] [1:5] [1:5] [1:5] [1:5]};
fem.bnd.expr = {'Vext' {'V0*sin(omega*t)' {} {} {} {}} ...
   'zeta' '-2*asinh(q/(2*sqrt(c+gamma)))' ...
   'Gammax' '-(1+2*Pe)*q*phiTx' 'Gammay' '-(1+2*Pe)*q*phiTy' ...
   'jx' '-(1+2*Pe)*w*phiTx' 'jy' '-(1+2*Pe)*w*phiTy' ...
   'us' '(zeta+Zeta)*phiTx' 'vs' '(zeta+Zeta)*phiTy'};
fem.bnd.weak = {{'phi_test*Jn' 'Jn_test*(Vext-phi-zeta+q*delta)' ...
   'c_test*Fn' 'Fn_test*(w-sqrt(q^2+4*(c+gamma))+sqrt(4*(c+gamma)))' ...
   'q_test*(qt+Jn)-epsilon*(qTx_test*jx+qTy_test*jy)' ...
   'w_test*(wt+Fn)-epsilon*(wTx_test*Gammax+wTy_test*Gammay)' ...
   'fx_test*(u-us)+fy_test*(v-vs)+u_test*fx+v_test*fy' ...
   'gamma_test*(gamma-Fntt*sqrt(epsilon/omega))'} {} {} {} {}};
fem.bnd.constr = {{} {'u' 'v'} {'u'} {'u' 'phi'} {'c-1'}};
fem.bnd.init = {{0 1 0 0 0 0 eps 0 0 0 0 0 0}};

% POINT CONSTRAINTS
% clamp weak constraint force where overruled by ideal constraints
fem.pnt.ind = {[1] [4]};
fem.pnt.shape = {[1:13] [1:13]};
fem.pnt.constr = {{'fx'} {'fx' 'fy'}};

% BUILD EXTENDED MESH
fem.xmesh = meshextend(fem);


% COMPUTE PERIODIC SOLUTION USING FEMFREQ
nt = 32; tlist = 2*pi*(0:nt-1)/nt;
fem.sol = femfreq(fem,'Tlist',tlist, ...
   'Solcomp',{'phi' 'c' 'q' 'w' 'Jn' 'Fn' 'gamma' 'u' 'v' 'p' 'fx' 'fy'}, ...
   'Symmetry',{'odd' 'const' 'odd' 'even' 'odd' 'even' 'even' ...
      'const' 'const' 'const' 'const' 'const'}, ...
   'Rule','leapfrog','Kernel',fkernel(0,nt,4));


% GET INDICES OF [Zeta gamma Jn c] IN GLOBAL SOLUTION VECTOR
iZeta = find(asseminit(fem,'Init',{'Zeta' 1},'Out','U'));
ndofs = flngdof(fem);
igamma = asseminit(fem,'Init',{'Zeta' 'gamma'},'U',(1:ndofs)','Out','U');
iJn = asseminit(fem,'Init',{'Zeta' 'Jn'},'U',(1:ndofs)','Out','U');
ic = asseminit(fem,'Init',{'Zeta' 'c'},'U',(1:ndofs)','Out','U');
igamma = igamma(iZeta); iJn = iJn(iZeta); ic = ic(iZeta);

% COMPUTE EXCESS POTENTIAL DROP ACROSS THE DIFFUSION LAYER
gamma = fem.sol.u(igamma,:); Jn = fem.sol.u(iJn,:); c = fem.sol.u(ic,1);
R = excessr(gamma,c);
Zeta = sqrt(fem.const.epsilon/fem.const.omega)*Jn.*R;
u0 = fem.sol.u; u0(iZeta,:) = Zeta;
fem.sol = femsol(u0,'tlist',tlist);
```

```
% COMPUTE INDIVIDUAL FLOW COMPONENTS
% flow due to slip induced in the Debye layer
ueo = femlin(fem,'Solcomp',{'u' 'v' 'p' 'fx' 'fy'}, ...
   'Const',{'Zeta' 0 'rho' 0},'U',fem.sol,'Out','U', ...
   'Pname','t','Plist',tlist,'Keep',{'K' 'N'});
% flow due to slip induced in the diffusion layer
uec = femlin(fem,'Solcomp',{'u' 'v' 'p' 'fx' 'fy'}, ...
   'Const',{'zeta' 0 'rho' 0},'U',fem.sol,'Out','U', ...
   'Pname','t','Plist',tlist,'Keep',{'K' 'N'});
% flow due to bulk electroconvection
ub = femlin(fem,'Solcomp',{'u' 'v' 'p' 'fx' 'fy'}, ...
   'Const',{'zeta' 0 'Zeta' 0},'U',fem.sol,'Out','U', ...
   'Pname','t','Plist',tlist,'Keep',{'K' 'N'});

% COMPUTE ANGULAR MOMENTUM FROM FLUID VORTEX
Leo = postint(fem,'x*v-y*u','U',mean(ueo,2));
Lec = postint(fem,'x*v-y*u','U',mean(uec,2));
Lb  = postint(fem,'x*v-y*u','U',mean(ub,2));
Ltot = Leo + Lec + Lb;
```

A.3 Periodic integro-differential equations in time

We wish to solve an integro-differential equation in time of the form

$$\mathbf{D}\,\partial_t\mathbf{U} + \frac{1}{T}\int_0^T \mathbf{G}(t,t')\mathbf{U}(t')\,\mathrm{d}t' = \mathbf{L} - \mathbf{N}^T\mathbf{\Lambda}, \tag{A.26}$$

subject to the constraint

$$\mathbf{0} = \mathbf{M}, \tag{A.27}$$

where the solution $\mathbf{U}(t)$ is periodic in time, i.e., $\mathbf{U}(t) = \mathbf{U}(t+T)$. Here $\mathbf{D}(\mathbf{U},t)$ is the mass matrix, $\mathbf{G}(t,t')$ the kernel matrix in the convolution, $\mathbf{L}(\mathbf{U},t)$ the residual vector, $\mathbf{M}(\mathbf{U},t)$ the constraint vector, $\mathbf{\Lambda}(t)$ a Lagrange multiplier for that constraint, and $\mathbf{N} = -\partial\mathbf{M}/\partial\mathbf{U}$ the matrix that couples the Lagrange multiplier into the problem. Apart from the convolution operator, this is identical to the generic form of a time-dependent problem after discretization in FEMLAB. However, we wish to use a relaxation method to compute the periodic solution instead of the standard time-steppig algorithm. Linearizing around a given trial solution $\mathbf{U}_0(t)$ we get

$$\mathbf{D}\,\partial_t\,\delta\mathbf{U} + \frac{1}{T}\int_0^T \mathbf{G}(t,t')\,\delta\mathbf{U}(t')\,\mathrm{d}t' = \mathbf{L}_0 - \mathbf{K}\,\delta\mathbf{U} - \mathbf{N}^T\mathbf{\Lambda}, \tag{A.28}$$

subject to

$$\mathbf{0} = \mathbf{M}_0 - \mathbf{N}\,\delta\mathbf{U}. \tag{A.29}$$

The derivative and integral terms of the trial solution are absorbed into the residual vector

$$\mathbf{L}_0 = \mathbf{L}(\mathbf{U}_0,t) - \mathbf{D}(\mathbf{U}_0,t)\,\partial_t\mathbf{U}_0 - \frac{1}{T}\int_0^T \mathbf{G}(t,t')\mathbf{U}_0(t')\,\mathrm{d}t', \tag{A.30}$$

and further $\mathbf{M}_0 = \mathbf{M}\big(\mathbf{U}_0,t\big)$ and $\mathbf{K} = -\partial\mathbf{L}/\partial\mathbf{U}$. We neglect the dependence of $\mathbf{D}$ and $\mathbf{N}$ on $\mathbf{U}$ when linearizing; this may cause problems in the Newton itertions if those are strongly nonlinear functions.

If we represent the temporal solution for $\mathbf{U}$ by N equispaced points in time $t_n = n\,T/N$ for $n = 1,2,\ldots,N$ then it can be written as a discrete Fourier transform

$$\mathbf{U}(t_n) = \sum_{m=1}^{N} \hat{\mathbf{U}}_m e^{im\omega\,t_n}, \tag{A.31}$$

where the coefficients

$$\hat{\mathbf{U}}_m = \frac{1}{N}\sum_{n=1}^{N} \mathbf{U}(t_n)e^{-im\omega\,t_n} \tag{A.32}$$

can be computed by a fast Fourier transform. By construction $\hat{\mathbf{U}}_m = \hat{\mathbf{U}}_{m+N}$, and since $\mathbf{U}$ is real we also have $\hat{\mathbf{U}}_m = \hat{\mathbf{U}}^*_{-m} = \hat{\mathbf{U}}^*_{N-m}$. The appropriate way of interpreting $\mathbf{U}(t)$ for $t \neq t_n$ is

$$\mathbf{U}(t) = \frac{1}{N} {\sum_{m=-N/2}^{N/2}}' \hat{\mathbf{U}}_m e^{im\omega t} = \sum_{n=1}^{N} \mathbf{U}(t_n) \left\{ \frac{1}{N} {\sum_{m=-N/2}^{N/2}}' e^{im\omega(t-t_n)} \right\}, \quad \text{(A.33)}$$

where the prime indicates that the terms $m = \pm N/2$ are counted half only. For an explanation why this is "appropriate", and for a nice introduction to spectral methods in general, see Ref. [68]. The function in the curly braces is the periodic sinc function $S_N(t - t')$

$$S_N(t-t') = \frac{1}{N} {\sum_{m=-N/2}^{N/2}}' e^{im\omega(t-t')} = \frac{1}{N} \frac{\sin[N\omega(t-t')/2]}{\tan[\omega(t-t')/2]}, \quad \text{(A.34)}$$

which allows the time detivative of $\mathbf{U}(t)$ to be evaluated as

$$\partial_t \mathbf{U}(t_n) = \sum_{m=1}^{N} S'_N(t_n - t_m)\mathbf{U}(t_m). \quad \text{(A.35)}$$

The convolution integral can be approximated as

$$\frac{1}{T} \int_0^T \mathbf{G}(t_n, t')\mathbf{U}(t')\,\mathrm{d}t' \approx \frac{1}{N} \sum_{m=1}^{N} \mathbf{G}(t_n, t_m)\mathbf{U}(t_m). \quad \text{(A.36)}$$

Substituting into the linearized problem we get

$$\sum_{m=1}^{N} \left[\mathbf{D}_n S'_N(t_n - t_m) + \frac{1}{N}\mathbf{G}_{nm} \right] \delta\mathbf{U}_m + \mathbf{K}_n \delta\mathbf{U}_n + \mathbf{N}_n^T \mathbf{\Lambda}_n = \mathbf{L}_{0,n}, \quad \text{(A.37)}$$

$$\mathbf{N}_n \delta\mathbf{U}_n = \mathbf{M}_{0,n}, \quad \text{(A.38)}$$

for $n = 1, 2, \ldots, N$, i.e., a simultaneous matrix problem for all the solution components $\delta\mathbf{U}_m$.

Unless N is small this problem will make most linear solvers choke; the point is that for harmonically driven weakly nonlinear systems the number of Fourier components required to represent the solution is often fairly small. If the problem cannot be solved accurately with just a few Fourier components, we could replace the spectral derivative with a low order finite-difference scheme, e.g., the leapfrog scheme $[\partial_t \mathbf{U}]_n \approx [\mathbf{U}_{n+1} - \mathbf{U}_{n-1}]/2\Delta t$, which would reduce the computational cost for a differential equation from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$ at the expense of the spectral accuracy. For comparison, the cost of one step forward in time by a standard timestepping algoritm is $\mathcal{O}(1)$ which is *much* more memory efficient but sometimes requires integrating for a long time before a periodic solution is reached. If the problem

involves convolution there is no obvious way to work around the nonlocal coupling in time; a pseudo-timestepping algorithm designed specifically for rapid convergence towards the periodic solution might provide a fairly inexpensive solution to the problem. However, we have not had time to look further into this subject.

For our particular application where the nonlocal coupling in time comes about from our semi-analytical model for the dynamics in the diffusion layer, the convolution kernel does not depend explicitly on time but only on the time delay, i.e., $\mathbf{G} = \mathbf{G}(t - t')$, and it is originally defined in terms of an infinite Fourier series

$$\mathbf{G}(t - t') = \sum_{n=-\infty}^{\infty} \hat{\mathbf{G}}_n e^{in\omega(t-t')}. \tag{A.39}$$

Substituting this and the truncated series Eq. (A.33) for $\mathbf{U}(t)$ into the convolution integral we obtain

$$\frac{1}{T}\int_0^T \mathbf{G}(t_n - t')\mathbf{U}(t')\,\mathrm{d}t' = \frac{1}{N}\sum_{m=1}^{N} \mathbf{G}_N(t_n - t_m)\mathbf{U}(t_m), \tag{A.40}$$

where

$$\mathbf{G}_N(t - t') = \sum_{k=-N/2}^{N/2\,\prime} \hat{\mathbf{G}}_k e^{ik\omega(t-t')}. \tag{A.41}$$

This essentially tells us how to truncate the infinite series Eq. (A.39) in a consistent way. If the spectral time derivative is replaced by the leapfrog rule, it is equivalent to replacing $S_N(t-t')$ by a piecewise linear interpolant, namely, the "hat" function $P_N(t - t') = \max\{1 - N|t - t'|, 0\}$. In that case it is more consistent to truncate the convolution kernel as

$$\mathbf{G}_N(t - t') = N\int_0^T \mathbf{G}(t - t'')P_N(t'' - t')\,\mathrm{d}t''. \tag{A.42}$$

Implementation: FEMFREQ

We have implemented the relaxation scheme in a function `femfreq`. It computes periodic solutions to simple first order problems in time with a syntax similar to FEMLAB's standard time-stepping algorithm `femtime`. For problems involving convolution we abuse the syntax for second order problems by interpreting the term $\mathbf{E}\,\partial_t^2\mathbf{U}$ instead as $\mathbf{E}\int_0^T \mathcal{G}(t-t')\mathbf{U}(t')\,\mathrm{d}t'$, where $\mathcal{G}(t-t')$ is a some scalar function supplied by the user via the property `'Kernel'`. This approach is significantly less flexible than a general $\mathbf{G}(t, t')$ that could depend explicitly on both time and spatial position, but it serves our purpose. We also support odd, even, and constant symmetries in the solution which allows the computational cost to be reduced. We do not include the code here but it is available at `http://www.mic.dtu.dk/mifts`.

Appendix B

Paper published in Phys. Rev. E

Title
AC electrokinetic micropumps: the effect of geometrical confinement, Faradaic current injection, and nonlinear surface capacitance

Authors
Laurits Højgaard Olesen, Henrik Bruus, and Armand Ajdari

Reference
Phys. Rev. E **73**, 056313 (2006) (16 pages)

ac electrokinetic micropumps: The effect of geometrical confinement, Faradaic current injection, and nonlinear surface capacitance

Laurits Højgaard Olesen,[1,2] Henrik Bruus,[1] and Armand Ajdari[2]
[1]*MIC, Department of Micro and Nanotechnology, Technical University of Denmark, DTU bldg. 345 east, DK-2800 Kongens Lyngby, Denmark*
[2]*Laboratoire de Physico-Chimie Théorique, UMR CNRS-ESPCI 7083, 10 rue Vauquelin, F-75231 Paris Cedex 05, France*
(Received 9 October 2005; published 31 May 2006)

Recent experiments have demonstrated that ac electrokinetic micropumps permit integrable, local, and fast pumping (velocities ~mm/s) with low driving voltage of a few volts only. However, they also displayed many quantitative and qualitative discrepancies with existing theories. We therefore extend the latter theories to account for three experimentally relevant effects: (i) vertical confinement of the pumping channel, (ii) Faradaic currents from electrochemical reactions at the electrodes, and (iii) nonlinear surface capacitance of the Debye layer. We report here that these effects indeed affect the pump performance in a way that we can rationalize by physical arguments.

DOI: 10.1103/PhysRevE.73.056313 PACS number(s): 47.65.−d, 47.32.−y, 85.90.+h

I. INTRODUCTION

Lab-on-a-chip systems require micropumps and valves to manipulate small volumes of a liquid sample [1]. Often large external pumps, such as syringe pumps, are used to deliver the necessary pressures, but for portable systems a number of on-chip micropumps have been developed over the last decade. The latter fall into two major categories: One category comprises mechanical actuation and deflectable membranes to create pumps and valves. These have proven to be versatile and simple to fabricate and operate, but there are some difficulties with further downscaling. The second category involves no moving parts but uses electric fields to induce electrokinetic pumping. However, classical dc electro-osmosis requires a relatively large voltage and field strength, which is undesirable. Moreover, when a dc current is drawn from electrodes integrated in a microfluidic system there are problems with bubble formation due to electrolysis.

ac electro-osmosis has recently been observed to induce fluid motion over pairs of microelectrodes [2–4]. Based on general symmetry arguments, Ajdari predicted that the same mechanism would generate a net flow over an asymmetric array of electrodes [5], which was soon after demonstrated experimentally by Brown *et al.* [6] and later by several other groups obtaining pumping velocities ~mm/s with driving voltage of a few volts [7–14].

Theoretically, Ramos *et al.* performed a more detailed linear response investigation of the pumping mechanism and found a specific pumping direction, a pumping velocity scaling as the square of the driving voltage, and a well-defined frequency for maximal pumping [15].

However, the above analysis is strictly valid only at low driving voltage $V_0 \ll k_B T/e \sim 25$ mV, whereas experiments are usually performed with at least a few volts. As a consequence, the theory disagrees with experimental observations in several ways: e.g., experimentally the pumping velocity does not always compare well with the predicted V_0^2 scaling, but looks more like a linear scaling [10,11] or tends to saturate at large voltage [13]. Moreover, Studer *et al.* found a reversal of the pumping direction at driving frequencies well above that for maximal pumping in the "forward" direction [11] and Ramos and co-workers observed reversal of the pumping direction on a traveling-wave device for driving voltages above 2 V at 1 kHz driving frequency [13,14]. The existing theory does not give many clues as to the mechanism for this reversal, except that in the original paper Ajdari predicted pumping in the "reverse" direction at low frequency when Faradaic electrode reactions were included in the model [5]. Indeed, several groups have observed bubble formation and electrode degradation at low frequency and high voltage which indicates that electrode reactions are actually taking place [8–11,14]. Lastochkin *et al.* observed fluid motion in the reverse direction of that usually expected for ac electro-osmosis, which was attributed to strong Faradaic current injection [16]. However, their experiments were performed with driving frequencies in the MHz range which is of the order of the Debye frequency for the electrolyte, whereas ac electro-osmosis generally occurs around the inverse RC time of the device.

These discrepancies between experimental observations and existing theory demonstrate the need for a more complete theoretical understanding. As a first step we here address a few generalizations of the existing linear response theory to the weakly nonlinear regime by taking into account, within the thin Debye layer approximation, the effect of Faradaic currents both in a linearized scheme and using the full nonlinear Butler-Volmer reaction kinetics. We also include the nonlinear surface capacitance of the Debye layer as described in Gouy-Chapman theory.

The paper is organized as follows: In Sec. II we fix the device geometry and in Sec. III we describe our model for the electrokinetic system. In Sec. IV we extend the linear analysis for low driving voltage to study in more detail the effects of the device geometry and vertical confinement of the system, and we investigate the effect of Faradaic current injection in a linearized scheme. This provides a firm starting point when in Sec. V we study the nonlinear model both with and without Faradaic current. In Sec. VI we summarize our results and compare to experiments reported in the literature, and finally, in Sec. VII we conclude the paper.

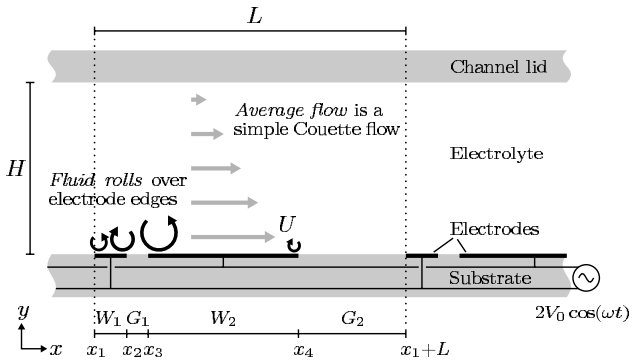


FIG. 1. Sketch of the device geometry. The interdigitated electrode array is biased with an ac voltage that induces a buildup of a Debye screening layer and electro-osmotic fluid motion.

II. DEVICE GEOMETRY

We consider a two-dimensional (2D) geometry like that shown in a side view in Fig. 1, similar to devices used experimentally [6–11,14]. It consists of a substrate on which an asymmetric array of interdigitated microelectrodes is deposited. Each electrode pair in the array consists of a narrow electrode of width W_1 and a wide electrode of width W_2 separated by a narrow and a wide gap G_1 and G_2, respectively, and the full period is $L=W_1+G_2+W_2+G_2$. Notice that in order to break the left-right symmetry of the array it is necessary that both $W_1 \neq W_2$ and $G_1 \neq G_2$. On top of the electrode array a microfluidic channel of height H is placed and we assume that the device extends sufficiently far into the third dimension that a 2D description in the xy plane is appropriate.

The channel is filled with an electrolyte, and the narrow and wide electrodes are biased with an ac voltage $V_{\text{ext}} = \pm V_0 \cos(\omega t)$, respectively; i.e., the applied voltage difference is $2V_0 \cos(\omega t)$. This induces the formation of a Debye layer on the electrodes that acts to partially screen the electric field. The nonuniform partial screening gives rise to tangential electric field in the Debye layer and electro-osmotic flow. This flow takes the form of fluid rolls above the electrode edges as indicated in Fig. 1.

III. ELECTROKINETIC MODEL

Our model for the electrohydrodynamics of the system is similar to that of Ramos *et al.* [15] and is based on the following classical approximations [17–20]: (i) The bulk electrolyte is assumed to be charge neutral with uniform salt concentration, such that the ionic transport can be described as Ohmic current. (ii) The Debye layer is assumed to be in local equilibrium with the electrolyte immediately outside the layer, such that the charge distribution and potential variation in the Debye layer can be described by Gouy-Chapman theory. (iii) Moreover, we assume that the thickness of the Debye layer is much smaller than the size of the electrodes, and we neglect surface diffusion and migration of charge. (iv) The bulk fluid motion is described by Stokes flow with a slip condition on the electrodes set by the electroosmotic flow induced in the Debye layer. (v) Finally, we also assume that the bulk concentration of the reactants in the Faradaic electrode reaction is constant; i.e., we neglect the effect of mass transfer.

Upon these approximations, the only dynamical variables we are left to consider are the potential distribution $\phi(x,y,t)$ in the bulk and the local surface charge density $q(x,t)$ accumulated in the Debye layer on the electrodes. The instantaneous value of q determines the potential drop from the electrode to the electrolyte immediately outside the Debye layer, whereas the time evolution of q is determined by the balance between the Ohmic current from the bulk and Faradaic current from electrochemical reactions on the electrodes.

We emphasize that this simple electrokinetic model is appropriate only in the weakly nonlinear regime of not-too-high driving voltage ($\lesssim$125 mV). Beyond this, the assumption of uniform bulk electrolyte concentration breaks down, the effect of mass transfer on the Faradaic reaction kinetics cannot be neglected, and the Debye layer may be driven out of local equilibrium. This is discussed in more detail in Sec. VI.

A. Bulk electrolyte

In the bulk electrolyte we assume charge neutrality such that the charge continuity equation reduces to

$$\nabla \cdot \mathbf{J} = 0. \tag{1}$$

The electric current $\mathbf{J}$ is described simply as the Ohmic current $\mathbf{J}=-\sigma\nabla\phi$ where σ is the conductivity of the electrolyte. Since we assume uniform salt concentration and conductivity throughout the bulk, we then simply end up with a Laplace problem for the potential $\phi(\mathbf{r},t)$.

B. Debye layer

The Debye layer is assumed to be in local equilibrium with the electrolyte immediately outside the layer. This implies that the driving frequency should be well below the Debye frequency $\omega_D=\sigma/\epsilon$ of the electrolyte and that the Faradaic current injection cannot be too strong. The total charge accumulated in the Debye layer can then be directly related to the potential drop $\zeta(x,t)$ across it [18]

$$q(x,t) = \int_{\text{Debye layer}} \rho(x,y,t)dy \tag{2}$$

$$= -\operatorname{sgn}(\zeta)\sqrt{2\epsilon k_B T \sum_n c_n^*(e^{-z_n e\zeta/k_B T}-1)}, \tag{3}$$

where ρ is the charge density, ϵ the permittivity of the solvent, k_B Boltzmann's constant, T the temperature, c_n^* the bulk concentration (number density) of the nth ionic species, and z_n its valence. In the Debye-Hückel limit $\zeta \ll k_B T/ze$, Eq. (3) can be linearized to

$$q = -\frac{\epsilon}{\lambda_D}\zeta, \tag{4}$$

where $\lambda_D=\sqrt{\epsilon k_B T/\Sigma_n c_n^* z_n^2 e^2}$ is the Debye length and ϵ/λ_D is the Debye layer capacitance (per unit area) at low voltage.

For larger potentials we focus on the simple case of a monovalent symmetric (1:1) electrolyte where Eq. (3) reduces to

$$q = -\frac{\epsilon}{\lambda_D}\frac{2k_BT}{e}\sinh\left(\frac{e\zeta}{2k_BT}\right). \tag{5}$$

As in the classical Gouy-Chapman-Stern model we assume that the compact (Stern) layer on the electrode simply gives rise to an additional surface capacitance C_s in series with the Debye layer [18]. Alternatively C_s could model an oxide layer grown intentionally on the electrodes to inhibit electrochemical reactions [4,12,20]. In either case we shall for simplicity assume that C_s is independent of potential. The total potential drop across the double layer (Debye and Stern) is

$$V_{\text{ext}} - \phi = \zeta - \frac{1}{C_s}q, \tag{6}$$

where V_{ext} is the electrode potential and ϕ is the potential in the electrolyte immediately outside the Debye layer. The overall capacitance (per unit area) of the double layer is defined as

$$C(\zeta) = -\frac{q}{V_{\text{ext}} - \phi}. \tag{7}$$

In the Debye-Hückel limit this reduces to

$$C_0 = \left[\frac{\lambda_D}{\epsilon} + \frac{1}{C_s}\right]^{-1} = \frac{1}{1+\delta\lambda_D}\frac{\epsilon}{}, \tag{8}$$

where we introduced the parameter $\delta = \epsilon/\lambda_D C_s$ for the surface capacitance ratio. At larger potential the Debye layer capacitance becomes very large and $C(\zeta)$ is dominated by the Stern layer only. Notice that $C(\zeta)$ always satisfies $C_0 \leqslant C(\zeta) \leqslant C_s$.

The charging of the Debye layer is due to both Ohmic current from the bulk and Faradaic current from electrochemical reactions at the electrode:

$$\partial_t q = -\mathbf{n}\cdot\mathbf{J} + j_{\text{ext}}, \tag{9}$$

where $\mathbf{n}$ is a unit normal pointing out of the electrode, $-\mathbf{n}\cdot\mathbf{J}$ is the Ohmic current running into the Debye layer, and j_{ext} is the Faradaic current from the electrode reaction. We neglect any surface diffusion and migration of charge in the Debye layer.

C. Insulating walls

On the channel substrate between the electrodes and on the channel lid we assume no normal current:

$$\mathbf{n}\cdot\mathbf{J} = 0. \tag{10}$$

This is justified because the amount of charge required to screen any field normal to the channel wall is both very small and builds up on the Debye time scale $\tau_D = \epsilon/\sigma$.

D. Electrochemistry

For the Faradaic electrode reaction we consider a simple one-step, one-electron redox process of the form

$$\mathrm{O} + \mathrm{e} \rightleftharpoons \mathrm{R}, \tag{11}$$

where $z_{\mathrm{O}} = z_{\mathrm{R}} + 1$ to ensure charge conservation. Activated-complex theory predicts an electric current density j_{ext} from this process given by [18]

$$j_{\text{ext}} = e(k_{\mathrm{R}}^{\ominus} c_{\mathrm{R}}^{s} e^{(1-\alpha)e\Delta\phi_s/k_BT} - k_{\mathrm{O}}^{\ominus} c_{\mathrm{O}}^{s} e^{-\alpha e\Delta\phi_s/k_BT}), \tag{12}$$

where $k_{\mathrm{O}}^{\ominus}$ and $k_{\mathrm{R}}^{\ominus}$ are the forward and backward standard rate constants for the reaction, c_{O}^{s} and c_{R}^{s} are the concentrations of the oxidized and reduced species directly at the electrode (Stern layer) surface, α is the transfer coefficient, and $\Delta\phi_s = -q/C_s$ is the potential drop across the Stern layer. Assuming local equilibrium, the concentrations directly at the surface are related to those immediately outside the Debye layer by a Boltzmann factor, $c_n^s = c_n e^{-z_n e\zeta/k_BT}$. Then Eq. (12) can be reexpressed as the current-overpotential equation

$$j_{\text{ext}} = j_0 e^{-(z_{\mathrm{O}}-\alpha)e\zeta/k_BT}\left(\frac{c_{\mathrm{R}}}{c_{\mathrm{R}}^{*}} e^{(1-\alpha)e\vartheta/k_BT} - \frac{c_{\mathrm{O}}}{c_{\mathrm{O}}^{*}} e^{-\alpha e\vartheta/k_BT}\right), \tag{13}$$

where $j_0 = ek^{\ominus}(c_{\mathrm{R}}^{*})^{\alpha}(c_{\mathrm{O}}^{*})^{1-\alpha}$ is the exchange current, $k^{\ominus}$ is the standard rate constant, c_{O} and c_{R} are the concentrations immediately outside the Debye layer, and c_{O}^{*} and c_{R}^{*} are the corresponding bulk values. $\vartheta = (V_{\text{ext}} - \phi) - (V_{\text{ext}} - \phi)^{\text{eq}}$ is the difference between the actual potential drop across the double layer and the thermal equilibrium for the given redox process and bulk concentration; ϑ is also termed the overpotential. Appreciating that c_n^s in Eq. (12) differs from c_n by a Boltzmann factor is known as the Frumkin correction to the standard Butler-Volmer equation.

Linearization for small ϑ and for c_{O} and c_{R} close to their bulk values yields

$$j_{\text{ext}} = j_0 e^{-(z_{\mathrm{O}}-\alpha)e\zeta^{\text{eq}}/k_BT}\left(\frac{c_{\mathrm{R}}}{c_{\mathrm{R}}^{*}} - \frac{c_{\mathrm{O}}}{c_{\mathrm{O}}^{*}} + \frac{e\vartheta}{k_BT}\right), \tag{14}$$

where ζ^{eq} is the intrinsic equilibrium potential drop across the Debye layer.

For simplicity we shall take $\zeta^{\text{eq}} = 0$ such that $\vartheta = V_{\text{ext}} - \phi$ in Eqs. (13) and (14), set the transfer coefficient α equal to 1/2, and assume that c_{O} and c_{R} are virtually at their bulk values; i.e., we neglect the effect of mass transfer. Moreover, for surface potentials outside the Debye-Hückel limit we shall focus on a monovalent symmetric (1:1) electrolyte with $z_{\mathrm{O}} = z_{\mathrm{R}} + 1 = 1$. Then Eqs. (13) and (14) reduce to

$$j_{\text{ext}} = j_0 e^{-e\zeta/2k_BT} 2\sinh\left[\frac{e(V_{\text{ext}} - \phi)}{2k_BT}\right] \tag{15}$$

and

$$j_{\text{ext}} = j_0\frac{e(V_{\text{ext}} - \phi)}{k_BT} = \frac{V_{\text{ext}} - \phi}{R_{\text{ct}}}, \tag{16}$$

respectively, defining the (area specific) charge-transfer resistance $R_{\text{ct}} = k_BT/j_0 e$ of units $[\Omega\mathrm{m}^2]$.

E. Fluid dynamics

Since we are considering a microsystem where the Reynolds number is usually very low—i.e., where viscosity

dominates over inertia—the fluid motion in the electrolyte is described by the Stokes equation

$$\rho_m \partial_t \mathbf{u} = -\nabla p + \eta \nabla^2 \mathbf{u}, \tag{17}$$

together with the incompressibility constraint

$$\nabla \cdot \mathbf{u} = 0. \tag{18}$$

Here ρ_m is the fluid mass density, $\mathbf{u}$ the velocity, p the pressure, and η the dynamic viscosity. On the insulating walls the no-slip boundary condition applies, whereas on the electrodes we impose a tangential slip condition based on the Helmholtz-Smoluchowski slip velocity:

$$u_s = -\frac{\epsilon \zeta}{\eta} E_t, \tag{19}$$

where $E_t = -\partial_x \phi$ is the tangential field. This classical result for the slip at a flat surface with an externally applied tangential field also holds in general for thin Debye layers in quasiequilibrium at a metal-electrolyte interface [21].

The fluid flow pattern is complex with rolls above the electrode edges as sketched in Fig. 1. However, we are mostly concerned with the net pumping $\langle Q \rangle = \int_0^H \langle u_x \rangle dy$. A Fourier analysis shows that *on average* in time and space the flow is a simple Couette flow driven by the average slip velocity U on the bottom wall [15]

$$U = \frac{1}{L}\left\{ \int_{x_1}^{x_2} \langle u_s \rangle dx + \int_{x_3}^{x_4} \langle u_s \rangle dx \right\}, \tag{20}$$

in terms of which $\langle Q \rangle = HU/2$—that is, assuming zero backpressure on the device. Here x_i denote the positions of the electrode edges (cf. Fig. 1) and the insulating walls do not contribute due to the no-slip condition.

Notice that u_s and hence U are determined entirely by the solution to the electrical problem. Therefore we can study pumping without resolving any details of the actual flow pattern above the electrodes—i.e., without solving Eq. (17)—which is very convenient.

F. Characteristic dimensions

Before going into a detailed analysis it is convenient to discuss the overall properties of the electrokinetic system. Very roughly speaking the system is equivalent to the simple *RC* circuit shown in Fig. 2. The characteristic length scale ℓ_0 of the device is the narrow electrode gap G_1, which determines the magnitude of the electric field strength. The characteristic time τ_0 is the *RC* time for charging the Debye layer through the bulk electrolyte $\tau_0 = R_0 C_0$, where R_0 is the (area specific) bulk electrolyte resistance. However, at high voltage the Stern layer dominates the capacitance of the double layer and the relaxation time changes to $\tau_\infty = R_0 C_s = \tau_0(1 + \delta^{-1})$. The characteristic time for (de)charging the Debye layer through the Faradaic reaction is $\tau_{ct} = R_{ct} C_0$. When the electrode reaction is very facile and $R_{ct} \ll R_0$ this can be significantly faster than the Ohmic charging, acting effectively as a "short-circuit" on the Debye layer. Finally, the characteristic fluid velocity obtained with $\zeta \sim V_0/(1+\delta)$ and $E_t \sim V_0/\ell_0$ is $u_0 = \epsilon V_0^2/\eta \ell_0 (1+\delta)$.

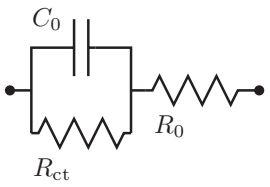


FIG. 2. Equivalent circuit diagram for the electrokinetic system: R_0 is the bulk electrolyte Ohmic resistance, C_0 is the double-layer capacitance, and R_{ct} is the charge-transfer resistance for the electrode reaction. The characteristic time for charging the Debye layer through the electrolyte is $\tau_0 = R_0 C_0$, and the time scale for decharging the layer through Faradaic electrode reaction is $\tau_{ct} = R_{ct} C_0$.

IV. LINEAR ANALYSIS

We first study the problem in the Debye-Hückel limit $\zeta \ll k_B T/ze$. Here the electrokinetic problem can be linearized which simplifies the analysis significantly. Indeed the entire problem is characterized by only two dimensionless groups measuring the ratios of the two relaxation times τ_0 and τ_{ct} and the period of the driving voltage. Of course, the solution also depends on the geometrical parameters W_1, W_2, G_1, G_2, and H, of which four can be varied independently once a characteristic length scale is fixed. We do not vary all these parameters simultaneously, though.

In the simplest case with negligible confinement ($H \gg L$) and no Faradaic electrode reaction we recover the earlier results of Ramos *et al.* We also determine the optimal parameters for the geometry of the electrode array. Then with the relative size of the electrodes fixed we consider how vertical confinement of the fluidic channel to $H \lesssim L$ affects the system and we investigate the behavior when electrode reactions occur. Finally, we discuss how part of the linear analysis can be extended to cover systems with arbitrary intrinsic zeta potential ζ^{eq} on the electrodes but low driving voltage $V_0 \ll (1+\delta) k_B T/ze$.

A. Dimensionless form for linear analysis

In preparation of our numerical analysis of the problem we rescale the variables based on the characteristic dimensions defined in Table I, denoting dimensionless variables by a tilde:

$$\mathbf{r} = \ell_0 \tilde{\mathbf{r}}, \quad t = \tau_0 \tilde{t}, \quad \omega = \tau_0^{-1} \tilde{\omega},$$

$$\phi = V_0 \tilde{\phi}, \quad q = C_0 V_0 \tilde{q}, \quad \mathbf{u} = u_0 \tilde{\mathbf{u}}. \tag{21}$$

In terms of these, the linearized relations (4), (6), (9), and (16) for the charging of the Debye layer on the electrodes reduce to

$$\partial_{\tilde{t}} \tilde{q} = \mathbf{n} \cdot \tilde{\nabla} \tilde{\phi} + K(\tilde{V}_{ext} - \tilde{\phi}), \tag{22}$$

$$\tilde{V}_{ext} - \tilde{\phi} = -\tilde{q}, \tag{23}$$

where $K = R_0/R_{ct} = \tau_0/\tau_{ct}$ is a measure of the facility of the electrode reaction—i.e., a Faradaic conductance. The har-

TABLE I. Characteristic dimensions of the electrokinetic system and typical experimental values for a 10^{-4} M KCl working electrolyte, assuming $\delta=0.1$.

Geometric length	ℓ_0	G_1	5 μm
Debye length	λ_D	$\sqrt{\epsilon k_B T/\Sigma_n c_n^* z_n^2 e^2}$	30 nm
Ohmic relaxation time	τ_0	$R_0 C_0$	70 μs
at large voltage	τ_∞	$R_0 C_s$	770 μs
Faradaic relax. time	τ_{ct}	$R_{ct} C_0$	— s
Debye relaxation time	τ_D	ϵ/σ	0.5 μs
Double-layer cpt.	C_0	$\epsilon/\lambda_D(1+\delta)$	20 mF m^{-2}
Capacitance ratio	δ	$\epsilon/\lambda_D C_s$	0.1 —
Bulk resistance	R_0	ℓ_0/σ	3.3 mΩ m^2
Charge-transfer rst.	R_{ct}	$k_B T/j_0 e$	— Ω m^2
Thermal voltage	—	$k_B T/e$	25 mV
Fluid velocity	u_0	$\epsilon V_0^2/\eta\ell_0(1+\delta)$	— ms^{-1}

monic time dependence is most conveniently dealt with by introducing complex variables

$$\tilde{\phi}(\tilde{\mathbf{r}},\tilde{t}) = \tfrac{1}{2}\hat{\phi}(\tilde{\mathbf{r}})e^{i\tilde{\omega}\tilde{t}} + \text{c.c.} \tag{24}$$

and similarly for $\tilde{V}_{ext}$ and $\tilde{q}$. Then $\tilde{q}$ in Eq. (22) can be eliminated using Eq. (23) to obtain

$$\mathbf{n}\cdot\tilde{\boldsymbol{\nabla}}\hat{\phi} = (i\tilde{\omega}+K)(\hat{\phi}-\hat{V}_{ext}) \tag{25}$$

for the charge balance in the Debye layer. On the insulating walls,

$$\mathbf{n}\cdot\tilde{\boldsymbol{\nabla}}\hat{\phi} = 0, \tag{26}$$

whereas in the bulk Eq. (1) reduces to a Laplace problem

$$-\tilde{\boldsymbol{\nabla}}^2\hat{\phi} = 0. \tag{27}$$

The time average electro-osmotic slip velocity can be expressed as

$$\langle\tilde{u}_s\rangle = \langle(\tilde{V}_{ext}-\tilde{\phi})\partial_{\tilde{x}}\tilde{\phi}\rangle = -\frac{1}{2}\partial_{\tilde{x}}\langle(\tilde{V}_{ext}-\tilde{\phi})^2\rangle = -\frac{1}{4}\partial_{\tilde{x}}|\hat{V}_{ext}-\hat{\phi}|^2, \tag{28}$$

where we used that $\partial_{\tilde{x}}\tilde{V}_{ext}=0$ across the electrodes. The spatial average is then [15]

$$\begin{aligned}\tilde{U} &= \frac{1}{\tilde{L}}\left\{\int_{\tilde{x}_1}^{\tilde{x}_2}\langle\tilde{u}_s\rangle d\tilde{x} + \int_{\tilde{x}_3}^{\tilde{x}_4}\langle\tilde{u}_s\rangle d\tilde{x}\right\} \\ &= -\frac{1}{4\tilde{L}}(|\hat{V}_{ext}-\hat{\phi}|^2_{\tilde{x}_2} - |\hat{V}_{ext}-\hat{\phi}|^2_{\tilde{x}_1} \\ &\quad + |\hat{V}_{ext}-\hat{\phi}|^2_{\tilde{x}_3} - |\hat{V}_{ext}-\hat{\phi}|^2_{\tilde{x}_4}).\end{aligned} \tag{29}$$

B. Validation of the numerical scheme

We solve Eqs. (25)–(27) numerically with the finite-element method, using the commercial software package FEMLAB. We use second-order Lagrange elements to approximate the solution and a nonuniform finite-element mesh with increased resolution at the electrode edges.

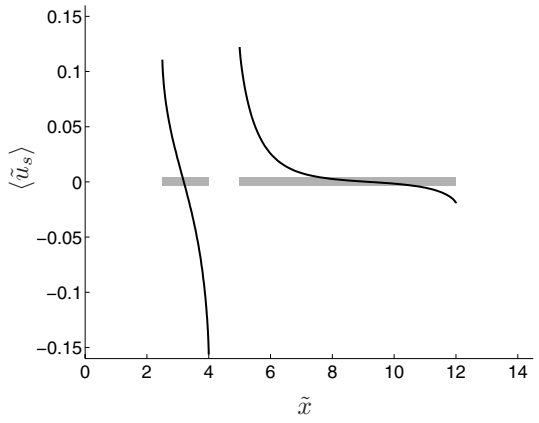


FIG. 3. Time average slip velocity $\langle\tilde{u}_s\rangle$ across the electrodes obtained for $\tilde{\omega}=1$ and $K=0$ (no Faradaic current) in a geometry with $\tilde{W}_1=1.5$, $\tilde{W}_2=7$, $\tilde{G}_2=5$, and $\tilde{H}\gg\tilde{L}$.

In the simple case without any Faradaic current, the Debye layer is charged by Ohmic current through the bulk only. Figure 3 shows the time average slip velocity $\langle\tilde{u}_s\rangle$ across the electrodes for $\tilde{\omega}=1$ and $K=0$ in a geometry with $\tilde{W}_1=1.5$, $\tilde{W}_2=7$, $\tilde{G}_2=5$, and $\tilde{H}\gg\tilde{L}$. The slip velocity is strongest close to the edges of the electrodes and is always directed from the edges towards the center of the electrodes. To understand this behavior, notice that close to the electrode edges the electric field and hence the Ohmic current are stronger, which makes the charging of the Debye layer faster and the screening of $\tilde{V}_{ext}$ more efficient. Then Eq. (28) clearly shows that $\langle\tilde{u}_s\rangle$ is always directed from regions where the screening is good and $|\hat{V}_{ext}-\hat{\phi}|^2$ is large towards regions where the screening is less efficient and $|\hat{V}_{ext}-\hat{\phi}|^2$ is smaller—i.e., from the electrode edges towards the center.

The slip velocity on the electrodes gives rise to bulk fluid motion with rolls above the electrode edges as sketched in Fig. 1. However, the net pumping is determined only by the average slip $\tilde{U}$ which is shown in Fig. 4 as a function of frequency $\tilde{\omega}$. When $\tilde{\omega}\ll 1$ the screening is almost complete at all times. Hence, there is no tangential field and no electroosmotic flow. Conversely, when $\tilde{\omega}\gg 1$ the driving is too fast for any significant screening to occur; i.e., there is no charge in the Debye layer and no flow. As a consequence, in Fig. 4 we see that the pumping velocity $\tilde{U}$ is maximized around $\tilde{\omega}\sim 1$, whereas it falls off as $\tilde{\omega}^2$ at low frequency and slightly faster than $\tilde{\omega}^{-2}$ at high frequency.

In the low-frequency limit $\tilde{\omega}\ll 1$ we can verify our numerical results analytically by expanding the problem in powers of $\tilde{\omega}$ [22]. It then turns out that the leading-order time average slip $\langle\tilde{u}_s\rangle$ scales as $\tilde{\omega}^2$ with a corresponding pumping velocity $\tilde{U}$ shown with a dashed line in Fig. 4. It is interesting to note that the overall magnitude of $\tilde{u}_s$ falls off only linearly with $\tilde{\omega}$ at low frequency. This results from the uniform surface charge density to zeroth order in $\tilde{\omega}$ interacting

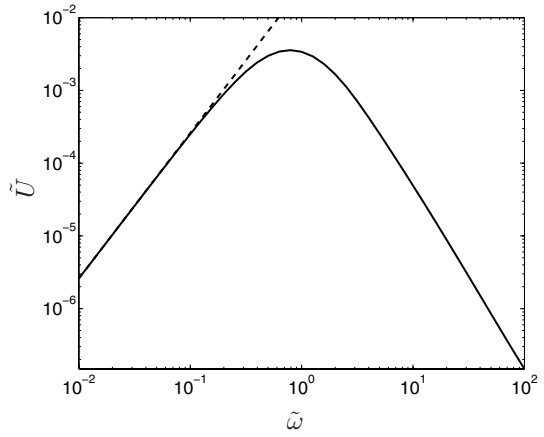


FIG. 4. Pumping velocity $\tilde{U}$ as a function of frequency $\tilde{\omega}$. Dashed line: analytical result at low frequency where $\tilde{U} \propto \tilde{\omega}^2$.

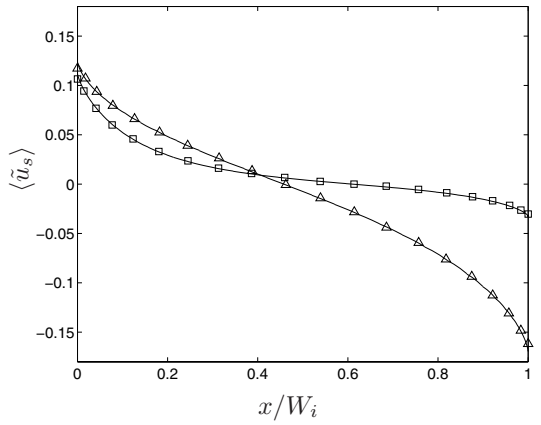


FIG. 5. Time average slip velocity $\langle \tilde{u}_s \rangle$ for both electrodes against the relative position on the electrode x/W_i. Symbols: results of Ramos *et al.* [15] for the narrow (△) and wide (□) electrodes, respectively. Solid lines: our numerical solution.

with the first-order electric field, which gives rise to oscillating flow but no time average. A similar expansion in the high-frequency limit is more difficult. Still, our numerical results show that the magnitude of $\tilde{u}_s$ falls off as $\tilde{\omega}^{-2}$ at high frequency for fixed $\tilde{x}$, but also that it remains finite in a narrow region of width $\tilde{\omega}^{-1}$ from the electrode edges. The latter appears to be an artifact due to our simplified geometry model with infinitely thin electrodes: For electrodes of finite thickness and radius of curvature $\tilde{R}$ at the edges, the magnitude of $\tilde{u}_s$ falls of as $\tilde{\omega}^{-2}$ everywhere on the electrodes for $\tilde{\omega} \gg \tilde{R}^{-1}$; see also [23].

One particular question that one might ask from Fig. 4 is why the maximum net pumping velocity $\tilde{U}_{\max} \sim 0.003$ is so low: If the chosen velocity scale u_0 is appropriate, should we not expect $\tilde{U}_{\max} \sim 1$? Starting from Fig. 3 we note that already the time average slip $\langle \tilde{u}_s \rangle$ is less than $O(1)$ because the double-layer potential drop $|\hat{V}_{\text{ext}} - \hat{\phi}|$ in Eq. (28) is only around half the total applied voltage $\hat{V}_{\text{ext}}$ at the relaxation frequency. The spatial average over the full period of the array includes both positive and negative values of $\langle \tilde{u}_s \rangle$ and passive sections of insulating channel wall, which reduces the value of $\tilde{U}$ further compared to unity. Interestingly, the average slip across only the narrow electrode,

$$\tilde{U}_1 = \frac{1}{\tilde{L}} \int_{\tilde{x}_1}^{\tilde{x}_2} \langle \tilde{u}_s \rangle d\tilde{x}, \tag{30}$$

changes sign for $\tilde{\omega} \sim 1$. At low frequency, $\tilde{U}_1$ and the corresponding average $\tilde{U}_2$ across the wide electrode are both positive with $\tilde{U}_1 \sim 0.1 \tilde{U}_2 \propto \tilde{\omega}^2$. Above the relaxation frequency we have $-\tilde{U}_1 \sim \tilde{U}_2 \propto \omega^{-1}$, although their sum $\tilde{U} = \tilde{U}_1 + \tilde{U}_2$ falls off as (or slightly faster than) $\tilde{\omega}^{-2}$.

In order to further validate our numerical scheme we compare to the results obtained by Ramos *et al.* using a boundary element method based on the Greens function for a unit charge line source [15]. This is shown in Fig. 5, where the time average slip $\langle \tilde{u}_s \rangle$ from both electrodes is shown on the same graph for $\tilde{\omega} = 1.24$ in a geometry with $\tilde{W}_1 = 1$, $\tilde{W}_2 = \tilde{G}_2 = 10/3$, and $\tilde{H} \gg \tilde{L}$. The results from the two different numerical schemes are hardly distinguishable, except close to the electrode edges where our finite-element solution displays a slightly larger slip velocity [24].

We also checked our results for convergence with mesh refinement. The potential $\hat{\phi}(\mathbf{r})$ is everywhere smooth, but the tangential field is singular at the electrode edges due to the abrupt change in boundary condition between Eqs. (25) and (26) [25]. Of course, our finite-element solution cannot represent an infinite slope, but by choosing a strongly nonuniform mesh with very fine resolution near the edges we obtain fairly good quality of the solution even close to the singularity. Moreover, we are mostly concerned with the net pumping $\tilde{U}$ which is an integral quantity and therefore less sensitive to the details at the electrode edges [26].

C. Optimal device geometry

An important question in relation to actual device design is what electrode geometry would maximize the pumping velocity. It is immediately clear that since the characteristic velocity u_0 is inversely proportional to the characteristic length scale ℓ_0, the pumping is increased by downscaling the overall size of the electrodes. We focus again on the limit $H \gg L$, where the results are independent of H. Still, there remains four different length scales describing the electrode geometry—namely, W_1, W_2, G_1, and G_2—of which three can be varied independently once a characteristic length scale is fixed.

Ramos *et al.* also investigated how the pumping velocity depends on the electrode geometry [15]. They rescaled the problem using the width of the wide electrode W_2 as their characteristic length scale, but found that the pumping velocity would increase without bounds as $G_1/W_2 \to 0$. Therefore they fixed the ratio G_1/W_2 at 0.1751, as in the experiment by Brown *et al.* [6], and found that the pumping velocity was then maximized when $W_1/W_2 = 0.24$ and $G_2/W_2 = 0.8$.

TABLE II. List of geometries used experimentally and our linear model prediction for their dimensionless maximum pumping velocity and optimal driving frequency. Of course, the actual pumping velocity depends on the absolute size of the device through u_0.

	$\tilde{W}_1$	$\tilde{W}_2$	$\tilde{G}_2$	$\tilde{U}_{\max}$	$\tilde{\omega}_{\max}$	Refs.
Brown *et al.*	0.93	5.7	3.5	0.00331	1.12	[6,7,11]
Studer *et al.*	0.67	4.33	2.67	0.00293	1.47	[8]
Mpholo *et al.*	1.0	5.0	3.0	0.00332	1.12	[9]
Debesset *et al.*	0.8	5.2	3.2	0.00317	1.26	[10]
Ramos *et al.*	1.37	5.71	4.57	0.00357	0.88	[15] (theory)
Global optimum	1.51	6.55	4.74	0.00359	0.80	Present work

However, the choice $G_1/W_2=0.1751$ is arbitrary. It is more appropriate to identify the narrow electrode gap G_1 as the characteristic length scale and look for an optimal device geometry as a function of $\tilde{W}_1$, $\tilde{W}_2$, and $\tilde{G}_2$ in the parameter ranges $0<\tilde{W}_1\leqslant\tilde{W}_2<\infty$ and $1\leqslant\tilde{G}_2<\infty$. Notice that when $\tilde{W}_2$ and $\tilde{G}_2$ attain their lower limits, the net pumping $\tilde{U}$ vanishes due to left-right symmetry; i.e., the optimum cannot be on the parameter range boundaries. Therefore we can use a simple unconstrained optimization algorithm such as the MATLAB function FMINSEARCH to determine the global optimum:

$$\tilde{U}_{\max}=\max_{[\tilde{\omega},\tilde{W}_1,\tilde{W}_2,\tilde{G}_2]}\{\tilde{U}(\tilde{\omega},K=0,\tilde{W}_1,\tilde{W}_2,\tilde{G}_2)\}. \quad (31)$$

In this way we find that the pumping velocity is maximized for $\tilde{W}_1=1.51$, $\tilde{W}_2=6.55$, and $\tilde{G}_2=4.74$, where $\tilde{U}_{\max}=0.003\,59$ at $\tilde{\omega}_{\max}=0.80$.

Given that we have tracked down the global optimum, it is natural to ask how much better this geometry performs as compared to that suggested by Ramos *et al.* and those used experimentally. This is summarized in Table II: Almost all experimental geometries come within 10% of the maximal pumping velocity, and the theoretical result of Ramos *et al.* is less than 1% off. Yet the frequency $\tilde{\omega}_{\max}$ at which pumping is maximized for the various geometries differs by almost a factor of 2.

The overall conclusion to be drawn is that the pumping velocity is fairly insensitive to the particular choice of electrode geometry. Henceforth, we shall therefore continue to use a geometry with rounded numbers $\tilde{W}_1=1.5$, $\tilde{W}_2=7$, and $\tilde{G}_2=5$ which is still close to optimal; see also Fig. 6.

D. Confined geometry

In several experimental studies the height of the pumping channel H has been comparable to the period of the electrode array L [8,10,11]. It is not immediately clear how this would affect the properties of the device compared to the simple case of no confinement $H\gg L$.

Figure 7 shows the pumping velocity $\tilde{U}$ as a function of frequency $\tilde{\omega}$ and degree of confinement L/H. Our previous results correspond to the base line $L/H\ll 1$, whereas the confinement becomes significant for $L/H\gtrsim 2$. In particular, the pumping velocity increases by roughly a factor of 2 and the optimal driving frequency $\tilde{\omega}_{\max}$ becomes proportional to $\tilde{H}$ for $L/H\gg 1$. The latter is easily understood from our simple circuit model of the system in Fig. 2: While the double-layer capacitance C_0 is largely independent of confinement, the resistance of the bulk electrolyte changes from R_0 to $R_0/\tilde{H}$ because the cross section of the conducting channel is decreased. The relaxation frequency therefore becomes $\tilde{H}/R_0C_0$—i.e., proportional to $\tilde{H}$.

Our findings here are important in relation to actual device design because they indicate that working with a strongly confined geometry does not, as one might have worried, mask the asymmetry and destroy the ability to pump. Thus the channel height H and the period of the electrode array L can be chosen independently. It is important to realize that while the pumping velocity U does not depend much on confinement, the maximal flow rate $Q_{\max}=HU/2$ that the pump can deliver (per unit width of the channel) and the maximal backpressure $\Delta p_{\max}=6\eta NLU/H^2$ it can sustain (where N is the total number of electrode pairs) certainly do.

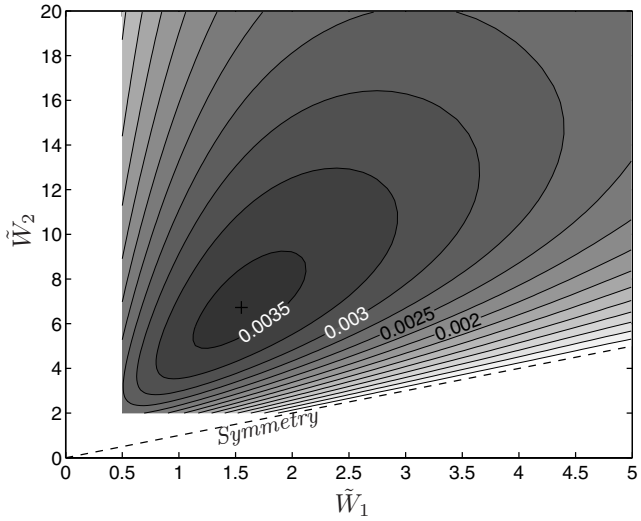


FIG. 6. Contour plot of the maximal pumping velocity over frequency $\max_{\tilde{\omega}}\{\tilde{U}(\tilde{\omega},\ldots)\}$ as a function of $\tilde{W}_1$ and $\tilde{W}_2$ for $\tilde{G}_2=5$. When $\tilde{W}_1=\tilde{W}_2$ the net pumping vanishes due to symmetry. There is a unique optimum close to $\tilde{W}_1=1.5$ and $\tilde{W}_2=7$ with $\tilde{U}_{\max}=0.00358$ and $\tilde{\omega}_{\max}=0.8$.

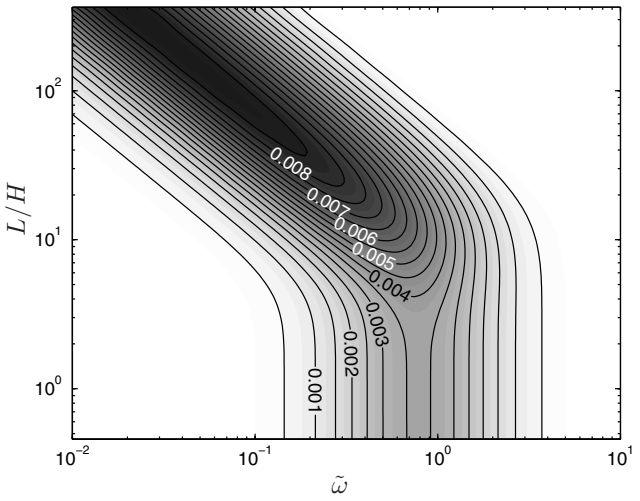


FIG. 7. Contour plot of pumping velocity $\tilde{U}$ as a function of frequency $\tilde{\omega}$ and degree of confinement L/H.

E. Faradaic current injection

When we include electrode reactions and Faradaic current in our model, one of the most important new features is that the complete screening of the electrodes at low frequency is unbalanced. When the reaction is slow and the Faradaic conductance $K=R_0/R_{\mathrm{ct}} \ll \tilde{\omega}$, the effect is small and only perturb the result from purely Ohmic charging slightly. Conversely, when $K \gg \tilde{\omega}$ the reaction is fast enough to establish a steady state where the Ohmic and Faradaic currents are balancing. For either $K \gg 1$ and/or $\tilde{\omega} \gg 1$ the Debye layer is effectively "short-circuited;" cf. Fig. 2. On the basis of these observations we would expect to see pumping even in the low-frequency limit when the Faradaic reaction speed is moderate and $K \sim 1$.

Figure 8 shows the pumping velocity $\tilde{U}$ as a function of $\tilde{\omega}$ and K. For $K \ll \tilde{\omega}$ we recover our previous results for pure

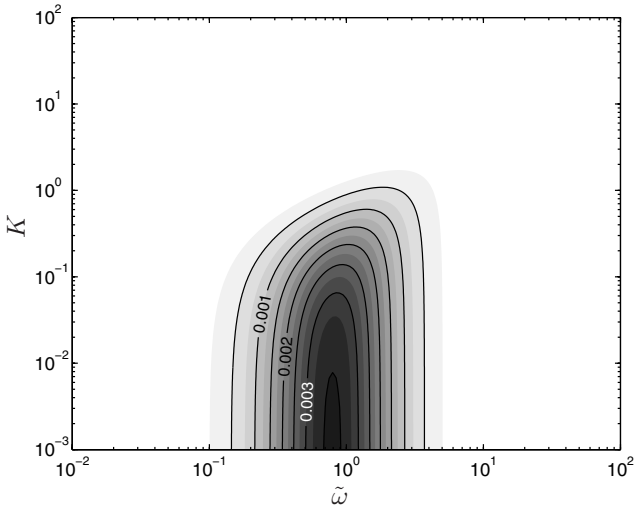


FIG. 8. Contour plot of the pumping velocity $\tilde{U}$ as a function of frequency $\tilde{\omega}$ and Faradaic conductance K.

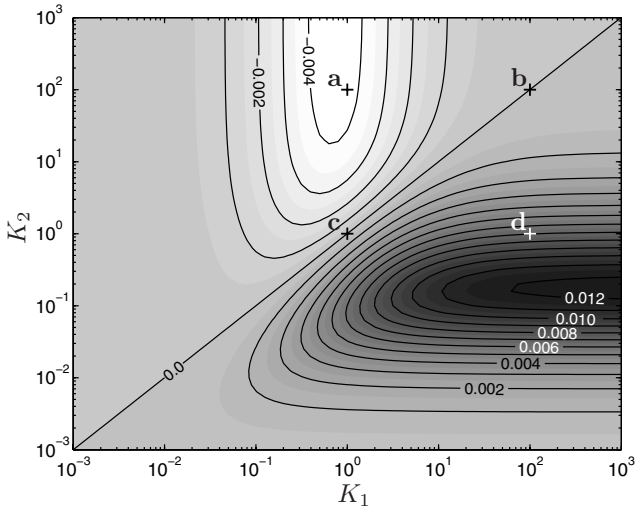


FIG. 9. Contour plot of $\tilde{U}$ in the zero-frequency limit $\tilde{\omega} \sim 0$ with asymmetric current injection: K_1 and K_2 denote the Faradaic conductance on the narrow and wide electrodes, respectively. The slip velocity $\langle \tilde{u}_s \rangle$ at the points (a)–(d) is shown in Fig. 10.

Ohmic charging. However, for $K \gg \tilde{\omega}$ we observe no pumping at all, not even for $K \sim 1$. Figure 9 shows the result of a more detailed investigation in the zero-frequency limit $\tilde{\omega} \sim 0$, where we consider a device with asymmetric surface properties on the electrodes, such that the Faradaic conductance takes the values K_1 and K_2 on the narrow and wide electrodes, respectively. The figure shows that whenever $K_1 \neq K_2$ the device does actually pump: When $K_1 > K_2$ it pumps in the forward direction—i.e., in the same direction as observed for purely Ohmic charging—but for $K_2 > K_1$ the pumping direction is reversed.

This reversal of the pumping direction can be understood qualitatively by inspecting the form of the time average slip $\langle \tilde{u}_s \rangle$ for the four different points marked (a) to (d) in Fig. 9, the result of which is shown in Fig. 10. When $K_1 = K_2 = 1$ (c) the slip velocity is clearly nonzero on both electrodes, but the average across the narrow electrode exactly cancels that across the wide electrode. When $K_1 \gg K_2 = 1$ (d) the Debye layer on the narrow electrode is essentially short-circuited by the fast electrode reaction, leaving no charge and no electroosmotic flow there except from in a narrow region of width $\sim K_1^{-1}$ from the electrode edges. Therefore the pumping velocity $\tilde{U}$ is dominated by the contribution from the wide electrode. Similarly, when $K_2 \gg K_1 = 1$ (a) the Debye layer on the wide electrode is essentially short-circuited and the pumping velocity is dominated by the contribution from the narrow electrode—which tends to be negative.

In conclusion, the absence of pumping in the low-frequency limit when $K_1 = K_2$ is a special case due to high symmetry; whenever $K_1 \neq K_2$ there is some pumping and this falls off linearly with K_i for $K_i \ll 1$. More generally we can prove that for a device with spatially varying capacitance ratio $\delta(\tilde{x})$ and Faradaic conductance $K(\tilde{x})$, there can be no pumping in the low-frequency limit when the product $[1 + \delta(\tilde{x})]K(\tilde{x})$ is a constant on all electrodes. The details of the proof are outlined in Appendix A.

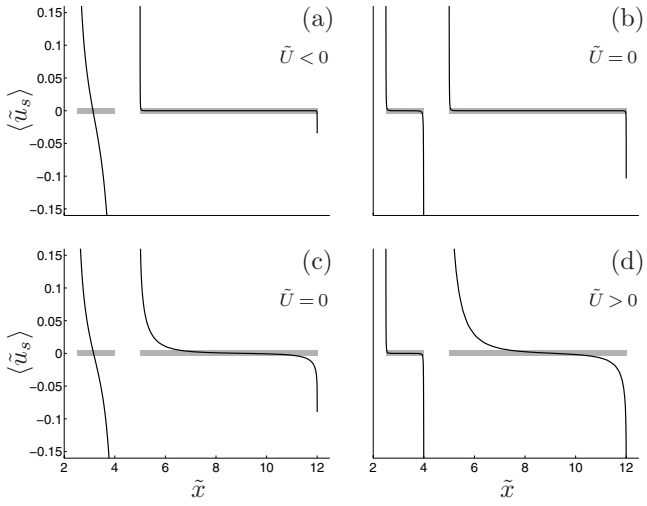


FIG. 10. Time average slip velocity $\langle \tilde{u}_s \rangle$ on the electrodes for the parameter values marked in Fig. 9: (a) $K_1=1$, $K_2=10^2$; (b) $K_1=K_2=10^2$; (c) $K_1=K_2=1$; (d) $K_1=10^2$, $K_2=1$.

F. Linear response at nonzero intrinsic zeta potential

Many electrode-electrolyte systems spontaneously form a Debye layer and possess an intrinsic zeta potential typically of the order $\zeta^{eq} \sim 100$ mV which is well beyond the Debye-Hückel limit $\zeta \ll k_B T/ze \sim 25$ mV. Hence, our foregoing linear analysis cannot be applied immediately to such systems. But then it is standard procedure in electrochemical impedance spectroscopy to study the linear response of the electrokinetic system to a small applied voltage [18].

Consider a system where the electrolyte forms a Debye layer on the electrodes with an intrinsic zeta potential ζ^{eq}. In equilibrium there is a uniform surface charge density q^{eq} given by Eq. (3) on the electrodes and a constant potential $\phi^{eq}=-\zeta^{eq}+q^{eq}/C_s$ in the bulk. If the electrode array is biased with a small-amplitude ac voltage $V_{ext}=\pm V_0 \cos(\omega t)$, we can linearize around equilibrium and writing $\phi=\frac{1}{2}(\phi^{eq}+\hat{\phi}' e^{i\omega t})$ +c.c. and similarly for q and ζ we recover the form of Eqs. (26) and (27) for the linear response $\hat{\phi}'$. Only now the characteristic time used for the rescaling in Eq. (21) becomes

$$\tau_0 = R_0 C_d, \tag{32}$$

where $C_d(\zeta^{eq})$ is the differential capacitance of the double layer:

$$C_d = -\frac{dq}{d(V_{ext}-\phi)}. \tag{33}$$

For a symmetric $z:z$ electrolyte this can be written as

$$C_d = \frac{\epsilon}{\lambda_D[\delta + \mathrm{sech}(ze\zeta^{eq}/2k_BT)]}. \tag{34}$$

The relaxation time for the Faradaic electrode reaction becomes

$$\tau_{ct} = R_{ct} C_d, \tag{35}$$

where $R_{ct}(\zeta^{eq})$ is the differential charge-transfer resistance given by

$$\frac{1}{R_{ct}} = \frac{dj_{ext}}{d(V_{ext}-\phi)} = \frac{j_0 e}{k_B T} e^{-(z_O-\alpha)e\zeta^{eq}/k_BT}, \tag{36}$$

and the Faradaic conductance is redefined as $K=R_0/R_{ct}(\zeta^{eq})$ accordingly. The interaction of the linear response tangential field with the equilibrium Debye layer charge gives rise to a linear response slip velocity

$$u_s' = \frac{\epsilon \zeta^{eq}}{\eta} \partial_x \phi'. \tag{37}$$

However, because ϕ' is harmonic in time, this does not give rise to any net pumping. The leading-order time average slip is to second order in V_0, where the interaction of the induced charge to first order with the tangential field gives rise to a slip velocity of

$$\langle u_s'' \rangle = \frac{\epsilon}{\eta} \langle \zeta' \partial_x \phi' \rangle = -\frac{\epsilon}{4\eta} \frac{d\zeta}{d(V_{ext}-\phi)} \partial_x |\hat{V}_{ext} - \hat{\phi}'|^2. \tag{38}$$

This is fully equivalent to Eq. (28), and the net pumping produced can be obtained directly from the analysis presented earlier in the text. Only now the characteristic fluid velocity becomes

$$u_0 = \frac{\epsilon V_0^2}{\eta \ell_0} \frac{d\zeta}{d(V_{ext}-\phi)}, \tag{39}$$

where, for a symmetric $z:z$ electrolyte,

$$\frac{d\zeta}{d(V_{ext}-\phi)} = \frac{1}{1+\delta \cosh(ze\zeta^{eq}/2k_BT)}. \tag{40}$$

Notice that $u_0(\zeta^{eq})$ falls off exponentially with ζ^{eq} because ζ' becomes negligible.

The linear response dominates over the higher-order components only at low applied voltage. Specifically, we must require that both $V_0 \ll C_d/[dC_d/d(V_{ext}-\phi)]$ and $V_0 \ll R_{ct}/[dR_{ct}/d(V_{ext}-\phi)]$, and a sufficient condition for this is $V_0 \ll (1+\delta)k_BT/ze$. However, because the time average slip velocity vanishes to first order in V_0, one cannot neglect the second-order tangential field $-\partial_x \phi''$, which gives rise to an additional component $\langle \vec{u}_s'' \rangle$ to the slip velocity given by

$$\langle \vec{u}_s'' \rangle = \frac{\epsilon \zeta^{eq}}{\eta} \partial_x \langle \phi'' \rangle. \tag{41}$$

In the absence of Faradaic currents, the time average potential $\langle \phi \rangle$ vanishes to second and all higher orders within our simple electrokinetic model, as is easily seen by taking the time average of Eq. (9). Otherwise, $\langle \phi'' \rangle$ is determined as a solution to the Laplace equation with the boundary condition

$$0 = \sigma \mathbf{n} \cdot \nabla \langle \phi'' \rangle - \frac{1}{R_{ct}} \langle \phi'' \rangle + \frac{1}{4} \frac{d^2 j_{ext}}{d(V_{ext}-\phi)^2} |\hat{V}_{ext} - \hat{\phi}'|^2 \tag{42}$$

at the electrodes, where

$$\frac{d^2 j_{\mathrm{ext}}}{d(V_{\mathrm{ext}}-\phi)^2} = -\frac{e}{k_B T}\frac{2}{R_{\mathrm{ct}}}\left[(z_{\mathrm{O}}-\alpha)\frac{d\zeta}{d(V_{\mathrm{ext}}-\phi)}+\alpha-\frac{1}{2}\right]. \tag{43}$$

This demonstrates that when Faradaic electrode reactions occur, the actual computation of the time average fluid flow and net pumping for finite intrinsic zeta potential (on the electrodes and/or the insulating channel walls) requires a separate and more elaborate study, which we have to leave open despite its experimental relevance.

V. NONLINEAR ANALYSIS

While theoretically it is convenient to work in the low-voltage limit where the system can be characterized by relatively few dimensionless parameters, most experimental work has been done with driving voltages of a few volts in order to obtain an appreciable fluid motion. However, even with the simple nonlinear model we consider here, the parameter space is large, and we therefore focus on the case of a monovalent symmetric (1:1) electrolyte with no intrinsic zeta potential formed on the electrodes—i.e., $\zeta^{\mathrm{eq}}=0$. Moreover, for the Faradaic electrode reaction we take $z_{\mathrm{O}}=z_{\mathrm{R}}+1=1$ and assume a symmetric transfer coefficient $\alpha=1/2$.

We first investigate the system behavior without the presence of the electrode reaction so that the only nonlinearity in our model arises from the nonlinear capacitance of the Debye layer. Then we turn on the electrode reaction and study the full nonlinear model.

A. Dimensionless form for nonlinear analysis

Beyond the Debye-Hückel limit, the smallest characteristic voltage of the electrokinetic system is the thermal voltage $k_B T/e \sim 25$ mV and not the driving voltage V_0. We therefore introduce a new rescaling for the electrical part of the problem:

$$\phi = \frac{k_B T}{e}\tilde{\phi}, \quad V_0 = \frac{k_B T}{e}\tilde{V}_0, \quad q = C_0\frac{k_B T}{e}\tilde{q}. \tag{44}$$

Apart from this, the scaling remains as in Eq. (21); in particular, we retain $u_0=\epsilon V_0^2/\eta\ell_0(1+\delta)$ for the characteristic fluid velocity. Then Eqs. (5), (6), (9), and (15), governing the charging of the Debye layer on the electrodes, reduce to

$$\tilde{V}_{\mathrm{ext}} - \tilde{\phi} = \tilde{\zeta} - \tilde{q}\delta/(1+\delta), \tag{45}$$

$$\partial_{\tilde{t}}\tilde{q} = \mathbf{n}\cdot\tilde{\boldsymbol{\nabla}}\tilde{\phi} + \tilde{j}_{\mathrm{ext}}, \tag{46}$$

where $\tilde{j}_{\mathrm{ext}}$ is the dimensionless Faradaic current given by

$$\tilde{j}_{\mathrm{ext}} = 2Ke^{-\tilde{\zeta}/2}\sinh[(\tilde{V}_{\mathrm{ext}}-\tilde{\phi})/2], \tag{47}$$

with $K=R_0/R_{\mathrm{ct}}$ as before, and where the zeta potential is directly related to the Debye layer charge by

$$\tilde{\zeta}(\tilde{q}) = -2\sinh^{-1}[\tilde{q}/2(1+\delta)]. \tag{48}$$

On the insulating walls we have again

$$\mathbf{n}\cdot\tilde{\boldsymbol{\nabla}}\tilde{\phi} = 0 \tag{49}$$

and, in the bulk,

$$-\tilde{\boldsymbol{\nabla}}^2\tilde{\phi} = 0. \tag{50}$$

The problem, Eqs. (45)–(50), is solved numerically with the finite-element method using FEMLAB. Due to the nonlinearity, it is not possible to solve the problem with a single complex variable for the base frequency component. Instead we employ a relaxation method, where we represent the periodic solution by a set of equispaced points over one period in time, and calculate the time derivative in Eq. (46) using the leapfrog finite-difference scheme; see Appendix B for details.

B. Nonlinear Debye layer capacitance

In the absence of Faradaic electrode reactions the only nonlinearity in the model arises from the nonlinear surface capacitance in the Debye layer. Figure 11 shows the pumping velocity $\tilde{U}$ as a function of the driving frequency $\tilde{\omega}$ and voltage $\tilde{V}_0$ for $\delta=0.1$ and $K=0$. At the base line for $\tilde{V}_0 \ll 1$ we recover the results from the linear analysis with $\tilde{U}$ peaking for $\tilde{\omega}\sim 1$. However, at higher voltage when the double-layer capacitance is dominated by the Stern layer, the frequency at which $\tilde{U}$ is maximized drops to $\tilde{\omega}\sim\tilde{\omega}_\infty$, where $\tilde{\omega}_\infty=\tau_0/\tau_\infty=\delta/(1+\delta)$ is the relaxation frequency at high voltage; cf. Table I.

Figure 11 also shows that the (dimensionless) pumping velocity $\tilde{U}$ falls off at high voltage. This is simply due to the fact that the (physical) electro-osmotic slip velocity $u_s \propto \zeta\partial_x\phi$ no longer scales as V_0^2. When the driving frequency is not too large, Eqs. (45) and (48) allow us to estimate the zeta potential from

$$\tilde{V}_0 \sim \tilde{\zeta} + 2\sinh(\tilde{\zeta}/2)\delta. \tag{51}$$

At low voltage this reduces to $\tilde{\zeta}\sim V_0/(1+\delta)$ whereas at high voltage $\tilde{\zeta}\sim 2\log(\tilde{V}_0/\delta)$. The transition between the two regimes occurs for $\tilde{V}_0\sim\delta$ when $\delta\gg 1$ and for $\tilde{V}_0\sim -2\log\delta$ when $\delta\ll 1$. Hence, for $\tilde{V}_0\gg\delta-2\log\delta$ we expect u_s to scale as $V_0\log V_0$ rather than V_0^2. This is confirmed by Fig. 12, which shows $\tilde{U}_{\max}$ as a function of $\tilde{V}_0$; the data are scaled with $\tilde{V}_0^2$ to recover the voltage dependence of the physical pumping velocity:

$$U = \frac{\epsilon}{\eta\ell_0(1+\delta)}\left(\frac{k_B T}{e}\right)^2\tilde{U}\tilde{V}_0^2. \tag{52}$$

We must emphasize at this point that the upper voltage limit $\tilde{V}_0=10^3$ in Figs. 11 and 12, corresponding to $V_0=25$ V, is far outside the range of validity of our simple electrokinetic model: In Sec. VI we discuss a number of strongly nonlinear phenomena that become significant for $V_0\gtrsim 125$ mV in a typical experiment. Beyond this voltage our results can therefore only provide some characteristic features of the system and should not be regarded as an accurate description.

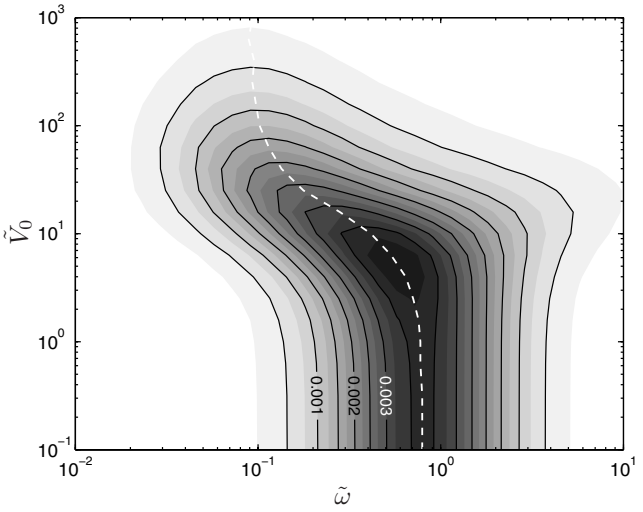


FIG. 11. Contour plot of the pumping velocity $\tilde{U}$ as a function of frequency $\tilde{\omega}$ and driving voltage $\tilde{V}_0$ with capacitance ratio $\delta=0.1$ and no Faradaic current, $K=0$. The white dashed line shows the frequency $\tilde{\omega}_{\max}$ at which $\tilde{U}$ is maximized as a function of $\tilde{V}_0$. At large driving voltage $\tilde{\omega}_{\max}$ shifts from ~ 1 to $\sim \tilde{\omega}_\infty$, reflecting the larger relaxation time when the double-layer capacitance is dominated by the Stern layer.

C. Nonlinear Faradaic current injection

While the nonlinearity induced by the Debye layer capacitance manifests itself relatively slowly as the voltage is increased, we expect a more dramatic effect from the Faradaic current due to the exponential voltage dependence in Eq. (47). Figure 13 shows the pumping velocity $\tilde{U}$ as a function of driving frequency $\tilde{\omega}$ and voltage $\tilde{V}_0$ for $\delta=0.1$ and $K=0.1$. At the base line for $\tilde{V}_0 \ll 1$ we recover the results from the linear analysis in Fig. 8 for $K=0.1$—i.e., no pumping in the low-frequency limit and maximal pumping for $\tilde{\omega} \sim 1$,

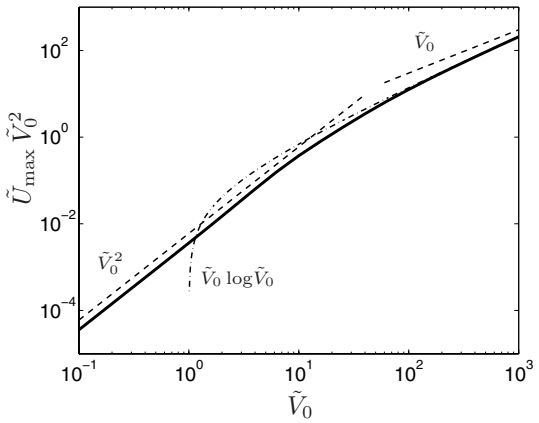


FIG. 12. Maximal pumping velocity $\tilde{U}_{\max}$ as a function of $\tilde{V}_0$, as evaluated along the white dashed line in Fig. 11. The data are scaled with $\tilde{V}_0^2$ to recover the physical dependence on driving voltage. For comparison, the dashed lines show linear and quadratic scaling with $\tilde{V}_0$, while the dash-dotted line is a fit to $\tilde{V}_0 \log \tilde{V}_0$.

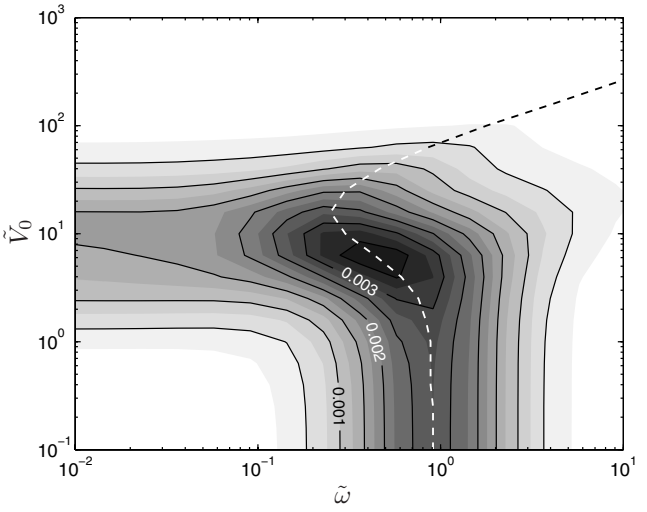


FIG. 13. Contour plot of the pumping velocity $\tilde{U}$ as a function of driving frequency $\tilde{\omega}$ and voltage $\tilde{V}_0$ with capacitance ratio $\delta=0.1$ and Faradaic conductance $K=0.1$. The dashed line shows the optimum driving frequency $\tilde{\omega}_{\max}$ as a function of $\tilde{V}_0$. Notice the pumping in the low-frequency limit for $\tilde{V}_0 \gtrsim 1$.

with a peak value of $\tilde{U}$ less than 0.003. However, for $\tilde{V}_0 \gtrsim 1$ we do observe pumping at low frequency, whereas $\tilde{U}$ falls off rapidly for all $\tilde{\omega}$ when $\tilde{V}_0 \gg 10$. The figure also shows that at low voltage the optimal driving frequency $\tilde{\omega}_{\max}$ follows the same pattern as in Fig. 11, but for $\tilde{V}_0 \gtrsim 10$ it bends sharply up towards higher frequency.

Qualitatively it is not so difficult to see how the pumping at low frequency arises: At low voltage the (magnitude of the) charge density in the Debye layer on the narrow electrode is a factor of W_2/W_1 larger than on the wide electrode to ensure overall charge conservation. Correspondingly, also the double-layer potential drop $\tilde{V}_{\text{ext}}-\tilde{\phi}$ is larger on the narrow electrode, such that the exponential increase in the Faradaic current [cf. Eq. (47)] starts off earlier there. Effectively, this makes the Faradaic conductance larger on the narrow electrode than on the wide one, and according to Fig. 9 from our linear analysis this should indeed give rise to pumping in the forward direction at low frequency.

Figure 14 shows in more detail the voltage dependence of the pumping velocity at low frequency $\tilde{\omega} \ll 1$. For $\tilde{V}_0 \lesssim 1$ the net pumping grows rapidly with $\tilde{V}_0$. Apart from the intrinsic $\tilde{V}_0^2$ scaling, the time average asymmetry in effective Faradaic conductance between the two electrodes grows roughly as $\tilde{V}_0^2$ as well, yielding an overall $\tilde{V}_0^4$ voltage dependence. At larger voltage we find that the pumping levels off to become almost independent of $\tilde{V}_0$. This behavior can be partially understood by considering the scaling of $\tilde{\zeta}$ and the tangential field in this limit: When the frequency is low enough that a steady state is established with the Ohmic and Faradaic currents balancing, we can write the applied voltage $\tilde{V}_0$ as the sum of the potential drop across the double layer $\tilde{V}_{\text{ext}}-\tilde{\phi}$ and the poten-

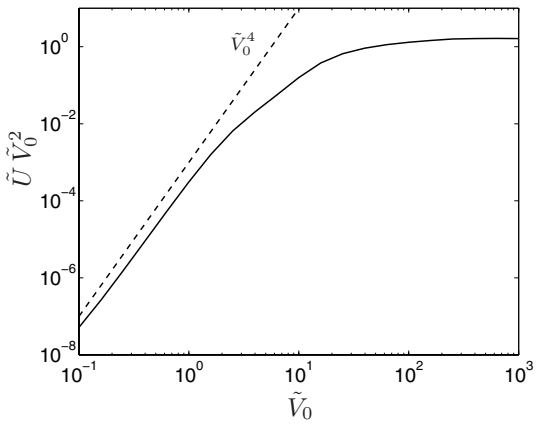


FIG. 14. Pumping velocity $\tilde{U}$ as a function of $\tilde{V}_0$ in the low-frequency limit $\tilde{\omega} \ll 1$. The data are scaled with $\tilde{V}_0^2$ to recover the physical dependence on driving voltage.

tial drop across the bulk electrolyte when an Ohmic current equal to $\tilde{j}_{\text{ext}}$ is passed through it—i.e.,

$$\tilde{V}_0 \sim (\tilde{V}_{\text{ext}} - \tilde{\phi}) + \tilde{j}_{\text{ext}} \sim (\tilde{V}_{\text{ext}} - \tilde{\phi}) + 2Ke^{-\tilde{\zeta}/2} \times \sinh[(\tilde{V}_{\text{ext}} - \tilde{\phi})/2]. \quad (53)$$

If we neglect for simplicity the $\tilde{\zeta}$-dependent exponential prefactor, we find immediately that $\tilde{V}_{\text{ext}} - \tilde{\phi} \sim \tilde{V}_0/(1+K)$ at low voltage $\tilde{V}_0 \ll K - 2\log K$, changing to $\tilde{V}_{\text{ext}} - \tilde{\phi} \sim 2\log(\tilde{V}_0/K)$ at higher voltage. In effect, both the tangential field $\partial_{\tilde{x}}\tilde{\phi} = -\partial_{\tilde{x}}(\tilde{V}_{\text{ext}} - \tilde{\phi})$ and the zeta potential $\tilde{\zeta} \approx (\tilde{V}_{\text{ext}} - \phi)/(1+\delta)$ go as $\log \tilde{V}_0$ at large voltage, and we would expect the (physical) electro-osmotic slip velocity $u_s \propto \zeta \partial_x \phi$ to grow as slowly as $(\log V_0)^2$ with driving voltage.

Finally, we note that since the electrode reaction is very fast at large voltage, the "low-frequency" limit where the Faradaic and Ohmic currents balance extends higher and higher in frequency as the driving voltage is increased. Therefore, the characteristic behavior from Fig. 14 with the pumping velocity leveling off dominates even for $\tilde{\omega} \sim 1$ when $\tilde{V}_0 \gtrsim 10^2$. And when the optimal driving frequency in Fig. 13 shifts towards higher frequency for $\tilde{V}_0 \gtrsim 10$, it is essentially a trace of the point at which the Faradaic current starts to dominate.

VI. DISCUSSION

Our numerical study of the simple electrokinetic model has brought about a number of interesting theoretical predictions that we briefly summarize below and compare to experimental observations reported in the literature. In particular, we comment on the extent to which our results agree or disagree with the experiments and discuss a number of (nonlinear) effects not included in our model that could account for some of the more exotic experimental observations.

A. Confinement

Our analysis has shown that vertical confinement of the pumping channel does not mask the asymmetry of the device and destroy the ability to pump, as one might have worried, but instead enhances the pumping velocity by roughly a factor of 2 and causes the optimal driving frequency to shift down inversely proportional to L/H. In all experimental studies to date, the confinement has been relatively weak with $L/H \lesssim 2$, whereas our results indicate that confinement becomes significant only for $L/H > 2$. It is important to keep in mind, though, that while the pumping velocity U does not depend much on confinement, the maximal flow rate $Q_{\max} = HU/2$ that the pump can deliver and the maximal back-pressure $\Delta p_{\max} \propto U/H^2$ it can sustain certainly do. The optimal choice of channel height therefore depends strongly on whether the pump is intended to work as a pressure or current generator [27].

B. Nonlinear double-layer capacitance

In the absence of Faradaic electrode reactions, where the sole nonlinearity in our model is due to the nonlinear capacitance of the Debye layer, we find that as the driving voltage V_0 is increased well beyond the thermal voltage $k_B T/e$, the relaxation frequency for the electrokinetic system, and hence also the optimal frequency for pumping, drops from $\omega_0 = 1/R_0 C_0$ to $\omega_\infty = 1/R_0 C_s = \omega_0 \delta/(1+\delta)$, where $\delta = \epsilon/\lambda_D C_s$ is the surface capacitance ratio. At the same time, the scaling of the pumping velocity with driving voltage changes from V_0^2 to $V_0 \log V_0$ because the zeta potential entering the Helmholtz-Smoluchowski slip velocity becomes only a small fraction of the overall potential drop across the double layer.

Experimentally, Brown *et al.* reported that the optimal driving frequency $f_{\max}$ dropped from 3 kHz at 0.2 V_{rms} to 1 kHz at 1.2 V_{rms} [6]. For their particular electrolyte and electrode geometry the relaxation frequency at low voltage was expected to be $f_0 = \omega_0/2\pi \simeq (1+\delta) \times 2$ kHz. The observed downshift in optimal frequency therefore fits well with our model predictions if we assume a capacitance ratio of $\delta \simeq 0.5$. Likewise, Studer *et al.* found the optimal driving frequency around 1 kHz for driving voltages larger than 1 V_{rms} with similar electrolyte and electrode geometry [11]. The electrode material in those experimental studies was uncoated gold and platinum, respectively, consistent with a large Stern layer capacitance and hence small δ.

Green *et al.* mapped out the entire fluid velocity field over a single pair of gold electrodes with titanium coating [4]. The titanium spontaneously forms a thin oxide layer which was intended to inhibit Faradaic electrode reactions in the experiment. The observed velocity magnitude and frequency dependence at 500 mV driving voltage was matched with linear theory predictions by assuming a capacitance ratio of $\delta = 3$. It is remarkable that this value was obtained for both a 0.15 mM and a 0.6 mM KCl solution: Naively, assuming a constant value for the Stern layer capacitance C_s one would expect $\delta = \epsilon/\lambda_D C_s$ to depend on the electrolyte concentration through λ_D [28].

Cahill *et al.* observed ac electroosmotic fluid motion due to traveling-wave electric fields induced on a four-phase in-

terdigitated electrode array and found an accurate V_0^2 scaling for 100–500 mV driving voltage, although the pumping velocity peak position and magnitude were smaller than the linear theory predictions [12]. In their experiments the electrodes were coated with a 50-nm-thick Teflon-like insulating layer, yielding a capacitance ratio of $\delta=75-150$ depending on the electrolyte concentration. Hence, they remained in the linear regime over the entire voltage range applied. Ramos *et al.* performed a similar study using titanium-coated gold electrodes [13]. Their optimal driving frequency around 2 kHz was matched with the linear theory by assuming $\delta=1.5$, while the maximum pumping velocity at low frequency was an order of magnitude below the prediction from linear theory.

Generally, the experimentally observed pumping velocities tend to be smaller than the linear theory prediction and sometimes resemble more an affine dependence than a quadratic scaling with V_0 [10,11]. This is at least in qualitative agreement with our nonlinear model results because it is difficult to distinguish V_0 from $V_0 \log V_0$ over the limited voltage range used experimentally.

C. Faradaic current injection

Our linear analysis at low driving voltage shows that when Faradaic electrode reactions occur, the pumping is strongly suppressed for $R_{ct} \ll R_0$ due to "short-circuiting" of the double layer, where R_{ct} is the charge-transfer resistance characterizing the electrode reaction and R_0 is the bulk electrolyte Ohmic resistance. When $R_{ct} \sim R_0$ the Faradaic current injection merely imbalances the otherwise perfect screening at low frequency and therefore induces electro-osmotic fluid motion. Moreover, we predict that the direction of the pumping in the low-frequency limit depends on the relative magnitude of R_{ct} on the narrow and wide electrodes, respectively: If R_{ct} is smaller on the narrow electrodes than on the wide ones, the pumping at low frequency is in the forward direction, whereas it is in the reverse direction otherwise. In the symmetric special case when R_{ct} is equal on both electrodes there is no net pumping at low frequency.

When the driving voltage is increased, the nonlinearity in the Faradaic current injection induces an effective asymmetry in the charge-transfer resistance and we obtain pumping in the forward direction at low frequency, even if the intrinsic surface properties are symmetric. Moreover, at even larger voltage our nonlinear analysis shows that the pumping velocity levels off to a constant value. Using the typical parameter values from Table I and $R_0/R_{ct}=0.1$ as in Fig. 14, the magnitude of the pumping velocity at this plateau is roughly $U \sim 100\ \mu\mathrm{m/s}$.

The observation of electrode degradation and bubble formation from electrolysis is experimental evidence that Faradaic reactions do occur above 1.5–4.0 V_{rms}, depending on the driving frequency and electrode material [8–11,14]. Moreover, Studer *et al.* observed that the pumping direction was reversed when the driving frequency was increased above 20 kHz and the voltage above 3 V_{rms} [11]. Ramos *et al.* also found reversal of the pumping direction on an asymmetric electrode array, as well as a similar reversal on a traveling-wave array for driving voltages above 2 V and frequencies around 1 kHz [13,14]. And Lastochkin *et al.* reported that the direction of the time average slip velocity on an array of T-shaped electrodes was reversed as compared to the direction usually expected for ac electroosmosis [16].

We note that the reversal of the pumping direction that we predict at low frequency is essentially due to an imbalance between net pumping contributions of opposite signs from the two electrodes. We never find reversal of the time average slip direction within our simple electrokinetic model. It is possible that the pumping reversal observed by Studer *et al.* could be due to imbalance between the net pumping contributions from the two electrodes: Although our nonlinear analysis predicts pumping in the forward direction, there could be other nonlinear effects leading to imbalance into the reverse direction. However, Ramos *et al.* observed pumping reversal on a traveling-wave array where all the electrodes are identical—this rules out any imbalance between the contributions from individual electrodes in their setup. And the reversal of the time average slip direction observed by Lastochkin *et al.* is certainly a different phenomenon.

Lastochkin *et al.* argued that strong Faradaic current injection in their system would reverse the polarity of the charge in the Debye layer and hence the direction of the slip. Based on the simple circuit model for the electrokinetic system from Fig. 2, it is difficult to see how the potential drop across (and hence the charge in) the double layer could change sign, regardless of how small R_{ct} becomes. However, since the driving frequency in their experiments was in the MHz range, which is of the order of the Debye frequency for the electrolyte, the Debye layer cannot be assumed to be in local equilibrium and Eqs. (4) and (16) are no longer valid.

Finally, we recall that for simplicity we have been focusing here on the case with no intrinsic zeta potential ζ^{eq} on the electrodes and/or the insulating channel walls. Our analysis in Sec. IV F indicates that when Faradaic electrode reactions are present, the behavior for nonzero ζ^{eq} needs to be studied more carefully, especially in the nonlinear regime.

D. Effect of mass transfer on reaction kinetics

One important effect that we have neglected in our model is the influence of mass transfer on the electrode reaction kinetics. When the oxidized and reduced species O and R are consumed or produced at the electrode by the Faradaic reaction, their local concentrations c_O and c_R become different from the bulk concentrations c_O^* and c_R^*, respectively. Now, the lower c_O drops at the electrode, the more difficult it becomes to run the reaction forward, and vice versa for c_R. This can be modeled as an additional impedance $Z(\omega)$, called the Warburg impedance, in series with the charge-transfer resistance R_{ct} in Fig. 2 [18]. The transport of O and R from the bulk to the surface is by diffusion, and the characteristic distance that the nth species can diffuse over one cycle of the driving voltage is $\ell_n(\omega)=\sqrt{D_n/\omega}$, where D_n is the diffusion constant. If this distance is small compared to the characteristic dimension of the electrodes ℓ_0, the diffusion process is essentially one dimensional and the (area specific) Warburg impedance takes the form

$$Z(\omega) = \frac{k_B T}{\sqrt{2} e^2} \left[\frac{1}{c_O^* \sqrt{D_O}} + \frac{1}{c_R^* \sqrt{D_R}} \right] \frac{1-i}{\sqrt{\omega}}. \tag{54}$$

Notice that $Z(\omega)$ increases at low frequency because the reactants need to diffuse longer. However, at very low frequency when $\ell(\omega) \gg \ell_0$—i.e., $\omega \ll D/\ell_0^2$—the diffusion layer extends much longer than the separation between the electrodes and the Warburg impedance levels off to a constant

$$Z_0 \sim \frac{k_B T}{e^2} \left[\frac{\ell_0}{c_O^* D_O} + \frac{\ell_0}{c_R^* D_R} \right]. \tag{55}$$

We note that $Z_0 \gtrsim R_0 = \ell_0/\sigma$, where the electrolyte conductivity is given by $\sigma = \Sigma_n D_i c_n^* z_n^2 e^2 / k_B T$. Therefore, even if $R_{ct} \ll R_0$, it is not possible to short-circuit the double layer for $\omega \ll D/\ell_0^2$, because it is the larger of R_{ct} and Z_0 that dominates. More generally, the effect of mass transfer is to make the short-circuiting of the double layer discussed in Secs. IV E and V C less dramatic. Still, we do expect the prediction of the pumping direction based on Fig. 9 to hold for $\tilde{\omega} \ll \omega_0$ and $R_{ct} \gtrsim R_0$.

As the driving voltage is increased, the Faradaic current can become so strong that the concentrations of the reactants at the electrodes differ significantly from their equilibrium values. Ultimately, as one of them approaches zero, the reaction stagnates. The current at which this occurs is termed the limiting current [18], and it is given by the maximal rate at which the reactants can be transported from the bulk to the electrodes by diffusion. Hence, for the forward reaction the limiting current is $j_O^{\lim} \sim e D_O c_O^* / \ell_O(\omega)$. Assuming the frequency is low enough that the Ohmic and Faradaic currents are balancing—i.e., $j_{ext} \sim \sigma V_0 / \ell_0$—we find that the limiting current is reached around $V_0 \sim e D_O c_O^* \ell_0 / \sigma \ell_O(\omega)$. For the typical parameter values from Table I and $\omega \sim \omega_0$, $D_O \sim 10^{-9}$ m^2/s, and $c_O^* \sim 0.1$ mM, this corresponds to $V_0 \sim 100$ mV. Beyond this voltage, the concentration c_O tends to zero at the electrode and the Faradaic current levels off at the limiting value. If there is no supporting electrolyte, this makes the conductivity in the diffusion layer differ significantly from the bulk value and the Debye layer is driven out of local equilibrium and expands in width to $\sim \ell_0^{1/3} \lambda_D^{2/3}$ [30,31]. However, if there is an excess of supporting electrolyte that does not participate in the electrode reaction, the diffusion layer should remain charge neutral and the conductivity largely unaffected. In this case the system will be dominated by capacitive charging and behave as if no Faradaic reaction is taking place. This underlines the potential role of the specifics of the electrolyte and electrode material, known to play a decisive role in a few electrically generated phenomena [32,33].

E. Surface conduction in the Debye layer

At large voltage the Debye layer can accumulate enough charge that the mean ion density, and hence the conductivity, becomes significantly larger than in the bulk electrolyte. In this case one cannot neglect lateral surface currents in Eq. (9). This is quantified by the Dukhin number $\mathrm{Du} = \sigma_s / \sigma \ell_0$, where σ_s is the surface conductivity in the Debye layer, incorporating both migration and electro-osmotic convection of charge, σ is the bulk conductivity, and ℓ_0 is the characteristic geometrical length scale. For a symmetric monovalent (1:1) electrolyte the Dukhin number can be expressed as [17]

$$\mathrm{Du} = \left[1 + \left(\frac{k_B T}{e} \right)^2 \frac{2\epsilon}{\eta D} \right] \frac{\lambda_D}{\ell_0} 4 \sinh^2 \left(\frac{e\zeta}{4 k_B T} \right). \tag{56}$$

Taking $\zeta \sim V_0/(1+\delta)$ and using typical parameter values from Table I, we find that surface conductance becomes significant and $\mathrm{Du} \gtrsim 1$ for $V_0 \gtrsim 250$ mV. Generally, we expect surface currents to smear out the charge distribution across the electrodes, reducing the tangential field and hence the electroosmotic slip and pumping velocity. We do not anticipate that this alone could be the mechanism responsible for reversal of the time average slip velocity or the direction of net pumping.

F. Local salt depletion at electrodes

Another issue relating to the massive accumulation of ions in the Debye layer at large voltage is discussed by Bazant *et al.* [19]—namely, that of where those ions come from. For a symmetric monovalent (1:1) electrolyte we define the excess amount of neutral salt absorbed in the Debye layer as

$$w(x,t) = \frac{1}{2} \int_{\text{Debye layer}} (c_+ + c_- - 2c^*) dy \tag{57}$$

$$= c^* \lambda_D 4 \sinh^2 \left(\frac{e\zeta}{4 k_B T} \right). \tag{58}$$

This salt is periodically taken up and released again between the Debye layer and a diffusion zone of width $\ell(\omega) = \sqrt{D/\omega}$. However, when the amount of salt absorbed in the Debye layer approaches the total amount $c^* \ell(\omega)$ available in the diffusion zone, we get local depletion of salt and our assumption of uniform electrolyte concentration throughout the bulk breaks down. This occurs for

$$\zeta \sim \frac{k_B T}{e} 2 \log(\ell(\omega)/\lambda_D) = \frac{k_B T}{e} \log(\omega_D/\omega), \tag{59}$$

where we wrote $\omega_D = D/\lambda_D^2$. Taking again $\zeta \sim V_0/(1+\delta)$ and using typical parameter values from Table I, we find that salt depletion around the relaxation frequency ω_0 becomes an issue for $V_0 \gtrsim 125$ mV.

There are several interesting phenomena arising when the approximation of uniform bulk electrolyte concentration and conductivity breaks down. Among other things, this would change the local Debye length as seen from the electrode surface and give rise to space charge like $\rho \sim \epsilon \partial_y \phi \partial_y \log(\sigma)$ in the diffusion layer [32]. While the amount of space charge in the diffusion layer would typically be much smaller than in the Debye layer, it interacts with the fluid in a place more remotely from the wall and is therefore more efficient in setting the fluid in motion globally. We are currently studying in more detail the role of this nonlinear mechanism for inducing fluid motion.

VII. CONCLUSION

We have extended existing theory for ac electrokinetic pumping to account for vertical confinement of the pumping channel, nonlinear surface capacitance of the double layer, and current injection from Faradaic electrode reactions in both a linear and a nonlinear scheme. For our particular model system of an asymmetric electrode array micropump we have obtained a number of results that we subsequently have interpreted using simple physical arguments. As these arguments are more general than the specific model, we expect our results to hold with some generality for other similar electrokinetic systems. We therefore believe that our work will be useful for design of electrokinetic micropumps.

Our results compare well with experiments in many aspects but there still remains unexplained phenomena—e.g., the reversal of the net pumping or the time average slip direction. This points out the need for further studies of (other) nonlinear phenomena to fully understand the complex electrohydrodynamic system.

APPENDIX A

We give here a proof that there can be no net pumping at low frequency when the electrode surface properties are symmetric and we stay within the Debye-Hückel limit. Consider a device where the capacitance ratio $\delta(\tilde{x})$ and inverse charge-transfer resistance $K(\tilde{x})$ vary across the electrodes. The electro-osmotic slip velocity becomes

$$\tilde{u}_s = \frac{1+\delta_0}{1+\delta(\tilde{x})}(\tilde{V}_{\text{ext}} - \tilde{\phi})\partial_{\tilde{x}}\tilde{\phi}, \tag{A1}$$

where δ_0 is the capacitance ratio used in the rescaling of the problem, and Eq. (25) describing the charge balance in the Debye layer generalizes to

$$\mathbf{n}\cdot\tilde{\boldsymbol{\nabla}}\hat{\phi} = \left(i\tilde{\omega}\frac{1+\delta_0}{1+\delta(\tilde{x})} + K(\tilde{x})\right)(\hat{\phi} - \hat{V}_{\text{ext}}). \tag{A2}$$

Combining these we can write the time average slip as

$$\langle\tilde{u}_s\rangle = \frac{1+\delta_0}{1+\delta(\tilde{x})}\frac{1}{2}\,\text{Re}[(\hat{V}_{\text{ext}} - \hat{\phi})\partial_{\tilde{x}}\hat{\phi}^*]$$
$$= -\frac{1}{2}\,\text{Re}\left[\frac{\mathbf{n}\cdot\tilde{\boldsymbol{\nabla}}\hat{\phi}}{i\tilde{\omega} + \kappa(\tilde{x})}\partial_{\tilde{x}}\hat{\phi}^*\right], \tag{A3}$$

where $\kappa(\tilde{x}) = K(\tilde{x})[1+\delta(\tilde{x})]/(1+\delta_0)$. In the high-symmetry case where $\kappa(\tilde{x})$ is a constant independent of $\tilde{x}$, the net pumping velocity $\tilde{U}$ can be manipulated as follows:

$$\tilde{U} = \frac{1}{\tilde{L}}\int_{\text{electrodes}}\langle\tilde{u}_s\rangle d\tilde{x} = \frac{1}{2\tilde{L}}\int_{\partial\mathcal{V}} -\text{Re}\left[\frac{\mathbf{n}\cdot\tilde{\boldsymbol{\nabla}}\hat{\phi}}{i\tilde{\omega}+\kappa}\partial_{\tilde{x}}\hat{\phi}^*\right]d\tilde{s}, \tag{A4}$$

where $\partial\mathcal{V}$ denotes the boundary of the computational domain and we used the fact that $\mathbf{n}\cdot\tilde{\boldsymbol{\nabla}}\hat{\phi}$ vanishes on insulating walls and cancels over periodic boundaries. Then, using Gauss' law and recalling that $\mathbf{n}$ is a unit normal pointing out of the electrodes and *into* $\mathcal{V}$, we find

$$\tilde{U} = \frac{1}{2\tilde{L}}\,\text{Re}\left[\frac{1}{i\tilde{\omega}+\kappa}\int_{\mathcal{V}}\tilde{\boldsymbol{\nabla}}\cdot(\tilde{\boldsymbol{\nabla}}\hat{\phi}\partial_{\tilde{x}}\hat{\phi}^*)d\tilde{\mathbf{r}}\right] \tag{A5}$$

$$= \frac{1}{2\tilde{L}}\,\text{Re}\left[\frac{1}{i\tilde{\omega}+\kappa}\int_{\mathcal{V}}\tilde{\boldsymbol{\nabla}}\hat{\phi}\cdot\tilde{\boldsymbol{\nabla}}(\partial_{\tilde{x}}\hat{\phi}^*)d\tilde{\mathbf{r}}\right] \tag{A6}$$

$$= \frac{1}{2\tilde{L}}\int_{\mathcal{V}}\frac{\frac{1}{2}\kappa\partial_{\tilde{x}}(\tilde{\boldsymbol{\nabla}}\hat{\phi}\cdot\tilde{\boldsymbol{\nabla}}\hat{\phi}^*) + \tilde{\omega}\,\text{Im}[\tilde{\boldsymbol{\nabla}}\hat{\phi}\cdot\tilde{\boldsymbol{\nabla}}(\partial_{\tilde{x}}\hat{\phi}^*)]}{\kappa^2+\tilde{\omega}^2}d\tilde{\mathbf{r}} \tag{A7}$$

$$= \frac{1}{2\tilde{L}}\frac{\tilde{\omega}}{\kappa^2+\tilde{\omega}^2}\,\text{Im}\left[\int_{\mathcal{V}}\tilde{\boldsymbol{\nabla}}\hat{\phi}\cdot\tilde{\boldsymbol{\nabla}}(\partial_{\tilde{x}}\hat{\phi}^*)d\tilde{\mathbf{r}}\right]. \tag{A8}$$

Here Eq. (A6) follows from $\tilde{\boldsymbol{\nabla}}^2\hat{\phi}=0$, Eq. (A7) follows from simple complex arithmetics, and finally, Eq. (A8) holds because the integral of the term $\partial_{\tilde{x}}(\tilde{\boldsymbol{\nabla}}\hat{\phi}\cdot\boldsymbol{\nabla}\hat{\phi}^*)$ cancels out due to periodicity.

The final result in Eq. (A8) clearly shows that regardless of κ, the net pumping $\tilde{U}$ goes to zero in the low-frequency limit $\tilde{\omega}\to 0$. The apparent singular behavior for $\kappa=0$ is only apparent because the screening is then complete and $\tilde{\boldsymbol{\nabla}}\hat{\phi}$ zero in the bulk. Hence, our analysis rules out any net pumping in the low-frequency limit of our simple model when the electrode surface properties are symmetric in the sense of making the parameter κ independent of $\tilde{x}$.

APPENDIX B

The nonlinear problem is converted to weak form by multiplying a test function $\varphi(\tilde{\mathbf{r}})$ for $\tilde{\phi}$ on Eq. (50), integrating over the computational domain $\mathcal{V}$, and using Gauss' law with Eqs. (46) and (49) to eliminate $\mathbf{n}\cdot\tilde{\boldsymbol{\nabla}}\tilde{\phi}$. Hence, for $\tilde{\phi}$ to be a weak solution we require that

$$\int_{\mathcal{V}}\tilde{\boldsymbol{\nabla}}\varphi\cdot\boldsymbol{\nabla}\tilde{\phi}d\tilde{\mathbf{r}} + \int_{\text{electrodes}}\varphi[\partial_{\tilde{t}}\tilde{q} - \tilde{j}_{\text{ext}}(\tilde{q})]d\tilde{x} = 0 \tag{B1}$$

for all φ and at all times. The constraint, Eq. (45), on the potential drop across the double layer is satisfied by multiplying a test function $\varrho(\tilde{x})$ for $\tilde{q}$ and requiring that

$$\int_{\text{electrodes}}\varrho[\tilde{V}_{\text{ext}} - \tilde{\phi} - \tilde{\zeta}(\tilde{q}) + \tilde{q}\delta/(1+\delta)]d\tilde{x} = 0 \tag{B2}$$

for all ϱ and at all times. The weak problem is discretized using the Galerkin method; i.e., we expand $\tilde{\phi}$ on a finite set of basis functions $\{\varphi_n(\tilde{\mathbf{r}})\}_{n=1}^{N_\phi}$ as

$$\tilde{\phi}(\tilde{\mathbf{r}},\tilde{t}) \approx \sum_{n=1}^{N_\phi}\tilde{\phi}_n(\tilde{t})\varphi_n(\tilde{\mathbf{r}}) \tag{B3}$$

and similarly for $\tilde{q}$ on $\{\varrho_n(\tilde{x})\}_{n=1}^{N_q}$, and we use those same basis sets as test functions in Eqs. (B1) and (B2). Further, the solution is periodic in time, so we discretize on a set of M equispaced points $\tilde{t}_m = m\Delta\tilde{t}$, $m=1,2,\ldots,M$, over one period, and use the leapfrog finite-difference scheme $\partial_{\tilde{t}}\tilde{q}_n(\tilde{t}_m) \approx [\tilde{q}_n(\tilde{t}_{m+1}) - \tilde{q}_n(\tilde{t}_{m-1})]/2\Delta\tilde{t}$ to evaluate the time derivative in Eq. (B1).

The full coupled problem for all M time steps is solved numerically by damped Newton iterations, using the commercial finite-element software FEMLAB to define the basis functions $\varphi_n(\tilde{\mathbf{r}})$ and $\varrho_n(\tilde{x})$ and to evaluate the Jacobian matrix of the discretized problem. In the frequency-voltage maps in Figs. 11 and 13 there are five data points per decade for both frequency and voltage. The convergence of the Newton iterations is accelerated by a continuation scheme where the converged solution at one frequency is used as initial guess for the next solution at higher frequency and the same voltage; at the lowest frequency, we use the solution from lower voltage as initial guess when stepping to higher voltage.

With nonlinear Debye layer capacitance but no Faradaic currents, we obtain fairly accurate results with a coarse time resolution of $M=16$: When compared to a fine resolution result at $M=64$, the maximal relative error on $\tilde{U}$ is less than 1% (3%) for $\tilde{V}_0 \leqslant 10^1$ (10^3), whereas the maximal relative error on $\tilde{q}$ is about twice as large. Figures 11 and 12 show the results for $M=64$. With Faradaic current injection, the charging and decharging of the Debye layer becomes very rapid at high voltage, which makes the solution more demanding: Comparing results for $M=64$ and $M=128$ we find that the maximal relative error on $\tilde{U}$ is less than 0.1% (1.5%) for $\tilde{V}_0 \leqslant 10^1$ (10^3), but that the corresponding error on $\tilde{q}$ is as large as 0.5% (20%). However, one should keep in mind that at high voltage our weakly nonlinear model also becomes physically inadequate for reasons discussed in Secs. VI D–VI F. Figures 13 and 14 show results for $M=128$.

[1] H. A. Stone, A. D. Stroock, and A. Ajdari, Annu. Rev. Fluid Mech. **36**, 381 (2004).

[2] N. G. Green, A. Ramos, A. Gonzalez, H. Morgan, and A. Castellanos, Phys. Rev. E **61**, 4011 (2000).

[3] A. Gonzalez, A. Ramos, N. G. Green, A. Castellanos, and H. Morgan, Phys. Rev. E **61**, 4019 (2000).

[4] N. G. Green, A. Ramos, A. Gonzalez, H. Morgan, and A. Castellanos, Phys. Rev. E **66**, 026305 (2002).

[5] A. Ajdari, Phys. Rev. E **61**, R45 (2000).

[6] A. B. D. Brown, C. G. Smith, and A. R. Rennie, Phys. Rev. E **63**, 016305 (2000).

[7] S. Debesset, C. J. Hayden, C. Dalton, J. C. T. Eijkel, and A. Manz (unpublished).

[8] V. Studer, A. Pepin, Y. Chen, and A. Ajdari, Microelectron. Eng. **61-62**, 915 (2002).

[9] M. Mpholo, C. G. Smith, and A. B. D. Brown, Sens. Actuators B **92**, 262 (2003).

[10] S. Debesset, C. J. Hayden, C. Dalton, J. C. T. Eijkel, and A. Manz, Lab Chip **4**, 396 (2004).

[11] V. Studer, A. Pepin, Y. Chen, and A. Ajdari, Analyst (Cambridge, U.K.) **129**, 944 (2004).

[12] B. P. Cahill, L. J. Heyderman, J. Gobrecht, and A. Stemmer, Phys. Rev. E **70**, 036305 (2004).

[13] A. Ramos, H. Morgan, N. G. Green, A. Gonzalez, and A. Castellanos, J. Appl. Phys. **97**, 084906 (2005).

[14] P. Garcia, A. Ramos, N. G. Green, and H. Morgan (unpublished).

[15] A. Ramos, A. Gonzalez, A. Castellanos, N. G. Green, and H. Morgan, Phys. Rev. E **67**, 056302 (2003).

[16] D. Lastochkin, R. Zhou, P. Wang, Y. Ben, and H.-C. Chang, J. Appl. Phys. **96**, 1730 (2004).

[17] R. J. Hunter, *Foundations of Colloid Science* (Oxford University Press, New York, 1991).

[18] A. Bard and L. Faulkner, *Electrochemical Methods* (Wiley, New York, 1980).

[19] M. Z. Bazant, K. Thornton, and A. Ajdari, Phys. Rev. E **70**, 021506 (2004).

[20] N. A. Mortensen, L. H. Olesen, L. Belmon, and H. Bruus, Phys. Rev. E **71**, 056306 (2005).

[21] I. Rubinstein and B. Zaltzman, Math. Models Meth. Appl. Sci. **11**, 263 (2001).

[22] L. H. Olesen (unpublished), following an approach initially suggested in Y. Ben (unpublished).

[23] F. Nadal, F. Argoul, P. Krestener, B. Pouligny, C. Ybert, and A. Ajdari, Eur. Phys. J. E **9**, 387–399 (2002).

[24] Ramos *et al.*, used the width W_2 of the wide electrode as their characteristic length scale rather than G_1. Hence in [15] both the dimensionless frequency and slip velocity are larger than ours by a factor of W_2/G_1.

[25] A well-known instance of this singular behavior is that at high frequency, where the screening is negligible, the potential close to the sharp electrode edge goes like $\phi \sim r^{1/2}\cos(\theta/2)$, and the electric field diverges as $|\nabla\phi| \sim r^{-1/2}$. Here r is the distance from the edge and θ the polar angle. At finite frequency the normal field is linked to the potential drop across the double layer through Eq. (25) and therefore cannot be singular. However, the tangential field diverges logarithmically. In particular, at low frequency it can be shown analytically that the potential to first order in ω goes like $\phi \sim \mathrm{Im}[\mathrm{Li}_2(e^{-2\pi(i(x-x_j)+y)/L})]$ close to the electrode edge [22] and the tangential field diverges as $|\partial_x\phi| \sim -\log|x-x_j|$. Here $\mathrm{Li}_2(z)=\Sigma_{n=1}^{\infty} z^n/n^2$ is the dilogarithm function.

[26] At high frequency $\tilde{u}_s$ is nonzero only in a region of width $\tilde{\omega}^{-1}$ from the electrode edges. Therefore our numerical solution becomes inaccurate for $\tilde{\omega} \gtrsim 10^3$ and likewise for $K \gtrsim 10^3$ in the case of Faradaic current injection.

[27] A. Ajdari, C. R. Phys. **5**, 539 (2004).

[28] Experimentally, it is often found even for perfectly polarizable solid electrodes that the double-layer impedance does not match that of a perfect capacitor [29]. Some authors therefore find it efficient to use the empirical so-called constant-phase element to model the double layer, but we prefer not to do so due to the lack of structural justification for it.

[29] Z. Kerner and T. Pajkossy, Electrochim. Acta **46**, 207211 (2000).

[30] W. H. Smyrl and J. Newman, Trans. Faraday Soc. **63**, 207 (1967).

[31] K. T. Chu and M. Z. Bazant, SIAM J. Appl. Math. **65**, 1485 (2005).

[32] P. J. Sides, Langmuir **19**, 2745 (2003).

[33] J. A. Fagan, P. J. Sides, and D. C. Prieve, Langmuir **21**, 1784 (2005).

Appendix C

Paper published in Int. J. Numer. Meth. Eng.

Title
A high-level programming-language implementation of topology optimization applied to steady-state Navier–Stokes flow

Authors
Laurits Højgaard Olesen, Fridolin Okkels, and Henrik Bruus

Reference
Int. J. Num. Meth. Eng. **65**, 975-1001 (2006) (27 pages)

INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2006; **65**:975–1001
Published online 1 September 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nme.1468

A high-level programming-language implementation of topology optimization applied to steady-state Navier–Stokes flow

Laurits Højgaard Olesen*,†, Fridolin Okkels and Henrik Bruus

MIC—Department of Micro and Nanotechnology, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark

SUMMARY

We present a versatile high-level programming-language implementation of non-linear topology optimization. Our implementation is based on the commercial software package FEMLAB, and it allows a wide range of optimization objectives to be dealt with easily. We exemplify our method by studies of steady-state Navier–Stokes flow problems, thus extending the work by Borrvall and Petersson on topology optimization of fluids in Stokes flow (*Int. J. Num. Meth. Fluids* 2003; **41**:77–107). We analyse the physical aspects of the solutions and how they are affected by different parameters of the optimization algorithm. A complete example of our implementation is included as FEMLAB code in an appendix. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: topology optimization; Navier–Stokes flow; inertial effects; FEMLAB

1. INTRODUCTION

The material distribution method in topology optimization was originally developed for stiffness design of mechanical structures [1] but has now been extended to a multitude of design problems in structural mechanics as well as to optics and acoustics [2–5]. Recently Borrvall and Petersson introduced the method for fluids in Stokes flow [6]. However, it is desirable to extend the method to fluids described in a full Navier–Stokes flow; a direction pioneered by the work of Sigmund and Gersborg-Hansen [7–9].

In the present work we present such an extension by introducing a versatile high-level programming-language implementation of non-linear topology optimization, based on the commercial software package FEMLAB. It has a wider range of applicability than the Navier–Stokes

*Correspondence to: Laurits Højgaard Olesen, MIC—Department of Micro and Nanotechnology, DTU Bldg. 345 East, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark.
†E-mail: lho@mic.dtu.dk

Contract/grant sponsor: The Danish Technical Research Council; contract/grant number: 26-03-0037

Received 16 November 2004
Revised 24 May 2005
Accepted 26 June 2005

problems studied here, and moreover it allows a wide range of optimization objectives to be dealt with easily.

Extending the topology optimization method to new physical domains generally involves some rethinking of the design problem and some 'trial and error' to determine suitable design objectives. It also requires the numerical analysis and implementation of the problem, e.g. using the finite element method (FEM). This process is accelerated a lot by using a high-level FEM library or package that allows different physical models to be joined and eases the tasks of geometry setup, mesh generation, and postprocessing. The disadvantage is that high-level packages tend to have rather complex data structure, not easily accessible to the user. This can complicate the actual implementation of the problem because the sensitivity analysis is traditionally formulated in a low-level manner.

In this work we have used the commercial finite-element package FEMLAB both for the solution of the flow problem and for the sensitivity analysis required by the optimization algorithm. We show how this sensitivity analysis can be performed in a simple way that is almost independent of the particular physical problem studied. This approach proves even more useful for multi-field extensions, where the flow problem is coupled to, e.g. heat conduction, convection–diffusion of solutes, and deformation of elastic channel walls in valves and flow rectifiers [10].

The paper is organized as follows: In Section 2 we introduce the topology optimization method for fluids in Navier–Stokes flow, and discuss the objective of designing fluidic devices or channel networks for which the power dissipation is minimized. In Section 3 we express the Navier–Stokes equations in a generic divergence form that allows them to be solved with FEMLAB. This form encompasses a wide range of physical problems. We also work out the sensitivity analysis for a class of integral-type optimization objectives in such a way that the built-in symbolic differentiation tools of FEMLAB can be exploited. In Section 4 we present our two numerical examples that illustrates different aspects and problems to consider: The first example deals with designing a structure that can guide the flow in the reverse direction of an applied pressure drop. The general outcome of the optimization is an *S*-shaped channel, but the example illustrates how the detailed structure depends on the choice of the parameters of the algorithm. The second example deals with a four terminal device where the fluidic channel design that minimizes the power dissipation shows a Reynolds number dependence. As the Reynolds number is increased a transition occurs between two topologically different solutions, and we discuss how the position of the transition depends on the choice of initial conditions. Finally in the appendix we include a transcript of our FEMLAB code required for solving the second numerical example. The code amounts to 111 lines—excluding the optimization algorithm that can be obtained by contacting Svanberg [11–13].

2. TOPOLOGY OPTIMIZATION FOR NAVIER–STOKES FLOW IN STEADY STATE

Although our high-level programming-language implementation is generally applicable we have chosen to start on the concrete level by treating the basic equations for our main example: the full steady-state Navier–Stokes flow problem for incompressible fluids.

We consider a given computational domain Ω with appropriate boundary conditions for the flow given on the domain boundary $\partial\Omega$. The goal of the optimization is to distribute a certain amount of solid material inside Ω such that the material layout defines a fluidic device or

channel network that is optimal with respect to some objective, formulated as a function of the variables, e.g. minimization of the power dissipated inside the domain.

The basic principle in the material distribution method for topology optimization is to replace the original discrete design problem with a continuous one where the material density is allowed to vary continuously between solid and void [2]. Thus in our flow problem we assume the design domain to be filled with some idealized porous material of spatially varying permeability. Solid wall and open channels then correspond to the limits of very low and very high permeability, respectively.

In the final design there should preferably be no regions at intermediate permeability since otherwise it cannot be interpreted as a solution to the original discrete problem. Alternatively it may be possible to fabricate the device from polymeric materials such as PDMS that naturally have a finite permeability to the fluid [14].

2.1. Governing equations for flow in idealized porous media

We assume that the fluid flowing in the idealized porous medium is subject to a friction force $\mathbf{f}$ which is proportional to the fluid velocity $\mathbf{v}$, c.f. Darcy's law. Thus $\mathbf{f} = -\alpha\mathbf{v}$, where $\alpha(\mathbf{r})$ is the inverse of the local permeability of the medium at position $\mathbf{r}$. These properties of the idealized porous medium may only be approximately valid for an actual medium. However, the assumptions are not in conflict with any fundamental physical law, and since the converged solutions contain only solid walls and open channels, the specific nature of the idealized porous medium is of no consequence.

The flow problem is described in terms of the fluid velocity field $\mathbf{v}(\mathbf{r})$ and pressure $p(\mathbf{r})$. The governing equations are the steady state Navier–Stokes equation and the incompressibility constraint

$$\rho(\mathbf{v}\cdot\nabla)\mathbf{v} = \nabla\cdot\boldsymbol{\sigma} - \alpha\mathbf{v} \tag{1}$$

$$\nabla\cdot\mathbf{v} = 0 \tag{2}$$

where ρ is the mass density of the fluid. For an incompressible Newtonian fluid the components σ_{ij} of the Cauchy stress tensor $\boldsymbol{\sigma}$ are given by

$$\sigma_{ij} = -p\delta_{ij} + \eta\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right) \tag{3}$$

where η is the dynamic viscosity. The formalism is valid in three dimensions, but for simplicity we shall consider only two-dimensional problems, i.e. we assume translational invariance in the third dimension and set $\mathbf{r} = (x_1, x_2)$ and $\mathbf{v} = (v_1(\mathbf{r}), v_2(\mathbf{r}))$. The boundary conditions will typically be either Dirichlet type specifying the velocity field $\mathbf{v}$ on the boundary or Neumann type specifying the external forces $\mathbf{n}\cdot\boldsymbol{\sigma}$.

It is convenient to introduce a design variable field $\gamma(\mathbf{r})$ controlling the local permeability of the medium. We let γ vary between zero and unity, with $\gamma = 0$ corresponding to solid material and $\gamma = 1$ to no material. Following Reference [6] we then relate the local inverse permeability $\alpha(\mathbf{r})$ to the design field $\gamma(\mathbf{r})$ by the convex interpolation

$$\alpha(\gamma) \equiv \alpha_{\min} + (\alpha_{\max} - \alpha_{\min})\frac{q[1-\gamma]}{q+\gamma} \tag{4}$$

where q is a real and positive parameter used to tune the shape of $\alpha(\gamma)$. Ideally, impermeable solid walls would be obtained with $\alpha_{\max} = \infty$, but for numerical reasons we need to choose a finite value for $\alpha_{\max}$. For the minimal value we choose $\alpha_{\min} = 0$.[‡]

For a given material distribution $\gamma(\mathbf{r})$ there are two dimensionless numbers characterizing the flow, namely the Reynolds number

$$Re = \frac{\rho \ell v}{\eta} \tag{5}$$

describing the ratio between inertia and viscous forces, and the Darcy number

$$Da = \frac{\eta}{\alpha_{\max} \ell^2} \tag{6}$$

describing the ratio between viscous and porous friction forces. Here ℓ is a characteristic length scale of the system and v a characteristic velocity.

Almost impermeable solid material is obtained for very low Darcy numbers, in practice $Da \lesssim 10^{-5}$. Further insight into the meaning of the Darcy number is gained by considering Poiseuille flow in a channel or slit of width ℓ between two infinite parallel plates of porous material. In this case the fluid velocity inside the porous walls decays on a length scale ℓ_{Da}, where $\ell_{Da} = \sqrt{Da}\,\ell = \sqrt{\eta/\alpha_{\max}}$. See also Section 4.1.1 for details on how the flow depends on Da.

2.2. Power dissipation

In the pioneering work by Borrvall and Petersson [6] the main focus was on minimizing the power dissipation in the fluid. The total power Φ dissipated inside the fluidic system (per unit length in the third dimension) is given by [15]

$$\Phi(\mathbf{v}, p, \gamma) = \int_\Omega \left[\frac{1}{2}\,\eta \sum_{i,j} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right)^2 + \sum_i \alpha(\gamma) v_i^2 \right] \mathrm{d}\mathbf{r} \tag{7a}$$

In steady-state this is equal to the sum of the work done on the system by the external forces and the kinetic energy convected into it

$$\Phi(\mathbf{v}, p, \gamma) = \int_{\partial\Omega} \sum_{i,j} \left[n_i \sigma_{ij} v_j - n_i v_i \left(\tfrac{1}{2}\,\rho v_j^2 \right) \right] \mathrm{d}s \tag{7b}$$

Here $\mathbf{n}$ is a unit outward normal vector such that $\mathbf{n} \cdot \boldsymbol{\sigma}$ is the external force acting on the system boundary and $\mathbf{n} \cdot \boldsymbol{\sigma} \cdot \mathbf{v}$ is the work done on the system by this force. Moreover, in the common case where the geometry and boundary conditions are such that the no-slip condition $\mathbf{v} = \mathbf{0}$

[‡]Borrvall and Petersson suggest a model for plane flow between two parallel surfaces of varying separation $h(\mathbf{r})$. The power dissipation due to out-of-plane shears is modelled by an absorption term $-\alpha\mathbf{v}$, where $\mathbf{v}(\mathbf{r})$ is the average velocity between the surfaces and $\alpha(\mathbf{r}) = 12\eta/h(\mathbf{r})^2$. In their model it is therefore natural to operate with a non-zero $\alpha_{\min} = 12\eta/h_{\max}^2$ in Equation (4).

applies on all external solid walls, while on the inlet and outlet boundaries $\mathbf{v}$ is parallel to $\mathbf{n}$ and $(\mathbf{n} \cdot \nabla)\,\mathbf{v} = 0$,[§] Equation (7b) reduces to

$$\Phi(\mathbf{v}, p, \gamma) = \int_{\partial\Omega} -\mathbf{n} \cdot \mathbf{v} \left(p + \tfrac{1}{2}\rho v^2\right) \mathrm{d}s \tag{7c}$$

Borrvall and Petersson showed that for Stokes flow with Dirichlet boundary conditions everywhere on the boundary $\partial\Omega$, the problem of minimizing the total power dissipation inside the fluidic device subject to a volume constraint on the material distribution is mathematically well-posed. Moreover it was proven that in the case where $\alpha(\gamma)$ is a linear function, the optimal material distribution is fully discrete-valued.

When $\alpha(\gamma)$ is not linear but *convex* then the solid/void interfaces in the optimal solution are not discrete zero/unity transitions but slightly smeared out. Convexity implies that the (negative value of the) slope of α at $\gamma = 0$ is larger than at $\gamma = 1$; therefore there will be a neighbourhood around the discrete interface where it pays to move material from the solid side to the void. Using the interpolation in Equation (4) we have $\alpha'(0) = (\alpha_{\min} - \alpha_{\max})(1 + q)/q$ and $\alpha'(1) = (\alpha_{\min} - \alpha_{\max})q/(1 + q)$. For large values of q the interpolation is almost linear and we expect almost discrete interfaces, whereas for small q we expect smeared out interfaces in the optimized solution.

Consider the case when Equation (7c) applies. If the system is driven with a prescribed flow rate then minimizing the total power dissipation is clearly equivalent to minimizing the pressure drop across the system. Conversely, if the system is driven at a prescribed pressure drop, then the natural design objective will be to maximize the flow rate which is equivalent to maximizing the dissipated power, c.f. Equation (7c). In either case the objective can be described as minimizing the *hydraulic resistance* of the system.

For problems with more complex design objectives, such as a minimax problem for the flow rate through several different outlets, there will typically be no analog in terms of total dissipated power. In such cases there is no guarantee for the existence of a unique optimal solution and one has to be extra careful when formulating the design problem.

3. GENERALIZED FORMULATION OF THE OPTIMIZATION PROBLEM

For a given material distribution we solve the Navier–Stokes flow problem using the commercial finite element software FEMLAB. It provides both a graphical front-end and a library of high-level scripting tools based on the MATLAB programming language, and it allows the user to solve a wide range of physical problems by simply typing in the strong form of the governing equations as text expressions. The equations must then comply with a generic divergence form that eases the conversion to weak form required for the finite element solution. However, that is not a severe constraint since this is the natural way of expressing most partial differential equations originating from conservation laws.

Since we have chosen fluidics as our main example, we begin by expressing the incompressible Navier–Stokes flow problem in divergence form. Then we state the optimization problem

[§]In particular this is the case when the inlets and outlets are chosen as straight channels sufficiently long that prescribing a parabolic Poiseuille profile can be justified, see Figures 1 and 6.

with a general form of the design objective function and perform the discretization and sensitivity analysis based on this generalized formulation. We stress that although for clarity our examples are formulated in two dimensions only, the method is fully applicable for 3D systems.

3.1. The flow problem in divergence form

We first introduce the velocity–pressure vector $\mathbf{u} = [v_1,\ v_2,\ p]$ and define for $i = 1, 2, 3$ the quantities $\mathbf{\Gamma}_i$ and F_i as

$$\mathbf{\Gamma}_1 \equiv \begin{bmatrix} \sigma_{11} \\ \\ \sigma_{21} \end{bmatrix}, \quad \mathbf{\Gamma}_2 \equiv \begin{bmatrix} \sigma_{12} \\ \\ \sigma_{22} \end{bmatrix}, \quad \mathbf{\Gamma}_3 \equiv \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix} \tag{8}$$

and

$$F_1 \equiv \rho(\mathbf{v}\cdot\nabla)v_1 + \alpha(\gamma)v_1, \quad F_2 \equiv \rho(\mathbf{v}\cdot\nabla)v_2 + \alpha(\gamma)v_2, \quad F_3 \equiv \nabla\cdot\mathbf{v} \tag{9}$$

Using this, Equations (1) and (2) can be written in divergence form as

$$\nabla\cdot\mathbf{\Gamma}_i = F_i \quad \text{in } \Omega, \text{ Governing equations} \tag{10a}$$

$$R_i = 0 \quad \text{on } \partial\Omega, \text{ Dirichlet b.c.} \tag{10b}$$

$$-\mathbf{n}\cdot\mathbf{\Gamma}_i = G_i + \sum_{j=1}^{3} \frac{\partial R_j}{\partial u_i}\,\mu_j \quad \text{on } \partial\Omega, \text{ Neumann b.c.} \tag{10c}$$

where $\mathbf{\Gamma}_i$ and F_i are understood to be functions of the solution $\mathbf{u}$, its gradient $\nabla\mathbf{u}$, and of the design variable γ. The quantity $R_i(\mathbf{u}, \gamma)$ in Equation (10b) describes Dirichlet type boundary conditions. For example, fluid no-slip boundary conditions are obtained by defining $R_1 \equiv v_1$ and $R_2 \equiv v_2$ on the external solid walls. The quantity $G_i(\mathbf{u}, \gamma)$ in Equation (10c) describe Neumann type boundary conditions, and μ_i denote the Lagrange multiplier necessary to enforce the constraint $R_i = 0$, e.g. the force with which the solid wall has to act upon the fluid to enforce the no-slip boundary condition. Of course, it is not possible to enforce both Dirichlet and Neumann boundary conditions for the same variable simultaneously. Only when the variable u_i is not fixed by any of the Dirichlet constraints R_j does the Neumann condition G_i come into play, as all $\partial R_j/\partial u_i$ vanish and the Lagrange multipliers μ_j are decoupled from Equation (10c). Inactive Dirichlet constraints can be obtained simply by specifying the zero-function $R_i \equiv 0$, that also satisfies Equation (10b) trivially.

3.2. The objective function

In general the design objective for the optimization is stated as the minimization of a certain *objective function* $\Phi(\mathbf{u}, \gamma)$. We shall consider a generic integral-type objective function of the form

$$\Phi(\mathbf{u}, \gamma) = \int_\Omega A(\mathbf{u}, \gamma)\,\mathrm{d}\mathbf{r} + \int_{\partial\Omega} B(\mathbf{u}, \gamma)\,\mathrm{d}s \tag{11}$$

In particular, we can treat the design objective of minimizing the power dissipation inside the fluidic domain by taking, c.f. Equation (7a)

$$A \equiv \frac{1}{2}\eta \sum_{i,j}\left(\frac{\partial v_i}{\partial x_j}+\frac{\partial v_j}{\partial x_i}\right)^2 + \sum_i \alpha(\gamma) v_i^2 \quad \text{in } \Omega \quad \text{and} \quad B \equiv 0 \quad \text{on } \partial\Omega \tag{12}$$

Alternatively, the objective of maximizing the flow out through a particular boundary segment $\partial\Omega_o$ is obtained by choosing

$$A \equiv 0 \quad \text{in } \Omega \quad \text{and} \quad B \equiv \begin{cases} -\mathbf{n}\cdot\mathbf{v} & \text{on } \partial\Omega_o \\ 0 & \text{on } \partial\Omega\backslash\partial\Omega_o \end{cases} \tag{13}$$

and objectives related to N discrete points $\mathbf{r}_k$ can be treated using Dirac delta functions as

$$A \equiv \sum_{k=1}^{N} A_k(\mathbf{u},\gamma)\delta(\mathbf{r}-\mathbf{r}_k) \quad \text{in } \Omega \quad \text{and} \quad B \equiv 0 \quad \text{on } \partial\Omega \tag{14}$$

Finally we stress that not all optimization objectives lend themselves to be expressed in the form of Equation (11)—an example of which is the problem of maximizing the lowest vibrational eigenfrequency in structural mechanics.

3.3. Optimization problem

The optimal design problem can now be stated as a continuous constrained non-linear optimization problem

$$\min_{\gamma} \quad \Phi(\mathbf{u},\gamma) \tag{15a}$$

$$\text{subject to:} \quad \int_\Omega \gamma(\mathbf{r})\mathrm{d}\mathbf{r} - \beta|\Omega| \leqslant 0, \quad \text{Volume constraint} \tag{15b}$$

$$: \quad 0 \leqslant \gamma(\mathbf{r}) \leqslant 1, \quad \text{Design variable bounds} \tag{15c}$$

$$: \quad \text{Equations (10a)–(10c)}, \quad \text{Governing equations} \tag{15d}$$

With the volume constraint we require that at least a fraction $1-\beta$ of the total volume $|\Omega|$ should be filled with porous material.

The very reason for replacing the original discrete design problem with a continuous one by assuming a porous and permeable material, is that it allows the use of efficient mathematical programming methods for smooth problems. We have chosen the popular method of moving asymptotes (MMA) [11, 12], which is designed for problems with a large number of degrees-of-freedom and thus well-suited for topology optimization [2]. It is a gradient-based algorithm requiring information about the derivative with respect to γ of both the objective function Φ and the constraints. Notice that for any γ the governing equations allow us to solve for $\mathbf{u}$; therefore in effect they define $\mathbf{u}[\gamma]$ as an implicit function. The gradient of Φ is then obtained using the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}[\Phi(\mathbf{u}[\gamma],\gamma)] = \frac{\partial\Phi}{\partial\gamma} + \int_\Omega \frac{\partial\Phi}{\partial\mathbf{u}}\cdot\frac{\partial\mathbf{u}}{\partial\gamma}\,\mathrm{d}\mathbf{r} \tag{16}$$

However, because $\mathbf{u}[\gamma]$ is implicit, it is impractical to evaluate the derivative $\partial\mathbf{u}/\partial\gamma$ directly. Instead, we use the adjoint method to eliminate it from Equation (16) by computing a set of Lagrange multipliers for Equations (10a)–(10c) considered as constraints [16]. For details see Section 3.4.

The optimization process is iterative and the kth iteration consists of three steps:

(i) Given a guess $\gamma^{(k)}$ for the optimal material distribution we first solve Equations (10a)–(10b) for $\mathbf{u}^{(k)}$ as a finite element problem using FEMLAB.
(ii) Next, the sensitivity analysis is performed where the gradient of the objective and constraints with respect to γ is evaluated. In order to eliminate $\partial\mathbf{u}/\partial\gamma$ from Equation (16) we solve the adjoint problem of Equations (10a)–(10c) for the Lagrange multipliers $\tilde{\mathbf{u}}^{(k)}$, also using FEMLAB.
(iii) Finally, we use MMA to obtain a new guess $\gamma^{(k+1)}$ for the optimal design based on the gradient information and the past iteration history.

Of the three steps, (i) is the most expensive computationalwise since it involves the solution of a non-linear partial differential equation.

3.4. Discretization and sensitivity analysis

The starting point of the finite element analysis is to approximate the solution component u_i on a set of finite element basis functions $\{\varphi_{i,n}(\mathbf{r})\}$

$$u_i(\mathbf{r}) = \sum_n u_{i,n}\,\varphi_{i,n}(\mathbf{r}) \tag{17}$$

where $u_{i,n}$ are the expansion coefficients. Similiarly, the design variable field $\gamma(\mathbf{r})$ is expressed as

$$\gamma(\mathbf{r}) = \sum_n \gamma_n\,\varphi_{4,n}(\mathbf{r}) \tag{18}$$

For our incompressible Navier–Stokes problem we use the standard Taylor–Hood element pair with quadratic velocity approximation and linear pressure. For the design variable we have chosen the linear Lagrange element.¶

The problem Equations (10a)–(10c) is discretized by the Galerkin method and takes the form

$$\mathbf{L}_i(\mathbf{U},\boldsymbol{\gamma}) - \sum_{j=1}^{3} \mathbf{N}_{ji}^{\mathrm{T}}\boldsymbol{\Lambda}_j = \mathbf{0} \quad \text{and} \quad \mathbf{M}_i(\mathbf{U},\boldsymbol{\gamma}) = \mathbf{0} \tag{19}$$

where $\mathbf{U}_i$, $\boldsymbol{\Lambda}_i$, and $\boldsymbol{\gamma}$ are column vectors holding the expansion coefficients for the solution $u_{i,n}$, the Lagrange multipliers $\mu_{i,n}$, and the design variable field γ_n, respectively. The column vector $\mathbf{L}_i$ contains the projection of Equation (10c) onto $\varphi_{i,n}$ which upon partial integration

¶Another common choice is the discontinuous and piecewise constant element for the design variable. Notice that for second and higher order Lagrange elements the condition $0 \leqslant \gamma_n \leqslant 1$ does not imply $0 \leqslant \gamma(\mathbf{r}) \leqslant 1$ for all $\mathbf{r}$ because of overshoot at sharp zero-to-unity transitions in γ. This in turn can result in negative α, c.f. Equation (4), which is unphysical and also destroys the convergence of the algorithm.

is given by

$$\mathbf{L}_{i,n} = \int_{\Omega} (\varphi_{i,n} F_i + \nabla \varphi_{i,n} \cdot \mathbf{\Gamma}_i)\, \mathrm{d}\mathbf{r} + \int_{\partial\Omega} \varphi_{i,n} G_i\, \mathrm{d}s \tag{20}$$

The column vector $\mathbf{M}_i$ contains the pointwise enforcement of the Dirichlet constraint Equation (10b)

$$\mathbf{M}_{i,n} = R_i(\mathbf{u}(\mathbf{r}_{i,n})) \tag{21}$$

Finally, the matrix $\mathbf{N}_{ij} = -\partial \mathbf{M}_i / \partial \mathbf{U}_j$ describes the coupling to the Lagrange multipliers in Equation (10c). The solution of the non-linear system in Equation (19) above corresponds to step (i) in kth iteration. The sensitivity analysis in step (ii) requires us to compute

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\gamma}}[\Phi(\mathbf{U}(\boldsymbol{\gamma}), \boldsymbol{\gamma})] = \frac{\partial \Phi}{\partial \boldsymbol{\gamma}} + \sum_{i=1}^{3} \frac{\partial \Phi}{\partial \mathbf{U}_i} \frac{\partial \mathbf{U}_i}{\partial \boldsymbol{\gamma}} \tag{22a}$$

which is done using the standard adjoint method [16]. By construction we have for any $\boldsymbol{\gamma}$ that $\mathbf{L}_i(\mathbf{U}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) - \sum_{j=1}^{3} \mathbf{N}_{ji}^{\mathrm{T}} \mathbf{\Lambda}_j(\boldsymbol{\gamma}) = \mathbf{0}$ and $\mathbf{M}_i(\mathbf{U}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) = \mathbf{0}$. Therefore also the derivative of those quantities with respect to $\boldsymbol{\gamma}$ is zero, and adding any multiple, say $\tilde{\mathbf{U}}_i$ and $\tilde{\mathbf{\Lambda}}_i$, of them to Equation (22a) does not change the result

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\gamma}}[\Phi(\mathbf{U}(\boldsymbol{\gamma}), \boldsymbol{\gamma})] &= \frac{\partial \Phi}{\partial \boldsymbol{\gamma}} + \sum_{i=1}^{3} \frac{\partial \Phi}{\partial \mathbf{U}_i} \frac{\partial \mathbf{U}_i}{\partial \boldsymbol{\gamma}} + \sum_{i=1}^{3} \left[\tilde{\mathbf{U}}_i^{\mathrm{T}} \frac{\partial}{\partial \boldsymbol{\gamma}} \left(\mathbf{L}_i - \sum_{j=1}^{3} \mathbf{N}_{ji}^{\mathrm{T}} \mathbf{\Lambda}_j \right) - \tilde{\mathbf{\Lambda}}_i^{\mathrm{T}} \frac{\partial}{\partial \boldsymbol{\gamma}} (\mathbf{M}_i) \right] \\
&= \frac{\partial \Phi}{\partial \boldsymbol{\gamma}} + \sum_{i=1}^{3} \left(\tilde{\mathbf{U}}_i^{\mathrm{T}} \frac{\partial \mathbf{L}_i}{\partial \boldsymbol{\gamma}} - \tilde{\mathbf{\Lambda}}_i^{\mathrm{T}} \frac{\partial \mathbf{M}_i}{\partial \boldsymbol{\gamma}} \right) + \sum_{i=1}^{3} \left[\frac{\partial \Phi}{\partial \mathbf{U}_i} + \sum_{j=1}^{3} \left(\tilde{\mathbf{U}}_j^{\mathrm{T}} \frac{\partial \mathbf{L}_j}{\partial \mathbf{U}_i} + \tilde{\mathbf{\Lambda}}_j^{\mathrm{T}} \mathbf{N}_{ji} \right) \right] \frac{\partial \mathbf{U}_i}{\partial \boldsymbol{\gamma}} \\
&\quad - \sum_{i=1}^{3} \left[\sum_{j=1}^{3} \tilde{\mathbf{U}}_j^{\mathrm{T}} \mathbf{N}_{ij}^{\mathrm{T}} \right] \frac{\partial \mathbf{\Lambda}_i}{\partial \boldsymbol{\gamma}}
\end{aligned} \tag{22b}$$

Here we see that the derivatives $\partial \mathbf{U}_i / \partial \boldsymbol{\gamma}$ and $\partial \mathbf{\Lambda}_i / \partial \boldsymbol{\gamma}$ of the implicit functions can be eliminated by choosing $\tilde{\mathbf{U}}_i$ and $\tilde{\mathbf{\Lambda}}_i$ such that

$$\sum_{j=1}^{3} (\mathbf{K}_{ji}^{\mathrm{T}} \tilde{\mathbf{U}}_j - \mathbf{N}_{ji}^{\mathrm{T}} \tilde{\mathbf{\Lambda}}_j) = \frac{\partial \Phi}{\partial \mathbf{U}_i} \quad \text{and} \quad \sum_{j=1}^{3} \mathbf{N}_{ij} \tilde{\mathbf{U}}_j = \mathbf{0} \tag{23}$$

where we introduced $\mathbf{K}_{ij} = -\partial \mathbf{L}_i / \partial \mathbf{U}_j$. This problem is the adjoint of Equation (19) and $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{\Lambda}}$ are the corresponding Lagrange multipliers.

In deriving Equation (22b) we implicitly assumed that $\mathbf{N}_{ij}$ is independent of $\boldsymbol{\gamma}$, i.e. that the constraint $R_i(\mathbf{u}, \boldsymbol{\gamma})$ is a linear function. If this is not true then the gradient $\partial \Phi / \partial \boldsymbol{\gamma}$ computed from Equation (22b) is not exact, which may lead to poor performance of the optimization algorithm if the constraints are strongly non-linear. In order to avoid such problems it is necessary to include the non-linear parts of the constraint vector $\mathbf{M}$ into $\mathbf{L}$ and move the corresponding Lagrange multipliers from $\mathbf{\Lambda}$ into $\mathbf{U}$. While this is beyond the scope of the divergence form

discussed in Section 3.1, it is certainly possible to deal with such problems in FEMLAB. Also the sensitivity analysis above remains valid since it relies only on the basic form of Equation (19) for the discretized problem.

3.5. Implementation aspects

We end this section by discussing a few issues on the implementation of topology optimization using FEMLAB.

Firstly there is the question of how to represent the design variable $\gamma(\mathbf{r})$. The governing equations as expressed by $\mathbf{\Gamma}_i$ and F_i in Equation (10a) depend not only on the solution $\mathbf{u}$ but also on γ, and the implementation should allow for this dependence in an efficient way. Here our simple and straightforward approach is to include γ as an extra dependent variable on equal footing with the velocity field and pressure, i.e. we append it to the velocity–pressure vector, redefining $\mathbf{u}$ as

$$\mathbf{u} \equiv [v_1, v_2, p, \gamma] \tag{24}$$

This was already anticipated when we denoted the basis set for γ by $\{\phi_{4,n}(\mathbf{r})\}$. By making γ available as a field variable we can take full advantage of all the symbolic differentiation, matrix, and postprocessing tools for analysing and displaying the material distribution. Appending γ to the list of dependent variables we are required to define a fourth governing equation. However, since we are never actually going to solve this equation, but rather update γ based on the MMA step, we simply define

$$\mathbf{\Gamma}_4 \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad F_4 \equiv 0, \quad G_4 \equiv 0, \quad R_4 \equiv 0. \tag{25}$$

It is crucial then that the finite element solver allows different parts of the problem to be solved in a decoupled manner, i.e. it must be possible to solve Equations (10a)–(10c) for u_i for $i = 1, 2, 3$ while keeping u_4, i.e. γ, fixed.

In FEMLAB the non-linear problem Equation (19) is solved using damped Newton iterations [17]. Therefore, the matrices $\mathbf{K}_{ij} = -\partial \mathbf{L}_i / \partial \mathbf{U}_j$ and $\mathbf{N}_{ij} = -\partial \mathbf{M}_i / \partial \mathbf{U}_j$ appearing in the adjoint problem Equation (23) are computed automatically as part of the solution process and can be obtained directly as MATLAB sparse matrices. They are given by

$$\mathbf{K}_{ij,nm} = -\int_\Omega \left(\varphi_{i,n} \left[\frac{\partial F_i}{\partial u_j} \varphi_{j,m} + \frac{\partial F_i}{\partial \nabla u_j} \cdot \nabla \varphi_{j,m} \right] + \nabla \varphi_{i,n} \right.$$
$$\left. \cdot \left[\frac{\partial \mathbf{\Gamma}_i}{\partial u_j} \varphi_{j,m} + \frac{\partial \mathbf{\Gamma}_i}{\partial \nabla u_j} \cdot \nabla \varphi_{j,m} \right] \right) \mathrm{d}\mathbf{r} - \int_{\partial\Omega} \varphi_{i,n} \frac{\partial G_i}{\partial u_j} \varphi_{j,m} \, \mathrm{d}s \tag{26}$$

and

$$\mathbf{N}_{ij,nm} = -\left. \frac{\partial R_i}{\partial u_j} \right|_{\mathbf{r}_{i,n}} \varphi_{j,m}(\mathbf{r}_{i,n}) \tag{27}$$

Regarding the right-hand side vector $\partial\Phi/\partial\mathbf{U}_i$ in Equation (23), notice that for a general objective as Equation (11), it has the form

$$\frac{\partial\Phi}{\partial u_{i,n}} = \int_\Omega \left(\frac{\partial A}{\partial u_i} + \frac{\partial A}{\partial \nabla u_i}\cdot\nabla\right)\varphi_{i,n}\,\mathrm{d}\mathbf{r} + \int_{\partial\Omega}\frac{\partial B}{\partial u_i}\,\varphi_{i,n}\,\mathrm{d}s \tag{28}$$

It is not in the spirit of a high-level finite element package to program the assembly of this vector by hand. Instead we employ the built-in assembly subroutine of FEMLAB. We construct a copy of the original problem sharing the geometry, finite element mesh, and degree-of-freedom numbering with the original. Only we replace the original fields $\mathbf{\Gamma}_i$, F_i, and G_i with

$$\tilde{\mathbf{\Gamma}}_i \equiv \frac{\partial A}{\partial\nabla u_i}, \quad \tilde{F}_i \equiv \frac{\partial A}{\partial u_i} \quad \text{and} \quad \tilde{G}_i \equiv \frac{\partial B}{\partial u_i} \tag{29}$$

Assembling the right-hand-side vector $\tilde{\mathbf{L}}_i$ with this definition yields exactly Equation (28), c.f. Equation (20). An extra convenience in FEMLAB is that we can rely on the built-in symbolic differentiation tools to compute the derivatives $\partial A/\partial u_i$, etc. In order to try out a new objective for the optimization problem, the user essentially only needs to change the text expressions defining the quantities A and B.

After solving the adjoint problem Equation (23) for $\tilde{\mathbf{U}}_i$ and $\tilde{\mathbf{\Lambda}}_i$ to eliminate $\partial\mathbf{U}_i/\partial\boldsymbol{\gamma}$ and $\partial\mathbf{\Lambda}_i/\partial\boldsymbol{\gamma}$ for $i=1,2,3$ in Equation (22b) we can evaluate the sensitivity

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\gamma}}[\Phi(\mathbf{U},\boldsymbol{\gamma})] &= \frac{\partial\Phi}{\partial\boldsymbol{\gamma}} + \sum_{j=1}^{3}\left(\frac{\partial\mathbf{L}_j}{\partial\boldsymbol{\gamma}}\right)^{\mathrm{T}}\tilde{\mathbf{U}}_j - \left(\frac{\partial\mathbf{M}_j}{\partial\boldsymbol{\gamma}}\right)^{\mathrm{T}}\tilde{\mathbf{\Lambda}}_j \\ &= \tilde{\mathbf{L}}_4 - \sum_{j=1}^{3}(\mathbf{K}_{j4}^{\mathrm{T}}\tilde{\mathbf{U}}_j - \mathbf{N}_{j4}^{\mathrm{T}}\tilde{\mathbf{\Lambda}}_j)\end{aligned} \tag{30}$$

where $\mathbf{K}_{i4} = -\partial\mathbf{L}_i/\partial\boldsymbol{\gamma}$, $\mathbf{N}_{i,4} = -\partial\mathbf{M}_i/\partial\boldsymbol{\gamma}$, and $\tilde{\mathbf{L}}_4 = \partial\Phi/\partial\boldsymbol{\gamma}$ in accordance with $\mathbf{U}_4 \equiv \boldsymbol{\gamma}$. Since the fourth variable γ is treated on equal footing with the other three variables, all expressions required to compute the matrices $\mathbf{K}_{i,4}$ and $\mathbf{N}_{i,4}$ come out of the standard linearization of the problem. This is yet another advantage of including γ as an extra dependent variable.

When dealing with a problem with a volume constraint as in Equation (15b), it is necessary to compute the derivative of the constraint with respect to $\boldsymbol{\gamma}$

$$\frac{\partial}{\partial\gamma_n}\left[\frac{1}{|\Omega|}\int_\Omega\gamma(\mathbf{r})\,\mathrm{d}\mathbf{r} - \beta\right] = \frac{1}{|\Omega|}\int_\Omega\varphi_{n,4}(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{31}$$

which can be obtained by assembling $\hat{\mathbf{L}}_4$ with $\hat{\mathbf{\Gamma}}_4 \equiv \mathbf{0}$, $\hat{F}_4 \equiv 1$, and $\hat{G}_4 \equiv 0$. In the appendix we have included a transcript of the code required to set up and solve the example from Section 4.2 below with FEMLAB. It amounts to 111 lines of code, of which the majority are spent on setting up the actual Navier–Stokes flow problem. Only a minor part goes to set up the adjoint problem and perform the sensitivity analysis. Moreover, this part contains almost no reference to the actual physical problem being solved, and therefore it should apply for any multi-field problem expressed in the divergence form Equations (10a)–(10c) with an objective function of the form

of Equation (11). The code example employs, but does not include, a MATLAB implementation of the MMA optimization algorithm [11–13].

3.5.1. Mesh dependence and regularization techniques. It is well known that many topology optimization problems have trouble with mesh dependence, e.g. in stiffness design of mechanical structures it often pays to replace a thick beam with two thinner beams for a given amount of material. As the finite element mesh is refined, smaller and smaller features can be resolved and therefore appear in the optimized structure. In that sense the flow problem that we consider here is atypical because it is generally unfavourable to replace a wide channel with two narrower channels; hence the proof for the existence of a unique optimal solution with respect to minimization of the total power dissipation in Reference [6].

The problem with mesh dependence can be overcome by various regularization techniques based on filtering of either the design variable $\gamma(\mathbf{r})$ or the sensitivity $\mathrm{d}\Phi/\mathrm{d}\gamma$ [2]. The regularization works by defining a certain length scale r_0 below which any features in $\gamma(\mathbf{r})$ or $\mathrm{d}\Phi/\mathrm{d}\gamma$ are smeared out by the filter; in both cases this results in optimized structures with a minimal feature size $\sim r_0$ independent of the mesh refinement. Unfortunately FEMLAB does not come with such a filter, and hence its implementation is an issue that has to be dealt with before our methodology here can be successfully applied to problems that display mesh dependence.

One strategy is to implement the convolution operation of the filter directly [2]. If the computational domain is rectangular and discretized by square finite elements this is both efficient and fairly easy to program, if not one simply uses a standard filter from the MATLAB Image Processing Toolbox. For an unstructured mesh of triangular elements the programming is more involved and slow in MATLAB due to the need to loop over the design variable nodes and searching the mesh for neighbouring nodes within the filter radius. Therefore an explicit matrix representation of the filter would often be preferred [18].

Another possible strategy is to solve an artificial diffusion problem for the design variable $\gamma(\mathbf{r})$ over some period in 'time' $\Delta t = r_0^2/k$ where k is the 'diffusion' constant. The diffusion equation could be included into the fields of Equation (29) that are otherwise unused, and the 'time' evolution solved using the built-in timestepper in FEMLAB. This procedure is equivalent to the action of a filter with Gaussian kernel of width r_0, and it conserves the total amount of material during the filter action. The same approach could be used to smooth out the sensitivity. However, because $\mathrm{d}\Phi/\mathrm{d}\gamma_n$ is sensitive to the local element size one would need to rescale it with $\int_\Omega \phi_{4,n}(\mathbf{r})\,\mathrm{d}\mathbf{r}$ before application of the filter—actually this is true for any filter acting on $\mathrm{d}\Phi/\mathrm{d}\gamma$ whenever the mesh is irregular and $\int_\Omega \phi_{4,n}(\mathbf{r})\,\mathrm{d}\mathbf{r}$ not constant for all n.

The major disadvantage of this strategy is that it involves solving a time evolution problem in each design iteration which could easily turn out to be the most time-consuming step. Alternatively the timestepping algorithm could be implemented by hand, e.g. deciding on the Crank–Nicholson algorithm with a fixed stepsize $\delta t \leqslant \Delta t$. The mass and stiffness matrices for the diffusion problem can be obtained from FEMLAB, and the corresponding iteration matrix need only be factorized once for the given stepsize and could thus be reused in all subsequent design iterations, making this approach relatively cheap, although more cumbersome than using the built-in timestepper.

3.5.2. Large-scale problems. For large-scale problems and three-dimensional modelling it is often necessary to resort to iterative linear solvers because the memory requirements of a direct matrix factorization becomes prohibitive. In that case the strategy we have outlined

here of obtaining the $\mathbf{K}$ and $\mathbf{N}$ matrices directly as sparse matrices in MATLAB and simply transposing $\mathbf{K}$ before the solution of the adjoint problem may not be practical. Alternatively, if the original physical problem is expressed in divergence form then the FEMLAB representation of that problem contains the symbolic derivatives of $\mathbf{\Gamma}_i$, F_i, and G_i appearing in Equation (26). These fields can be transposed and set in the auxiliary copy of the original problem such that it effectively defines $\tilde{\mathbf{K}}_{ij} = \mathbf{K}_{ji}^{\mathrm{T}}$, while retaining the definitions in Equation (29) for the right-hand side vector $\tilde{\mathbf{L}}_i$. Then the adjoint problem Equation (23) can be solved *without* direct handling of the matrices in MATLAB, and using the same iterative solver algorithm as would be employed for the original physical problem. Ultimately we still require an explicit representation of the matrices $\mathbf{K}_{i,4}$ and $\mathbf{N}_{i,4}$ to evaluate the sensitivity $\mathrm{d}\Phi/\mathrm{d}\boldsymbol{\gamma}$ in Equation (30).

From our point of view the major advantage of using FEMLAB in its present stage of development for topology optimization is not in solving large scale problems, though, but rather in the ease of implementation and the ability to handle problems with coupling between several physical processes.

4. NUMERICAL EXAMPLES

In this section we present our results for topology optimization of Navier–Stokes flow for two particular model systems that we have studied. These systems have been chosen because they illustrate the dependence of the solution on the two dimensionless numbers Re and Da, measuring the importance of the inertia of the fluid and the permeability of the porous medium, respectively, relative to viscosity. Moreover we discuss the dependence of the solution on the initial condition for the material distribution.

For simplicity and clarity we have chosen to consider only two-dimensional model systems. We note that the dimensionality of the problems has no fundamental consequence for the method and the numerics, but only affects computer memory requirements and the demand for CPU time. Our 2D examples can therefore be viewed as idealized test cases for our implementation of topology optimization. Yet, the 2D models are not entirely of academic interest only as they represent two limits of actual 3D systems. Due to planar process technology many contemporary lab-on-a-chip systems have a flat geometry with typical channel heights of about 10 μm and widths of 1 mm, i.e. an aspect ratio of 1:100 [14]. One limit is the case where the channel width is constant and the channel substrate and lid are patterned with a profile that is translation invariant in the transverse (width) direction. In the limit of infinitely wide channels the 2D-flow in the plane perpendicular to the width-direction is an exact solution, while it remains an excellent approximation in a 1:100 aspect ratio channel. This is the model system we have adopted for the numerical examples in the present work. The other important limit is when the channel width is not constant, but the channel height is sufficiently slowly varying that the vertical component of the fluid velocity can be neglected. Then writing the Navier–Stokes equation for the velocity averaged in the vertical (height) direction, the out-of-plane shear imposed by the channel substrate and lid gives rise to an absorption term $-\alpha\mathbf{v}$. This approach was studied by Borrvall and Petersson [6], see also the footnote in Section 2.1. Thus, if one is interested in optimizing the height-averaged flow field in a flat channel the 2D model is sufficient.

When solving the Navier–Stokes flow problem we use the standard direct linear solver in FEMLAB in the Newton iterations. Typically we have around 6000 elements in the mesh,

corresponding to 30 000 degrees-of-freedom. The constrained optimization problem is solved using a MATLAB implementation of the MMA algorithm kindly provided by Svanberg [11, 13], except that we modified the code to use the globally convergent scheme described in Reference [12]. The example script included in the appendix employs only the basic algorithm `mmasub`, though. The design iterations are stopped when the maximal change in the design field is $\|\gamma^{(k+1)}-\gamma^{(k)}\|_\infty \leqslant 0.01$, at which point we typically have $|\Phi^{(k+1)}-\Phi^{(k)}| < 10^{-5}$.

4.1. Example: a channel with reverse flow

Our first numerical example deals with the design of a structure that at a particular point inside a long straight channel can guide the flow in the opposite direction of the applied pressure drop. The corresponding problem with a prescribed flow rate was first suggested and investigated by Gersborg-Hansen [8]. We elaborate on it here to illustrate the importance of the choice of permeability for the porous medium.

The computational domain is shown in Figure 1. It consists of a long straight channel of height ℓ and length $L = 10\ell$; the actual design domain, inside which the porous material is distributed, is limited to the central part of length 5ℓ. The boundary conditions prescribe a pressure drop of Δp from the inlet (left) to the outlet (right), and no-slip for the fluid on the channel side walls.

The optimization problem is stated as a minimization of the horizontal fluid velocity at the point $\mathbf{r}^*$ at the centre of the channel, i.e. the design objective is

$$\Phi = v_1(\mathbf{r}^*) \tag{32}$$

In terms of the general objective Equation (11) this is obtained with $A \equiv v_1(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}^*)$ and $B \equiv 0$. There is no explicit need for a volume constraint because neither of the extreme solutions of completely filled or empty can be optimal. When the design domain is completely filled with porous material we expect a flat flow profile with magnitude below $\Delta p/(5\ell\alpha_{\max})$. In the other extreme case when the channel is completely devoid of porous material the solution is simply a parabolic Poiseuille profile with maximum

$$v_0 = \frac{\ell^2}{8\eta}\frac{\Delta p}{L} \tag{33}$$

However, a structure that *reverses* the flow such that $v_1(\mathbf{r}^*)$ becomes negative will be superior to both these extreme cases in the sense of minimizing Φ.

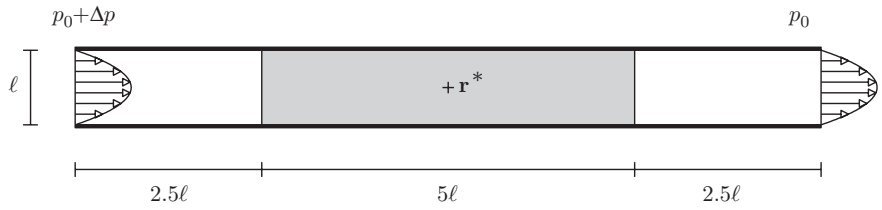


Figure 1. Computational domain for the reverse flow example. The design domain (grey) has length 5ℓ and height ℓ, and the fluid enters and leaves the design domain through leads of length 2.5ℓ. The boundary conditions prescribe a pressure drop of Δp across the system, and the design objective is to reverse the flow direction at the point $\mathbf{r}^*$ at centre of the channel.

4.1.1. Reverse flow in the Stokes limit, $Re = 0$. We first consider the Stokes flow limit of small Δp where the inertial term becomes negligible. The problem is then linear and the solution is characterized by a single dimensionless parameter, namely the Darcy number Da, Equation (6). We have solved the topology optimization problem for different values of Da. The initial condition for the material distribution was $\gamma^{(0)} = 1$, and the parameter q determining the shape of $\alpha(\gamma)$ in Equation (4) was set to $q = 0.1$. Anticipating that the structural details close to $\mathbf{r}^*$ should be more important than those further away we chose a non-uniform finite element mesh with increased resolution around $\mathbf{r}^*$.

Figure 2 shows the optimal structures obtained for $Da = 10^{-3}$, 10^{-4}, 10^{-5} and 10^{-6}. They all consist of two barriers defining an S-shaped channel that guides the fluid in the reverse direction of the applied pressure drop. At $Da = 10^{-3}$ the two barriers are rather thick but leaky with almost all the streamlines penetrating them; as the Darcy number is decreased the optimal structures become thinner and less penetrable. This result can be interpreted as a trade-off between having either thick barriers or wide channels. Thick barriers are necessary to force the fluid into the S-turn, while at the same time the open channel should be as wide as possible in order to minimize the hydraulic resistance and maximize the fluid flow at the prescribed pressure drop.

Notice that if we had chosen to prescribe the flow rate through the device rather than the pressure drop, then the optimal solution would have been somewhat different. When the flow rate is prescribed, it pays to make the gap between the barriers very small and the barriers very thick in order to force the fixed amount of fluid flow through the narrow contraction. The optimal structure is therefore one with a very large hydraulic resistance. In Reference [8], this problem was circumvented by adding a constraint on the maximal power dissipation allowed at the given flow rate.

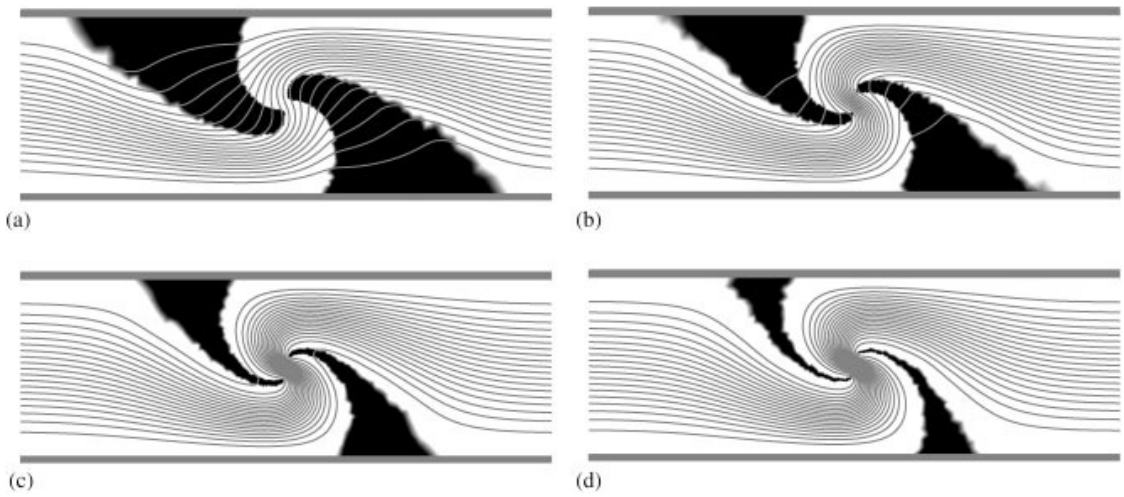


Figure 2. Optimized structures (black) and streamlines at 5% intervals for Stokes flow ($Re = 0$) at Darcy numbers decreasing from 10^{-3} to 10^{-6}. Only the central part of length 3ℓ of the design domain is shown. The structures consist of two barriers defining an S-shaped channel that reverses the flow at the central point $\mathbf{r}^*$. As the Darcy number is decreased, the optimized structures become thinner and less permeable: (a) $Da = 10^{-3}$; (b) $Da = 10^{-4}$; (c) $Da = 10^{-5}$; and (d) $Da = 10^{-6}$.

In order to validate the optimality of the structures computed by the topology optimization we do as follows: For each of the optimized structures from Figure 2 we freeze the material distribution and solve the flow problem for a range of Darcy numbers. The resulting family of curves for $v_1(\mathbf{r}^*)$ vs Da is shown in Figure 3 where it is seen that each of the four structures from Figure 2 do indeed perform better in minimizing $v_1(\mathbf{r}^*)$ than the others at the value of Da for which they are optimized.

For $Da \lesssim 10^{-5}$ the optimal value of $v_1(\mathbf{r}^*)$ tends to saturate because the thin barriers are then almost completely impermeable and the open channel cannot get much wider. In this limit the thickness of the optimized barrier structures approach the mesh resolution as seen in Figure 2(d). When the optimal barrier thickness gets below the mesh size we have observed the appearance of artificial local optima for the barrier structure. The problem is that the thin barriers cannot continuously deform into another position without going through an intermediate structure with barriers that are thicker by at least one mesh element. Depending on the initial condition, the optimization algorithm can therefore end up with a sub-optimal structure. We have tried to work around this problem by decreasing the value of q in order to smear out the solid/void interfaces and thus reduce the cost of going through the intermediate structure. This did not work out well; the reason may be that the smearing property of a convex $\alpha(\gamma)$ was derived for the objective of minimizing the power dissipation subject to a volume constraint. In the present example we are dealing with a different objective and have no volume constraint. However, when the barrier structures are resolved with at least a few elements across them the artificial local optima tend to be insignificant. Thus the problem can be avoided by choosing a sufficiently fine mesh, or by adaptively refining the mesh at the solid/void interfaces.

Returning to Figure 3 we notice that as Da increases all the structures perform poorly in minimizing $v_1(\mathbf{r}^*)$, as they all approach v_0. Extrapolating this trend one might suspect that the S-turn topology will cease to be optimal somewhere above $Da = 10^{-3}$ simply because the

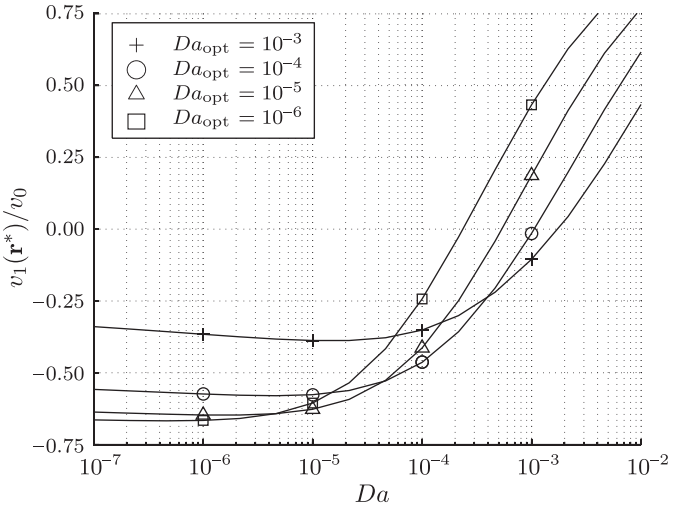


Figure 3. Comparing the performance of the structures from Figure 2 optimized at Da_{opt} for different values of Da. The objective $v_1(\mathbf{r}^*)$ is normalized with the velocity in an empty channel, v_0, c.f. Equation (33).

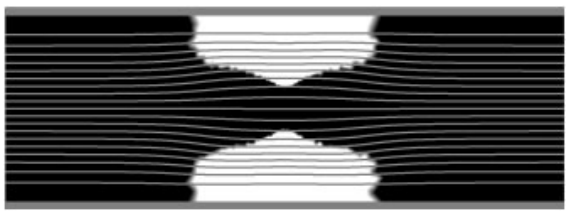


Figure 4. Optimized structure (black) and streamlines for Stokes flow at $Da = 10^{-2}$; only the central part of length 3ℓ. The design domain is completely filled with porous material, except immediately above and below $\mathbf{r}^*$ where two empty regions emerge. These voids divert the flow away from $\mathbf{r}^*$, resulting in a low velocity $v_1(\mathbf{r}^*) = 0.1v_0$.

porous material becomes too permeable to make reversal of the flow direction possible. We have tested this hypothesis by performing an optimization at $Da = 10^{-2}$, resulting in the structure shown in Figure 4 where the value of the objective is $v_1(\mathbf{r}^*) = 0.1v_0$. It is seen to display a different topology from those of Figure 2, with the design domain is almost completely filled with porous material blocking the flow through the channel. Only immediately above and below the point $\mathbf{r}^*$ we see two empty regions emerging that act to guide the flow away from $\mathbf{r}^*$.

Actually, in all four cases from Figure 2, starting from an empty channel the design iterations initially converge towards a symmetric structure blocking the flow like that in Figure 4. However, at a certain point in the iterations an asymmetry in the horizontal plane is excited and the structure quickly changes to the two-barrier S-geometry. Whether the optimization converge to an S- or an inverted S-turn depends on how the asymmetry is excited from numerical noise or irregularity in the finite element mesh; in fact the structure in Figure 2(b) originally came out as an inverted S but was mirrored by hand before plotting it to facilitate comparison with the three other structures.

4.1.2. Reverse flow at finite Reynolds number. We now consider flow at finite Reynolds number, characterized by the two dimensionless numbers Re and Da. The geometry and boundary conditions remain unchanged, for convenience we introduce a non-dimensional pressure drop $\Delta\tilde{p} = \Delta p\, \rho\, \ell^2/\eta^2$, and finally we fix the Darcy number at $Da = 10^{-5}$. We note from Figure 2 that this Darcy number allows some but not much fluid to penetrate the walls. We have nevertheless chosen this Darcy number for practical reasons, as the walls are 'solid' enough and a lower value (more 'solid' wall) would increase the calculation time.

We have solved the topology optimization problem for different values of $\Delta\tilde{p}$, always using an empty channel as initial condition. The results are shown in Figures 5(a)–(c) for $\Delta\tilde{p} = 0.2, 0.5$, and 1.0×10^5, where only a few streamlines are seen to penetrate the barriers. For comparison we also consider the flow field obtained when the structure optimized for Stokes flow at $Da = 10^{-5}$ is frozen and exposed to the three different elevated pressure drops. This is shown in Figures 5(d)–(f): As the pressure drop is increased, more and more streamlines penetrate the barriers. Moreover, we find a recirculation region emerging behind the second barrier which reduces the pressure drop over the neck between the barriers.

Returning to Figures 5(a)–(c), we find that the structures that have been optimized for the corresponding pressure drops are generally thicker than that optimized for Stokes flow, which reduces the number of streamlines penetrating them. Also a beak-like tip grows on the second

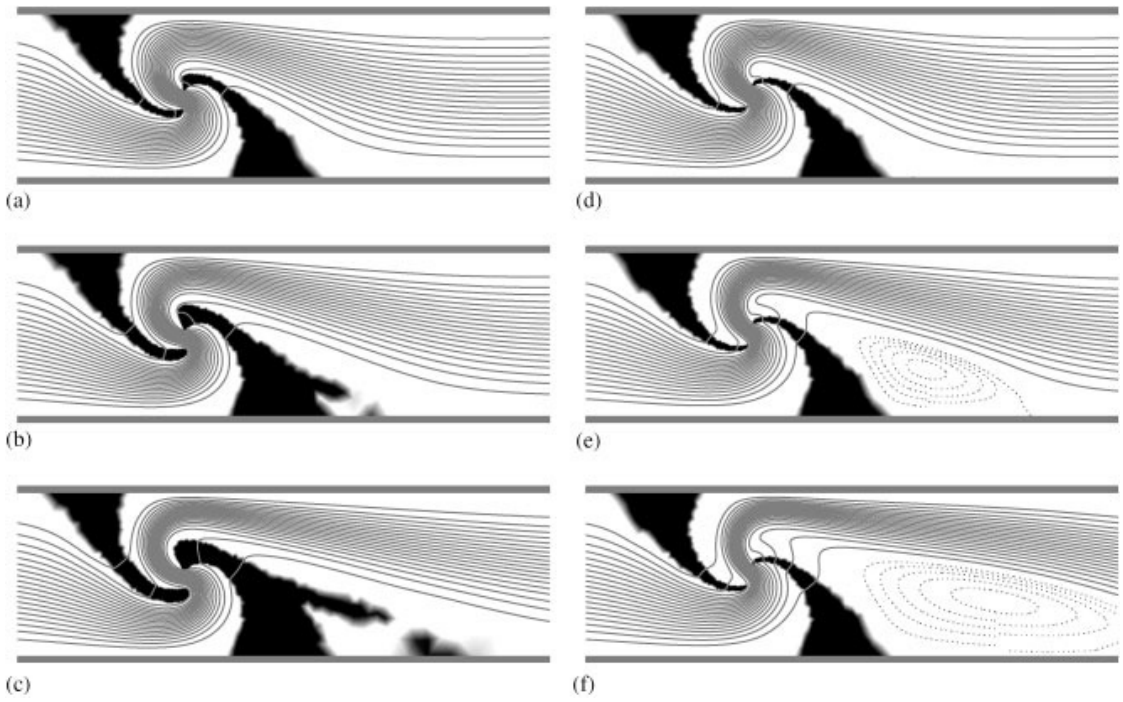


Figure 5. Optimized structures (black) and streamlines for Navier–Stokes flow; only a part of length 3.25ℓ near the centre of the channel is shown. Panels (a)–(c) to the left show the optimized structures for different values of the control parameter $\Delta\tilde{p} = \Delta p \rho \ell^2/\eta^2$. For comparison the flow field when the optimized structure from Figure 2(c) is frozen and exposed to the elevated pressure drops is shown in panels (d)–(f) to the right. The Reynolds number is defined as $Re = \rho \ell v_{\max}/\eta$ where $v_{\max}$ is the maximal velocity measured at the inlet; note that for a particular value of $\Delta\tilde{p}$, the Reynolds number is not fixed but differs slightly between left and right column: (a) $\Delta\tilde{p} = 0.2 \times 10^5$ $[Re = 23]$; (b) $\Delta\tilde{p} = 0.5 \times 10^5$ $[Re = 42]$; (c) $\Delta\tilde{p} = 1.0 \times 10^5$ $[Re = 64]$; (d) $\Delta\tilde{p} = 0.2 \times 10^5$ $[Re = 26]$; (e) $\Delta\tilde{p} = 0.5 \times 10^5$ $[Re = 47]$; and (f) $\Delta\tilde{p} = 1.0 \times 10^5$ $[Re = 71]$.

barrier that acts to bend the fluid stream down. Finally, on the back of the second barrier a wing- or spoiler-like structure appears that removes the recirculation.

In summary, our first example has demonstrated that our implementation of topology optimization works, but that the optimal design and performance may depend strongly on the choice of the Darcy number. In particular, the zero *Da* limit solution contains zero thickness and yet impermeable barriers deflecting the fluid. In order to approximate this solution at finite *Da* and on a finite resolution mesh it is important to choose the Darcy number small enough that even thin barriers can be almost impermeable, but large enough to avoid difficulties with artificial local optima in the discretized problem when the barrier thickness decreases below the mesh resolution.

4.2. Example: a four-terminal device

Our second numerical example deals with minimization of the power dissipation in a four-terminal device subject to a volume constraint. The problem is found to exhibit a discrete

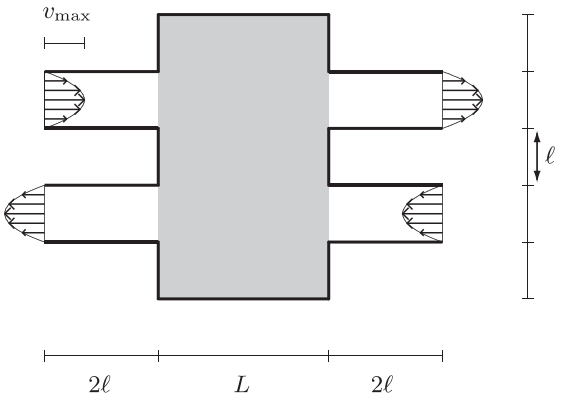


Figure 6. Schematic illustration of the four-terminal device. Two inlet and two outlet leads (white areas) of height ℓ and length 2ℓ are attached to the design domain (grey) of height 5ℓ and length L. The flow is characterized by the Reynolds number $Re = \rho \ell v_{\max}/\eta$, where $v_{\max}$ is the maximal velocity at the inlets.

change in optimal topology driven by the inertial term. The four-terminal device is related to one considered by Borrvall and Petersson for Stokes flow in Reference [6]; the present example demonstrates that the optimization algorithm has difficulties in finding the optimal topology when there are two strong candidates for the global optimum.

The computational domain, shown in Figure 6, consists of a rectangular design domain (grey) to which two inlet and two outlet leads (white) are attached symmetrically. The boundary conditions prescribe parabolic profiles for the flow at the inlets, zero pressure and normal flow at the outlets, and no-slip on all other external boundaries. Choosing the height ℓ of the leads as our characteristic length scale, we define the Reynolds number as $Re = \rho \ell v_{\max}/\eta$, where $v_{\max}$ is the maximal velocity at the inlets. The Darcy number is fixed at $Da = 10^{-4}$ to obtain reasonably small leakage through the porous walls.

The optimization problem is stated as a minimization of the total power dissipation inside the computational domain, given by Equation (7a), subject to the constraint that at most a fraction $\beta = 0.4$ of the design domain should be without porous material, c.f. Equation (15b).

Figures 7(a) and (b) shows the two optimal structures obtained for $Re = 20$ and 200, respectively, in a geometry with $L = 3.5\ell$. At $Re = 20$ the optimal structure turns out to be a pair of U-turns connecting the inlets to the outlets on the same side of the design domain, while at $Re = 200$ the optimal structure is a pair of parallel channels. In order to minimize the power dissipation at low Re, the channel segments should be as short and as wide as possible, which favours the U-turns in Figure 7(a). However, as the Reynolds number is increased, the cost of bending the fluid stream grows. When inertia dominates, larger velocity gradients appear in the long 'outer lane' of the U-turn. This increases the dissipation compared to low Re, where more fluid flows in the shorter 'inner lane'. At a certain point it will exceed the dissipation in the parallel channels solution

$$\Phi_0 = \frac{96}{9}\left(4 + \frac{L}{\ell}\right)\eta v_{\max}^2 \tag{34}$$

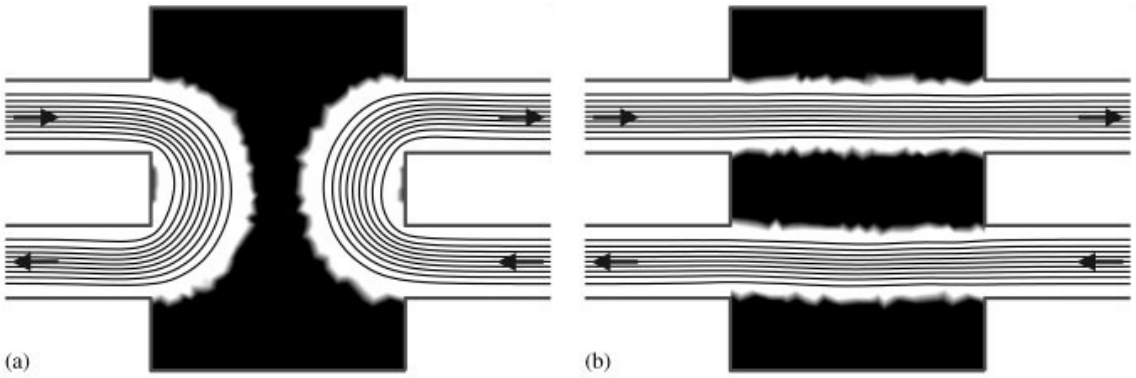


Figure 7. Optimal structures (black) and streamlines at 10% intervals for the four-terminal device at Reynolds number $Re = 20$ and 200, respectively, in a geometry with $L = 3.5\ell$: (a) $Re = 20$; and (b) $Re = 200$.

as estimated from Poiseuille flow in two straight channels, each of length $L + 4\ell$ and height ℓ. This number is independent of inertia due to translation symmetry, and we use Φ_0 as a natural unit of power dissipation (per unit length in the third dimension) in the following.

Clearly the Reynolds number at which the transition between the two classes of solutions occurs will depend strongly on the ratio L/ℓ. For short lengths $L \lesssim 2\ell$ the parallel channels solution is expected to be optimal at all Re, whereas for long lengths $L \gtrsim 3\ell$ the U-turn solution should be significantly better than the parallel channels solution at low Re.

4.2.1. Dependence on the Reynolds number. In the following we investigate more closely the transition between the U-turns and the parallel channels solution as a function of the Reynolds number for the particular geometry $L = 3\ell$. The topology optimization problem is solved for different Re in the range 0–200, using a homogeneous material distribution $\gamma^{(0)} = 0.4$ as initial condition. For the parameter q determining the shape of $\alpha(\gamma)$ in Equation (4) we use a two-step solution procedure as suggested in Reference [1]. First the problem is solved with $q = 0.01$ in order to obtain a solution with slightly smeared-out solid/void interfaces. Next this material distribution is used as initial guess for an optimization with $q = 0.1$ which generates fully discrete solid/void interfaces at the resolution of our finite element mesh.

Figure 8(a) shows the result for the normalized power dissipation Φ/Φ_0 obtained as a function of Re. At low Reynolds numbers the optimized solutions correctly come out as U-turns with a power dissipation Φ that is clearly less than Φ_0. However, at high Reynolds numbers $Re > 90$ the method fails because the optimized solutions continue to come out as U-turns even though this yields $\Phi/\Phi_0 > 1$. For $Re \geqslant 160$ the solution jumps from the simple U-turns to a hybrid structure, as shown in the inset. The full lines in Figure 8(a) show the result when the material distributions optimized for $Re = 0$, 50, and 180, respectively, are frozen and the power dissipation evaluated at different Re. It is seen that the optimized solutions, marked ($\circ$), all fall on or below the full lines which confirms that they are indeed superior to the other solutions of the U-turn family. This also holds for $Re > 90$, except for the hybrid structures at $Re \geqslant 160$, that are actually inferior to the U-turns. Moreover, at $Re = 160$ the optimized solution falls slightly above that optimized at $Re = 180$. This could be an indication

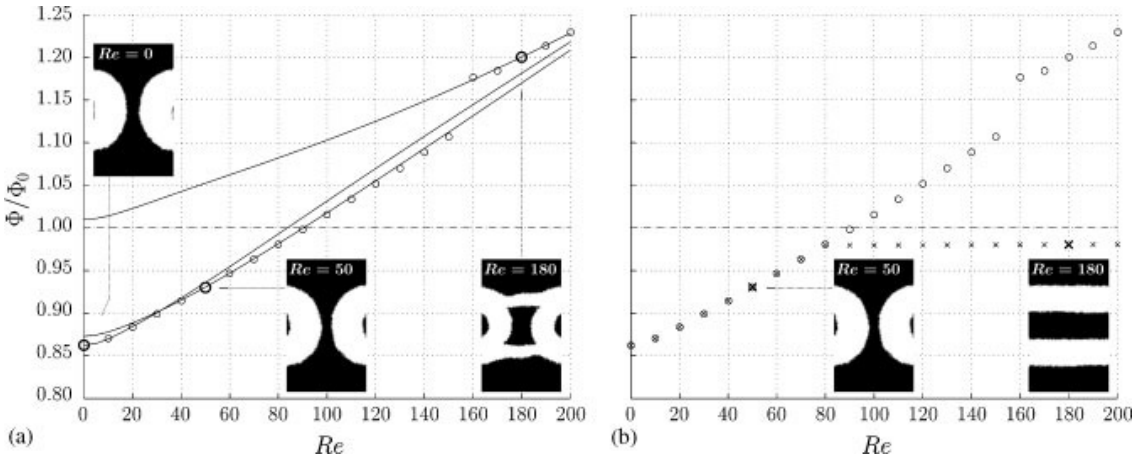


Figure 8. Power dissipation Φ in structures optimized for different Reynolds numbers; normalized with the Poiseuille flow result Φ_0 (dashed line): (a) markers ($\circ$) show results when $\gamma^{(0)} = 0.4$ is used as initial condition, failing to find the optimal solution for $Re > 90$. Full lines show the performance of the structures optimized at $Re = 0$, 50, and 180, when evaluated at different Reynolds numbers. As expected, all points fall on or below the full lines, except the hybrid solutions for $Re \geqslant 160$; and (b) comparison between the two different initial conditions $\gamma^{(0)} = 0.4$ ($\circ$) and $\gamma^{(0)} = 1$ ($\times$), showing the success of the empty channel initial condition in finding the optimal solution. The crosses ($\times$) fall slightly below $\Phi/\Phi_0 = 1$ due to leakage through the porous walls (see the text).

that the hybrid structures are not local optima in design space after all, but rather a very narrow saddle point that the optimization algorithm has a hard time getting away from.

The difficulty is that the two families of solutions, the U-turns and the parallel channels, are both deep local minima for the power dissipation in design space. Using $\gamma^{(0)} = 0.4$ as initial condition, the initial permeability is everywhere very low, such that the porous friction almost completely dominates the inertia and viscous friction in the fluid. Therefore the iteration path in design space is biased towards low Reynolds numbers and the U-turns solution.

In order to circumvent this problem we have tried using a completely empty design domain with $\gamma^{(0)} = 1$ as initial condition. This should remove the bias towards the U-turns and allow the optimization algorithm to take inertia into account from iteration one. The result is shown in Figure 8(b). For $Re \leqslant 80$ the solutions are still U-turns, whereas for $Re \geqslant 90$ they come out as parallel channels. Notice that Φ/Φ_0 for the parallel channels solution is actually slightly less than unity, namely 0.98. This is due to a small amount of fluid seeping through the porous walls defining the device, which lowers the hydraulic resistance compared to the Poiseuille flow result derived for solid walls.[‖]

Strictly speaking the initial condition $\gamma^{(0)} = 1$ is not a feasible solution because it violates the volume constraint that at least a fraction $1 - \beta = 0.6$ of the design domain should be filled

[‖]The flow in a straight channel of height ℓ bounded by two porous walls of thickness ℓ can easily be found analytically. At $Da = 10^{-4}$ the hydraulic resistance of this system is 94% of that for a channel of height ℓ bounded by solid walls, and it approaches this zero Da limit only as $\sqrt{Da}$. When $L = 3\ell$ we therefore expect a power dissipation $\Phi/\Phi_0 = (3 \times 0.94 + 4)/7 = 0.97$ for the parallel channels solution, including the leads. This is close to the observed 0.98.

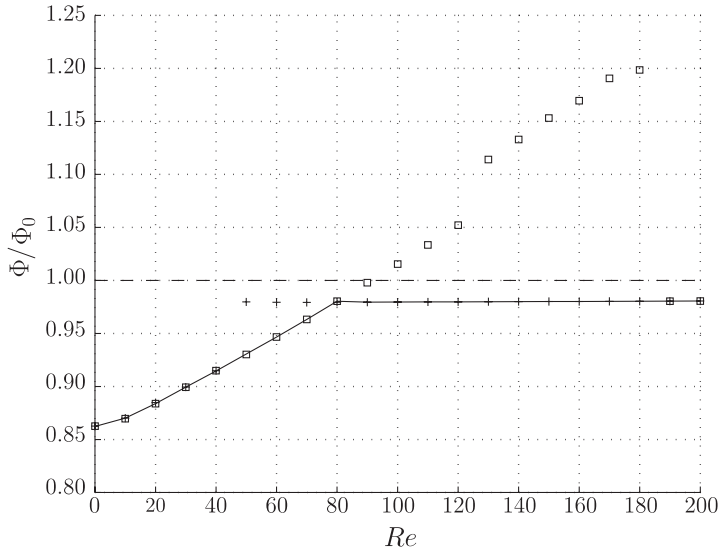


Figure 9. Comparison between structures optimized with the initially non-feasible material distribution $\gamma^{(0)} = 1$ for different penalty parameters in the MMA optimization algorithm, revealing the difficulty in choosing the condition for finding the global optimum. Full line: the successful result from Figure 8(b) with moderate penalty; (+) lower penalty yielding wrong result for $40 < Re < 80$; (□) higher penalty yielding wrong results for $80 < Re < 190$.

with porous material. However, the MMA optimization algorithm penalizes this and reaches a feasible solution after a few iterations. This is controlled by choosing a penalty parameter. If the penalty for violating the constraint is small, the material is added slowly and only where it does not disturb the flow much. If the penalty is large, the material is added quickly and almost homogeneously until the constraint is satisfied. The successful result from Figure 8(b) was obtained with a moderate penalty. In Figure 9 this is compared with results for smaller and larger penalty parameters, respectively. The figure shows that with the small penalization, the solution jumps to the parallel channels already at $Re = 50$ which is not optimal. For the large penalization, the solution does not jump until $Re \geqslant 190$. Also we observe hybrid structures similar to those in Figure 8(a) for $Re \geqslant 130$. We have thus not full control over the convergence towards the global optimum.

4.2.2. Discussion of problems with local optima. Further insight into the problem of local versus global optima is gained by inspecting the flow field in the initial material distribution $\gamma^{(0)}$. This is shown in Figure 10 for the Stokes flow limit, $Re = 0$. The streamlines are drawn as 10% contours of the streamfunction, and Figure 10(a) shows that for $\gamma^{(0)} = 0.4$ the streamline density is largest between the two leads on the same side of the design domain. Based on the sensitivity $\mathrm{d}\Phi/\mathrm{d}\gamma$ the optimization algorithm therefore decides to remove material from these strong-flow regions in order to reduce the porous friction. The iteration path in design space is therefore biased towards the U-turn solution. This remains true even at finite Reynolds numbers as long as the porous friction initially dominates inertia.

Figure 10(b) shows that when $\gamma^{(0)} = 1$ the streamline density is largest between the leads on the opposite side of the design domain. Because the volume constraint is violated the

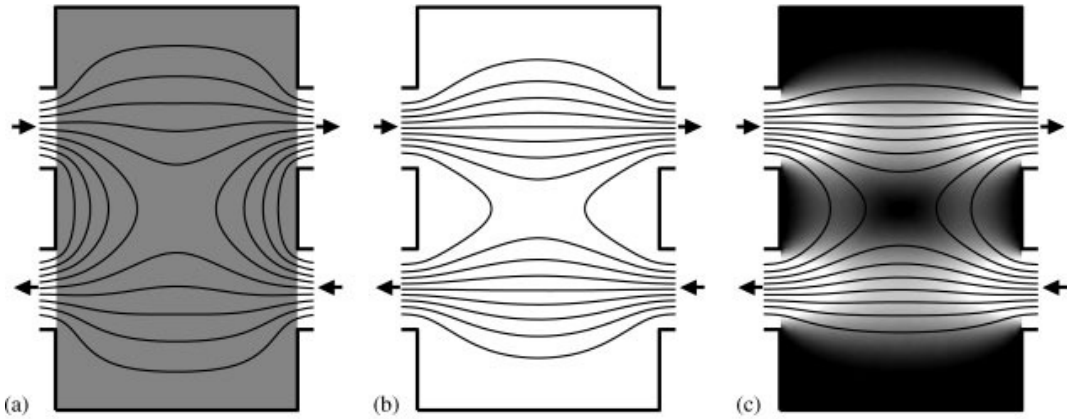


Figure 10. Flow distributions at $Re=0$, $L=3\ell$ and $q=0.01$: (a) initial design field $\gamma^{(0)}=0.4$; (b) initial design field $\gamma^{(0)}=0.4$; and (c) the optimal design field $\gamma^{(*)}$ obtained at $Da=10^{-2}$.

optimization algorithm has to place material somewhere, which it does in the weak-flow regions. The solution is therefore biased towards the parallel channels. Indeed if the penalty is chosen very small, the optimized solution comes out as parallel channels even for Stokes flow at $L=3\ell$, which is far from optimal. When the penalty is larger and the material is added faster, we move away from this adiabatic solution and closer to the situation for $\gamma^{(0)}=0.4$.

The additional complexity associated with making a proper choice of the penalty parameter is somewhat inconvenient. We have therefore attempted to construct a more convex problem by increasing the initial permeability. This can be done either by increasing the Darcy number, or by decreasing the parameter q, c.f. Equation (4). Figure 10(c) shows the optimal solution $\gamma^{(*)}$ obtained for $Da=10^{-2}$ and $q=0.01$ at $Re=0$. At this level the problem is convex because the solution is independent of the initial condition. Using this material distribution as initial guess and gradually decreasing the permeability to $Da=10^{-4}$ and $q=0.1$ we correctly end up in the U-turn solution. However, it is evident from Figure 10(c) that $\gamma^{(*)}$ has a fair amount of parallel channels nature. Using the same procedure of gradually decreasing the permeability at higher Reynolds numbers therefore result in a transition to the parallel channels solution already for $Re\geqslant 30$, which is not optimal. Moreover, when the Reynolds number is increased and the inertia starts to play in, the system tends to loose convexity even at the initial high permeability.

In summary, the topology optimization has difficulties in finding the global optimum for the problem. There are two strong candidates for the optimal structure, and the solution found is sensitive to the initial condition for the material distribution. Using an empty channel as initial condition, the method is able to find the correct solution for all Reynolds numbers. However, this successful result depends on a particular choice of the penalty parameter in the MMA algorithm. By using a high initial permeability of the porous medium, it is possible to convexify the problem at low Reynolds numbers, but continuation of this solution to the desired low permeability does not generally lead to the global minimum of the non-convex problem.

In the original paper Reference [6] it was argued that in Stokes flow the true optimal design should be rather insensitive to the choice of the Darcy number, although the dissipated power

may deviate quite a lot from the zero Da limit. In our work we have observed that the actual solution found by the topology optimization may depend a great deal on the choice of the Darcy number, whereas the dissipated power should approach the zero Da limit roughly as $\sqrt{Da}$.

5. CONCLUSION

Based on the work of Borrvall and Petersson [6] we have extended the topology optimization of fluid networks to cover the full incompressible Navier–Stokes equations in steady-state. Our implementation of the method is based on the commercial finite element package FEMLAB, which reduces the programming effort required to a minimum. Formulating the problem in terms of a general integral-type objective function and expressing the governing equations in divergence form makes the implementation very compact and transparent. Moreover, the code for performing the sensitivity analysis should remain almost the same for any problem expressed in this way, whereas that required for describing the physical problem of course changes. Topology optimization of multi-field problems can therefore be dealt with almost as easy as a single realization of the underlying physical problem.

We would like to mention that our methodology is not as such restricted to the (large) class of physical problems that can be expressed in divergence form. FEMLAB also allows problems to be stated directly in weak form, e.g. for systems with dynamics at the boundaries. This does in fact not invalidate the sensitivity analysis worked out in Section 3.4, since this analysis only relies on the basic structure of the discretized non-linear problem and the availability of the Jacobian matrix. It is therefore possible to apply our methodology to even larger classes of physical problems than the ones comprised by the divergence form.

Our implementation of topology optimization has been tested on two fluidics examples in 2D, both illustrating the influence of different quantities and conditions on the efficiency of the optimization method.

The first example, a channel with reversed flow, illustrates the influence of the Reynolds number Re and the Darcy number Da on the solutions. We have shown that the choice of Da has a strong impact on the solution when the structure contains barriers to deflect the fluid stream.

The second example, minimization of the power dissipation in a four-terminal device, reveals the problems of determining the global minimum when two strong minima are competing. This problem is highly non-convex, and we have shown that the solution depends on the initial condition. For an initial homogeneous material distribution, the porous friction dominates and the solution does not come out as the global optimum in all cases. Using an empty channel as the initial state, inertia plays a role from the beginning, and better results can be obtained. However, this initial condition in fact violates the volume constraint, and the part of the optimization routine correcting this depends on a penalty factor. Unfortunately, the particular value chosen for this factor strongly influences the results. Increasing the Darcy number makes the problem more convex, but continuation from large to small Da, i.e. from high to low permeability of the porous material, does not generally end up in the global optimum.

In conclusion, we have shown that our implementation of topology optimization is a useful tool for designing fluidic devices.

APPENDIX

```
% FEMLAB CODE FOR THE 4-TERMINAL DEVICE EXAMPLE OF SECTION 4.2
clear fem femadj
% DEFINE REYNOLDS NUMBER, DARCY NUMBER, LENGTH OF DESIGN DOMAIN, AND VOLUME FRACTION
Re = 50;
Da = 1e-4;
L0 = 3.0;
beta = 0.4;
% DEFINE GEOMETRY, MESH, AND SUBDOMAIN/BOUNDARY GROUPS [SEE FIGURE 6]
fem.geom = rect2(0,L0,0,5) + rect2(-2,0,1,2) + rect2(-2,0,3,4) + rect2(L0,L0+2,1,2) ...
   + rect2(L0,L0+2,3,4);
fem.mesh = meshinit(fem,'Hmaxsub',[3 0.125]);
% subdomain groups  1:design domain   2:inlet/outlet leads
fem.equ.ind = {[3] [1 2 4 5]};
% boundary groups  1:walls   2:inlets   3:outlets   4:interior
fem.bnd.ind = {[2:3 5:8 10 12:14 16:18 20:22] [4 23] [1 24] [9 11 15 19]};
% DEFINE SPACE CO-ORDINATES, DEPENDENT VARIABLES, AND SHAPE FUNCTIONS
fem.sdim = {'x' 'y'};
fem.dim = {'u' 'v' 'p' 'gamma'};
fem.shape = [2 2 1 1];
% DEFINE CONSTANTS
fem.const.rho = 1;
fem.const.eta = 1;
fem.const.umax = Re;
fem.const.alphamin = 0;
fem.const.alphamax = 1/Da;
fem.const.q = 0.1;
Phi0 = 96*fem.const.eta*(L0+4)*fem.const.umax^2/9;
% DEFINE EXPRESSIONS ON SUBDOMAIN AND BOUNDARY GROUPS
fem.equ.expr = {'A' 'eta*(2*ux*ux+2*vy*vy+(uy+vx)*(uy+vx))+alpha*(u*u+v*v)' ...
   'alpha' {'alphamin+(alphamax-alphamin)*q*(1-gamma)/(q+gamma)' '0'}};
fem.bnd.expr = {'B' '0'};

% DEFINE GOVERNING EQUATIONS AND INITIAL CONDITIONS [SEE EQUATIONS (8) AND (9)]
fem.form = 'general';
fem.equ.shape = {[1:4] [1:3]};           % only define gamma on subdomain group 1
fem.equ.ga = {{{'-p+2*eta*ux' 'eta*(uy+vx)'} {'eta*(uy+vx)' '-p+2*eta*vy'} {0 0} {0 0}}};
fem.equ.f = {{'rho*(u*ux+v*uy)+alpha*u' 'rho*(u*vx+v*vy)+alpha*v' 'ux+vy' 1}};
fem.equ.init = {{0 0 0 beta}};
% DEFINE BOUNDARY CONDITIONS
fem.bnd.shape = {[1:3]};                  % do not define gamma on any boundaries
fem.bnd.r = {{'u' 'v' 0 0} ...            % walls:      no-slip
   {'u*nx+4*umax*s*(1-s)' 'v' 0 0} ...  % inlets:    parabolic profile
   {0 'v' 0 0} ...                        % outlets:   normal flow
   {0 0 0 0}};                            % interior:  nothing
fem.bnd.g = {{0 0 0 0}};                  % zero prescribed external forces everywhere
% PERFORM LINEARIZATION, DEGREE-OF-FREEDOM ASSIGNMENT, AND ASSEMBLE INITIAL CONDITION
fem = femdiff(fem);
fem.xmesh = meshextend(fem);
fem.sol = asseminit(fem);

% DEFINE STRUCTURE FOR COMPUTING RIGHT-HAND-SIDE IN ADJOINT PROBLEM [SEE EQUATION (29)]
femadj = fem;
femadj.equ.ga = {{{'diff(A,ux)' 'diff(A,uy)'} {'diff(A,vx)' 'diff(A,vy)'} ...
   {'diff(A,px)' 'diff(A,py)'} {'diff(A,gammax)' 'diff(A,gammay)'}}};
femadj.equ.f = {{'diff(A,u)' 'diff(A,v)' 'diff(A,p)' 'diff(A,gamma)'}};
femadj.bnd.g = {{'diff(B,u)' 'diff(B,v)' 'diff(B,p)' 'diff(B,gamma)'}};
femadj.xmesh = meshextend(femadj);
% GET INDICES OF DESIGN VARIABLE IN THE GLOBAL SOLUTION VECTOR (fem.sol.u)
i4 = find(asseminit(fem,'Init',{'gamma' 1},'Out','U'));
% COMPUTE VOLUME BELOW DESIGN VARIABLE BASIS FUNCTIONS
L = assemble(fem,'Out',{'L'});
Vgamma = L(i4);
Vdomain = sum(Vgamma);
% GET INDICES OF VELOCITY-PRESSURE VARIABLES
i123 = find(asseminit(fem,'Init',{'u' 1 'v' 1 'p' 1},'Out','U'));

% DEFINE VARIABLES AND PARAMETERS FOR MMA OPTIMIZATION ALGORITHM [SEE REFERENCES [11,12,13]]
a0 = 1;
a = 0;
c = 20;
d = 0;
xmin = 0;
xmax = 1;
xold = fem.sol.u(i4);
xolder = xold;
```



```
low = 0;
upp = 1;

% DESIGN LOOP FOR THE ACTUAL TOPOLOGY OPTIMIZATION
for iter = 1:100
    % SOLVE NAVIER-STOKES FLOW PROBLEM TO UPDATE VELOCITY AND PRESSURE
    fem.sol = femnlin(fem,'Solcomp',{'u' 'v' 'p'},'U',fem.sol.u);
    % SOLVE ADJOINT PROBLEM FOR LAGRANGE MULTIPLIERS
    [K N] = assemble(fem,'Out',{'K' 'N'},'U',fem.sol.u);
    [L M] = assemble(femadj,'Out',{'L' 'M'},'U',fem.sol.u);
    femadj.sol = femlin('In',{'K' K(i123,i123)' 'L' L(i123) 'M' zeros(size(M)) 'N' N(:,i123)});
    % SENSITIVITY ANALYSIS
    gamma = fem.sol.u(i4);
    Phi = postint(fem,'A','Edim',2) + postint(fem,'B','Edim',1);
    dPhidgamma = L(i4) - K(i123,i4)'*femadj.sol.u;
    % PERFORM MMA STEP TO UPDATE DESIGN FIELD
    x = gamma;
    f = Phi/Phi0;                  g = gamma'*Vgamma/Vdomain - beta;
    dfdx = dPhidgamma/Phi0;       dgdx = Vgamma'/Vdomain;
    d2fdx2 = zeros(size(gamma)); d2gdx2 = zeros(size(gamma'));
    [xnew,y,z,lambda,ksi,eta,mu,zeta,s,low,upp] = mmasub(1,length(gamma),iter, ...
       x,xmin,xmax,xold,xolder,f,dfdx,d2fdx2,g,dgdx,d2gdx2,low,upp,a0,a,c,d);
    xolder = xold; xold = x; gamma = xnew;
    % TEST CONVERGENCE
    if iter >= 100 | max(abs(gamma-xold)) < 0.01
       break
    end
    % UPDATE DESIGN VARIABLE
    u0 = fem.sol.u; u0(i4) = gamma;
    fem.sol = femsol(u0);
    % DISPLAY RESULTS FOR EACH ITERATION STEP
    disp(sprintf('Iter.:%3d   Obj.: %8.4f   Vol.: %6.3f   Change: %6.3f', ...
       iter,f,xold'*Vgamma,max(abs(xnew-xold))))
    postplot(fem,'arrowdata',{'u' 'v'},'tridata','gamma','trimap','gray')
    axis equal; shg; pause(0.1)
end
```

ACKNOWLEDGEMENTS

We are grateful to Ole Sigmund and Allan Gersborg-Hansen for illuminating discussions on the topology optimization method, and to Krister Svanberg for providing us with the MATLAB code for the MMA algorithm. This work is partly supported by The Danish Technical Research Council, Grant No. 26-03-0037.

REFERENCES

1. Bendsøe MP, Kikuchi N. Generating optimal topologies in structural design using a homogenization method. *Computer Methods in Applied Mechanics and Engineering* 1988; **71**(2):197–224.
2. Bendsøe MP, Sigmund O. *Topology Optimization-Theory, Methods and Applications*. Springer: Berlin, 2003.
3. Eschenauer HA, Olhoff N. Topology optimization of continuum structures: a review. *Applied Mechanics Reviews* 2001; **54**(4):331–389.
4. Jensen JS. Phononic band gaps and vibrations in one- and two-dimensional mass-spring structures. *Journal of Sound and Vibration* 2003; **266**(5):1053–1078.
5. Jensen JS, Sigmund O. Systematic design of photonic crystal structures using topology optimization: low-loss waveguide bends. *Applied Physics Letters* 2004; **84**(12):2022–2024.
6. Borrvall T, Petersson J. Topology optimization of fluids in Stokes flow. *International Journal for Numerical Methods in Fluids* 2003; **41**:77–107.
7. Sigmund O, Gersborg-Hansen A, Haber RB. Topology optimization for multiphysics problems: a future FEMLAB application? In *Proceedings for the Nordic Matlab Conference* 2003, Gregersen L (ed.). Comsol A/S: Søborg, Denmark, 2003; 237–242.
8. Gersborg-Hansen A. Topology optimization of incompressible Newtonian flows at moderate Reynolds numbers. *M.Sc. Thesis*, Technical University of Denmark, Department of Mechanical Engineering, Lyngby, 2003.
9. Gersborg-Hansen A, Sigmund O, Haber RB. Topology optimization of channel flow problems. *Structural and Multidisciplinary Optimization* 2005. (Published online: DOI 10.1007/s00158-004-0508-7.)

10. Okkels F, Olesen LH, Bruus H. Applications of topology optimization in the design of micro- and nanofluidic systems. In *Proceedings of Nanotech*, vol. 1. Anaheim: CA, U.S.A., 2005; 575–578.
11. Svanberg K. The method of moving asymptotes—a new method for structural optimization. *International Journal for Numerical Methods in Engineering* 1987; **24**:359–373.
12. Svanberg K. A class of globally convergent optimization methods based on conservative convex separable approximations. *SIAM Journal on Optimization* 2002; **12**(2):555–573.
13. A MATLAB implementation, `mmasub`, of the MMA optimization algorithm [11] can be obtained (free of charge for academic purposes) from Krister Svanberg, KTH, Sweden. E-mail: krille@math.kth.se
14. Geschke O, Klank H, Telleman P (eds). *Microsystem Engineering of Lab-on-a-chip Devices*. Wiley-VCH Verlag: Weinheim, 2004.
15. Landau LD, Lifshitz EM. *Course of Theoretical Physics*: *Fluid Mechanics*, vol. 6 (2nd edn). Butterworth and Heinemann: Oxford, 2000.
16. Michaleris P, Tortorelli DA, Vidal CA. Tangent operators and design sensitivity formulations for transient non-linear coupled problems with applications in elastoplasticity. *International Journal for Numerical Methods in Engineering* 1994; **37**(14):2471–2500.
17. FEMLAB Reference Manual. COMSOL AB: Stockholm, 2004.
18. A simple implementation of a cone filter for dealing with mesh dependence problems on an unstructured triangular mesh in two dimensions with linear Lagrange elements for the design variable can be downloaded from http://www.mic.dtu.dk/research/MIFTS

Bibliography

[1] A. Manz, N. Graber, and H. M. Widmer, 'Miniaturized total chemical analysis systems. A novel concept for chemical sensing', Sens. Actuator B-Chem. **1**, 244 (1990).

[2] D. R. Reyes, D. Iossifidis, P. A. Auroux, and A. Manz, 'Micro total analysis systems. 1. Introduction, theory, and technology', Anal. Chem. **74**, 2623 (2002).

[3] H. A. Stone, A. D. Stroock, and A. Ajdari, 'Engineering flows in small devices: Microfluidics toward a lab-on-a-chip', Annu. Rev. Fluid Mech. **36**, 381 (2004).

[4] T. M. Squires and S. R. Quake, 'Microfluidics: Fluid physics at the nanoliter scale', Rev. Mod. Phys. **77**, 977 (2005).

[5] M. A. Unger, H.-P. Chou, T. Thorsen, A. Scherer, and S. R. Quake, 'Monolithic microfabricated valves and pumps by multilayer soft lithography', Science **288**, 113 (2000).

[6] D. J. Laser and J. G. Santiago, 'A review of micropumps', J. Micromech. Microeng. **14**, R35 (2004).

[7] S. H. Yao and J. G. Santiago, 'Porous glass electroosmotic pumps: theory', J. Colloid Interface Sci. **268**, 133 (2003).

[8] S. H. Yao, D. E. Hertzog, S. L. Zeng, J. C. Mikkelsen, and J. G. Santiago, 'Porous glass electroosmotic pumps: design and experiments', J. Colloid Interface Sci. **268**, 143 (2003).

[9] S. L. Zeng, C.-H. Chen, J. C. Mikkelsen, and J. G. Santiago, 'Fabrication and characterization of electroosmotic micropumps', Sens. Actuator B-Chem. **79**, 107 (2001).

[10] D. S. Reichmuth, G. S. Chirica, and B. J. Kirby, 'Increasing the performance of high-pressure, high-efficiency electrokinetic micropumps using zwitterionic solute additives', Sens. Actuator B-Chem. **92**, 37 (2003).

[11] A. Brask, J. P. Kutter, and H. Bruus, 'Long-term stable electroosmotic pump with ion exchange membranes', Lab Chip **5**, 730 (2005).

[12] A. Brask, 'Electroosmotic micropumps', PhD thesis, MIC – Dept. of Micro and Nanotechnology, DTU (2005). `http://www.mic.dtu.dk/mifts`

[13] Y. Takamura, H. Onoda, H. Inokuchi, S. Adachi, A. Oki, and Y. Horiike, 'Low-voltage electroosmosis pump for stand-alone microfluidics devices', Electrophoresis **24**, 185 (2003).

[14] A. Ajdari, 'Pumping liquids using asymmetric electrode arrays', Phys. Rev. E **61**, R45 (2000).

[15] V. Studer, A. Pépin, Y. Chen, and A. Ajdari, 'An integrated AC electrokinetic pump in a microfluidic loop for fast and tunable flow control', Analyst **129**, 944 (2004).

[16] A. Bard and L. Faulkner, *Electrochemical Methods* (Wiley, New York, 1980).

[17] R. J. Hunter, *Foundations of Colloid Science* (Oxford University Press, New York, 1991).

[18] V. N. Shilov and S. S. Dukhin, 'Theory of low-frequency dispersion of dielectric permittivity in suspensions of spherical colloidal particles due to double-layer polarization', Colloid J. USSR **32**, 245 (1970).

[19] N. I. Gamayunov, V. A. Murtsovkin, and A. S. Dukhin, 'Pair interaction of particles in electric-field. 1. Features of hydrodynamic interaction of polarized particles', Colloid J. USSR **48**, 197 (1986).

[20] N. I. Gamayunov and V. A. Murtsovkin, 'Motion of disperse particles in a uniform alternating electric field', Colloid J. USSR **49**, 543 (1987).

[21] T. M. Squires and M. Z. Bazant, 'Breaking symmetries in induced-charge electro-osmosis', J. Fluid Mech. **560**, 65 (2006).

[22] A. Ramos, H. Morgan, N. G. Green, and A. Castellanos, 'Ac electrokinetics: a review of forces in microelectrode structures', J. Phys. D Appl. Phys. **31**, 2338 (1998).

[23] A. González, A. Ramos, N. G. Green, A. Castellanos, and H. Morgan, 'Fluid flow induced by nonuniform ac electric fields in electrolytes on microelectrodes. II. A linear double-layer analysis', Phys. Rev. E **61**, 4019 (2000).

[24] N. G. Green, A. Ramos, A. González, H. Morgan, and A. Castellanos, 'Fluid flow induced by nonuniform ac electric fields in electrolytes on microelectrodes. I. Experimental measurements', Phys. Rev. E **61**, 4011 (2000).

[25] A. B. D. Brown, C. G. Smith, and A. R. Rennie, 'Pumping of water with ac electric fields applied to asymmetric pairs of microelectrodes', Phys. Rev. E **63**, 016305 (2000).

[26] V. Studer, A. Pépin, C. Y, and A. Ajdari, 'Fabrication of microfluidic devices for AC electrokinetic fluid pumping', Microelectron. Eng. **61**, 915 (2002).

[27] S. Debesset, C. J. Hayden, C. Dalton, J. C. T. Eijkel, and A. Manz, 'A circular AC electroosmotic micropump for chromatographic applications', in 'Proc. MicroTAS 2002', 655 – 657 (Kluwer, 2002).

[28] M. Mpholo, C. G. Smith, and A. B. D. Brown, 'Low voltage plug flow pumping using anisotropic electrode arrays', Sens. Actuator B-Chem. **92**, 262 (2003).

[29] S. Debesset, C. J. Hayden, C. Dalton, J. C. T. Eijkel, and A. Manz, 'An AC electroosmotic micropump for circular chromatographic applications', Lab Chip **4**, 396 (2004).

[30] P. Garcia, A. Ramos, N. G. Green, and H. Morgan, 'Experiments on AC electrokinetic pumping of liquids using arrays of microelectrodes', in 'Proc. IEEE Int. Conf. Dielectric Liquids', 181 – 184 (IEEE, 2005).

[31] P. Garcia, A. Ramos, N. G. Green, and H. Morgan, 'Experiments on AC electrokinetic pumping of liquids using arrays of microelectrodes', IEEE Trans. Dielect. El. In. **13**, 670 (2006).

[32] M. Gregersen, 'AC asymmetric electrode micropumps', MSc thesis, MIC – Dept. of Micro and Nanotechnology, DTU (2005). `http://www.mic.dtu.dk/mifts`

[33] A. Ramos, A. González, A. Castellanos, N. G. Green, and H. Morgan, 'Pumping of liquids with ac voltages applied to asymmetric pairs of microelectrodes', Phys. Rev. E **67**, 056302 (2003).

[34] A. Ramos, H. Morgan, N. G. Green, A. Gonzalez, and A. Castellanos, 'Pumping of liquids with traveling-wave electroosmosis', J. Appl. Phys. **97**, 084906 (2005).

[35] D. Lastochkin, R. H. Zhou, P. Wang, Y. X. Ben, and H. C. Chang, 'Electrokinetic micropump and micromixer design based on ac faradaic polarization', J. Appl. Phys. **96**, 1730 (2004).

[36] T. M. Squires and M. Z. Bazant, 'Induced-charge electro-osmosis', J. Fluid Mech. **509**, 217 (2004).

[37] M. Z. Bazant, K. Thornton, and A. Ajdari, 'Diffuse-charge dynamics in electrochemical systems', Phys. Rev. E **70**, 021506 (2004).

[38] K. T. Chu, 'Asymptotic analysis of extreme electrochemical transport', PhD thesis, Dept. Mathematics, Massachusetts Institute of Technology (2005). `http://www.princeton.edu/~ktchu`

[39] K. T. Chu and M. Z. Bazant, 'Nonlinear electrochemical relaxation around conductors', Phys. Rev. E **74**, 011501 (2006).

[40] I. Rubinstein and L. Shtilman, 'Voltage against current curves of cation-exchange membranes', J. Chem. Soc. Faraday Trans. 2 **75**, 231 (1979).

[41] S. S. Dukhin, 'Electrokinetic phenomena of the 2nd kind and their applications', Adv. Colloid Interface Sci. **35**, 173 (1991).

[42] I. Rubinstein and B. Zaltzman, 'Electro-osmotic slip of the second kind and instability in concentration polarization at electrodialysis membranes', Math. Mod. Meth. Appl. S. **11**, 263 (2001).

[43] I. Rubinstein, B. Zaltzman, and I. Lerman, 'Electroconvective instability in concentration polarization and nonequilibrium electro-osmotic slip', Phys. Rev. E **72**, 011505 (2005).

[44] L. H. Olesen, H. Bruus, and A. Ajdari, 'AC electrokinetic micropumps: The effect of geometrical confinement, Faradaic current injection, and nonlinear surface capacitance', Phys. Rev. E **73**, 056313 (2006).

[45] M. Z. Bazant and Y. Ben, 'Theoretical prediction of fast 3D AC electro-osmotic pumps', Lab Chip **6**, 1455 (2006).

[46] L. H. Olesen, F. Okkels, and H. Bruus, 'A high-level programming-language implementation of topology optimization applied to steady-state Navier-Stokes flow', Int. J. Num. Meth. Eng. **65**, 975 (2006).

[47] COMSOL Multiphysics 3.2, the general purpose finite element software formerly known as FEMLAB, by COMSOL A/B, Sweeden.

[48] N. A. Mortensen, L. H. Olesen, L. Belmon, and H. Bruus, 'Electrohydrodynamics of binary electrolytes driven by modulated surface potentials', Phys. Rev. E **71**, 056306 (2005).

[49] N. A. Mortensen, L. H. Olesen, and H. Bruus, 'Transport coefficients for electrolytes in arbitrarily shaped nano and micro-fluidic channels', New J. Phys. **8**, 37 (2006).

[50] L. H. Olesen, F. Okkels, and H. Bruus, 'Topology optimization of Navier–Stokes flow in microfluidics', in 'Proc. ECCOMAS 2004', volume 2, 224 (2004).

[51] F. Okkels, L. H. Olesen, and H. Bruus, 'Applications of topology optimization in the design of micro- and nanofluidic devices', in 'Proc. Nanotech 2005', volume 1, 575–578 (2005).

[52] L. H. Olesen, H. Bruus, and A. Ajdari, 'AC electrokinetic micropumps: The effect of geometrical confinement, Faradaic current injection, and nonlinear surface capacitance', in 'Proc. MicroTAS 2005', volume 1, 560–562 (2005).

[53] R. F. Probstein, *PhysicoChemical Hydrodynamics, an introduction* (John Wiley and Sons, New-York, 1994).

[54] J. Newman, *Electrochemical systems* (Prentice-Hall, Inc., Eaglewood Cliffs, New Jersey, 1991).

[55] I. Borukhov, D. Andelman, and H. Orland, 'Steric effects in electrolytes: A modified Poisson–Boltzmann equation', Phys. Rev. Lett. **79**, 435 (1997).

[56] M. S. Kilic, M. Z. Bazant, and A. Ajdari, 'Steric effects in the dynamics of electrolytes at large applied voltages: I. Double-layer charging', Phys. Rev. E **75**, 021502 (2007).

[57] M. S. Kilic, M. Z. Bazant, and A. Ajdari, 'Steric effects in the dynamics of electrolytes at large applied voltages: II. Modified Poisson–Nernst–Planck equations', Phys. Rev. E **75**, 021503 (2007).

[58] S. S. Dukhin, 'Nonequilibrium electric surface phenomena', Adv. Colloid Interface Sci. **44**, 1 (1993).

[59] M. V. Dyke, *Perturbation methods in fluid mechanics* (The Parabolic Press, Stanford, 1975).

[60] W. H. Smyrl and J. Newman, 'Double layer structure at limiting current', Trans. Faraday Soc. **63**, 207 (1967).

[61] E. J. Hinch, J. D. Sherwood, W. C. Chew, and P. N. Sen, 'Dielectric response of a dilute suspension of spheres with thin double-layers in an asymmetric electrolyte', J. Chem. Soc. Faraday Trans. 2 **80**, 535 (1984).

[62] E. H. B. DeLacey and L. R. White, 'The polarization impedance on an ideally polarizable plane electrode', J. Chem. Soc. Faraday Trans. 2 **78**, 457 (1982).

[63] V. N. Shilov and S. S. Dukhin, 'Theory of polarization of diffuse part of a thin double layer at a spherical particle in an alternating electric field', Colloid J. USSR **32**, 90 (1970).

[64] S. S. Dukhin and V. N. Shilov, 'Theory of static polarization of diffuse part of thin electric double layer of spherical particles', Colloid J. USSR **31**, 564 (1969).

[65] B. V. Deryagin and S. S. Dukhin, 'Theory of surface conductance', Colloid J. USSR **31**, 277 (1969).

[66] K. T. Chu and M. Z. Bazant, 'Electrochemical thin films at and above the classical limiting current', SIAM J. Appl. Math. **65**, 1485 (2005).

[67] A. Gonzalez, A. Ramos, and A. Castellanos, 'Pumping of liquids using travelling wave electroosmosis: a nonlinear analysis', in 'Proc. IEEE Int. Conf. Dielectric Liquids', 221 – 224 (IEEE, 2005).

[68] L. N. Trefethen, *Spectral methods in matlab* (SIAM, Philadelphia, 2000).

[69] J. P. Urbanski, J. A. Levitan, M. Z. Bazant, and T. Thorsen, 'Fast AC electroosmotic pumps with non-planar electrodes', Appl. Phys. Lett. **89**, 143508 (2006).

[70] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C, the art of scientific computing* (Cambridge University Press, New-York, 1992).

[71] K. Svanberg, 'The method of moving asymptotes - a new method for structural optimization', Int. J. Numer. Meth. Eng. **24**, 359 (1987).

[72] K. Svanberg, 'A Class of Globally Convergent Optimization Methods Based on Conservative Convex Separable Approximations', SIAM J. Optim. **12**, 555 (2002).

[73] A MATLAB implementation, `mmasub`, of the MMA optimization algorithm can be obtained (free of charge for academic purposes) from Krister Svanberg, KTH, Sweden. Email: `krille@math.kth.se`.

[74] M. P. Bendsøe and N. Kikuchi, 'Generating optimal topologies in structural design using a homogenization method', Comp. Meth. Appl. Mech. Engng. **71**, 197 (1988).

[75] M. P. Bendsøe and O. Sigmund, *Topology Optimization: Theory, Methods and Applications* (Springer, 2003).

[76] H. A. Eschenauer and N. Olhoff, 'Topology optimization of continuum structures: a review', Appl. Mech. Rev. **54**, 331 (2001).

[77] J. S. Jensen, 'Phononic band gaps and vibrations in one- and two-dimensional mass-spring structures', J. Sound Vib. **266**, 1053 (2003).

[78] J. S. Jensen and O. Sigmund, 'Systematic design of photonic crystal structures using topology optimization: Low-loss waveguide bends', Appl. Phys. Lett. **84**, 2022 (2004).

[79] T. Borrvall and J. Petersson, 'Topology optimization of fluids in Stokes flow', Int. J. Numer. Meth. Fluids **41**, 77 (2003).

[80] A. Gersborg-Hansen, 'Topology optimization of incompressible Newtonian flows at moderate Reynolds numbers', MSc thesis, Dept. of Mechanical Engineering, DTU (2003).

[81] F. Okkels and H. Bruus, 'Scaling behavior of optimally structured catalytic microfluidic reactors', Phys. Rev. E **75**, 016301 (2007).

[82] P. Michaleris, D. A. Tortorelli, and C. A. Vidal, 'Tangent operators and design sensitivity formulations for transient nonlinear coupled problems with applications to elastoplasticity', Int. J. Numer. Meth. Eng. **37**, 2471 (1994).

[83] J. P. Urbanski, T. Thorsen, J. A. Levithan, and M. Z. Bazant, 'Experiments on nonlinear electrokinetic pumps in microfluidics', Talk presented at American Physical Society MAR06 meeting, Baltimore, MD.

[84] D. Burch, unpublished (2006).