

**Learning and investigating
phenomenological models for active matter**

by

Rohit B. Supekar

B.Tech., Indian Institute of Technology Madras (2015)

S.M., Massachusetts Institute of Technology (2017)

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2021

© Massachusetts Institute of Technology 2021. All rights reserved.

Author
Department of Mechanical Engineering
July 8, 2021

Certified by.....
Jörn Dunkel
Associate Professor of Mathematics
Thesis Supervisor

Accepted by
Nicolas G. Hadjiconstantinou
Graduate Officer, Department of Mechanical Engineering

Dedicated to
my mother, Jayashree Supekar, who is a symbol of resilience and kindness,
and to my father, Balasaheb Supekar, for his life-long hard work.

Learning and investigating phenomenological models for active matter

by

Rohit B. Supekar

Submitted to the Department of Mechanical Engineering
on July 8, 2021, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Mechanical Engineering

Abstract

Active matter, such as a suspension of swimming cells or colloidal particles, or even a flock of birds, consists of self-propelled units that convert stored or ambient free energy into motion. Recent advances in high-resolution imaging techniques and particle-based simulation methods have enabled the precise characterization of the collective dynamics in such biological and engineered active systems. In parallel, data-driven algorithms for learning interpretable continuum models have shown promising potential for the recovery of underlying partial differential equations (PDEs) from continuum simulation data. Motivated by these advancements, in this thesis we analytically and numerically investigate phenomenological models in the context of active fluids, and subsequently leverage recent model learning algorithms to infer continuum models directly from microscopic simulation and experimental video data.

First, we consider idealized two-dimensional swimmers with a fixed body shape (‘squirmers’) to understand the impact of yield stress on the swimming characteristics of micro-organisms. Using the Bingham constitutive law, we compute numerical flow fields around squirmers and determine their swimming speeds. Our findings demonstrate how yield stress localizes the flow and makes jet-based propulsion energetically more efficient than tangential ciliary motions. Additionally, for the related problem of non-squirming translating cylinders, we derive and provide numerical evidence for previously unestablished analytical solutions using slipline theory from ideal plasticity.

Second, we explore how a phenomenological continuum model for active suspensions can provide insights into other pattern-forming systems. Motivated in part by the complex flow patterns observed in planetary atmospheres, we investigate generalized Navier–Stokes (GNS) equations that couple nonlinear advection with a generic linear instability. This analytically tractable minimal model for fluid flows driven by internal active stresses has recently been shown to permit exact solutions on a stationary 2D sphere. Here, we extend the analysis to linearly driven flows on rotating spheres. We derive exact solutions of the GNS equations corresponding to time-independent zonal jets and superposed westward-propagating Rossby waves, qualita-

tively similar to those seen in planetary atmospheres. Direct numerical simulations with large rotation rates obtain statistically stationary states close to these exact solutions. The measured phase speeds of waves in the GNS simulations agree with our analytical predictions for Rossby waves.

Finally, we address the challenge of learning macroscopic hydrodynamic equations for active matter directly from microscopic data. Here, we present a framework that leverages spectral basis representations and sparse regression algorithms to discover PDE models from microscopic simulation and experimental data, while incorporating the relevant physical symmetries. We illustrate the practical potential through applications to a chiral active particle model mimicking swimming cells and to recent experiments of engineered active particles. In both cases, our scheme learns hydrodynamic equations that quantitatively reproduce the self-organized collective dynamics observed in the simulations and experiments. This inference framework makes it possible to measure a large number of hydrodynamic parameters in parallel and directly from video data.

Overall, this thesis shows how phenomenological models for active matter can offer dynamical insights that can be further leveraged to inform data-driven equation discovery.

Thesis Supervisor: Jörn Dunkel

Title: Associate Professor of Mathematics

Acknowledgments

First and foremost, I want to convey heartfelt gratitude to my advisor, Professor Jörn Dunkel for his guidance and mentorship in both academic and personal settings! His enthusiasm for science is infectious, and he has been very generous with his time for many scientific as well as philosophical discussions with me. Jörn, thank you for giving me a chance to pursue a Ph.D. with you and I am truly grateful for your support towards my career beyond MIT. My sincere thanks to Professor Neil Balmforth for his advice and collaboration on one of the projects in this thesis. I will never forget our research discussions in numerous coffee shops at Woods Hole and Vancouver! Many thanks to my thesis committee members, Professor Gareth McKinley and Professor Ken Kamrin, for their keen interest in my research and for many insightful questions that have helped me dive deeper into my research results.

I owe the deepest gratitude to my research collaborators and mentors – Alexander Mietke, Vili Heinonen, Keaton Burns, Duncan Hewitt, Henrik Ronellenfitsch, Boya Song, and Alasdair Hastewell. Their help has been invaluable in making this thesis possible and I feel lucky to have learned so much from them. I have also been fortunate to be a part of a very collaborative group with diverse research interests – thanks to all the group members for their camaraderie!

During my time at MIT, I have had the opportunity to be a teaching assistant for numerous semesters. I am grateful to Professor Gareth McKinley for giving me a chance to teach the graduate course in Fluid Mechanics (2.25). It offered me a unique perspective on the subject and was one of my finest experiences at MIT. Subsequently, I was fortunate to be a TA for Dynamics and Controls (2.003) with Professor Alberto Rodriguez for several semesters. Alberto, thank you so much for looking out for me and sometimes going out of your way in helping me secure a TA position. Beyond the teaching assistantships, I want to acknowledge the Geophysical Fluid Dynamics (GFD) summer fellowship [1] which supported the research in Chapter 2, the MathWorks Engineering Fellowship that funded my final year, and MIT Supercloud [2] for providing computational resources. I also want to convey my

thanks to Leslie Regan and Saana McDaniel at the MechE graduate office for helping me through various logistical challenges.

Life in Boston would not have been the same without the companionship and support of all the amazing friends I made here. Kasturi, thank you so much for your unparalleled support through the ups and downs over the last many years. You are a true inspiration to me, and I cherish the many miles we have run together! I must give a special shout-out to ‘Dynamically Qualified (DQ)’ – Saviz Mowlavi, Cameron McBride, Essie Yu, and Boyu Fan – for the weekly meetups where unfiltered discussions about life and school were a norm rather than the exception. Graduate school would have been incredibly difficult without them. I am also thankful for the friendship of the residents (and affiliates) of the research ‘Nexus’ of building 3, Jerry Wang, Rachel Kurchin, Boyu Fan, Essie Yu, Margaux Filippi, Nisha Chandramoorthy, and Ruth Musgrave. I miss our hours-long discussions on wide-ranging topics that have hopefully not disturbed anyone else in the Nexus. The friends I made at the 2018 GFD program deserve a very special mention for all the shared memories at ‘The Barn’ and Walsh cottage.

When I migrated over to the Mathematics Department, the 2-333C office would have been incomplete without the joyful conversations with Sam Turton, Miles Couchman, and Alasdair Hastewell. The ‘long’ walks to the Stata food truck would have felt very lonely without them. And speaking of food, I cannot forget to thank Arkopal Dutt and Ravi Kishore for all the pleasant conversations while we religiously made our way through the Cafe Spice food. Thanks to the other frequent visitors, Karthik Gopalakrishnan, Nigamaa Nayakanti, and Rushina Shah, who were certainly more open to other options! Many thanks to Karthik for introducing me to the joy of skiing! And Saviz, thanks for watching me fall endlessly as I gradually got better at the sport! Kasturi, Saviz and Arko, I am so grateful for your advice and countless conversations over many months in 2018, which undoubtedly influenced the subsequent trajectory of my graduate studies.

My running club – Community Running – made me a better runner and the Monday/Wednesday track workouts kept me sane. I am deeply grateful to the coaches

for giving me a chance to participate in the 2018 Boston Marathon. Despite the cold weather, torrential rain, and heavy winds on race day, it will remain as one of my fondest memories in the city. Beyond running, I have also enjoyed being a part of the student activities at MIT. I am grateful to the members of Sangam, GAME, the Tang hall student government, and the MIT SIAM student chapter for all the fun times we have shared in planning several events.

It would be surreal to not mention the COVID-19 pandemic, which overshadowed the final 1.5 years of my graduate studies. During the remote work environment, it was possible to remain productive only because of regular video calls with the DQ crew and many shared meals with Arkopal, Saviz, Cindy, and Kasturi at 25 Andrew St. For their support during those trying times, I am truly thankful!

And as I write this section, I am reminded of the memorable vacations I took with my friends from IIT Madras who have moved to the States. Super thanks to Abhishek Namballa, Abhishek Sivaram, Adarsh Chavakula, Addarsh Chandrasekar, Neha Ashok, Kalyan Vasudev, and Karthik Murugadoss for re-living ‘insti’ memories from time to time, despite having moved on to different stages of life.

Finally, I would not be where I am today without the support of my parents and sister. Thanks to them for everything! It has been more than a decade since I started moving to different cities and countries for education. I am sure that the distance has been difficult for them and I am grateful for their understanding and support!

Contents

1	Introduction	31
1.1	Active matter	31
1.2	Models for swimming micro-organisms	33
1.3	Hydrodynamic models for active matter	34
1.4	Learning continuum models from data	35
1.5	Thesis contributions	36
1.6	Outline	37
2	Translating and squirming cylinders in a viscoplastic fluid	41
2.1	Introduction	41
2.2	Mathematical formulation	43
2.2.1	Boundary conditions	45
2.2.2	Numerical method	47
2.2.3	Ideal plasticity	48
2.3	Revisiting flow around a no-slip cylinder	50
2.3.1	Randolph and Houlsby’s slipline solution	50
2.3.2	The residual plugs	52
2.4	Flow past a partially rough cylinder	55
2.4.1	The Martin and Randolph slipline solution and upper bound	57
2.4.2	Lower bound and torque balance	59
2.5	Viscoplastic squirmers	61
2.5.1	Newtonian limit	61
2.5.2	Symmetries	62

2.5.3	Numerical results	63
2.5.4	Boundary layer theory	69
2.6	Conclusions	72
3	Linearly forced fluid flow on a rotating sphere	77
3.1	Introduction	77
3.2	Generalized Navier-Stokes model for linearly driven flow	79
3.2.1	Planar geometry	79
3.2.2	On a rotating sphere	81
3.3	Exact time-dependent solutions	85
3.3.1	Global solutions	86
3.3.2	β -plane solutions	88
3.4	Simulations	88
3.5	Conclusions	92
4	Learning hydrodynamic equations for active matter from particle simulations and experiments	93
4.1	Introduction	93
4.2	Learning framework	96
4.2.1	Active particle simulations	96
4.2.2	Hydrodynamic fields	97
4.2.3	Spatiotemporal representation and differentiation	98
4.2.4	Inference of hydrodynamic equations	100
4.2.5	Validation of learned models	105
4.3	Learning from experimental data	108
4.3.1	Coarse-graining and spectral representation of experimental data	108
4.3.2	Physics-informed library	108
4.3.3	Learned hydrodynamic equations and validation	109
4.3.4	Predicting collective roller dynamics in confinement	111
4.4	Conclusions	111

5	Summary and outlook	113
5.1	Conclusions	113
5.2	Future extensions	115
	Appendix A	117
A.1	Viscoplastic boundary layers for a no-slip cylinder	117
A.1.1	Beyond the rotating plug	117
A.1.2	Underneath the rotating plug	118
A.2	Slipline results for a partially rough cylinder	122
A.2.1	The drag force and lower bound	122
A.2.2	Angular momentum balance	123
A.2.3	The net dissipation rate and upper bound	126
	Appendix B	131
B.1	Formulas for Λ , τ and κ	131
B.2	Total angular momentum	131
	Appendix C	133
C.1	Methods	133
C.1.1	Particle simulations	133
C.1.2	Kernel coarse-graining with periodic and non-periodic boundaries	134
C.1.3	Spectral representation	135
C.1.4	Sparse regression	135
C.1.5	Linear dependencies of the library terms	137
C.1.6	Continuum simulations	138
C.2	Analytic coarse-graining of the particle model	139
C.2.1	Dynamic equation of the one-particle probability density . . .	139
C.2.2	Dynamic equations from conventional kernel coarse-graining .	141
C.3	Parameters of learned models	143

List of Figures

- 1-1 Biological and engineered active matter. (a) A dense bacterial suspension in a microfluidic chamber exhibits collective chaotic motion with a prominent length scale (reproduced from Wensink et al. [3]). (b) Two-dimensional array of vortices visible in time-averaged snapshots of swimming sperm cells (from sea urchins) due to hydrodynamic entrainment (from Riedel et al. [4]; reproduced with permission from AAAS). (c) Colloidal particles submerged in a dielectric fluid undergo electro-hydrodynamic instabilities and start to roll over a surface, thus mimicking active particles [5]. Over a million micro-rollers are confined within a 5 mm x 5 mm square domain, resulting in spontaneous symmetry breaking and the formation of four inter-woven density shocks due to the reflections from the boundaries (Credits: Alexandre Morin, Delphine Geyer, and Denis Bartolo). (d) Colloidal micro-magnets suspended in water rotate about their axes due to an external rotating magnetic field. They hydrodynamically interact with each other to form macro-droplets containing hundreds of micro-magnets (from Soni et al. [6]; reproduced with permission from Springer Nature). Scale bars: (a) 10 μm , (b) 200 μm , (c) 1 mm, and (d) 100 μm 32

- 2-1 Comparison of the numerical solution and the slipline pattern for a fully rough (no-slip) cylinder. (a) Color plot of the logarithm of the strain rate ($\log_{10}(\dot{\gamma})$) from the numerical solution at $\text{Bi} = 2^{14} \gg 1$, together with streamlines (blue). (b) The slipline pattern made up of α - and β -lines (red and blue; § 2.2.3) obtained by Randolph and Houlsby [7] (§ 2.3.1). In the white region, the α -lines form circular arcs with center A, and in the green region, the β -lines form the involutes to the circle. Arrows indicate the direction of motion of the cylinder. In (a) and (b), the overall patterns and the rigid plugs (with zero strain rate; shaded black) match except for the rigidly rotating ones on the top and bottom of the cylinder. These residual plugs appear due to the inconsistency of pressure change in the thin viscoplastic boundary layer above the cylinder and the overlying plastic solution. As described in § 2.3.2, they eventually vanish, albeit very slowly, with increasing Bingham number Bi , thus confirming that the viscoplastic solution converges to the plastic one. 51
- 2-2 Pressure variation over the cylinder scaled by Bi , for the Bingham numbers indicated. The dashed line corresponds to the pressure of the slipline solution given by (2.18); the numerical profile approaches this solution as the Bingham number increases. The inset shows overlapping profiles of p/Bi against $(\theta - \pi/2)\text{Bi}^{1/7}$, which confirms the scaling predicted from the boundary layer theory in appendix A.1. 53
- 2-3 Scaling data for the residual rotating plug against Bi , showing (a) the plug radius $y_p - 1$ where $(0, y_p)$ is the the top of the plug, (b) it's rotation rate ω , (c) the boundary layer thickness at $\theta = \frac{1}{2}\pi$ and $\frac{1}{3}\pi$, and (d) the angular size of the smoothing region, estimated by the location $\hat{\theta}_*$ for which $p = \frac{1}{2}\text{Bi}$. All these plug characteristics follow the scalings predicted by the boundary layer theory of appendix A.1 (dashed lines). 54

2-4 (a) Drag coefficient C_d against Bi for computations using a no-slip boundary condition (red circles) or the plastic slip law in (2.7) with $\varrho = 1$ (blue stars). The dashed line shows (2.19). (b) Plots of $\log_{10} \dot{\gamma}$ and streamlines for two solutions at $\text{Bi} = 2^8$; the upper half shows the no-slip cylinder, and the lower half the cylinder with (2.7) and $\varrho = 1$. Due to the absence of the no-slip condition and the resulting viscoplastic boundary layer against the cylinder, the rigidly rotating plug disappears in (b, bottom) as compared to (b, top) [§ 2.3.2]. . . 54

2-5 Numerical solutions (left) showing $\log_{10} \dot{\gamma}$ with streamlines (light blue) and slipline solutions [right; same color scheme as in figure 2-1] for (a) $\varrho = 0$ and (b) $\varrho = 0.5$, both at $\text{Bi} = 2^{12}$. In the slipline construction, the angle β_2 can be variable and dictates the angular extent of the circular fan with center P . Here, its value is set according to (2.21) [63.0° in (a) and 69.2° in (b)], where the lower and upper bounds match [see figure 2-6]. At these values, the slipline pattern compares well with the numerical solution, with the velocity jumps [along arc GD in (a, left) and arcs BFH and CDG in (b, left)] replaced by viscoplastic shear layers [bright yellow regions in (a, b; left)]. This provides numerical evidence that the slipline pattern is the true plastic solution. The primed points, B' to F' , are the reflections of points, B to F , about the y -axis. 56

2-6 Upper (blue) and lower (black) bounds derived from the Martin and Randolph [8] slipline solution for the values of ϱ indicated; the circles indicate the minimum of the upper bound, which coincides with the intersection of the lower and upper bounds. For β_2 less than this minimum, the lower bound exceeds the upper bound and is thus spurious (shown dashed). For $\beta_2 > \frac{1}{2}\pi$, the velocity and stress fields become inconsistent, as in the original Randolph and Houlsby [7] construction (the corresponding bounds are shown by dotted lines). The red stars show the extrapolated values for $\text{Bi} \rightarrow \infty$ of the drag and angle β_2 from numerical computations. The excellent match between the numerical computations (red stars) and the intersection of the lower and upper bound curves indicates that the slipline solution at the corresponding β_2 is the true plastic solution. [The angle β_2 is estimated from the numerical computations by matching the rotation rate of the rigid plugs with the expression $\sin \beta_2 / \lambda$. This avoids tracing the yield surfaces which are sensitive to numerical errors.] 58

2-7 Squirmer solutions for $(n, a) = (1, 0)$ showing $\log_{10} \dot{\gamma}$ overlain by streamlines (blue), with $\text{Bi} = 1$ (a–c) and $\text{Bi} = 2^8$ (d–f), for $U_s \text{Bi}^{1/2} \approx 0$ (a, d), 0.1 (b, e), 0.38 (c, f), respectively. In (f), the full circle has the scale of the axes as in (d) and (e), whereas the quarter circle shown in the inset is a magnification to highlight the thin boundary layer. (g) Numerical drag coefficients C_d against scaled translation speed $\text{Bi}^{1/2} U_s$ match well with the asymptotic predictions for $\text{Bi} \gg 1$ from the Randolph and Houlsby slipline solution (§3) and the boundary-layer analysis of § 2.5.4 (dashed black lines). The $C_d = 0$ line (dotted) corresponds to the swimming speed of force-free squirmers. 64

2-8	<p>Squirmers solutions showing $\log_{10} \dot{\gamma}$ at $\text{Bi} = 2^8$ for (a-c) $n = 3, m = 0$ and (d-f) $n = 5, m = 0$ with the imposed swimming speed (a,d) $U_s = -0.015$, (b,e) $U_s = 0$ and (c,f) $U_s = 0.015$. (g) The variation of the drag coefficient C_d with U_s for $n = 3$ (blue stars) and $n = 5$ (red squares). Since $n > 1$ and $m = 0$, $C_d = 0$ [dashed line in (g)] is necessary for $U_s = 0$ due to multiple lines of reflection symmetry (see § 2.5.2).</p>	64
2-9	<p>Force-free squirmer solutions showing showing $\log_{10} \dot{\gamma}$ (same color bar as in figure 2-8; streamlines, blue) for surface velocities with $n = 1, m = 2$. (a-d) $\text{Bi} = 1$ and (e-h) $\text{Bi} = 2^8$, for (a,e) $a = 0$, (b,f) $a = 0.25$, (c,g) $a = 1$ and (d,h) $a = 2$. (i) The corresponding swimming speeds. (j) Rescaled swimming speeds for $a = 0$ and $a = 0.25$ asymptotically approach predictions of the boundary-layer theory (§2.5.4, dashed) as $\text{Bi} \rightarrow \infty$.</p>	66
2-10	<p>Force-free squirmer solutions with $(n, m) = (1, 3)$ and $\text{Bi} = 2^8$ showing $\log_{10} \dot{\gamma}$ for (a) $a = 0.25$, (b) $a = -0.25$, (c) $a = 1$, (d) $a = -1$, (e) $a = 2$ and (f) $a = -2$. (g-i) Variation of the swimming speed with Bi. In (g), the asymptotic prediction (dashed) from (2.36) for $a = 0.25$ matches well with the numerical data.</p>	66
2-11	<p>Force-free squirmer solutions showing $\log_{10} \dot{\gamma}$ for $(n, m) = (2, 3)$ and $a = 1$ for (a) $\text{Bi} = 1$ and (b) $\text{Bi} = 2^8$. (c) The variation of the non-zero swimming speed U_s with Bi. In contrast, these swimming modes have a zero swimming speed in the Newtonian limit ($\text{Bi} \rightarrow 0$).</p>	68
2-12	<p>Force-free squirmer solutions showing $\log_{10} \dot{\gamma}$ for the normal surface velocity condition (2.11) with (a) $a = 0.1$, (b) 0.25 and (c) 1, at $\text{Bi} = 2^8$. (d) The swimming speed approaches the Newtonian limit of 0.5 as $\text{Bi} \rightarrow 0$ and a finite value dependent on a as $\text{Bi} \rightarrow \infty$.</p>	68

2-13 Boundary-layer thicknesses of squirmers for (a) $(n, a) = (1, 0)$, (b) $(n, m, a) = (1, 2, 0.25)$, and (c) $(n, m, a) = (1, 3, 0.25)$, with $\text{Bi} = 2^6$ (black), $\text{Bi} = 2^8$ (blue), $\text{Bi} = 2^{10}$ (red) and $\text{Bi} = 2^{12}$ (green). The boundary-layer theory predictions from (2.31)[dashed lines] match well with the numerical profiles except near $\theta = 0, \pi$ 68

3-1 Statistically stationary states of the normalized vorticity $\zeta\tau$ from simulations (a-i) for $\kappa\Lambda = 1$ become more zonal (or banded) as the rotation rate $\Omega\tau$ increases. At the highest rotation rate $\Omega\tau = 500$, the width of the alternating zonal jets is determined by the parameter R/Λ that represents the ratio of the radius of the sphere and the diameter of the vortices forced by the GNS dynamics. The main characteristics of these flow patterns at high rotation rate are captured by spherical harmonics $Y_\ell^0(\theta, \phi)$ that solve the dynamical equations. Matching the length scale R/Λ gives $\ell = 6$ (j), $\ell = 11$ (k), and $\ell = 21$ (l) for $R/\Lambda = 2, 4$, and 8 , respectively. 80

3-2 The growth rate Ξ of spherical harmonic modes $Y_\ell^m(\theta, \phi)$ in (3.5) plotted as a function of the wavenumber ℓ . The parameters used to make this plot are $((\tau/R^2)\Gamma_0, (\tau/R^4)\Gamma_2, (\tau/R^6)\Gamma_4) \simeq (1.43 \times 10^{-2}, -4.86 \times 10^{-5}, 3.72 \times 10^{-8})$ which correspond to $R/\Lambda = 8$ and $\kappa\Lambda = 1$. The grey region indicates the active bandwidth where $\Xi > 0$ and energy is injected. Λ is the diameter of the vortices forced by the mode with the maximum growth rate $1/\tau$, and κ is the active bandwidth i.e. $\kappa = (\ell_+ - \ell_-)/R$ where $\Xi(\ell_\pm) = 0$ 82

3-3 Data from steady-state solutions at $t/\tau = 15$ for the highest rotation rate $\Omega\tau = 500$. The rows correspond to $R/\Lambda = 2$ (a–d), 4 (e–h) and 8 (i–l). Panels (a, e, i) show Mercator projections of the dimensionless vorticity $\zeta\tau$. Panels (b, f, j) show the zonal-mean azimuthal velocities $\langle v_\phi \rangle_\phi / (R/\tau)$. Panels (c, g, k) show spherical harmonic decomposition of dimensionless vorticity $\zeta\tau$ with marker size indicating amplitude and color indicating phase. All plots indicate the existence of dominant zonal jets with $m = 0$ and ℓ 's within the active band indicated in grey. These modes are close to the exact solutions in figure 3-1(j–l). Panels (d, h, l) show time-variation of the energy of all the modes (black), active $m = 0$ modes (red), and all other modes (green); the energy contained in the active $m = 0$ accounts for most of the total energy in the statistically stationary state. 84

3-4 Energy spectra for (a) $R/\Lambda = 2$, (b) $R/\Lambda = 4$ and (c) $R/\Lambda = 8$. The grey shaded region indicates the active bandwidth where the spectra show a peak. The spectra for $(R/\Lambda, \Omega\tau) = (4, 250), (4, 500)$ and $(8, 500)$ have been obtained from an ensemble average of 10 simulations with random initial conditions. 87

3-5 Phase speed of the spherical harmonic modes (ℓ, m) in the forcing bandwidth, normalized by the analytical phase speed in (3.11), for $\Omega\tau = 500$ and different values of R/Λ . The dotted line indicates the value 1 for comparison. 87

3-6 (a–d) Time-space diagrams of the deviation of vorticity, $\zeta - \langle \zeta \rangle$, where $\langle \cdot \rangle$ is the average over time and space, indicate that the phase speed of the westward propagating Rossby waves in the local β -plane changes with latitude. (e–l) Logarithm of the power spectral density, $S = |\tau \hat{\zeta}(k_x, \sigma)|^2$, where $\hat{\zeta}$ is the discrete Fourier transform, at different northern (e–h) and southern (i–l) latitudes. The grey regions indicate the forcing bandwidth with $k_- < |k_x| < k_+$ justifying the rapid decay of the power spectral density for $|k_x| > k_+$. The white lines in each panel show the analytical dispersion relation from (3.14) with $|\mathbf{k}| = k_+$ and $|\mathbf{k}| = k_-$; the one with the steeper slope corresponds to k_- . These predictions capture the variance of power spectral density. 89

4-1 Learning hydrodynamic models from particle simulations and experiments. (a) Inputs are time-series data for particle positions $\mathbf{x}_i(t)$, particle orientations $\mathbf{p}_i(t) = (\cos \theta_i, \sin \theta_i)^\top$, etc., measured in simulations or experiments with microscale resolution (§ 4.2.1). (b) Spatial kernel coarse-graining of the discrete microscopic variables provides continuous hydrodynamic fields, such as the density $\rho(t, \mathbf{x})$ or the polarization density $\mathbf{p}(t, \mathbf{x})$ (§ 4.2.2). (c) Coarse-grained fields are sampled on a spatiotemporal grid and projected onto suitable spectral basis functions. Systematic spectral filtering (compression) ensures smoothly interpolated hydrodynamic fields, enabling efficient and accurate computation of spatiotemporal derivatives (§ 4.2.3). (d) Using these derivatives, a library of candidate terms $C_l(\rho, \mathbf{p})$ and $\mathbf{C}_l(\rho, \mathbf{p})$ consistent with prior knowledge about conservation laws and broken symmetries is constructed. A sparse regression algorithm determines subsets of relevant phenomenological coefficients a_l and b_l (§ 4.2.4). The resulting hydrodynamic models are sparse and interpretable, and their predictions can be directly validated against analytic coarse-graining results (§ 4.2.5) or experiments (§ 4.3). Bottom: Snapshots illustrating the workflow for microscopic data generated from simulations of chiral active Brownian particles [equation (4.1)]. 95

4-2 Learning mass conservation dynamics. (a) Top: Time evolution of positions and orientations of 12,000 particles following the dynamics in equations (4.1). Bottom: Coarse-grained density ρ (color code) and polarization field \mathbf{p} (arrows). Starting from random initial conditions ($t = 0$), a long-lived vortex pattern with well-defined handedness emerges ($t = 1250$). Training data were randomly sampled from the time window $t \in [40, 400]$, contained within the gray box. Domain size: 100×100 . (b) Slices through the power spectrum $S_{x;n,\mathbf{q}} = |\mathbf{e}_x \cdot \hat{\mathbf{p}}_{n,\mathbf{q}}|^2$ for different values of the Chebyshev polynomial order $n \in \{0, 300, 600\}$, corresponding to modes with increasing temporal frequencies. The rightmost panel depicts the total spectral power $\sum_{\mathbf{q}} S_{x;n,\mathbf{q}}$ [see equation (4.3b)] of each Chebyshev mode n . The slowly decaying long tail of fast modes indicates a regime in which fluctuations dominate over a smooth signal. The cut-off $n_0 = 600$ removes these modes, in line with the goal to learn a hydrodynamic model for the slow long-wavelength modes. (c) Kymographs of the spectral derivatives $\partial_t \rho$ and $-\nabla \cdot \mathbf{p}$ at $y = 50$, obtained from the spectrally truncated data. (d) Mass conservation in the microscopic system restricts the physics-informed candidate library to terms that can be written as divergence of a vector field. (e) Learned phenomenological coefficients a_l of PDEs with increasing complexity (decreasing sparsity) (appendix C.3). PDE 1 (\blacktriangleleft) is given by $\partial_t \rho = a_1 \nabla \cdot \mathbf{p}$ with $a_1 = -0.99$. As PDE 1 is the sparsest PDE that agrees well with analytic coarse-graining results (table 4.1), it is selected for the hydrodynamic model. 99

4-3 Learning polarization dynamics. (a) Same particle dynamics as in figure 4-2a for visual reference. (b) Top: Coarse-grained density and polarization field as in figure 4-2a. Bottom: Magnitude $|\mathbf{p}|$ of the coarse-grained polarization field. Emerging vortices ($t = 400, 1250$) appear as ring-like patterns in $|\mathbf{p}|$. Training data were randomly sampled from the time window $t \in [40, 400]$, enclosed within the gray box. **c**: Physics-informed candidate library (with $b_1 = -D_r$) including terms constructed from $\mathbf{p}_\perp = (-p_y, p_x)^\top$, which are allowed due to the chirality of the microscopic system. **d**: Learned phenomenological coefficients b_l of PDEs with increasing complexity (appendix C.3). For all PDEs, learned coefficients of the linear terms \mathbf{p}_\perp and $\nabla\rho$ compare well with analytic predictions (table 4.1, appendix C.2.2). **e**: Simulation of the final hydrodynamic model (PDE 8 for the polarization dynamics and PDE 1 in figure 4-2e for the density dynamics). Starting from random initial conditions ($t = 0$), long-lived vortex states emerge on a similar time scale, with similar spatial patterns, and with comparable density and polarization amplitudes as in the coarse-grained microscopic model data (**b**). Hydrodynamic models with PDEs sparser than PDE 8 do not form stable vortex patterns. 102

4-4 Learning from active polar particle experiments. (a) Snapshot of particle positions and velocity components of $\sim 2,200$ spontaneously moving Quincke rollers in a microfluidic channel [9]. Scale bar, $200\ \mu\text{m}$. (b) Coarse-grained density field $\rho(t, \mathbf{x})$, expressed as the fraction of area occupied by the rollers with diameter $D_c = 4.8\ \mu\text{m}$, and components $v_x(t, \mathbf{x})$ and $v_y(t, \mathbf{x})$ of the coarse-grained velocity field ($\sigma = 45\ \mu\text{m}$). 5×10^5 randomly sampled data points from ~ 580 such snapshots over a time duration of $1.4\ \text{s}$ were used for the learning algorithm. (c) Physics-informed candidate libraries for the density and velocity dynamics, $\{\bar{C}_l(\rho, \mathbf{v})\}$ and $\{\bar{\mathbf{C}}_l(\rho, \mathbf{v})\}$, respectively [equation (4.5)]. These are the same libraries as shown in Figs. 4-2e and 4-3d, but without the chiral terms and replacing $\mathbf{p} \rightarrow \mathbf{v}$. (d) Learned phenomenological coefficients c_l and d_l of the four sparsest PDEs for the density (left) and velocity (right) dynamics. The coefficients are non-dimensionalized with length scale σ and time scale σ/v_0 , where $v_0 = 1.2\ \text{mm s}^{-1}$ is the average roller speed. PDE 1 for density dynamics corresponds to $\partial_t \rho = c_3 \nabla \cdot (\rho \mathbf{v})$ with $c_3 \simeq -0.95$. PDE 2 for the velocity dynamics is shown in equation (4.6b). Learned coefficients compare well with the values reported in Ref. [9] (table 4.2). (e) Simulation snapshot at $t = 1.8\ \text{s}$ of the learned hydrodynamic model (PDEs marked by \blacktriangleleft in (a)) in a doubly periodic domain. Spontaneous flow emerges from random initial conditions, and exhibits density and velocity fluctuations that show similar spatial patterns and amplitudes as seen in the experiments (a). (f) Simulation snapshots at $t = 18.5\ \text{s}$ of the same hydrodynamic model as in (e) on a square domain with reflective boundary conditions. The model predicts the emergence of a vortex-like flow permeated by density shock waves. This prediction agrees qualitatively with experimental observations (rightmost panel) of Quincke rollers in a $5\ \text{mm} \times 5\ \text{mm}$ confinement with average density $\rho_0 \approx 0.1$ (Image credits: Alexandre Morin, Delphine Geyer, and Denis Bartolo). Scale bars, $200\ \mu\text{m}$ (simulation) and $1\ \text{mm}$ (experiment). 107

4-5	The learned model accurately predicts collective Quincke roller speeds v_0 at different average area fractions ρ_0 . Although equation (4.6) was learned from a single experiment (Supplementary Movie S2 in Ref. [9]) at fixed average area fraction $\rho_0 = 0.11$ (filled black circle), the model prediction $v_0(\rho_0) = \sqrt{-(d_1 + d_2\rho_0)/d_3}$ (solid line) with inferred parameters d_1, d_2, d_3 (table C.4), agrees well the experimentally measured speed values (red symbols) reported in Supplementary figure 4 of Ref. [9].	110
A-1	A magnification of the region underneath the upper rotating plug for a no-slip cylinder with $\text{Bi} = 2^{12}$. The color shading shows $\log_{10} \dot{\gamma}$, and the blue lines are streamlines. The α -lines (black) of the slipline solution are selected to coincide with the streamlines at $\theta = \frac{1}{2}\pi$ and they match with the streamlines except near the boundary layer attached to the cylinder.	118
A-2	Rescaled boundary layer profiles ($\Xi(\Theta) - \Xi(0)$) with the boundary-layer theory prediction from (A.9)[dashed]. The vertical dashed-dotted line marks the angular location Θ_* where the rotating plug separates from the boundary layer for $\text{Bi} = 4^7$. Within those limits, the numerical profiles match well with the boundary-layer theory predictions. . . .	121
C-1	Distribution of particle speeds v_i and rotation rates Ω_i used to simulate Eqs. (4.1). Those parameters were drawn from a heuristic distribution $p(v_i, \Omega_i)$ that is explained in more detail in Sec. C.1.1. The white marker and dashed lines depict the mean velocity $\langle v_i \rangle_p \simeq 1$ and $\langle \Omega_i \rangle_p \simeq 0.45$	134

List of Tables

4.1	Parameters of the hydrodynamic model learned for the microscopic dynamics in equation (4.1) and values predicted by analytic coarse-graining (appendix C.2.2). $\langle \cdot \rangle_p$ denotes averages over the distribution $p(v_i, \Omega_i)$ of particle velocities v_i and rotation rates Ω_i (appendix C.1.1).	106
4.2	Parameters of the learned hydrodynamic model for the Quincke roller system are close to values expected from analytic coarse-graining (*) and reported in Ref. [9] for experiments performed at mean area fraction $\rho_0 \approx 0.11$.	109
C.1	Parameters a_l of the density dynamics PDE (Fig. 4-2d) learned from simulations the microscopic active particle system in Eq. (4.1). The sparsest model (\blacktriangleleft) agrees well with the analytic coarse-graining prediction (Tab. 4.1).	144
C.2	Parameters b_l of the nine sparsest PDEs for the polarization dynamics (Fig. 4-3c), learned from simulations of the microscopic system in Eq. (4.1). PDE 8 (\blacktriangleleft) reproduces the characteristic vortex dynamics as in the microscopic simulations (Fig. 4-3a,b,e) and the coefficients of the linear terms compare well with analytic coarse-graining predictions (Tab. 4.1).	144

- C.3 Parameters c_i of the PDE for the density dynamics (Fig. 4-4c) learned from experimental data for self-propelled Quincke rollers (Supplementary Movie S2 of Ref. [9]). The dimensions of the coefficients are such that $[\mathbf{v}] = \text{mm/s}$ and $[\rho] = 1$, where the density ρ represents the area fraction of rollers of diameter $D_c = 4.8 \mu\text{m}$. The four sparsest PDEs are shown corresponding to the cut-off $n_0 \in \{50, 100\}$ above which the temporal Chebyshev modes in Eq. (4.3) are set to zero to ignore high frequencies. The sparsest PDEs (\blacktriangleleft) have coefficients close to each other and agree well with the mass conservation equation obtained from analytic coarse-graining (Ref. [9]). 145
- C.4 Parameters d_i of the PDE for the velocity dynamics (Fig. 4-4c) learned from experimental data for self-propelled Quincke rollers (Supplementary Movie S2 of Ref. [9]). The dimensions of the coefficients are such that $[\mathbf{v}] = \text{mm/s}$ and $[\rho] = 1$, where the density ρ represents the area fraction of rollers of diameter $D_c = 4.8 \mu\text{m}$. The four sparsest PDEs are shown corresponding to the cut-off $n_0 \in \{50, 100\}$ above which the temporal Chebyshev modes in Eq. (4.3) are set to zero to ignore high frequencies. The sparsest PDEs which reproduce the experimental observations (\blacktriangleleft) have coefficients that are close to each other for different values of n_0 , and they agree well with corresponding values reported in Ref. [9] (Tab. 4.2). 145

Chapter 1

Introduction

1.1 Active matter

Active matter systems are composed of self-driven units that convert stored or ambient free energy into motion, often within a surrounding fluid medium [10, 11, 12]. In nature, such systems can be found at a variety of physical length and time scales, from tens of meters in flocks of birds [13] and schools of fish [14], to micro-meters in dense suspensions of sperm cells [4] or bacteria [3]. In addition to living matter, active fluids have also been engineered, for example, by exploiting electro-hydrodynamic instabilities of colloidal particles submerged in a dielectric fluid [5] or by imposing an external rotating magnetic field on a collection of colloidal magnets [6]. Inspired by natural swarms, researchers have also created them artificially by assembling thousands of autonomous robots that cooperate only through local interactions [15].

At high particle concentrations in both living and non-living active matter, coherent large-scale collective motion has been observed as a result of interactions between the particles, mediated by the surrounding medium (figure 1-1). The most vivid examples are collections of birds or fish that exhibit ordered motion, or ‘flocking’ [16]. A dense bacterial suspension leads to chaotic flows with characteristic length and time scales, popularly referred to as ‘active turbulence’ [17, 18, 19] (figure 1-1a). Sperm cells get hydrodynamically entrained in a suspension to form a 2D array of vortices [4] (figure 1-1b). Active colloidal particles have also been observed to spon-

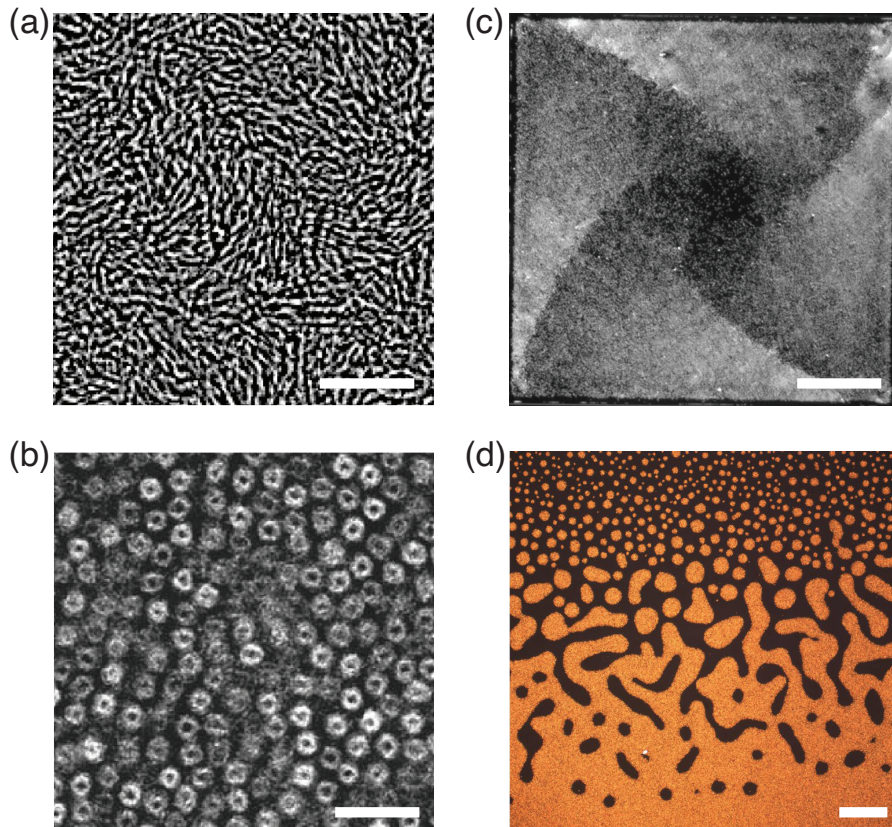


Figure 1-1: Biological and engineered active matter. (a) A dense bacterial suspension in a microfluidic chamber exhibits collective chaotic motion with a prominent length scale (reproduced from Wensink et al. [3]). (b) Two-dimensional array of vortices visible in time-averaged snapshots of swimming sperm cells (from sea urchins) due to hydrodynamic entrainment (from Riedel et al. [4]; reproduced with permission from AAAS). (c) Colloidal particles submerged in a dielectric fluid undergo electrohydrodynamic instabilities and start to roll over a surface, thus mimicking active particles [5]. Over a million micro-rollers are confined within a 5 mm x 5 mm square domain, resulting in spontaneous symmetry breaking and the formation of four interwoven density shocks due to the reflections from the boundaries (Credits: Alexandre Morin, Delphine Geyer, and Denis Bartolo). (d) Colloidal micro-magnets suspended in water rotate about their axes due to an external rotating magnetic field. They hydrodynamically interact with each other to form macro-droplets containing hundreds of micro-magnets (from Soni et al. [6]; reproduced with permission from Springer Nature). Scale bars: (a) 10 μm , (b) 200 μm , (c) 1 mm, and (d) 100 μm .

taneously break symmetry under confinement (figure 1-1c) or form chiral macrodroplets (figure 1-1d). Such suspensions of living and non-living active particles are also referred to as ‘active fluids’ [20]. A key question for these fluids is how one can construct continuum models that capture their phenomenology and elucidate the physical principles that arise from particle interactions [10]. It is also a major challenge to quantify [21] and measure the bulk hydrodynamic parameters [22] and relate them to microscopic properties [9].

1.2 Models for swimming micro-organisms

To understand collective behavior, it is important to characterize the microscopic flows that are generated by individual swimming organisms or engineered active particles. These flows affect other particles in their neighborhood and hence influence the collective behavior. Micro-organisms swim in the world of low-Reynolds number, where viscous effects dominate the fluid or particle inertia [23]. They typically propel themselves through tiny protrusions known as flagella or cilia, albeit through a variety of different physical motions. For example, sperm cells have a single flagella to which they induce a wave-like deformation to swim [24]. *Chlamydomonas* algae have two flagella that beat together much like the human breast stroke [25].

From an applied mathematician’s perspective, swimmers can be minimally represented by force dipoles (or stresslets), since their net force must be zero in the absence of inertia [23]. Based on the nature of the force dipoles, they can be broadly categorized into ‘pushers’ and ‘pullers’. Pushers have a dipole of forces pointing away from the body so that they repel fluid along the body axis and draw it in from the sides; examples include spermatozoa and *E. coli*. Pullers have dipole forces pointing inwards, thus pushing fluid away from the sides; a prominent example is *Chlamydomonas*. Another prominent swimming mechanism is exhibited by a multi-cellular spherical colony of *Volvox carteri* via hundreds of cilia beating synchronously to form metachronal waves [26]. To model such spherical swimming colonies and other organisms, the so-called idealized ‘squirmers’ models consider cylinders or spheres with

an imposed steady surface velocity [27, 28, 29]. This mimics the coordinated beating of cilia and the specific form of the imposed velocity can be varied to model both pushers and pullers.

Apart from the kinematics of the swimming mode, the rheology of the surrounding medium also plays a key role in the swimming characteristics and the resulting flows [23]. Micro-organisms often swim in fluids with a variety of non-Newtonian properties. Sperm cells must swim through mucus which is known to have visco-elasto-plastic characteristics [30]. Marine invertebrates such as Kinorhyncha, also known as ‘mud dragons’, have evolved to navigate muddy sediments [31] which are viscoplastic.

While most squirmer models have been considered in Newtonian fluids, there has also been focus on the effects of viscoelasticity [32, 33]. In contrast, relatively less attention has been given to how the presence of a yield stress can affect the swimming behavior. To address this gap in the literature and motivated by recent related viscoplastic locomotion problems [34, 35], in Chapter 2 we investigate the motion of pusher and puller squirmers and translating cylinders in viscoplastic fluids.

1.3 Hydrodynamic models for active matter

To obtain a continuum hydrodynamic description of active systems, a popular approach is to start with their microscopic description and subsequently use tools from statistical physics to coarse-grain the microscopic dynamics [36, 37, 11]. If it is possible to postulate the microscopic physics, several follow-up assumptions typically need to be made to arrive at the continuum description for the long-wavelength, slow-timescale hydrodynamics [38, 39]. An alternative approach is to simply write down continuum equations that govern the macroscopic fields by including all possible terms that respect physical symmetries in the system. Typically, a large number of such terms are possible due to the absence of generic conservation laws, with the exception of particle number conservation in some systems. This approach was pioneered by Toner and Tu [40], who proposed a continuum extension of the agent-based

model by Vicsek et al. [41]. The Toner-Tu model provided a means to explain flocking of active particles as well as fluctuations around such ordered states. Such a phenomenological approach has also been successful in constructing models for active turbulence [17, 19]. One class of these non-equilibrium models is the generalized Navier-Stokes (GNS) equations which couple nonlinear advection with a generic linear instability [42]. Due to the success of GNS in modeling pattern formation in active suspensions [18, 43], it is a natural question whether such models with active energy injection can be utilized in other pattern forming systems. Motivated further by recent studies which show that non-equilibrium approaches can provide insights into the dynamics of planetary flows and atmospheres [44, 45], in Chapter 3 we investigate a rotating two-dimensional GNS model on a sphere.

1.4 Learning continuum models from data

Due to advances in microscopy techniques [46, 47] and agent-based computational modeling [48], active matter systems can now be observed and analyzed at remarkable spatiotemporal resolutions. In parallel with this progress, machine learning methods have shown potential in applications to active matter [49] as well as fluid mechanics [50]. There have been largely two streams of methodologies to incorporate machine learning into the mathematical modeling of physical systems. The first approach is to approximate the dynamical equations or the system’s state partially or entirely by black-box neural networks [51, 52, 53, 54]. This methodology has also been employed for solving inverse problems using Physics-Informed Neural Networks (PINNs) [55] which incorporate the underlying governing equations – if known *a priori* – as constraints. The second approach is to discover the governing ordinary or partial differential equations [56, 57], which is particularly useful when the system’s governing equations are not known. This methodology is highly interpretable since specific terms in the discovered equations can be connected back to the underlying physics.

Learning differential equations from data goes back more than two decades when Bär et al. [58] demonstrated the recovery of reaction-diffusion equations from simu-

lated data. Subsequently, using symbolic regression, Schmidt and Lipson [59] showed how analytic expressions for conserved quantities such as the Hamiltonian of a double pendulum system can be learned from experimental data. More recently, Brunton et al. [56] proposed the Sparse Identification of Nonlinear Dynamics (SINDy) framework that performs sparse regression on a pre-defined candidate library of terms to recover the underlying differential equations. To enhance its noise robustness, this framework has been further combined with artificial neural networks [60, 61], weak formulations [62, 63], and stability selection [64, 65]. These seminal studies, however, have primarily focused on the recovery of equations from continuum simulation data and it remains to be seen how successful they can be for systems where the governing equations are not known *a priori*. Due to the inherent complexities of first principle derivations in active matter systems, such equation discovery methods could provide an accelerated and informative pathway for modeling them [49]. In line with this motivation, we present a learning framework in Chapter 4 that successfully translates both microscopic simulation and experimental data into interpretable and predictive hydrodynamic models.

1.5 Thesis contributions

This thesis contributes to answering the following fundamental questions:

1. How do individual micro-organisms swim in complex fluids? By studying idealized squirmer models in viscoplastic fluids, what is the impact of yield stress on swimming speed and the corresponding flow fields?
2. To what extent can recently proposed phenomenological models for active fluids be used to describe other pattern-forming systems? Specifically, what insights can be drawn from the generalized Navier-Stokes equations when they are coupled with a curved geometry in a rotating frame?
3. How can we make use of recent machine learning methods to learn phenomenological continuum models for active matter directly from experimental video

data? In particular, how can equation discovery methods be integrated with physical constraints and our knowledge of (broken) symmetries to learn active hydrodynamic models that are not known *a priori*?

1.6 Outline

Chapter 2

To study the effect of a yield stress on the swimming characteristics of micro-organisms, we consider the Bingham model for viscoplastic creeping flow around squirming cylinders. We begin by first addressing two related problems for non-squirming cylinders. First, translating cylinders with no-slip surfaces appear to generate adjacent rotating plugs in the limit where the translation speed becomes vanishingly small. In this plastic limit, analytical results are available from slipline theory of perfect plasticity which indicate that no such plugs should exist. Using a combination of numerical computations and asymptotic analysis, we show that the plugs of the viscoplastic theory actually disappear in the plastic limit, albeit very slowly (§ 2.3). Second, when the boundary condition on the cylinder is replaced by one that permits sliding, the plastic limit corresponds to a partially rough cylinder. In this case, a plasticity solution has not been previously established; we provide evidence from numerical computations and slipline theory that a previously proposed upper bound by Martin and Randolph [8] is actually the true plastic solution (§ 2.4). Finally, we consider how a prescribed surface velocity field can propel cylindrical squirmers through a viscoplastic fluid (§ 2.5). We determine swimming speeds, contrast the results with those from the corresponding Newtonian problem, and discuss implications of flow localization on collective dynamics and the energetics of different swimming modes.

Chapter 3

We study a continuum phenomenological model for active fluids to investigate its general utility in pattern forming systems. In particular, we investigate higher-order extensions of the Navier-Stokes equations which couple nonlinear advection

with active energy injection through a generic linear instability. These generalized Navier-Stokes (GNS) [18] equations present a minimal model for fluid flows driven by active stresses, are analytically tractable, and have been recently shown to permit exact solutions on a stationary 2D sphere [66]. In this chapter, motivated by planetary atmospheres, we extend the analysis to linearly driven flows on rotating spheres (§ 3.2). Exact solutions with time-independent zonal jets superimposed with westward-propagating Rossby waves are derived on the sphere as well as on the local β -plane (§ 3.3). These solutions are qualitatively similar to those seen in planetary atmospheres. Using a spectral numerical solver, we also perform direct numerical simulations (§ 3.4); at large rotation rates, the resulting statistically stationary states are close to the exact solutions of the model. The measured phase speeds of waves in the numerical solutions agree with our analytical predictions for Rossby waves. Our results show that the GNS framework can be a useful minimal model for describing pattern forming flows on rotating spheres, such as planetary atmospheres.

Chapter 4

We present a learning framework that addresses the challenge of learning macroscopic hydrodynamic equations for active matter directly from particle data. Our framework leverages spectral basis representations and sparse regression algorithms [56] to discover PDE models from microscopic simulation and experimental data, while incorporating the relevant physical symmetries. The practical potential of this approach is first demonstrated for a chiral active particle model mimicking swimming cells (§ 4.2.1). The learned equations quantitatively reproduce the self-organized vortices observed in the microscopic simulations. Furthermore, the key linear coefficients of the model compare well with those obtained from coarse-graining procedures (§ 4.2.5). We then apply the framework to experimental data from recent colloidal micro-roller experiments by Geyer et al. [9] (§ 4.3). Here, the learned models reproduce the phenomenology of sound waves in the experimental data and the coefficients compare well with estimates from the experimental study. The learned equations are further analyzed to obtain the density-dependent speed of the rollers,

which matches very well with measurements from different experiments (§ 4.3.3). To further test the predictability of the learned model, we simulate it in a different geometry and find that the resulting phenomena is remarkably similar to analogous experiments (§ 4.3.4). These results show that our inference framework can be used to learn interpretable and predictive hydrodynamic models by incorporating the relevant physical constraints and domain knowledge of active systems. The framework also makes it possible to measure a large number of hydrodynamic parameters in parallel and directly from video data.

Chapter 5

We present the conclusions of this work (§ 5.1) and discuss avenues for future research (§ 5.2).

Chapter 2

Translating and squirming cylinders in a viscoplastic fluid

The contents of this chapter have been published in the article: R. Supekar, D. R. Hewitt, and N. J. Balmforth, Translating and squirming cylinders in a viscoplastic fluid, *J. Fluid Mech.*, 882:A11, 2020 [67]. This research was initiated when the author of the thesis participated in the *Geophysical Fluid Dynamics* summer fellowship program [1] in 2018.

2.1 Introduction

Slow viscous flow around a cylinder is a classical problem in fluid mechanics, associated with Stokes' paradox and its resolution by the inclusion of weak inertial terms in the far field. The analogous problem for non-Newtonian fluids has also played a role in understanding viscoelastic extensional flow [68, 69] and how a yield stress localizes deformation and provides drag for viscoplastic fluids [70, 71] and granular materials [72, 73]. The latter developments connect with soil mechanics and the problem of the critical load required to shift a circular pile through a plastic medium [7, 8]. The purpose of the current chapter is to explore further the viscoplastic version of the problem and analyze flows of yield-stress fluids around cylinders. Our eventual goal is to understand how the yield stress impacts the swimming behavior of micro-

organisms (§1.2). To gain preliminary insight, we begin by considering two related problems.

The first is a short reconsideration of the relatively classical problem of the motion of a cylinder with a no-slip surface through a viscoplastic medium. This problem has been approached previously using variational methods [71], numerical computation [74, 75, 76, 77], and laboratory experiments [78], and has applications to the sedimentation of particles through a viscoplastic medium [79]. In the limit of vanishing flow speeds, one expects that this viscoplastic problem reduces to that for the critical load on a circular pile in an ideal cohesive plastic medium. For that critical load problem, Randolph and Houlsby [7] provided an analytical solution using the method of sliplines (the characteristics of the stress field), the no-slip condition corresponding to a fully rough surface. The critical loads found for viscoplastic computations in the limit of no motion do indeed appear to agree with predictions by Randolph and Houlsby [7]. However, the computed velocity field is not consistent with the slipline solution, containing some unexpected rotating plugs [75, 77]. This is a concern because the viscoplastic problem is only expected to reduce to one of perfect plasticity outside any boundary layers wherein viscous effects remain important. The residual plugs are attached to such boundary layers at the surface of the cylinder, perhaps reflecting a pervasive viscous effect. We dissect this issue in order to show that the residual plugs disappear in the plastic limit, and thereby demonstrate that there is no conflict with perfect plasticity.

The second problem we address concerns the motion of cylinders whose surface permits some degree of slip. This situation has also been considered in plasticity theory, with Randolph and Houlsby [7] searching for the critical load on cylinders with partially rough surfaces. Importantly, the ability of the material to slide over the cylinder demands modifications to the slipline field. Unfortunately, the construction provided by Randolph and Houlsby leads to stress and velocity fields that are inconsistent with one another, implying that their slipline field cannot correspond to the true plastic solution [80, 8]. To shed more light on this issue, we consider viscoplastic flow around cylinders with boundary conditions that allow slip, with a

view to approaching the perfectly plastic limit. In so doing, we provide evidence for what is the true plastic solution for these partially rough cylinders. The situation also corresponds to a flow problem wherein sliding is possible or if a thin weakened layer exists sheathing the cylinder, exactly as commonly assumed to explain effective slip [81] and already studied in the context of viscoplastic flow around cylinders [76].

Finally, we consider the locomotion of cylindrical ‘squirmers’ in a viscoplastic fluid. Squirmers are a popular idealization of swimming micro-organisms that have fixed shape but propel themselves using a prescribed surface velocity field that represents the action of ciliary motion [82, 83, 84, 29]. Although most such models are based on spheres, cylindrical squirmers have been considered in Newtonian fluids, to study their interaction with walls or other swimmers [32, 85], or viscoelastic and power-law fluids, to determine their performance in an idealized physiological ambient [86, 33]. The idealized geometry in these cases allows for a first discussion of the complicating additional physics. Our goal here is to explore how these simplified model swimmers perform in a viscoplastic fluid, following on from related locomotion problems in which a yield stress was demonstrated to dramatically alter the swimming dynamics [34, 35]. Thus, we explore the impact of a yield stress on squirming locomotion, exploiting the results for translating cylinders to understand the exposed flow patterns. Our findings address the broader goal of this thesis to understand collective dynamics of swimmers as discussed in § 2.6.

2.2 Mathematical formulation

Neglecting inertia and gravity, we consider a cylinder of radius \mathcal{R} moving through an incompressible Bingham fluid (*e.g.* Balmforth et al. [79]) with a characteristic speed \mathcal{U} . To obtain a dimensionless set of equations, we use \mathcal{R} and \mathcal{U} to remove the dimensions of length and velocity, respectively. Pressure and stresses are scaled by the characteristic viscous stress $\mu\mathcal{U}/\mathcal{R}$ where μ is the (plastic) viscosity of the fluid. In the polar coordinate system (r, θ) with the origin at the center of the cylinder, the governing equations for the dimensionless fluid velocity $(u(r, \theta), v(r, \theta))$ and pressure

$p(r, \theta)$ are given by

$$\frac{1}{r} \partial_r (ru) + \frac{1}{r} \partial_\theta v = 0, \quad (2.1a)$$

$$\partial_r p = \frac{1}{r} \partial_r (r\tau_{rr}) + \frac{1}{r} \partial_\theta \tau_{r\theta} - \frac{\tau_{\theta\theta}}{r}, \quad (2.1b)$$

$$\frac{1}{r} \partial_\theta p = \frac{1}{r^2} \partial_r (r^2 \tau_{r\theta}) + \frac{1}{r} \partial_\theta \tau_{\theta\theta}, \quad (2.1c)$$

where τ_{ij} is the deviatoric stress tensor. We use the Bingham constitutive law,

$$\begin{aligned} \tau_{ij} &= \left(1 + \frac{\text{Bi}}{\dot{\gamma}}\right) \dot{\gamma}_{ij} & \text{for } \tau > \text{Bi}, \\ \dot{\gamma}_{ij} &= 0 & \text{for } \tau \leq \text{Bi}, \end{aligned} \quad (2.2)$$

where

$$\{\dot{\gamma}_{ij}\} = \begin{bmatrix} 2\partial_r u & \partial_r v + (\partial_\theta u - v)/r \\ \partial_r v + (\partial_\theta u - v)/r & 2(\partial_\theta v + u)/r \end{bmatrix}, \quad (2.3)$$

$\dot{\gamma} = \sqrt{\frac{1}{2} \sum_{j,k} \dot{\gamma}_{jk}^2}$ and $\tau = \sqrt{\frac{1}{2} \sum_{j,k} \tau_{jk}^2}$ denote the second tensor invariants. The dimensionless yield stress, or Bingham number, is

$$\text{Bi} = \frac{\tau_Y \mathcal{R}}{\mu \mathcal{U}}. \quad (2.4)$$

The drag force on the cylinder in the x -direction plays an important role, and is defined by

$$F_x = \oint d\theta [(\tau_{rr} - p) \cos \theta - \tau_{r\theta} \sin \theta]_{r=1} \equiv \oint d\theta [2\tau_{rr} \cos \theta + \partial_r (r\tau_{r\theta}) \sin \theta]_{r=1}. \quad (2.5)$$

The plastic drag coefficient C_d is related to this force by $C_d = -F_x/(2\text{Bi})$. Although this coefficient is strictly only relevant in the plastic limit $\text{Bi} \gg 1$, the implied rescaling of F_x is convenient for a wider range of Bi , leading us to use it as a measure of the drag for more general parameter settings.

2.2.1 Boundary conditions

Translating cylinder with a rough or no-slip surface

For a no-slip cylinder moving in the x -direction with unit speed (*i.e.* dimensional speed \mathcal{U}), we impose

$$(u, v) = (\cos \theta, -\sin \theta) \quad \text{at} \quad r = 1. \quad (2.6)$$

Both velocity conditions cannot be applied in ideal plasticity. Instead, a prescribed normal velocity forces plastic deformation with tangential slip along the boundary of the cylinder. At finite, but large Bingham number, one expects any such slip to become smoothed over viscous boundary layers wherein the shear stress dominates the other stress components. If this turns out to be the case, no slip is equivalent to the local stress condition $|\tau_{r\theta}| \sim \text{Bi}$, which is the fully rough surface condition used in plasticity theory.

Translation with slip

If the surface of the cylinder is partially rough, with a roughness factor $\varrho \in [0, 1]$, the boundary condition to be imposed is [7, 8]

$$u = \cos \theta \quad \text{and} \quad \tau_{r\theta} = \varrho \text{Bi} \operatorname{sgn}(y) \quad \text{at} \quad r = 1. \quad (2.7a, b)$$

Setting $\varrho = 1$ corresponds to a fully rough cylinder, and $\varrho = 0$ to a perfectly smooth, or free slip, cylinder. Although it is not necessarily a natural boundary condition for a fluid, the second condition in (2.7) is equivalent to the rate-independent limit of the Mooney-type slip-law

$$v(r = 1, \theta) + \sin \theta = A(|\tau_{r\theta}| - \tau_w)^q \operatorname{sgn}(\tau_{r\theta}), \quad (2.8)$$

for some parameters A , q and wall stress threshold $\tau_w = \varrho \text{Bi}$. Such slip laws are common when modelling effective slip due to surface interactions in many suspen-

sions (*e.g.* Barnes [81]; see also Ozogul et al. [76]).

Squirming surface motions

For a model squirmer, we again impose the surface velocity, this time in the frame of the cylinder, and select \mathcal{U} as its characteristic scale. The speed of the cylinder with respect to the ambient fluid then becomes U_s . We consider purely tangential squirring motions and set

$$(u, v) = (U_s \cos \theta, V_p(\theta) - U_s \sin \theta) \quad \text{at} \quad r = 1, \quad (2.9)$$

where $(0, V_p)$ represents the prescribed surface velocity. For specific examples, we adopt previously employed models of treadmilling cilia given by

$$V_p(\theta) = \sin n\theta + a \sin m\theta, \quad (2.10)$$

with integers n and $m \neq 1$. Notable conventional models include the simplest case, with $(n, a) = (1, 0)$, or employ $(n, m) = (1, 2)$ with $a < 0$ giving a ‘pusher’ and $a > 0$ a ‘puller’ (based on the distribution of $V_p(\theta)$). Note that, although one can generate solutions for any U_s , the swimming speed of a free locomotor is set by the requirement that the net force on the cylinder in the x -direction should vanish; *i.e.* $F_x = 0$ in (2.5).

Finally, we also consider a limited number of examples in which we replace (2.9) with a squirring motion normal to the cylinder surface,

$$(u, v)|_{r=1} = (U_s \cos \theta - U_p(\theta), -U_s \sin \theta), \quad \text{with} \quad U_p = \cos n\theta + a \cos m\theta. \quad (2.11)$$

Although Blake also considered normal surface velocities, he took these as components of propagating wave-like motions, unlike the steady model in (2.11), which is closer to the propulsion mechanism discussed by Spagnolie and Lauga [87].

2.2.2 Numerical method

We solve the governing equations using the augmented Lagrangian scheme summarized by Hewitt and Balmforth [34]. In brief, after the elimination of the pressure from the momentum equations (2.1b)-(2.1c) and the introduction of a stream function $\psi(r, \theta)$ such that

$$(u, v) = \left(\frac{1}{r} \partial_\theta \psi, -\partial_r \psi \right), \quad (2.12)$$

we must solve the biharmonic-like problem

$$\nabla^4 \psi = \text{Bi} \left[\left(\frac{1}{r} \partial_r r \partial_r + \frac{2}{r} \partial_r - \frac{1}{r^2} \partial_\theta^2 \right) \frac{\dot{\gamma}_{r\theta}}{\dot{\gamma}} - \frac{2}{r} \left(\partial_r + \frac{1}{r} \right) \partial_\theta \left(\frac{\dot{\gamma}_{rr}}{\dot{\gamma}} \right) \right], \quad (2.13)$$

over the yielded regions $\dot{\gamma} > 0$. This is achieved by means of an iterative scheme in which one solves, at each step, a linear biharmonic equation over the whole domain (both yielded and plugged) and a nonlinear algebraic problem that incorporates the constitutive law.

We work on the domain $0 \leq \theta \leq \pi$, with symmetry conditions at $\theta = 0, \pi$. The stress invariant decays away from the cylinder, and must eventually fall below the yield stress. We therefore choose a sufficiently large computational domain to contain all the yielded fluid, and set $(u, v) = (0, 0)$ at the edge. If both velocity components are also specified on the surface of the cylinder, the boundary conditions there can be implemented directly in terms of the stream function and its derivatives. The boundary condition in (2.7b), however, imposes the shear stress, which is problematic as the iterative solution of (2.13) requires conditions involving the streamfunction. To surmount this difficulty, we replace (2.7b) by the condition $\dot{\gamma}_{r\theta} = \varrho \dot{\gamma} \text{sgn}(y)$ at $r = 1$, which reduces to (2.7b) where the fluid surface is yielded. If, however, the boundary is plugged, the two conditions are not equivalent. To avoid this inconsistency, in the corresponding computations we used a common regularized constitutive model $\tau_{ij} = \dot{\gamma}_{ij} [1 + \dot{\gamma}^{-1} \text{Bi} (1 - e^{-m\dot{\gamma}})]$, which reproduces the Bingham law in (2.2) for $\dot{\gamma} \gg m^{-1}$, with $m = 10^4$ (this choice of m was sufficiently high that the solutions match those for the unregularized law over the yielded regions, and are insensitive to the precise value

of m). Now the fluid is forced to yield everywhere, the boundary is never plugged, and the alternative boundary condition is always equivalent to (2.7b).

The linear biharmonic equation is solved by exploiting a Fourier sine series in θ , and second-order finite differences in the radial direction. The numerical resolution was chosen to be sufficient to resolve the smallest scales of the problem: the radial grid size was at most 0.003, and at least 512 Fourier modes in θ were used. In some of our computations at the highest Bingham numbers, we used a stretched grid in the radial direction to enhance the resolution in boundary layers near the cylinder's surface.

2.2.3 Ideal plasticity

In the limit $\text{Bi} \rightarrow \infty$, one expects that the viscous stresses become insignificant in comparison to the yield stress outside any boundary layers, implying that yielded material deforms at the yield stress, with $\tau_{ij} = \text{Bi}\dot{\gamma}_{ij}/\dot{\gamma}$. In Cartesian coordinates (x, y) , the stress components can then be written in terms of a local slip angle ϑ as $(\tau_{xx}, \tau_{xy}) = \text{Bi}(-\sin 2\vartheta, \cos 2\vartheta)$. Upon substituting the stress components into the momentum equations ($\nabla \cdot \boldsymbol{\tau} = \nabla p$), the equations are hyperbolic in p and ϑ with the characteristics of the stress field following the sliplines [88],

$$\alpha\text{-lines:} \quad \frac{dy}{dx} = \tan \vartheta, \quad p + 2\text{Bi}\vartheta = \text{constant}, \quad (2.14)$$

$$\beta\text{-lines:} \quad \frac{dy}{dx} = -\cot \vartheta, \quad p - 2\text{Bi}\vartheta = \text{constant}. \quad (2.15)$$

The angle ϑ is the anti-clockwise angle of the α -line as measured from the x -axis. The sliplines are a set of mutually orthogonal lines along which the shear stress is the maximum and the normal stresses are zero. In other words, if $\mathbf{R}(\vartheta)$ denotes the rotation matrix, then,

$$\mathbf{R}(\vartheta) \boldsymbol{\tau} \mathbf{R}(\vartheta)^\top = \begin{bmatrix} 0 & \pm\text{Bi} \\ \pm\text{Bi} & 0 \end{bmatrix}. \quad (2.16)$$

The components of the velocity field along the sliplines (u_α, u_β) satisfy

$$\frac{\partial u_\alpha}{\partial s_\alpha} = \frac{\partial u_\beta}{\partial s_\beta} = 0, \quad (2.17)$$

where s_α and s_β are the arclengths along the respective sliplines. That is, the component of the velocity directed along a particular slipline must be constant.

The plasticity problem can also be formulated in variational terms to establish the following two useful results [88]: first, if the velocity field is not simultaneously calculated, slipline fields that satisfy (2.14) and (2.15), together with any stress boundary conditions, constrain the true solution by providing strict lower bounds on the drag force on the cylinder. Second, trial velocity fields that satisfy the surface velocity and incompressibility conditions, but not the stress relations, place upper bounds on the drag force (given that the associated dissipation rate must balance the power input required to overcome the drag). Such upper bounds can be improved by posing trial velocity fields guided by the slipline fields. Indeed, if the lower and upper bounds then match, the stress and the velocity fields must correspond to those of the actual solution. Note that, in the slipline stress analysis, one must further demonstrate that there is an admissible stress distribution inside any rigid plugs that satisfies both the force balance equations and yield criterion ($\tau < \text{Bi}$).

Randolph and Houlsby [7] exploited these bounding principles for a fully rough cylinder driven through a perfectly plastic medium. In particular, they constructed a slipline solution and a matching velocity field for which the upper and lower bounds agreed. They further showed that an admissible stress distribution could be found for all the unyielded regions. Hence, their construction provides the true plastic solution. For partially rough cylinders, however, their trial velocity field was not consistent with the slipline solution over part of the yielded region, and the correct computation of the upper bound leaves a mismatch with the lower bound [80]. This led Martin and Randolph [8] to suggest an alternative trial velocity field, associated with a different slipline solution, that lay closer to, but not coincident with the lower bound. The true solution for partially rough cylinders has therefore not been previously identified.

2.3 Revisiting flow around a no-slip cylinder

In this section, we analyze the viscoplastic flow around a no-slip or fully rough cylinder, focussing on high Bingham number. Figure 2-1 shows a numerical solution for $\text{Bi} = 2^{14}$. Plotted is the strain rate, with the regions shaded black corresponding to the plugs, together with the Randolph and Houlsby [7] slipline solution. Three types of plugs appear in the numerical solution, as found previously [74, 75, 77]: first, the ambient medium plugs up sufficiently far from the cylinder to localize the flow. Second, triangular plugs are attached to the front and back of the cylinder. Finally, two plugs with almost semi-circular shape rotate rigidly near the top and bottom of the cylinder. Only the first two types of plugs feature in the perfectly plastic solution; the rigidly rotating plugs lie in the region of perfectly plastic deformation in the slipline solution where there is always shear.

2.3.1 Randolph and Houlsby's slipline solution

In detail and for the upper half of the solution, the slipline pattern (figure 2-1b) consists of a semi-circular centered fan at the top of the cylinder with center A at $(0, 1)$ and radius $1 + \pi/4$. The β -lines form the spokes and α -lines form the circular arcs. The α -lines are continued below the line AD by the involutes of the cylinder, and the β -lines become tangents. The construction of the involutes ensures that the stress field satisfies the fully-rough boundary condition, $\tau_{r\theta} = \text{Bi}$, on the cylinder surface. The limiting β -lines BC and $B'C'$ intersect the x -axis at 45° , as demanded by symmetry, which isolates the triangular plugs capping the front and back of the cylinder. The α -line $CDGD'C'$ determines the outermost yield surface.

The velocity field associated with the slipline pattern is directed purely along the α -lines (and so, in this case, the streamlines are α -lines): the involutes beginning along BC have $v_\alpha = 1/\sqrt{2}$, whereas those that begin at the cylinder along AB have $v_\alpha = \cos \theta$. At the base of both sets of sliplines, there is a velocity jump tangential to ABC . Similarly, along the outermost yield surface $CDGD'C'$, another velocity jump arises. In the viscoplastic computation, all these discontinuities become broadened

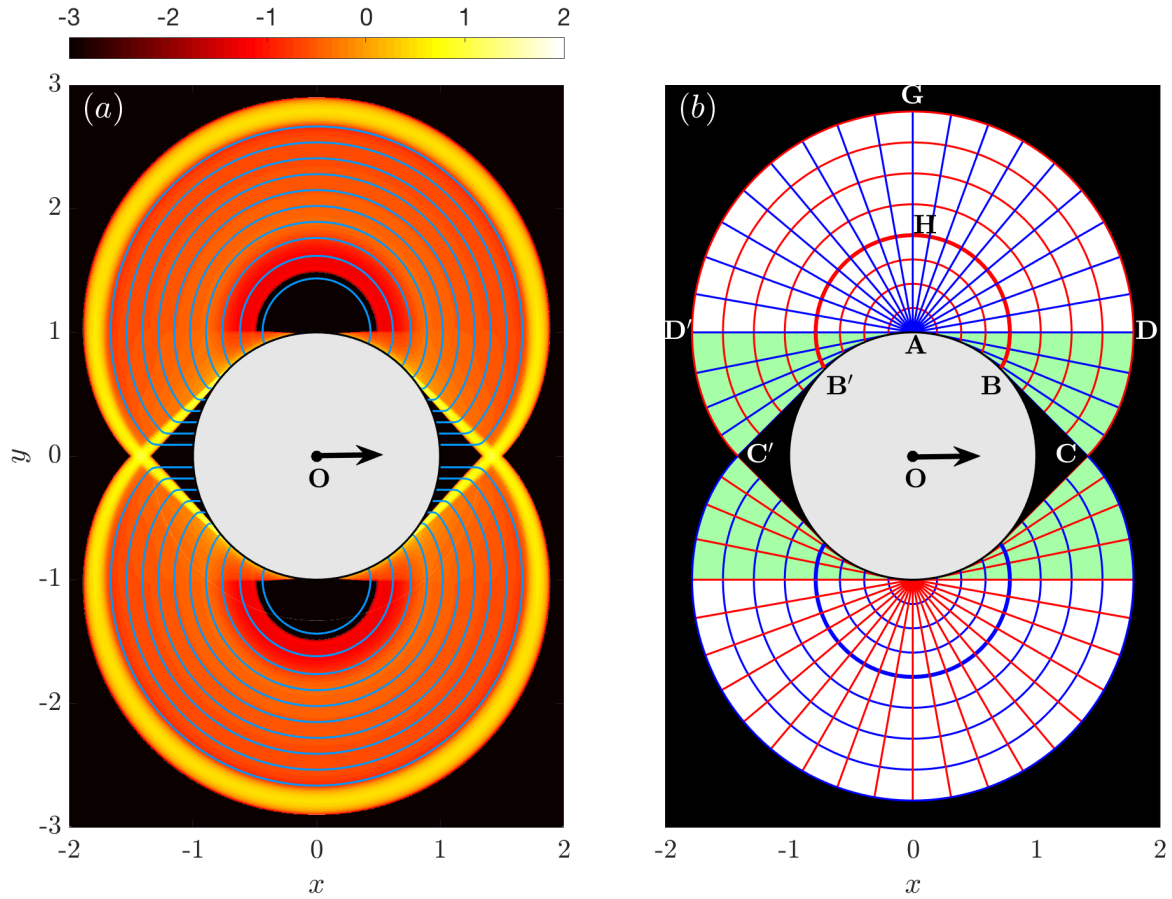


Figure 2-1: Comparison of the numerical solution and the slipline pattern for a fully rough (no-slip) cylinder. (a) Color plot of the logarithm of the strain rate ($\log_{10}(\dot{\gamma})$) from the numerical solution at $\text{Bi} = 2^{14} \gg 1$, together with streamlines (blue). (b) The slipline pattern made up of α - and β -lines (red and blue; § 2.2.3) obtained by Randolph and Houlsby [7] (§ 2.3.1). In the white region, the α -lines form circular arcs with center A, and in the green region, the β -lines form the involutes to the circle. Arrows indicate the direction of motion of the cylinder. In (a) and (b), the overall patterns and the rigid plugs (with zero strain rate; shaded black) match except for the rigidly rotating ones on the top and bottom of the cylinder. These residual plugs appear due to the inconsistency of pressure change in the thin viscoplastic boundary layer above the cylinder and the overlying plastic solution. As described in § 2.3.2, they eventually vanish, albeit very slowly, with increasing Bingham number Bi , thus confirming that the viscoplastic solution converges to the plastic one.

into thin boundary layers with enhanced shear rate (see figure 2-1a and figure A-1 for a magnification of the boundary layer attached to the cylinder). The thickness of these layers is expected to scale with either $\text{Bi}^{-1/3}$ or $\text{Bi}^{-1/2}$ [89], but otherwise they leave no enduring viscous disfigurement of the plastic solution.

Along the β -lines, the Riemann invariant is $p - 2\text{Bi}\vartheta$. If we set $p = 0$ along the vertical symmetry line at $x = 0$, this implies $p = 2\text{Bi}(\pi - \vartheta)$ throughout the deformed region, and so the pressure on the surface of the cylinder is given by

$$p(1, \theta) = \begin{cases} 2\text{Bi}(\pi - \theta), & \pi/4 < \theta < \pi/2, \\ 2\text{Bi}(\pi - \theta) - 2\text{Bi}\pi, & \pi/2 < \theta < 3\pi/4. \end{cases} \quad (2.18)$$

which jumps by $2\pi\text{Bi}$ at the centre of the fan.

With the stresses implied by the slipline pattern, we may integrate over the contour $CBAB'C'$ to determine the net horizontal force on the upper half of the cylinder F_x (although the stress field is not prescribed over the plugs, the net force on these regions must vanish, and so the horizontal force along BC or $B'C'$ must equal that along the corresponding plugged section of the cylinder's surface). We then find the drag coefficient [7],

$$C_d = -\frac{F_x}{2\text{Bi}} = 2(\pi + 2\sqrt{2}) \simeq 11.94. \quad (2.19)$$

2.3.2 The residual plugs

The rotating plugs of the viscoplastic computation in figure 2-1a are centred at the fans of the slipline solution and are attached to the viscous boundary layer buffering the cylinder surface. Since the viscous stress is prominent in that boundary layer, the question arises as to how the pressure jump at the centre of the fan becomes smoothed and whether this prompts a permanent adjustment of the slipline solution that explains the rotating plugs. Indeed, both Tokpavi et al. [75] and Chaparian and Frigaard [77] have suggested that these features are permanent for $\text{Bi} \rightarrow \infty$. Such a conclusion is problematic as it implies that the viscoplastic theory does not converge to perfect plasticity.

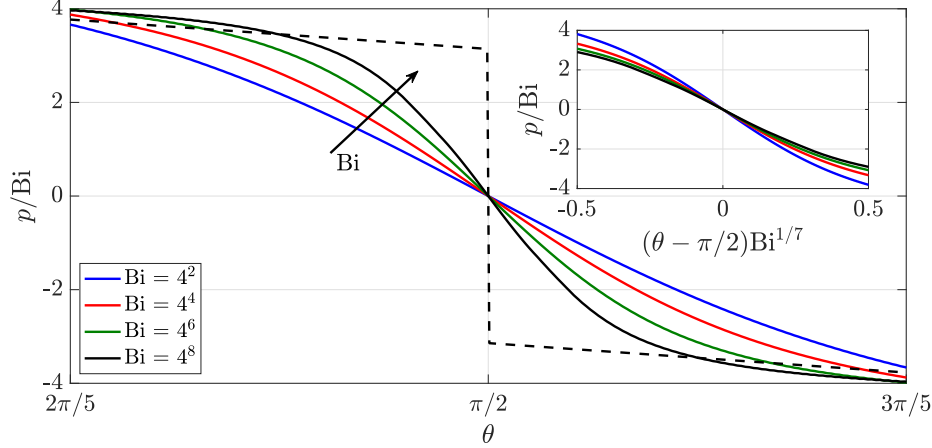


Figure 2-2: Pressure variation over the cylinder scaled by Bi , for the Bingham numbers indicated. The dashed line corresponds to the pressure of the slipline solution given by (2.18); the numerical profile approaches this solution as the Bingham number increases. The inset shows overlapping profiles of p/Bi against $(\theta - \pi/2)\text{Bi}^{1/7}$, which confirms the scaling predicted from the boundary layer theory in appendix A.1.

The current computations suggest an alternative perspective: the rotating plugs correspond to a persistent effect that arises from the pressure discontinuity of the slipline solution at the centre of the centred fans. Because fluid flows through the pressure gradient here, the discontinuity must necessarily become smoothed by viscous stresses over a narrow window of angles θ surrounding A . The angular scale of this smoothing regions turns out to be relatively wide (in comparison to the viscous boundary layers), scaling very weakly with Bi ; see figures 2-2 and 2-3d. Moreover, to accommodate the smoothing of the pressure over this scale (which is too wide to allow any viscous adjustments), the overlying plastic flow must plug up, thereby creating the persistent features. Crucially, the size of the plug therefore asymptotically decreases to zero, albeit extremely slowly, as $\text{Bi} \rightarrow \infty$ (see figure 2-3a; we find, in particular, that the radius decreases like $\text{Bi}^{-3/28}$). Consequently, the drag coefficient should approach the prediction in (2.19) for $\text{Bi} \rightarrow \infty$, as illustrated by the numerical results (figure 2-4a).

A number of other numerical results are shown in figure 2-3, including the rotation rate of the plug and the thickness of the boundary layer against the cylinder. Notably, directly under the plug, the boundary layer is thinned by the presence of the

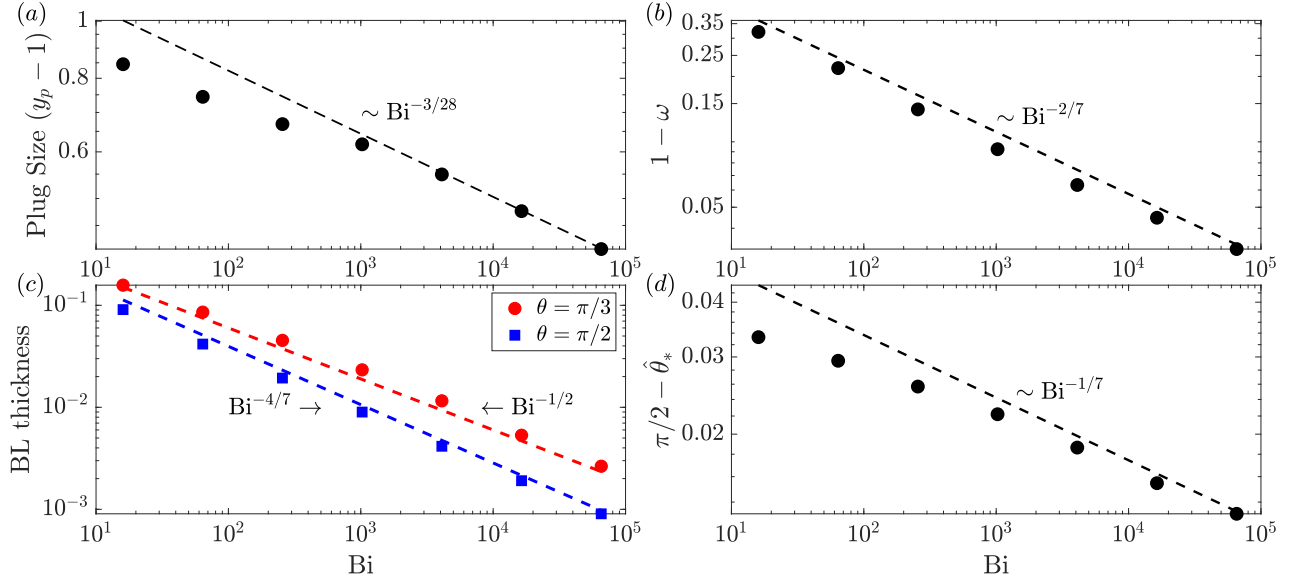


Figure 2-3: Scaling data for the residual rotating plug against Bi , showing (a) the plug radius $y_p - 1$ where $(0, y_p)$ is the the top of the plug, (b) it's rotation rate ω , (c) the boundary layer thickness at $\theta = \frac{1}{2}\pi$ and $\frac{1}{3}\pi$, and (d) the angular size of the smoothing region, estimated by the location $\hat{\theta}_*$ for which $p = \frac{1}{2}Bi$. All these plug characteristics follow the scalings predicted by the boundary layer theory of appendix A.1 (dashed lines).

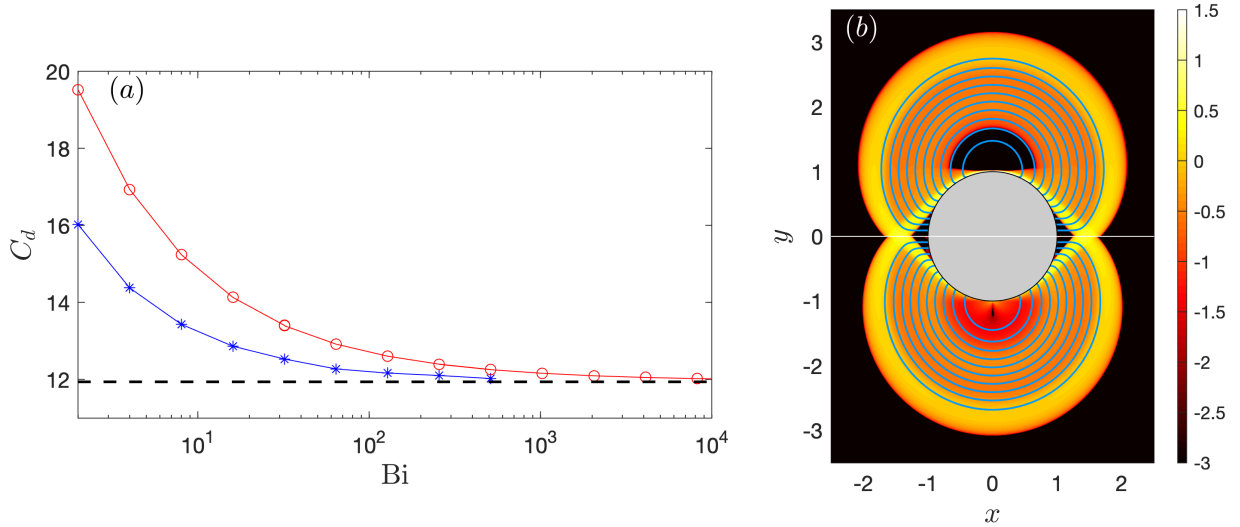


Figure 2-4: (a) Drag coefficient C_d against Bi for computations using a no-slip boundary condition (red circles) or the plastic slip law in (2.7) with $\varrho = 1$ (blue stars). The dashed line shows (2.19). (b) Plots of $\log_{10} \dot{\gamma}$ and streamlines for two solutions at $Bi = 2^8$; the upper half shows the no-slip cylinder, and the lower half the cylinder with (2.7) and $\varrho = 1$. Due to the absence of the no-slip condition and the resulting viscoplastic boundary layer against the cylinder, the rigidly rotating plug disappears in (b, bottom) as compared to (b, top) [§ 2.3.2].

smoothing region, scaling with $\text{Bi}^{-4/7}$. Beyond this region, the boundary layer has a thickness of $O(\text{Bi}^{-1/2})$, as expected from viscoplastic boundary-layer theory [89]. The thinner boundary layer scaling near A is in disagreement with the conclusions of Tokpavi et al. [75], although the difference between $-1/2$ and $-4/7$ is small (these authors actually find a scaling of -0.53). A boundary-layer theory in support of the observed scalings and the overall phenomenology of the rotating plug is provided in appendix A.1.

A key feature of the boundary-layer theory is that the slowly converging scalings of the rotating plug arise from the conspiracy between the flow within the boundary layer and the overlying plastic deformation. The important role played by the boundary layer therefore implies that the passage to the plastic limit should be different if that sharp feature is not present. Indeed, when we recompute the solutions using the slip law outlined in §2.2.1 (with $\varrho = 1$, corresponding to a fully rough surface), the boundary layers against the cylinder are removed as all the tangential slip that is required for the adjacent perfectly plastic deformation can be taken up along the boundary itself. No slowly shrinking plugs then appear at the center of the fans whatsoever and the convergence to the plastic limit is noticeably accelerated (see figure 2-4).

2.4 Flow past a partially rough cylinder

Numerical solutions for partially rough cylinders, with boundary condition (2.7), are shown in figures 2-5 and 2-6. The first of these figures displays strain-rate plots for two sample solutions with different roughness factors $\varrho = 0$ (free slip) and $\varrho = \frac{1}{2}$. Aside from viscoplastic shear layers that smooth out the velocity jumps, these numerical solutions are very like the slipline solution proposed by Martin and Randolph [8] which are also plotted in the figure and described in more detail below. Notably, the solutions now contain rigidly rotating plugs that are permanent features in the plastic limit $\text{Bi} \rightarrow \infty$, and which attach directly on to the sliding cylinder surface.

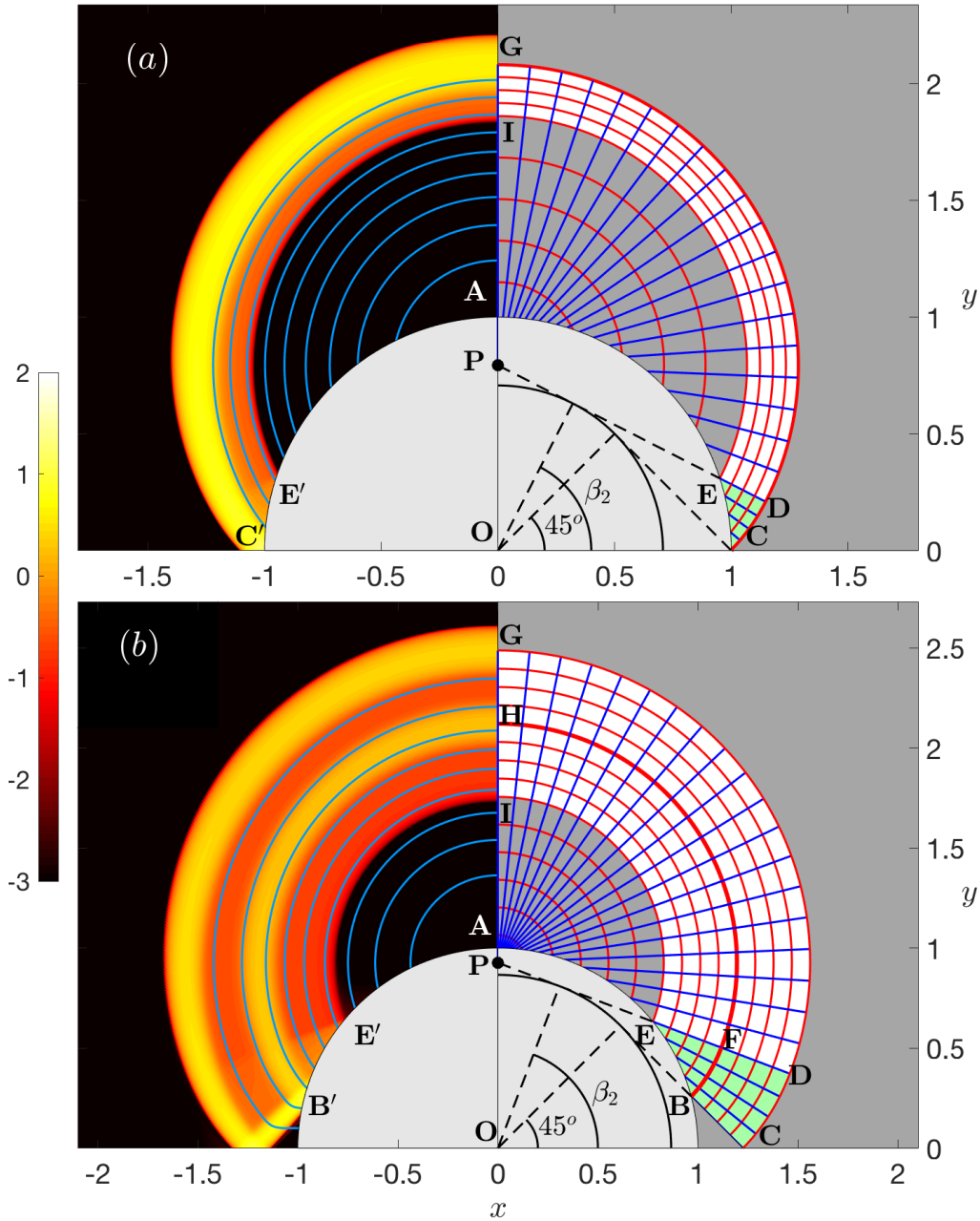


Figure 2-5: Numerical solutions (left) showing $\log_{10} \dot{\gamma}$ with streamlines (light blue) and slipline solutions [right; same color scheme as in figure 2-1] for (a) $\varrho = 0$ and (b) $\varrho = 0.5$, both at $\text{Bi} = 2^{12}$. In the slipline construction, the angle β_2 can be variable and dictates the angular extent of the circular fan with center P . Here, its value is set according to (2.21) [63.0° in (a) and 69.2° in (b)], where the lower and upper bounds match [see figure 2-6]. At these values, the slipline pattern compares well with the numerical solution, with the velocity jumps [along arc GD in (a, left) and arcs BFH and CDG in (b, left)] replaced by viscoplastic shear layers [bright yellow regions in (a, b; left)]. This provides numerical evidence that the slipline pattern is the true plastic solution. The primed points, B' to F' , are the reflections of points, B to F , about the y -axis.

2.4.1 The Martin and Randolph slipline solution and upper bound

As for Randolph and Houlsby's slipline field shown in figure 2-1, the pattern proposed by Martin and Randolph [8] consists of centred fans and involutes that leave a triangular plug at the front and back of the cylinder. However, the centres of fans are now displaced off the surface and the cores of the fans are replaced by the rigidly rotating plugs. Focussing on the upper right half of the pattern, the fan is centred at point P and occupies the region $EFDD'E'$ in figure 2-5. The rotating plug spans $AEIE'$. The involutes that extend the α -lines from the fan into $EBCDF$ correspond to β -lines that are tangent to an inner circle centred at O with radius

$$\lambda = \cos \left(\frac{\cos^{-1} \varrho}{2} \right). \quad (2.20)$$

This choice for λ ensures that the α -lines meet the surface of the cylinder at an angle $(\pi/4 - \Delta/2)$, where $\Delta = \sin^{-1} \varrho$, in line with the boundary condition (2.7b) on EB , $|\tau_{r\theta}|/\text{Bi} = \varrho$. In other words, the unwrapping of the β -lines from the inner circle ensures that the slip condition is satisfied along the yielded boundary of the cylinder, and follows Randolph and Houlsby's original generalization of figure 2-1 for $\varrho < 1$. The main difference between their generalization and the construction of Martin and Randolph is the introduction of the rigidly rotating plugs at the cores of the fans. Such plugs are permitted because any slipline can be taken as a yield surface and the normal velocity across the sliding, unyielded boundary can be made continuous by demanding that the rotation rate of the plugs is $\sin \beta_2/\lambda$, where $\pi - \beta_2$ dictates the angular extent of the fan (the angle between IPE). The introduction of the plugs then shifts the centre of the fan P so that it lies a vertical distance $\lambda/\sin \beta_2$ above O .

Again, the velocity field over the plastic region is prescribed by $v_\beta = 0$ and matching v_α with the normal velocity to the contour EBC . Velocity jumps thereby occur along the α -lines BFH and CDG , which broaden into the prominent viscoplastic shear layers of the computations in figure 2-5, and fluid slides along the cylinder boundary AEB .

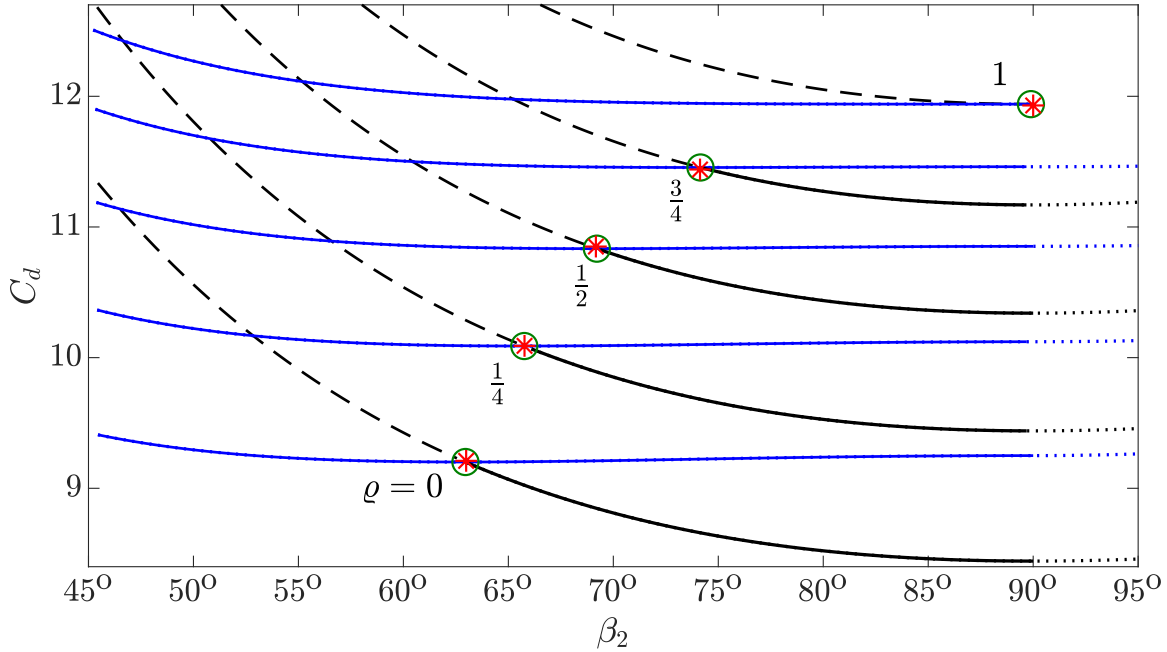


Figure 2-6: Upper (blue) and lower (black) bounds derived from the Martin and Randolph [8] slipline solution for the values of ϱ indicated; the circles indicate the minimum of the upper bound, which coincides with the intersection of the lower and upper bounds. For β_2 less than this minimum, the lower bound exceeds the upper bound and is thus spurious (shown dashed). For $\beta_2 > \frac{1}{2}\pi$, the velocity and stress fields become inconsistent, as in the original Randolph and Houlsby [7] construction (the corresponding bounds are shown by dotted lines). The red stars show the extrapolated values for $\text{Bi} \rightarrow \infty$ of the drag and angle β_2 from numerical computations. The excellent match between the numerical computations (red stars) and the intersection of the lower and upper bound curves indicates that the slipline solution at the corresponding β_2 is the true plastic solution. [The angle β_2 is estimated from the numerical computations by matching the rotation rate of the rigid plugs with the expression $\sin \beta_2 / \lambda$. This avoids tracing the yield surfaces which are sensitive to numerical errors.]

Martin and Randolph [8] treat the angle β_2 as an optimization parameter that can be adjusted to vary the upper bound on the drag force computed from the net dissipation rate incurred by the velocity field. The smallest possible drag coefficient provides the best upper bound, as illustrated in figure 2-6, which plots the upper bound against β_2 for a number of choices of the roughness factor ϱ . (Martin and Randolph [8] also include the inclination of the triangular plug and the radius λ as further optimization parameters; these turn out to be optimized by the choices of 45° and (2.20), respectively, both of which are in any case demanded by the boundary and symmetry conditions.) Note that, as $\beta_2 \rightarrow \frac{1}{2}\pi + \cos^{-1} \lambda$, the rotating plug disappears and the slipline construction reduces to that of Randolph and Houlsby [7].

2.4.2 Lower bound and torque balance

The stress field of the slipline solution can also be used to compute the drag force which, in principle, sets a complementary lower bound. For this task, we again set $p = 0$ along the y-axis, implying $p = 2\text{Bi}\pi - 2\text{Bi}\vartheta$ in the regions of deformation. With that pressure field and the known slipline angle, we may then calculate the drag force on the cylinder by integrating over the contour $IEBC$. The details of this calculation are provided in appendix A.2.1. This calculation is incomplete, however, because we do not extend the stress field into the plugs to demonstrate that an admissible solution that satisfies the yield condition exists there. Nevertheless, the construction of Randolph and Houlsby can be used to find admissible stress fields for both the triangular plugs at the front and back of the cylinder and the surrounding stagnant plug. The only missing piece of the puzzle is therefore the establishment of an admissible stress distribution for the rotating plugs.

Modulo this limitation, the implied lower bound on C_d is also plotted in figure 2-6 for comparison with the upper bound. The lower bound passes through the upper bound at exactly its minimum. That is, the upper bound and lower bounds match each other at the optimal choice for β_2 , which suggests that the corresponding slipline solution is the actual true solution. However, for smaller values of β_2 (indicated by dot-dashed lines in the figure), the lower bound calculation yields a higher value than

the upper bound, which is not possible. This flaw must have its origin in the lack of an admissible stress field for the rotating plugs.

The slipline construction also shares the same issue of incompatibility suffered by the original Randolph and Houlsby solution (which must be the case, given that the construction reduces to this solution in the limit $\beta_2 \rightarrow \frac{1}{2}\pi + \cos^{-1} \lambda$): for $\beta_2 > \frac{1}{2}\pi$ the shear stresses along the α -lines are not consistent with the corresponding shear rates everywhere throughout the deforming region. Nevertheless, this inconsistency does not affect the optimal solution for β_2 .

More light can be shed on the rotating plug by considering the balance of torques acting on this region (appendix A.2.2). In particular, and with reference to figure 2-5b, the upper plug is the crescent formed from the two circular arcs EAE' and EIE' . Along these arcs, the shear stresses are ϱBi and $-\text{Bi}$, respectively, which imply the torques $T_{EAE'} = 2\varrho \text{Bi}(\frac{1}{2}\pi - \theta_E)$ and $T_{EIE'} = -2r_{EI}^2 \text{Bi}(\pi - \beta_2)$, acting about the centres of the respective circles (*i.e.* P and O). Here, $\theta_E = \beta_2 - \frac{1}{4}\pi + \frac{1}{2}\Delta$ is the polar angle of point E and $r_{EI} = \lambda \cot \beta_2 + \sqrt{1 - \lambda^2}$ is the radius of the circular arc EIE' . In addition to these torques, the difference between the two horizontal forces on the arcs provides a moment that also acts on the plug. This moment is $-\lambda F_{EIE'}/\sin \beta_2$, if $F_{EIE'}$ is the horizontal force on the section EIE' , which must be equal and opposite to the force on EAE' if the plug is in force balance. But $F_{EAE'} = 2F_{AE}$, where $F_{AE} = -r_{EI} \text{Bi}[2(\pi - \beta_2) \cos \beta_2 + \sin \beta_2]$ is the force on the section AE (see appendix A.2.1). Hence, the rotating plug is free of torques if $T_{EIE'} - T_{EAE'} - 2\lambda F_{AE}/\sin \beta_2 = 0$, or

$$[(\pi - \beta_2)(\sqrt{1 - \lambda^2} - \lambda \cot \beta_2) - \lambda](\lambda \cot \beta_2 + \sqrt{1 - \lambda^2}) + \varrho(\frac{3}{4}\pi - \beta_2 + \frac{1}{2}\Delta) = 0. \quad (2.21)$$

This condition picks out a unique value for β_2 which coincides exactly with the optimal value. We conclude that there cannot be an admissible stress field for the rotating plug, except potentially at the torque-free value of β_2 . Thus, Martin and Randolph's slipline field with this choice is the only candidate for the true plastic solution. This

conclusion is supported by the numerical computations, which match the predictions of the slipline theory for a variety of choices for the roughness parameter ϱ (see figures 2-5 and 2-6), and which explicitly construct admissible stress solutions for the rotating plug. Thus, the combination of slipline theory and numerical computation provides evidence that the slipline patterns of figure 2-5 are the actual perfectly plastic solutions.

2.5 Viscoplastic squirmers

We now consider models for swimming micro-organisms driven by ciliary surface motions in a yield-stress fluid. More specifically, we adopt prescribed surface velocity patterns to drive locomotion, as outlined in §2.7, focussing primarily on tangential motions with $V_p = \sin n\theta + a \sin m\theta$ and $m > 1$. We also briefly consider swimming patterns comprising a normal surface velocity.

2.5.1 Newtonian limit

As described by Blake [83], it is straightforward to solve the Stokes problem in the Newtonian limit to furnish the streamfunction,

$$\psi = \psi_0 = U_s r \sin \theta + \frac{1}{2}(r^{-n} - r^{2-n}) \sin n\theta + \frac{1}{2}a(r^{-m} - r^{2-m}) \sin m\theta \quad (2.22)$$

for the prescribed tangential surface motion. This result incorporates the Stokes paradox in that there is no bounded solution for $r \rightarrow \infty$ unless $U_s = \frac{1}{2}$ and $n = 1$. Moreover, the solution implies that F_x in (2.5) is identically zero. A similar result holds if the normal surface velocity is prescribed [83].

Following Hewitt and Balmforth [35], we may proceed beyond this leading solution and compute the correction prompted by the yield stress using perturbation theory. In the vicinity of the cylinder ($r = O(1)$) this correction is forced by the need to match the solution with that in the far field ($r \gg 1$), as in the classical resolution of the Stokes paradox by the inclusion of inertia. Here, however, the far field region

is controlled by the yield stress. In particular, balancing the two sides of (2.13) for $r \gg 1$, we must have that $\psi = O(r^2 \text{Bi})$ in the far field. But, provided $U_s \neq \frac{1}{2}\delta_{n1}$, ψ_0 grows like r . Hence, the far-field balance demands that $r = O(\text{Bi}^{-1})$. Moreover, the yield stress eventually arrests motion here, limiting the flow to a yielded region with a radial extent of $O(\text{Bi}^{-1})$.

The correction to the near-field solution again satisfies the biharmonic equation and is proportional to $2r \log r - r + r^{-1}$, in view of the boundary conditions on the cylinder and the need to discard terms that grow any more rapidly with r [90]. This correction breaks asymptotic order and becomes of comparable size to ψ_0 as one enters the far field, leading to the estimate, $\psi - \psi_0 = O((U_s - \delta_{n1})/\log \text{Bi}^{-1})$. Hence, the horizontal drag force, which is dictated by the correction, can be calculated and the match to the far field demands

$$F_x = -4\pi \frac{(U_s - \frac{1}{2}\delta_{n1})}{\log \text{Bi}^{-1}}. \quad (2.23)$$

Evidently, the cylinder is force-free to leading order when $U_s = \frac{1}{2}\delta_{n1}$, which is the locomotion speed of a free swimmer in the limit $\text{Bi} \rightarrow 0$. Thus, in the Newtonian limit, surface motions without a $\sin \theta$ component cannot swim. Note that, because the streamfunction decays more rapidly when $U_s - \frac{1}{2}\delta_{n1} \rightarrow 0$, all the preceding scalings must change for the force-free case.

2.5.2 Symmetries

With the surface velocity condition $V_p = \sin n\theta + a \sin m\theta$, the problem can inherit spatial symmetries that constrain the solutions. First, if the driving angular velocity pattern contains multiple lines of reflection symmetry, then the problem with $U_s = 0$ is invariant under a set of finite angular rotations. This implies that the driving pattern possesses no preferred swimming direction and force-free states with $U_s = 0$ exist whatever the Bingham number. Single mode patterns with $n > 1$ and $a = 0$ are of this type.

Second, when m is even and n is odd, the equations and boundary conditions

are invariant under the transformation, $(U_s, a, \theta, u, v, \psi) \rightarrow (U_s, -a, \pi - \theta, -u, v, \psi)$. This implies that, given a solution with a certain U_s and $V_p = \sin n\theta + a \sin m\theta$, one can generate another solution with the same translation speed but driving pattern $V_p = \sin n\theta - a \sin m\theta$; the two solutions are symmetrical under reflection about $\theta = \frac{1}{2}\pi$. Consequently, the force-free swimming speed is independent of the sign of a . Similarly, the transformation $(U_s, a, \theta, u, v, \psi) \rightarrow (-U_s, -a, \pi - \theta, u, -v, -\psi)$ again leaves the system invariant if m is odd and n is even. Thus, in this case, for a swimmer with (a, U_s) , there is another with $(-a, -U_s)$.

Finally, for the alternative driving pattern $U_p = \cos n\theta + a \cos m\theta$ in (2.11) with $(n, a) = (1, 0)$, if we set $U_s = 1 - \hat{U}_s$, the boundary condition becomes $(u, v)_{r=1} = (\hat{U}_s \cos \theta, \sin \theta - \hat{U}_s \sin \theta)$. But this is identical to the driving pattern in (2.9) for $V_p = \sin \theta$ and translation speed \hat{U}_s . In other words, these particular squirmer solutions are identical, except for the switch in translation speed.

2.5.3 Numerical results

To gain a broader perspective on the problem, we now solve the system of equations numerically, calculating solutions with different (fixed) translation speeds U_s and Bingham numbers Bi , for a variety of driving patterns. To begin, we consider the simplest case, with $(n, a) = (1, 0)$ (figure 2-7). The most obvious feature of the computed flow patterns is their similarity to those around translating cylinders: in all but the example with highest translation speed in figure 2-7, the flow is localized to a region with a radial extent that is comparable to the diameter of the cylinder, and prominent recirculation cells with embedded rotating plugs appear above and below. In fact, at higher Bingham number, the organization of the flow looks identical to the Randolph and Houlsby slipline pattern (figure 2-1), with simply a different velocity distribution along the α -lines. Notably, the tangential surface forcing strengthens the boundary layers which are now able to adjust the plastic deformation beyond. As a result, the rotating plugs continue to widen with increasing Bi and become permanent in the plastic limit. The collapse to the Randolph and Houlsby stress field is reflected in the drag coefficient, which equilibrates to $C_d = -(2\pi + 4\sqrt{2})$ for

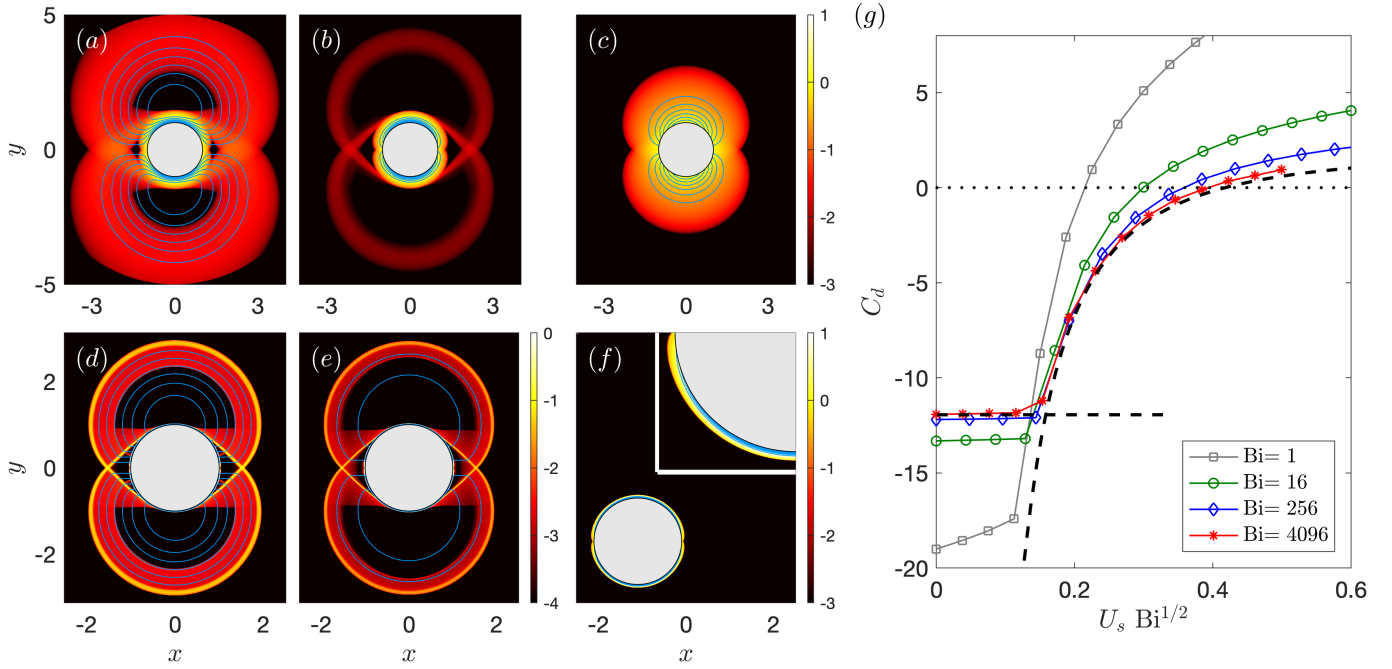


Figure 2-7: Squirmer solutions for $(n, a) = (1, 0)$ showing $\log_{10} \dot{\gamma}$ overlain by streamlines (blue), with $\text{Bi} = 1$ (a–c) and $\text{Bi} = 2^8$ (d–f), for $U_s \text{Bi}^{1/2} \approx 0$ (a, d), 0.1 (b, e), 0.38 (c, f), respectively. In (f), the full circle has the scale of the axes as in (d) and (e), whereas the quarter circle shown in the inset is a magnification to highlight the thin boundary layer. (g) Numerical drag coefficients C_d against scaled translation speed $\text{Bi}^{1/2} U_s$ match well with the asymptotic predictions for $\text{Bi} \gg 1$ from the Randolph and Houlsby slipline solution (§3) and the boundary-layer analysis of § 2.5.4 (dashed black lines). The $C_d = 0$ line (dotted) corresponds to the swimming speed of force-free squirmers.

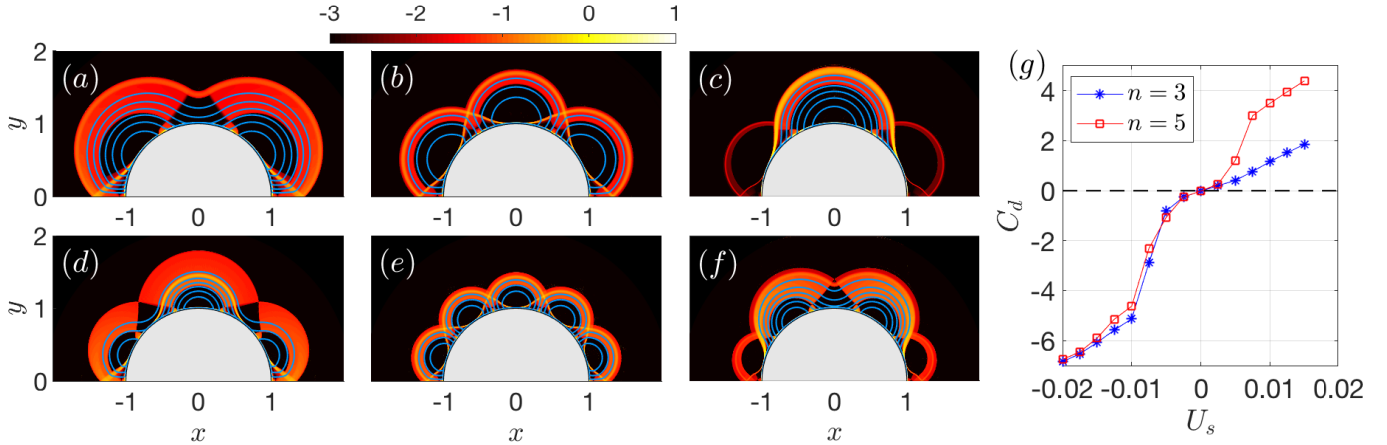


Figure 2-8: Squirmer solutions showing $\log_{10} \dot{\gamma}$ at $\text{Bi} = 2^8$ for (a–c) $n = 3, m = 0$ and (d–f) $n = 5, m = 0$ with the imposed swimming speed (a,d) $U_s = -0.015$, (b,e) $U_s = 0$ and (c,f) $U_s = 0.015$. (g) The variation of the drag coefficient C_d with U_s for $n = 3$ (blue stars) and $n = 5$ (red squares). Since $n > 1$ and $m = 0$, $C_d = 0$ [dashed line in (g)] is necessary for $U_s = 0$ due to multiple lines of reflection symmetry (see § 2.5.2).

$U_s \lesssim 0.16\text{Bi}^{-1/2}$ at high Bi (figure 2-7g). With higher translation speeds, however, the extent of the yielded region abruptly decreases, with all deformation becoming consumed by the boundary layer around the cylinder for $\text{Bi} \gg 1$ (figure 2-7f). The switch in flow pattern prompts a fall in the magnitude of the drag coefficient, which passes through zero at a critical value, U_s^{ff} , corresponding to the locomotion speed of a swimmer moving under its own power.

Flow fields around single-mode squirmers with higher n are shown in figure 2-8. As expected from the symmetries of the problem, the drag coefficient for these solutions vanishes only for $U_s = 0$, precluding locomotion at any Bi (see §2.5.2). In the force-free states, the flow patterns take the form of a straightforward geometrical generalization of the Randolph and Houlsby slipline field, containing a network of $2n$ centred fans with angular extent $\pi(\frac{1}{2} + n^{-1})$ and triangular plugs attached to the cylinder. These solutions become distorted by translation, but the patchwork of attached plugs and rotating fans persists, with broader seams of more complicated plastic deformation. Again, the fans contain persistent plugs, sometimes becoming displaced from the cylinder surface in the manner of the Martin and Randolph slipline field.

The force-free swimming states can be computed directly by employing an interval-bisection algorithm to vary U_s until $C_d = 0$. Figure 2-9(a–h) shows the output of this algorithm for both the simple swimmer with $(n, a) = (1, 0)$, and for pushers and pullers with $(n, m) = (1, 2)$ and varying a . As discussed in §2.5.2, since n is odd and m is even in this case, the solutions with $a < 0$ are reflections of those with $a > 0$ about the vertical axis, with the same swimming speed. Thus, as in the Newtonian limit, these pushers and pullers always travel at equal speeds. In all cases, the swimming speed converges to the Newtonian limit $U_s = \frac{1}{2}$ for $\text{Bi} \rightarrow 0$ (see §2.5.1 and Ref. [83]). At large yield stress, the swimming speeds instead decline as $U_s \sim \text{Bi}^{-1/2}$ (see figure 2-9i). The corresponding force-free flow patterns remain confined to the surface boundary layers at low values of a . But when this parameter is larger, the states becomes less confined and again adopt a wider scale pattern of plastic deformation with the form of a patchwork of triangular plugs and fans,

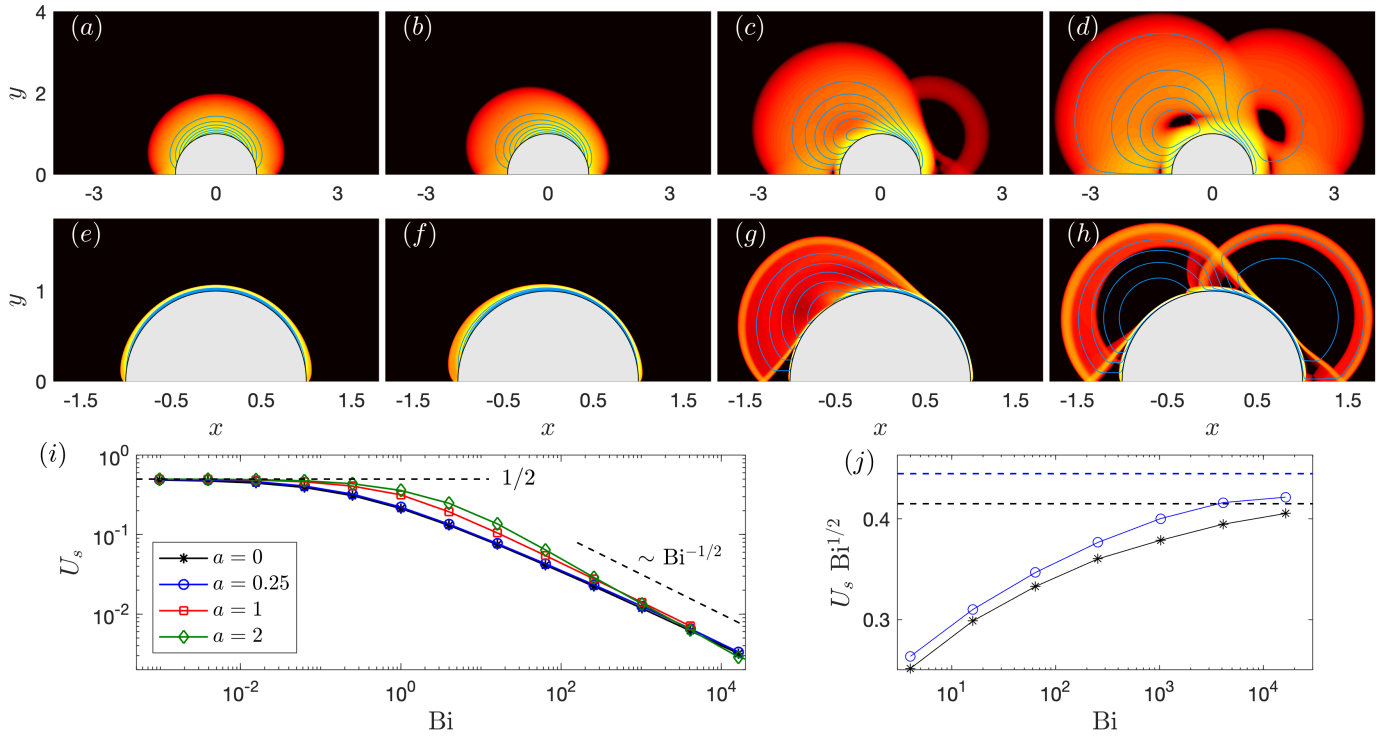


Figure 2-9: Force-free squirmer solutions showing $\log_{10} \dot{\gamma}$ (same color bar as in figure 2-8; streamlines, blue) for surface velocities with $n = 1$, $m = 2$. (a-d) $\text{Bi} = 1$ and (e-h) $\text{Bi} = 2^8$, for (a,e) $a = 0$, (b,f) $a = 0.25$, (c,g) $a = 1$ and (d,h) $a = 2$. (i) The corresponding swimming speeds. (j) Rescaled swimming speeds for $a = 0$ and $a = 0.25$ asymptotically approach predictions of the boundary-layer theory (§2.5.4, dashed) as $\text{Bi} \rightarrow \infty$.

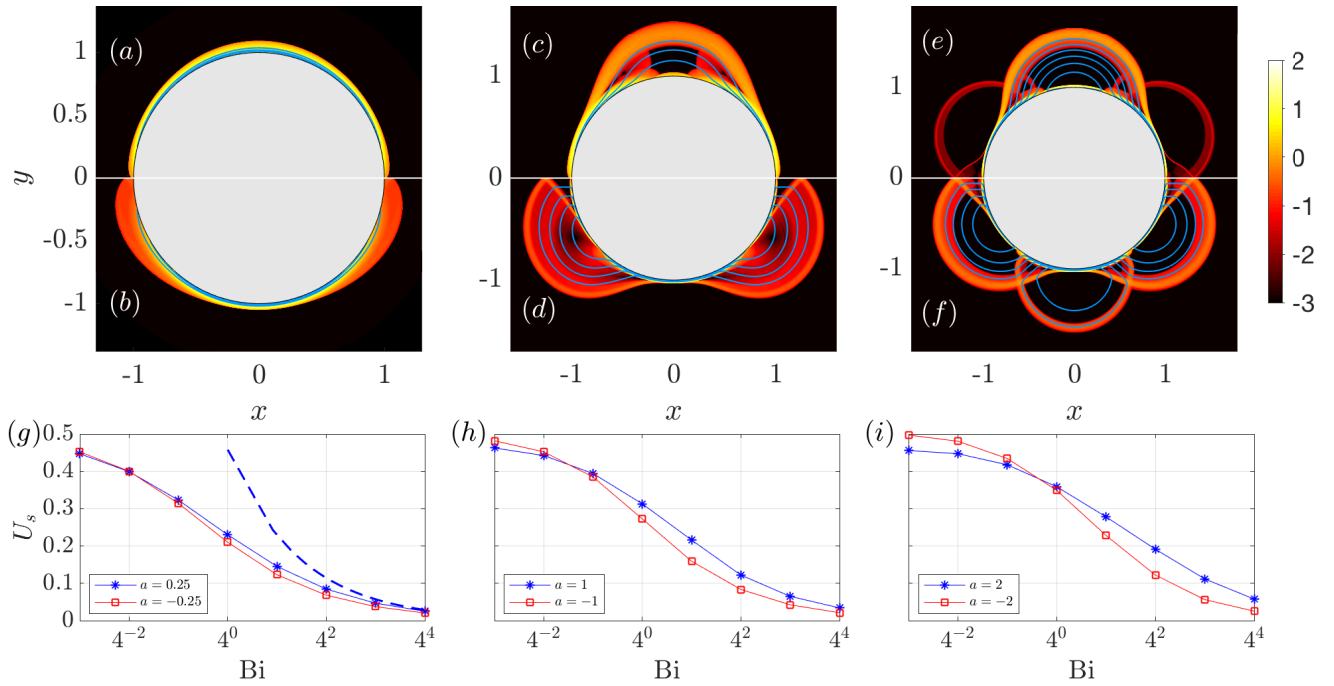


Figure 2-10: Force-free squirmer solutions with $(n, m) = (1, 3)$ and $\text{Bi} = 2^8$ showing $\log_{10} \dot{\gamma}$ for (a) $a = 0.25$, (b) $a = -0.25$, (c) $a = 1$, (d) $a = -1$, (e) $a = 2$ and (f) $a = -2$. (g-i) Variation of the swimming speed with Bi . In (g), the asymptotic prediction (dashed) from (2.36) for $a = 0.25$ matches well with the numerical data.

much like the slipline solutions of §3 and §4 (see figure 2-9h). Such patterns are not, however, the only possibility; figure 2-9g, for example, displays a swimmer in which closer examination reveals curved sliplines that peel off the surface boundary layer, and incorporate a non-circular fan pinned at the centre of circulation.

Force-free pushers and pullers with $(n, m) = (1, 3)$ are shown in figure 2-10(a–f). In this case, since m and n are both odd, the $a \rightarrow -a$ symmetry is lost (although the flow patterns remain symmetrical about the y -axis) and the swimming speed depends on the sign of a . Regardless of this, however the flow is again confined to the surface boundary layers for lower values of a , and features larger-scale plastic deformation for higher a , with the swimming speed scaling as $U_s \sim \text{Bi}^{-1/2}$ in the plastic limit and converging to $U_s = \frac{1}{2}$ for $\text{Bi} \rightarrow 0$.

A less expected result is shown in figure 2-11, which displays flow fields and swimming speeds for a swimmer with $(n, m) = (2, 3)$. In the Newtonian limit, such a mixed-mode driving pattern cannot provide propulsion as it does not contain a $\sin \theta$ component. Moreover, swimming is not possible with either of the $n = 2$ and $m = 3$ components individually. With a finite yield-stress, however, propulsion becomes possible, and the swimmer reaches a maximum speed at an intermediate value of Bi (figure 2-11c; again, the symmetry of the driving pattern implies that U_s does not depend on the sign of a).

Finally, we report an example exploiting the prescribed normal surface velocity in (2.11), rather than tangential motions. As argued in §2.5.2, the simplest example of this model with $n = 1$ and $a = 0$ is equivalent to the squirmer with the tangential surface velocity $V_p = \sin \theta$, but for the switch in translation speed $U_s \rightarrow 1 - U_s$. Hence, all the results in figure 2-7 and 2-9 immediately carry over, although the switch in U_s implies a very different limit for the force-free swimming speed for $\text{Bi} \gg 1$. Additional results for $m = 2$ and varying a are displayed in figure 2-12. When a is not small, the swimmer is no longer equivalent to a squirmer with tangential surface velocity; larger-scale patterns of plastic deformation develop with both curved sliplines, centred fans and constant-stress triangles. The swimming speed now converges to an a -dependent constant for $\text{Bi} \rightarrow \infty$ (the Newtonian limit is again $U_s \rightarrow \frac{1}{2}$).

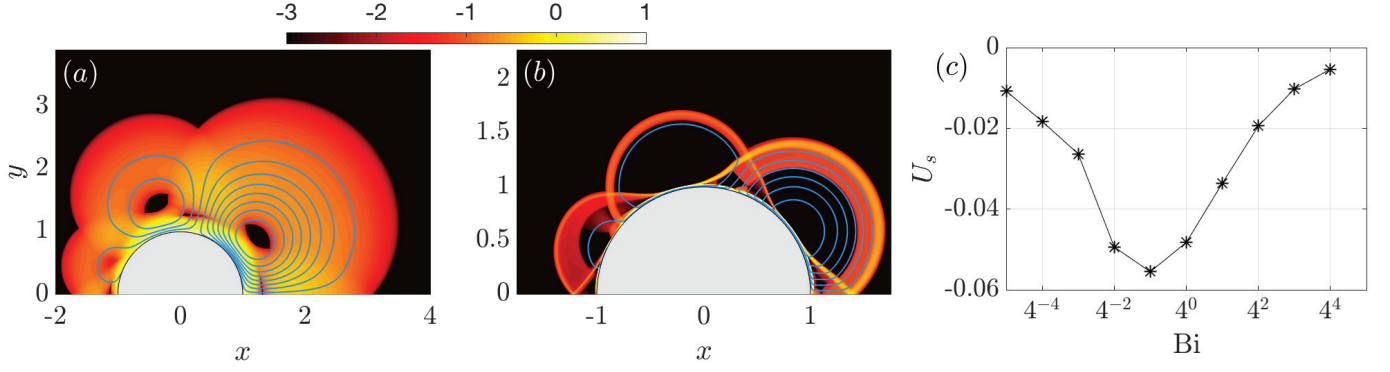


Figure 2-11: Force-free squirmer solutions showing $\log_{10} \dot{\gamma}$ for $(n, m) = (2, 3)$ and $a = 1$ for (a) $\text{Bi} = 1$ and (b) $\text{Bi} = 2^8$. (c) The variation of the non-zero swimming speed U_s with Bi . In contrast, these swimming modes have a zero swimming speed in the Newtonian limit ($\text{Bi} \rightarrow 0$).

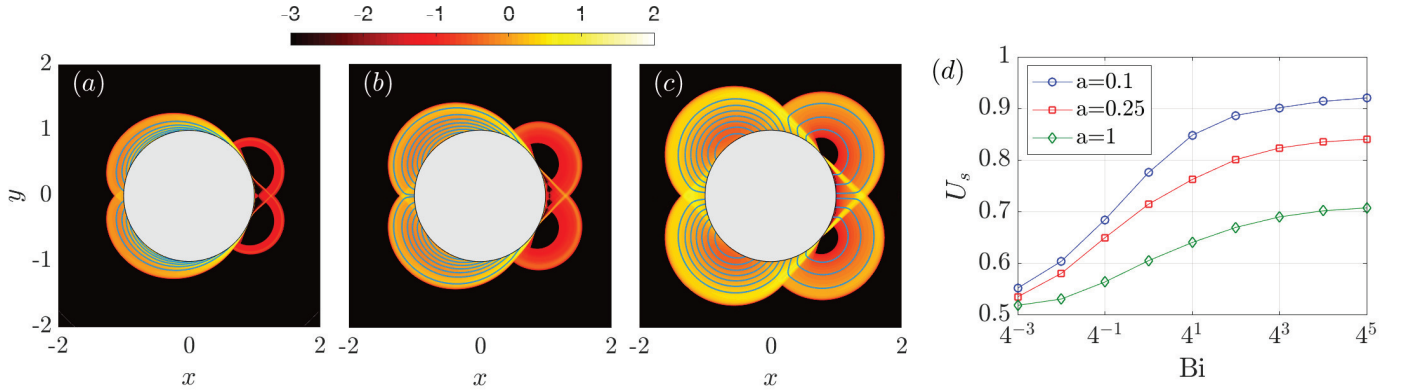


Figure 2-12: Force-free squirmer solutions showing $\log_{10} \dot{\gamma}$ for the normal surface velocity condition (2.11) with (a) $a = 0.1$, (b) 0.25 and (c) 1, at $\text{Bi} = 2^8$. (d) The swimming speed approaches the Newtonian limit of 0.5 as $\text{Bi} \rightarrow 0$ and a finite value dependent on a as $\text{Bi} \rightarrow \infty$.

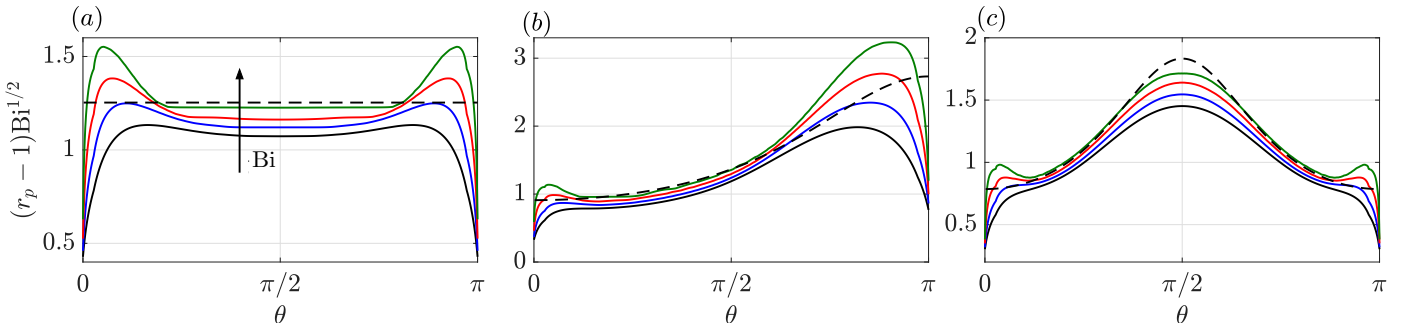


Figure 2-13: Boundary-layer thicknesses of squirmers for (a) $(n, a) = (1, 0)$, (b) $(n, m, a) = (1, 2, 0.25)$, and (c) $(n, m, a) = (1, 3, 0.25)$, with $\text{Bi} = 2^6$ (black), $\text{Bi} = 2^8$ (blue), $\text{Bi} = 2^{10}$ (red) and $\text{Bi} = 2^{12}$ (green). The boundary-layer theory predictions from (2.31)[dashed lines] match well with the numerical profiles except near $\theta = 0, \pi$.

2.5.4 Boundary layer theory

Boundary-layer structure

When flow becomes confined to the boundary layer attached to the cylinder, as in figures 2-9(e,f) and 2-10(a,b), we rescale the variables to describe that narrow region (*cf.* appendix A.1):

$$r = 1 + \text{Bi}^{-1/2}\eta, \quad (2.24a)$$

$$[u, U_s] = \text{Bi}^{-1/2}[U(\eta, \theta), U_1], \quad (2.24b)$$

$$v = V(\eta, \theta), \quad (2.24c)$$

$$p = \text{Bi}P(\eta, \theta). \quad (2.24d)$$

At leading order, local force balance then demands,

$$\partial_\eta P = 0 \quad \text{and} \quad \partial_\eta^2 V + 2 \operatorname{sgn}(\partial_\eta V) = P_\theta. \quad (2.25)$$

In terms of the rescaled variables, the boundary conditions in equation (2.9) become

$$V(0, \theta) \sim V_p(\theta) \quad \text{and} \quad U(0, \theta) = U_1 \cos \theta \quad (2.26)$$

whereas the match to the surrounding plug at the yield surface, $\eta = \eta_b(\theta)$, demands that

$$V(\eta_b, \theta) = \partial_\eta V(\eta_b, \theta) = U(\eta_b, \theta) = 0. \quad (2.27)$$

We therefore find the velocity profile,

$$V = V_p(\theta) \left(1 - \frac{\eta}{\eta_b}\right)^2, \quad (2.28)$$

with

$$\partial_\theta P = 2 \left(\frac{|V_p|}{\eta_b^2} - 1 \right) \operatorname{sgn}(V_p). \quad (2.29)$$

The integral of the leading-order continuity equation, $\partial_\theta V + \partial_\eta U \sim 0$ across the boundary layer now furnishes

$$\frac{\partial}{\partial \theta} \left[\frac{\eta_b V_p(\theta)}{3} \right] - U_1 \cos \theta = 0. \quad (2.30)$$

Focussing on surface motions that are up-down antisymmetric with $V_p(0) = V_p(\pi) = 0$, we now find the boundary-layer profile,

$$\eta_b(\theta) = \frac{3U_1 \sin \theta}{V_p(\theta)}. \quad (2.31)$$

This thickness is only positive when the angular surface motion $V_p(\theta)$ is directed opposite to sense of translation, $\text{sgn}(U_s)$. Otherwise, the balances implied by the scalings in (2.24) are not consistent, which we interpret to signify that the flow cannot be confined to the surface boundary layer. Indeed, the numerical solutions in §2.5.3 display larger-scale flow patterns at large Bi whenever $U_s/V_p < 0$.

Figure 2-13 compares the predictions of (2.31) with some of the measured yield surfaces of the numerical solution with confined flow patterns. In the simple case with $V_p(\theta) = \sin \theta$ (figure 2-13a), the boundary layer has the constant width, $\eta_b = (r_p - 1)\text{Bi}^{1/2} = \sqrt{\pi/2}$, where r_p is the radius of the yield surface. The yield surfaces of the computations are indeed relatively flat and compare well with the prediction, except close to the front and back of the cylinder where the boundary layer thickness sharply declines over further “corner regions.” For squirmers with $V_p = \sin \theta + a \sin m\theta$, the boundary layer thickness varies with position; again the predictions match well with numerical results except for the adjustments at the front and back (figure 2-13(b,c)).

For $V_p U_1 < 0$, the emergence of plastic deformation outside the boundary layer (with $(u, v) \sim O(\text{Bi}^{-1/2})$) modifies the final boundary condition in (2.26), and therefore the flux balance in (2.30). The resulting flow into or out of the boundary layer then maintains a boundary layer of finite thickness. Importantly, however, the scalings of the boundary layer in (2.24) do not change although one must now complete the solution by matching to the adjacent region of perfectly plastic deformation.

Drag force and swimming speed

Given the surface pressure $\text{Bi}P$ and tangential shear stress $\text{Bi} \text{sgn}(\partial_\eta V) = -\text{Bi} \text{sgn}(V_p)$, the net horizontal force on the swimmer is given by

$$F_x = -\text{Bi} \int_{-\pi}^{\pi} [P \cos \theta - \text{sgn}(V_p) \sin \theta] d\theta = 2\text{Bi} \int_0^{\pi} \left(\frac{2|V_p|^3}{9U_1^2 \sin \theta} - \sin \theta \right) \text{sgn}(V_p) d\theta. \quad (2.32)$$

The drag therefore vanishes for

$$U_s^{ff} = \frac{\text{Bi}^{-1/2}}{3} \sqrt{2 \int_0^{\pi} d\theta \frac{V_p(\theta)^3}{\sin \theta} \left[\int_0^{\pi} d\theta \sin \theta \text{sgn}(V_p) \right]^{-1/2}}, \quad (2.33)$$

furnishing an asymptotic prediction of the locomotion speed of the force-free swimmer.

The drag from (2.32) also increases with decreasing translation speed U_s , diverging for $U_s \rightarrow 0$, and must therefore exceed that associated with the Randolph and Houlsby slipline solution below some threshold in U_s . We interpret the cross-over to correspond to the switch from the confined flow pattern to larger-scale plastic deformation. The deconfinement of the flow must therefore occur for

$$U_s = U_s^{ff} \left\{ 1 + 2(\pi + 2\sqrt{2}) \left[\int_0^{\pi} d\theta \sin \theta \text{sgn}(V_p) \right]^{-1} \right\}^{-1/2}. \quad (2.34)$$

For the simplest case with $V_p = \sin \theta$ ($n = 1$, $a = 0$), the drag coefficient implied by (2.32) is

$$C_d = -\frac{F_x}{2\text{Bi}} = 2 - \frac{\pi}{9U_1^2}, \quad (2.35)$$

which is compared to the numerical results in figure 2-7g; the switch in flow pattern for (2.34) also matches satisfyingly with the abrupt drop in the magnitude of C_d in the numerical solutions. The corresponding force-free swimmer has $U_s^{ff} = \sqrt{\pi/18\text{Bi}}^{-1/2}$, which again compares well with the numerical results (figure 2-7g).

For squirmers with $V_p = \sin \theta + a \sin m\theta$, the swimming speed is

$$U_s^{ff} = \begin{cases} \frac{1}{3}\text{Bi}^{-1/2} \sqrt{\frac{1}{2}\pi(1 + 3a^2)} & m \text{ even,} \\ \frac{1}{3}\text{Bi}^{-1/2} \sqrt{\frac{1}{2}\pi(1 + 3a^2 + a^3)} & m \text{ odd,} \end{cases} \quad (2.36)$$

provided the boundary layer thickness remains finite everywhere, which demands that $a < m^{-1}$ for m even and $a \lesssim \sin(3\pi/2m)$ for m odd. Figures 2-9j and 2-10g include the predictions in (2.36).

Note that the prediction for the swimming speed in (2.33) relies on the solutions in (2.29) and (2.31), which fail when flow is no longer confined to the boundary layer. Nevertheless, because the scalings of the problem do not change in that situation, the swimming speed still scales with $\text{Bi}^{-1/2}$, as seen in the numerical computations (e.g. figure 2-9i). The match to the surrounding plastic deformation, however, determines the coefficient U_1 .

2.6 Conclusions

In this chapter, we have investigated viscoplastic flows around cylinders, with an emphasis on the limit of large yield stress. For a translating cylinder with a no-slip surface, we compared analytical plasticity solutions based on slipline theory with viscoplastic computations. Significant differences between the two arise due to the presence of rigidly rotating plugs above and below the cylinder in the computations, which are not present in the slipline solutions. These plugs ride on top of the viscous boundary layer that shrouds the cylinder, leading one to wonder whether they interfere with the plastic limit of the viscoplastic model. By performing a suite of careful computations and developing a boundary layer theory, we showed that such features do actually disappear in the plastic limit ($\text{Bi} \rightarrow \infty$), implying that viscoplasticity does converge to ideal plasticity.

We then modified the boundary condition on the translating cylinder to allow for partial slip over its surface. This situation corresponds to partially rough cylinders in ideal plasticity, for which the slipline solution has not been previously identified,

with an original construction proposed by Randolph and Houlsby [7] having been shown to be inconsistent for partial slip [8, 80]. Instead, we found our computations matched with an alternative slipline pattern proposed by Martin and Randolph [8] as an upper bound solution based on its velocity field. This alternative pattern contains genuine rigid plugs rotating above and below the cylinder, attached to, and sliding over the surface. Delving further into Martin and Randolph’s slipline construction, we found that the stress solution suggests a lower bound that matches the upper bound provided the rotating plugs are free of any net torque. This implies that the slipline field actually provides the true plastic solution. However, the slipline theory is incomplete in this example because no stress field is provided for the plugs that matches the partially rough surface conditions and is consistent with the yield condition. Nevertheless, the computations do explicitly construct an acceptable stress field for the plugs, providing numerical evidence for the conclusion that Martin and Randolph’s slipline pattern is the true plastic solution.

The slipline patterns of the translating cylinders provide a set of tools to understand viscoplastic flow around cylinders with a variety of other surface conditions. In the third thread of this study, we applied this idea to models of cylindrical “squirmers” swimming in yield-stress fluid. The slipline patterns do indeed characterize many of the flow structures seen around such model micro-organisms when we approach the plastic limit. However, we also found that flow can become consumed into the viscous boundary layers against the cylinder surfaces, allowing us to analytically construct the swimming states. We provided the viscoplastic analogues of squirming ‘pushers’ and ‘pullers’, for which the driving surface velocity is concentrated either to the back or front of the cylinder. While these squirmers have identical swimming speeds, as in the Newtonian limit, we also identified driving surface velocity patterns for which this symmetry is not preserved if the fluid has a yield stress. We also provided examples of swimming patterns that would be immobile in the Newtonian limit, but may swim when there is yield stress because of the non-linearity of the fluid rheology.

For squirmers driven by a prescribed tangential surface velocity in Newtonian fluid, without considering the specific details of the surface velocity pattern, the swimming

speed \mathcal{U}_s scales with the characteristic speed of the driving surface velocity \mathcal{U} and the power input per unit swimming speed and cylinder length is $\mathcal{P} \sim \mu\mathcal{U}$. In the opposite, plastic limit, the swimming speed and power turn out to scale as

$$\mathcal{U}_s \sim \mathcal{U}^{3/2} \sqrt{\frac{\mu}{\tau_Y \mathcal{R}}} \quad \text{and} \quad \mathcal{P} \sim \frac{\mathcal{R}\mathcal{U}\tau_{r\theta}}{\mathcal{U}_s} \sim \frac{(\tau_Y \mathcal{R})^{3/2}}{(\mu\mathcal{U})^{1/2}}, \quad (2.37)$$

if \mathcal{R} is the cylinder radius and μ and τ_Y are the fluid (plastic) viscosity and yield stress. The decaying dependence on yield stress and presence of the viscosity is symptomatic of the viscous boundary layers against the cylinder surface which activate locomotion (the $\text{Bi}^{-1/2}$ layers in §5). Note that the effective viscosity of the medium in the plastic limit is $\mu_{\text{eff}} \sim \tau_Y \mathcal{R}/\mathcal{U} \gg \mu$ (being given by the relatively large yield stress). Therefore, the scaling of the swimming speed is $\mathcal{U}_s \sim \mathcal{U} \sqrt{\mu/\mu_{\text{eff}}}$, which implies that the swimmer moves much slower in the viscoplastic medium than in a Newtonian fluid with the same effective viscosity μ_{eff} , for a given surface velocity pattern. Moreover, the input power per unit length and swimming speed is $\mathcal{P} \sim \mu_{\text{eff}}\mathcal{U} \sqrt{\mu_{\text{eff}}/\mu}$, rather larger than the Newtonian equivalent ($\mu_{\text{eff}}\mathcal{U}$).

Thus, swimming by tangential squirming motions in a nearly plastic medium is relatively inefficient, primarily as a result of the lubricating effect of the viscous boundary layers against the cylinder's surface. The situation is quite different if swimming is driven by normal surface motions. The swimming speed in the plastic limit then remains of order \mathcal{U} , as in the Newtonian limit, and the corresponding power input per unit swimming speed and length is given by $\mathcal{P} \sim \tau_Y \mathcal{R} \sim \mu_{\text{eff}}\mathcal{U}$. Now the swimming speed and power input are comparable to those for motion through Newtonian fluid with viscosity μ_{eff} (a situation shared by the viscoplastic version of Taylor's swimming sheet, considered by Hewitt and Balmforth [34]).

Another key feature of the swimming dynamics is that the yield stress always limits flow to within a yield surface that lies at a finite distance from the squirmer. This has important implications for the induced transport of nutrients or other tracers and the hydrodynamic interactions and collective dynamics of multiple swimmers [29, 23]. Finally, we add a cautionary note that our modelling of swimming micro-organisms as

cylinders with prescribed surface motions is somewhat restrictive, limiting the quantitative application of our results. In particular, we neglect all effects of viscoplasticity on the imposed surface velocity pattern, and a real concern is that the cilia responsible for driving these motions may themselves clog up under the action of the yield stress. However, the qualitative results of our work constitute a first step towards understanding the effect of a yield stress on swimming micro-organisms and their collective dynamics.

Chapter 3

Linearly forced fluid flow on a rotating sphere

The contents of this chapter have been published as the article: R. Supekar, V. Heinonen, K. J. Burns, & J. Dunkel, Linearly forced fluid flow on a rotating sphere, *J. Fluid Mech.*, 892:A30, 2020 [91].

3.1 Introduction

Turbulence is often described as the last unsolved problem in classical physics [92]. In recent years, considerable progress has been made in the modelling of stationary turbulence, which requires a driving force to continually balance kinetic energy losses due to viscous dissipation [93]. Theoretical and computational studies of turbulence phenomena typically focus on external driving provided by a random forcing [94], boundary forcing [95] or Kolmogorov forcing [96]. A fundamentally different class of internal driving mechanisms, less widely explored in the turbulence literature so far, is based on linear instabilities [97, 98, 99, 100, 18, 101, 102]. The profound mathematical differences between external and internal driving were emphasized by Arnold [103] in the context of classical dynamical systems described by ordinary differential equations. Specifically, he contrasted the externally forced Kolmogorov hydrodynamic system with the internally forced Lorenz system, the latter providing a simplified

model of atmospheric convection [104]. From the broader fluid-mechanical perspective, Arnold’s analysis raises the interesting question of how internally driven flows behave in rotating frames like the atmospheres of planets or stars. Models for such internally driven flows have been recently investigated in the context of ‘active turbulence’ [17, 18]. In this chapter, we ask: can these phenomenological models be utilized to minimally model pattern-forming flows on a rotating sphere?

To answer the above question, we investigate an analytically tractable minimal model for linearly forced quasi-2D flow on a rotating sphere. The underlying generalized Navier-Stokes (GNS) equations describe internally driven flows through higher-order hyperviscosity-like terms in the stress tensor [105, 18], and the associated GNS triad dynamics is structurally similar to the Lorenz system [101]. GNS-type models have been studied previously as effective phenomenological descriptions for seismic wave propagation [105, 98], magnetohydrodynamic flows [106] and active fluids [43, 18, 19]. A key difference compared with scale-free classical turbulence is that GNS flows can exhibit characteristic spatial and temporal scales that reflect the internal forcing mechanisms.

Remarkably, the minimal GNS model studied below permits nontrivial analytical solutions. Exact stationary solutions reported previously include 3D Beltrami flows [18] and 2D vortex lattices [43]. Furthermore, Mickelin et al. [66] recently explored GNS flows on 2D curved surfaces and constructed stationary solutions for the case of a non-rotating sphere. Here, we generalize their work by deriving exact time-dependent solutions for GNS flows on rotating spheres, and by comparing them with direct numerical simulations (figure 3-1). We shall see that these exact GNS solutions correspond to Rossby waves propagating along alternating zonal jets, qualitatively similar to the large-scale flow patterns seen in planetary atmospheres [107, 108]. This highlights how such minimal models originally proposed for active fluids can provide insights into generic pattern forming systems at different scales.

Our study complements recent work which showed that non-equilibrium approaches can provide analytical insights into the dynamics of planetary flows [45] and atmospheres [44]. In view of the recent successful application of phenomenological GNS

models to active fluids [17, 18], the results below can also help advance the understanding of active matter propagation on curved surfaces [109, 110, 111, 112] and in rotating frames [113].

3.2 Generalized Navier-Stokes model for linearly driven flow

After briefly reviewing the GNS equations in § 3.2.1, we derive the corresponding vorticity-stream function formulation on a rotating sphere in § 3.2.2.

3.2.1 Planar geometry

The GNS equations for an incompressible fluid velocity field $\mathbf{v}(\mathbf{x}, t)$ with pressure field $p(\mathbf{x}, t)$ read [18, 43]

$$\nabla \cdot \mathbf{v} = 0, \quad (3.1a)$$

$$\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla p + \nabla \cdot \boldsymbol{\sigma}, \quad (3.1b)$$

where the higher-order stress tensor

$$\boldsymbol{\sigma} = (\Gamma_0 - \Gamma_2 \nabla^2 + \Gamma_4 \nabla^4) [\nabla \mathbf{v} + (\nabla \mathbf{v})^\top] \quad (3.1c)$$

accounts for both viscous damping and linear internal forcing. Transforming to Fourier space, the divergence of the stress tensor gives the dispersion relation

$$\xi(k) = -k^2(\Gamma_0 + \Gamma_2 k^2 + \Gamma_4 k^4) \quad (3.2)$$

where k is the magnitude of the wave vector \mathbf{k} . Fixing hyper-viscosity parameters $\Gamma_0 > 0$, $\Gamma_4 > 0$ and $\Gamma_2 < -2\sqrt{\Gamma_0 \Gamma_4}$, the growth rate $\xi(k)$ is positive between the two real roots k_- and k_+ . Hence, Fourier modes in the active band $k \in (k_-, k_+)$ are linearly unstable, corresponding to active energy injection into the fluid. The distance

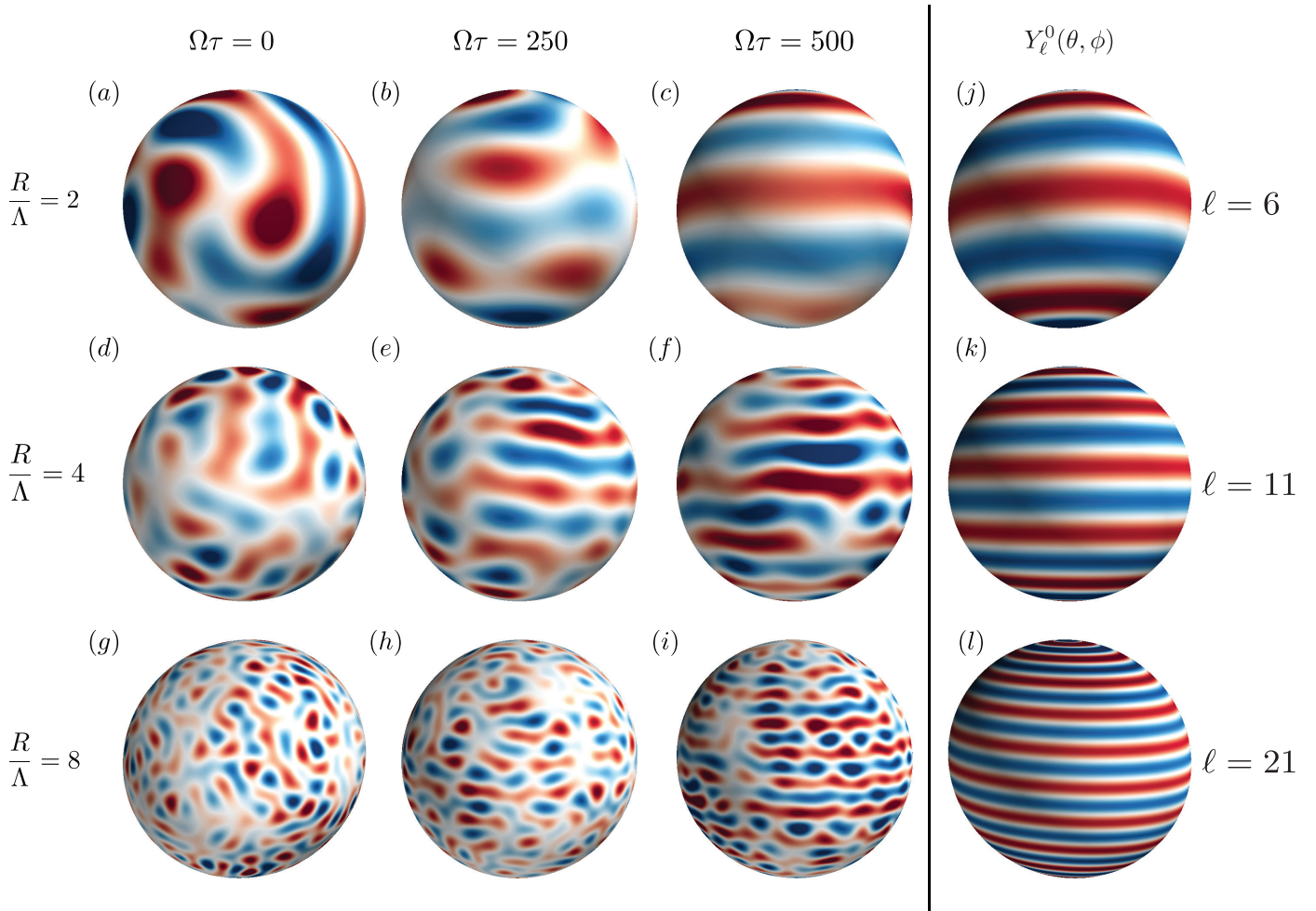


Figure 3-1: Statistically stationary states of the normalized vorticity $\zeta\tau$ from simulations (a-i) for $\kappa\Lambda = 1$ become more zonal (or banded) as the rotation rate $\Omega\tau$ increases. At the highest rotation rate $\Omega\tau = 500$, the width of the alternating zonal jets is determined by the parameter R/Λ that represents the ratio of the radius of the sphere and the diameter of the vortices forced by the GNS dynamics. The main characteristics of these flow patterns at high rotation rate are captured by spherical harmonics $Y_\ell^0(\theta, \phi)$ that solve the dynamical equations. Matching the length scale R/Λ gives $\ell = 6$ (j), $\ell = 11$ (k), and $\ell = 21$ (l) for $R/\Lambda = 2, 4$, and 8 , respectively.

between the neutral modes k_{\pm} defines the active bandwidth $\kappa = k_+ - k_-$.

Unstable bands are a universal feature of stress tensors exhibiting positive dispersion $\xi(k) > 0$ for some k . Polynomial GNS models of the type (3.1) were first studied in the context of seismic wave propagation [105, 98] and can also capture essential statistical properties of dense microbial suspensions [17, 18]. Since non-polynomial dispersion relation produce qualitatively similar flows [101, 102], we focus here on stress tensors of the generic polynomial form (3.1c).

Exact steady-state solutions of (3.1), corresponding to ‘zero-viscosity’ states, can be written as superpositions of modes \mathbf{k} with $|\mathbf{k}| = k_+$ or $|\mathbf{k}| = k_-$ [18]. Simulations of (3.1) with random initial conditions converge to statistically stationary states with highly dynamical vortical patterns that have a characteristic diameter $\sim \Lambda = \pi/k_*$, where k_* is the most unstable wavenumber, corresponding to the maximum of $\xi(k)$ [101, 114]. Inverse energy transport in 2D can bias the dominant vortex length scale towards larger values $\sim \pi/k_-$ [19].

3.2.2 On a rotating sphere

We generalize planar 2D GNS dynamics (3.1) to a sphere with radius R rotating at rate Ω . To this end, we adopt a co-rotating spherical coordinate system (θ, ϕ) where θ is the co-latitude and ϕ is the longitude. Following Mickelin et al. [66], we find the rotating GNS equations in vorticity-stream function form

$$\nabla^2 \psi = -\zeta, \quad (3.3a)$$

$$\partial_t \zeta + J(\psi, \zeta) = F(\nabla^2 + 4K)(\nabla^2 + 2K)\zeta + 2\Omega K \partial_\phi \psi, \quad (3.3b)$$

where ζ is the vorticity in the rotating frame, and $K = 1/R^2$ denotes the Gaussian curvature of the sphere. The active stress operator F has the polynomial form

$$F(x) = \Gamma_0 - \Gamma_2 x + \Gamma_4 x^2. \quad (3.4)$$

The Laplacian ∇^2 on the sphere is defined by $\nabla^2 = K (\cot \theta \partial_\theta + \partial_\theta^2 + (\sin \theta)^{-2} \partial_\phi^2)$

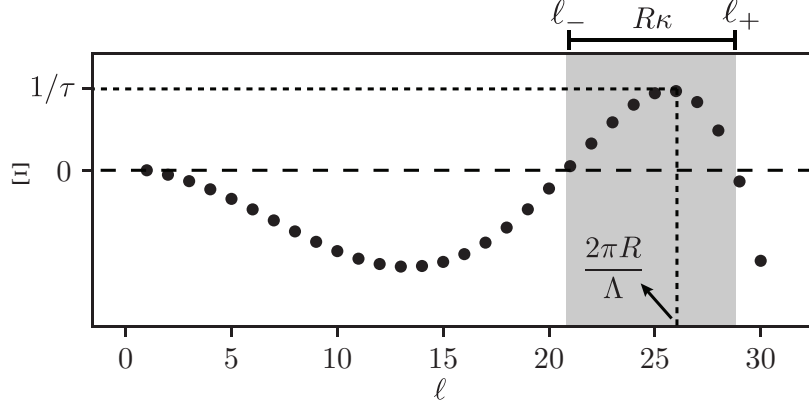


Figure 3-2: The growth rate Ξ of spherical harmonic modes $Y_\ell^m(\theta, \phi)$ in (3.5) plotted as a function of the wavenumber ℓ . The parameters used to make this plot are $((\tau/R^2)\Gamma_0, (\tau/R^4)\Gamma_2, (\tau/R^6)\Gamma_4) \simeq (1.43 \times 10^{-2}, -4.86 \times 10^{-5}, 3.72 \times 10^{-8})$ which correspond to $R/\Lambda = 8$ and $\kappa\Lambda = 1$. The grey region indicates the active bandwidth where $\Xi > 0$ and energy is injected. Λ is the diameter of the vortices forced by the mode with the maximum growth rate $1/\tau$, and κ is the active bandwidth i.e. $\kappa = (\ell_+ - \ell_-)/R$ where $\Xi(\ell_\pm) = 0$.

and $J(\psi, \zeta) = K(\sin \theta)^{-1} (\partial_\phi \psi \partial_\theta \zeta - \partial_\theta \psi \partial_\phi \zeta)$ is the determinant of the Jacobian of the mapping $(R\phi \sin \theta, R\theta) \mapsto (\psi, \zeta)$ from the tangent space of the sphere to the vectors (ψ, ζ) . The velocity components can be recovered from the stream function ψ by $(v_\phi, v_\theta) = (-\partial_\theta \psi / R, \partial_\phi \psi / R \sin \theta)$. In the non-rotating limit $\Omega \rightarrow 0$, (3.3b) reduces to the model studied by Mickelin et al. [66]. We note that (3.3b) is an internally forced extension of the unforced barotropic vorticity equation $\partial_t \zeta + J(\psi, \zeta) = 2\Omega K \partial_\phi \psi$ which has been widely studied in earth science since the pioneering work of Charney et al. [115]. Below we will show that the GNS model (3.3b) is analytically tractable, permitting exact traveling wave solutions that are close to the complex flow states observed in simulations.

Dimensionless parameters

We assess the linear behaviour of (3.3b) using spherical harmonics $Y_\ell^m(\theta, \phi)$, the eigenfunctions of the Laplacian operator on the sphere. With $\delta = K(\ell(\ell + 1) - 4)$, the linear growth rate of a spherical harmonic mode due to F is

$$\Xi(\ell) = -(\delta + 2K) F(-\delta) = -(\delta + 2K) (\Gamma_0 + \Gamma_2 \delta + \Gamma_4 \delta^2) \quad (3.5)$$

which is the spherical analog of (3.2). Using this relation, characteristic length and time scales, Λ and τ , for vortices forced by the GNS dynamics, along with the bandwidth of the forcing κ , can be expressed in terms of $\Gamma_0, \Gamma_2, \Gamma_4$ and R (figure 3-2 and appendix B.1). We use these scales to define the essential dimensionless parameters: R/Λ is the ratio between the radius of the sphere and the characteristic vortex scale, $\kappa\Lambda$ compares the forcing bandwidth to the characteristic vortex scale, and $\Omega\tau$ is dimensionless rotation rate. We note that since $\Xi(\ell = 1) = 0$, the GNS dynamics do not force the $\ell = 1$ mode which ensures that the total angular momentum is conserved (see appendix B.2). Finally, we define the Rossby number in terms of the characteristic flow speed $\mathcal{U} = \Lambda/\tau$ and the dominant length scale $\mathcal{L} = \Lambda$ as

$$Ro = \frac{\mathcal{U}}{\Omega\mathcal{L}} = \frac{1}{\Omega\tau}. \quad (3.6)$$

β -plane equations

When the vortical patterns are much smaller than the radius of the sphere ($R/\Lambda \gg 1$), one can linearize around a reference co-latitude θ_0 to produce a local model. We define metric coordinates in the directions of increasing ϕ and decreasing θ , respectively, by $x = R \sin(\theta_0)\phi$ and $y = R(\theta_0 - \theta)$. In these coordinates the dynamical equations are

$$\nabla_c^2 \psi = -\zeta, \quad (3.7a)$$

$$\partial_t \zeta + J_c(\psi, \zeta) = \nabla_c^2 F(\nabla_c^2) \zeta + \beta \partial_x \psi, \quad (3.7b)$$

where $J_c(\psi, \zeta) = \partial_y \psi \partial_x \zeta - \partial_x \psi \partial_y \zeta$ is the Cartesian Jacobian determinant, ∇_c^2 is the Cartesian Laplacian and the namesake β parameter is given by $\beta = 2\Omega \sin \theta_0/R$. The β -plane equations preserve the effect of a varying Coriolis parameter $2\Omega \cos \theta$ while simplifying the spatial operators. Rotational effects are accounted for by the term proportional to β .

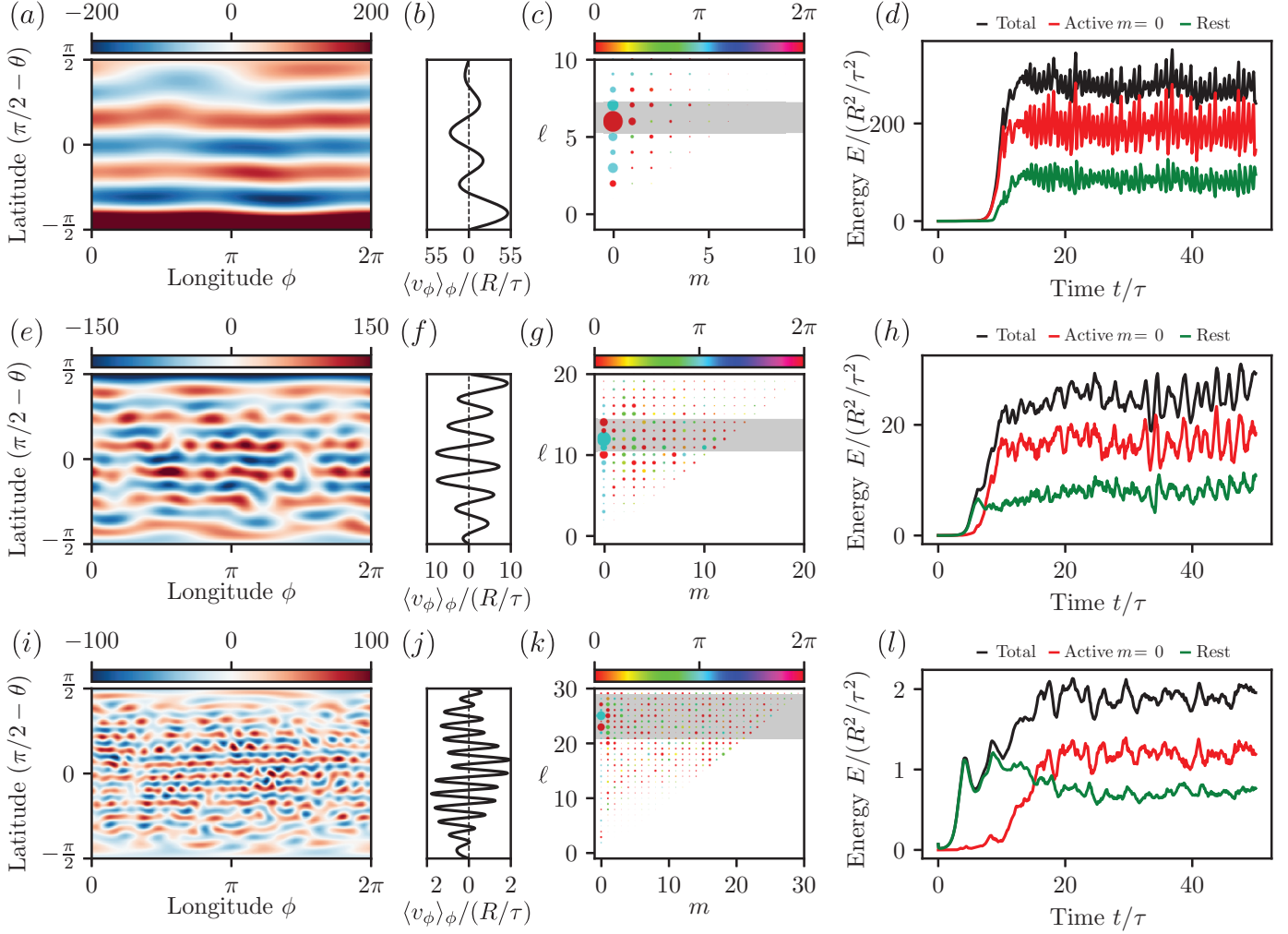


Figure 3-3: Data from steady-state solutions at $t/\tau = 15$ for the highest rotation rate $\Omega\tau = 500$. The rows correspond to $R/\Lambda = 2$ (a–d), 4 (e–h) and 8 (i–l). Panels (a, e, i) show Mercator projections of the dimensionless vorticity $\zeta\tau$. Panels (b, f, j) show the zonal-mean azimuthal velocities $\langle v_\phi \rangle_\phi / (R/\tau)$. Panels (c, g, k) show spherical harmonic decomposition of dimensionless vorticity $\zeta\tau$ with marker size indicating amplitude and color indicating phase. All plots indicate the existence of dominant zonal jets with $m = 0$ and ℓ 's within the active band indicated in grey. These modes are close to the exact solutions in figure 3-1(j–l). Panels (d, h, l) show time-variation of the energy of all the modes (black), active $m = 0$ modes (red), and all other modes (green); the energy contained in the active $m = 0$ accounts for most of the total energy in the statistically stationary state.

Characteristic length and time scales

The turbulent outer layers of rotating stars and planets ubiquitously contain east-west (zonal) jets of various scales and strengths. Determining physical processes that generate and maintain these jets is an important problem in planetary science. A variety of theories have arisen describing jet formation, particularly on the β -plane, including the arrest of the inverse cascade in rotating turbulence by nonlinear Rossby waves [116, 117], and as a bifurcation in the statistical dynamics of the zonal flow as a function of the intensity of background homogeneous turbulence [118, 119]. These theories predict that jet formation may depend on a variety of length and timescales, such as the Rhines scale $L_R = \sqrt{\mathcal{U}/\beta}$, the scale at which small-scale forcing injecting energy at a rate ϵ is effected by rotation $L_\epsilon = (\epsilon/\beta^3)^{1/5}$ [120], and the growth rate of unstable perturbations to the zonal flow in statistical models [121]. The GNS model investigated here provides a simplified setting for examining the dynamics of jets by parameterizing the jet-formation physics, rather than attempting to resolve the details of the underlying formation processes. If one is interested in matching the effective GNS parameters to specific length and velocity scales of more detailed models, then guidance can be drawn from the observation that typical zonal jets in the GNS model have width $\sim \Lambda$ and r.m.s. velocity $\sim \Lambda\sqrt{\Omega/\tau}$ in the rotation dominated regime $\Omega\tau > 1$.

3.3 Exact time-dependent solutions

Exact solutions can be constructed on the sphere as well as on the local β -plane. Although not stable, these solutions will provide an intuitive understanding of the numerical results in § 3.4, similar to the role of exact coherent structures [122, 123] in classical turbulence.

3.3.1 Global solutions

Exact time-dependent solutions to (3.3b) can be constructed as superpositions of normal spherical harmonic modes $Y_\ell^m(\theta, \phi)$ as

$$\left[\psi(\theta, \phi, t), \zeta(\theta, \phi, t) \right] = \left[1, \ell_\pm(\ell_\pm + 1) \right] \left(\psi_j(\theta) + \psi_w(\theta, \phi, t) \right), \quad (3.8)$$

with

$$\psi_j(\theta) = \mathcal{A}_0 Y_{\ell_\pm}^0(\theta), \quad \psi_w(\theta, \phi, t) = \text{Re} \left[\sum_{m=1}^{\ell_\pm} \mathcal{A}_m Y_{\ell_\pm}^m(\theta, \phi) \exp(-i\sigma_m t) \right], \quad (3.9)$$

where \mathcal{A}_m for $m = 0, 1, \dots$ are constants, ℓ_\pm are the roots of

$$F(-\ell_\pm(\ell_\pm + 1) + 4) = 0 \quad (3.10)$$

and σ_m satisfies the dispersion relation

$$c_p^\phi = \frac{\sigma_m}{m} = \frac{-2\Omega}{\ell_\pm(\ell_\pm + 1)}. \quad (3.11)$$

These solutions to (3.3b) are possible because the Jacobian determinant J vanishes if the stream function is a superposition of spherical harmonics Y_ℓ^m with fixed ℓ . We choose ℓ to be one of the roots of the polynomial F given in equation (3.10). Equation (3.11) describes the dispersion of normal-mode Rossby-Haurwitz waves [124, 125, 126]. These are well-known solutions of the barotropic vorticity equation [127] and propagate in the direction opposite to the sphere's rotation with phase speed c_p^ϕ . Overall, the exact solutions in (3.8) are a combination of time-independent zonal jets (ψ_j) and time-varying Rossby-Haurwitz waves (ψ_w). The time-independent zonal jets are spherical harmonics $Y_\ell^0(\theta)$ which consist of alternating crests and troughs. A selection of such modes corresponding to different ℓ 's are shown in figure 3-1(j-1).

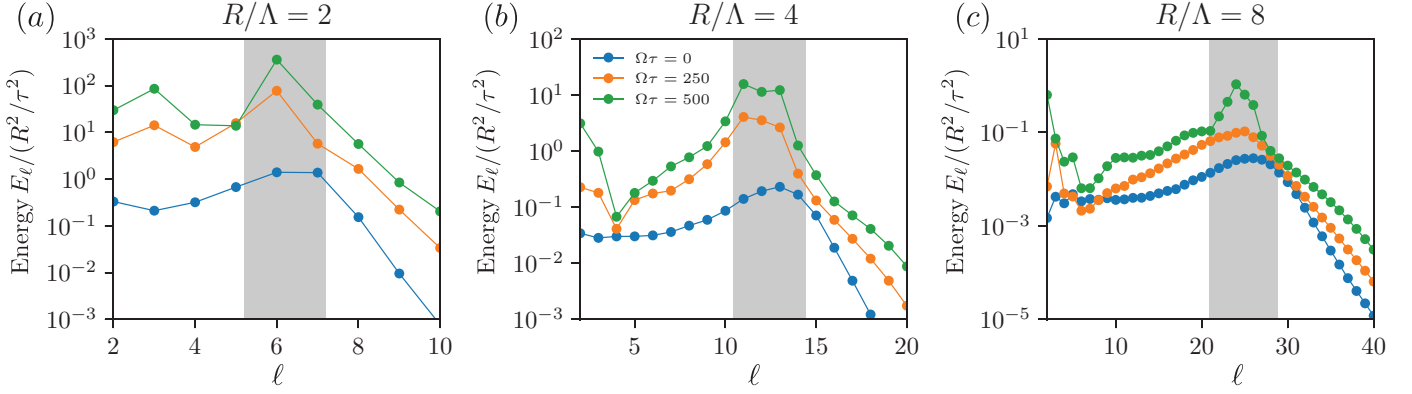


Figure 3-4: Energy spectra for (a) $R/\Lambda = 2$, (b) $R/\Lambda = 4$ and (c) $R/\Lambda = 8$. The grey shaded region indicates the active bandwidth where the spectra show a peak. The spectra for $(R/\Lambda, \Omega\tau) = (4, 250), (4, 500)$ and $(8, 500)$ have been obtained from an ensemble average of 10 simulations with random initial conditions.

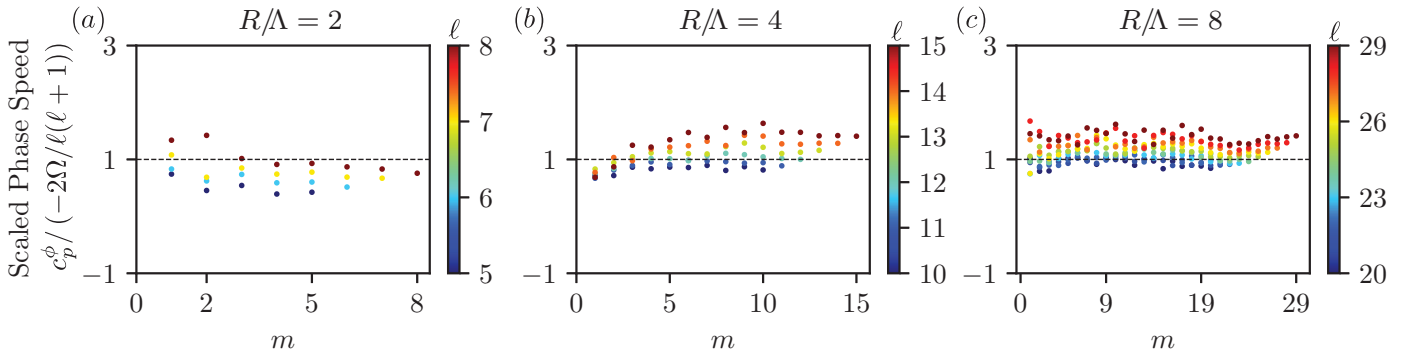


Figure 3-5: Phase speed of the spherical harmonic modes (ℓ, m) in the forcing bandwidth, normalized by the analytical phase speed in (3.11), for $\Omega\tau = 500$ and different values of R/Λ . The dotted line indicates the value 1 for comparison.

3.3.2 β -plane solutions

Similar to the procedure on the full sphere, exact solutions to (3.7b) can be constructed by considering superpositions of Fourier modes with wave vectors \mathbf{k} that correspond to the neutral modes of the pattern forming operator. Hence, the exact solutions are

$$\left[\psi(x, y, t), \zeta(x, y, t) \right] = \left[1, k_{\pm}^2 \right] \left(\psi_j(y) + \psi_w(x, y, t) \right), \quad (3.12)$$

with

$$\psi_j(y) = \text{Re} [\mathcal{A}_0 \exp(ik_{\pm}y)], \quad \psi_w(x, y, t) = \text{Re} \left[\sum_{|\mathbf{k}|=k_{\pm}} \mathcal{A}_{\mathbf{k}} \exp(i(\mathbf{k} \cdot \mathbf{x} - \sigma_{\mathbf{k}}t)) \right], \quad (3.13)$$

where $\mathcal{A}_{\mathbf{k}}$ are constants, k_{\pm} are the positive roots of $F(-k_{\pm}^2) = 0$, and $\sigma_{\mathbf{k}}$ satisfies the Rossby-wave dispersion relation [128]

$$c_p^x = \frac{\sigma_{\mathbf{k}}}{k_x} = \frac{-\beta k_x}{|\mathbf{k}|^2}. \quad (3.14)$$

Here again, solutions are a combination of a time-independent zonal flow ($\psi_j(y)$) and time-varying Rossby waves ($\psi_w(x, y, t)$). For the parameters at which the β -plane approximation holds, the expression for the phase speed (3.14) provides an explicit dependence on the co-latitude θ through the parameter $\beta = 2\Omega \sin \theta_0/R$. The dynamics at the poles are similar to those on a non-rotating flat plane and the non-inertial effects matter the most at the equator.

3.4 Simulations

Direct numerical simulations of (3.3b) were performed using a spectral code based on the open-source Dedalus framework [129]. The code uses a pseudo-spectral method with a basis of spin-weighted spherical harmonics [130, 131]; see appendix of Ref. [66]. A spectral expansion with a cut-off $\ell_{max} = 256$ suffices to obtain converged solutions.

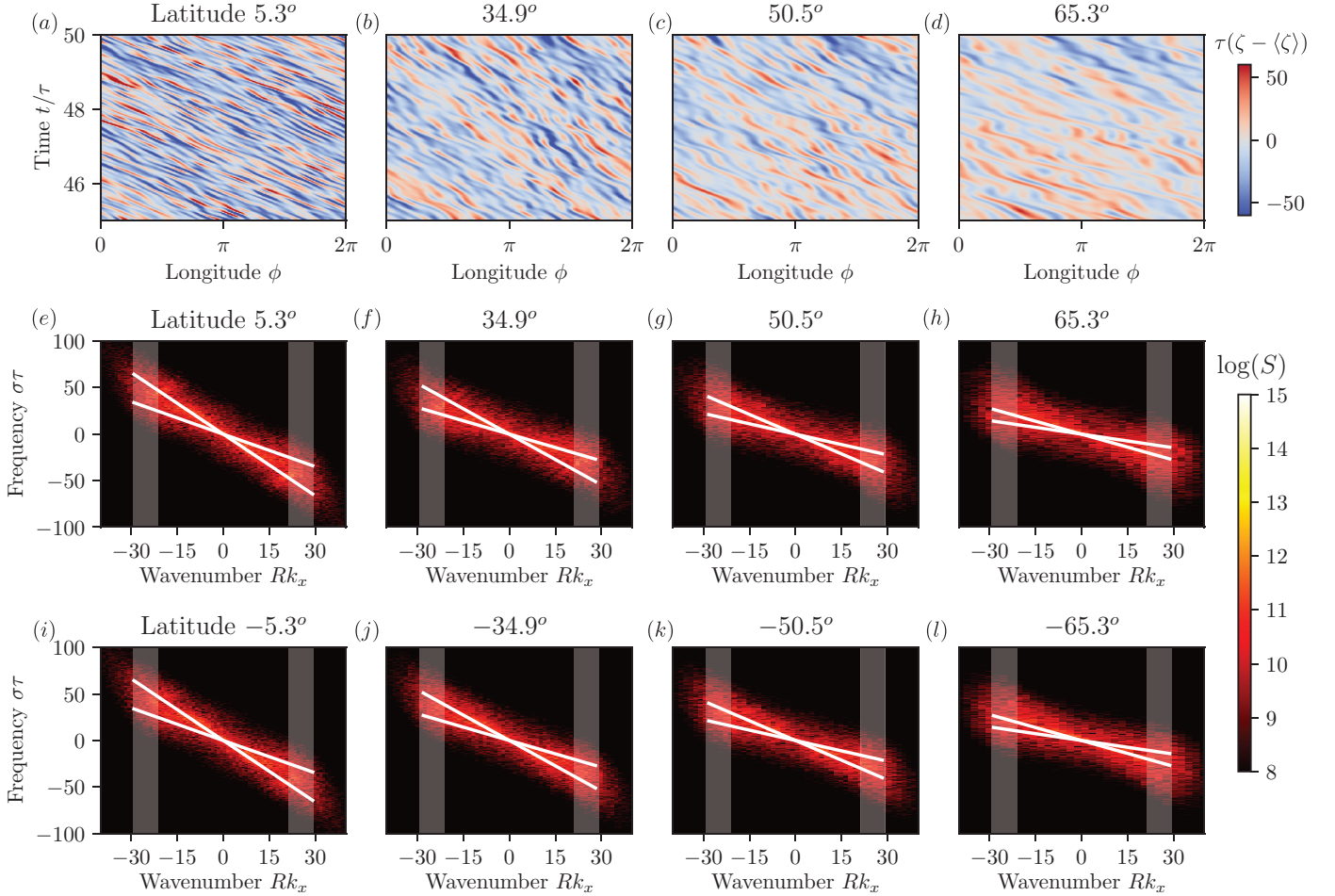


Figure 3-6: (a–d) Time-space diagrams of the deviation of vorticity, $\zeta - \langle \zeta \rangle$, where $\langle \cdot \rangle$ is the average over time and space, indicate that the phase speed of the westward propagating Rossby waves in the local β -plane changes with latitude. (e–l) Logarithm of the power spectral density, $S = |\tau \hat{\zeta}(k_x, \sigma)|^2$, where $\hat{\zeta}$ is the discrete Fourier transform, at different northern (e–h) and southern (i–l) latitudes. The grey regions indicate the forcing bandwidth with $k_- < |k_x| < k_+$ justifying the rapid decay of the power spectral density for $|k_x| > k_+$. The white lines in each panel show the analytical dispersion relation from (3.14) with $|\mathbf{k}| = k_+$ and $|\mathbf{k}| = k_-$; the one with the steeper slope corresponds to k_- . These predictions capture the variance of power spectral density.

The simulations are initialized with a random stream function and evolved from time $t/\tau = 0$ to $t/\tau = 50$. In all simulations, we vary the parameters R/Λ and $\Omega\tau$ for fixed dimensionless bandwidth $\kappa\Lambda = 1$. Narrow-band driving with $\kappa\Lambda \ll 1$ leads to ‘burst’ dynamics [66] whereas broad-band driving $\kappa\Lambda \gg 1$ leads to classical turbulence [93]. The simulations settle onto statistically stationary flow states after initial relaxation periods during which the active stresses continuously inject energy until the forcing and dissipation balance. The analysis below focuses on the statistically stationary states.

Figures 3-1(a-i) show snapshots of the dimensionless relative vorticity $\zeta\tau$ for a range of R/Λ and $\Omega\tau$ at $t/\tau = 15$. In the non-rotating case $\Omega\tau = 0$, we attain solutions akin to those obtained by Mickelin et al. [66]. When the dimensionless rotation rate $\Omega\tau$ is increased, the flow becomes zonal, that is, the ϕ -variation in the vorticity field decreases. At the highest rotation rate $\Omega\tau = 500$, the vorticity field contains alternating bands of high and low vorticity with a characteristic width. For comparison, we plot the steady state solutions $Y_{[\ell-1]}^0(\theta)$, where $[\cdot]$ is the ceiling function, in figures 3-1 (j-l). This corresponds to the smallest ℓ inside the active band. The formation of zonal flows in our model is consistent with the view [132, 133] that such flow structures can be described within a generic pattern formation framework.

To better visualize the banded solutions for high rotation rates, we plot Mercator projections of the vorticity in figure 3-3 (panels a, e, i). The banded nature of the vorticity is also reflected in the alternating structure of the mean azimuthal velocity $\langle v_\phi \rangle_\phi$ (panels b, f, j). The predominant scales in the flow field can be measured using the spherical harmonic decomposition of the relative vorticity. Since $\zeta\tau$ is a real field, we plot only the coefficients with positive m in panels c, g, k. The largest modes have $m = 0$ with ℓ values in the active band of the GNS model (indicated in grey). Panels (d-l) of figure 3-3 show the total energy and the energy contained in the active $m = 0$ modes as a function of time, calculated from the spherical harmonic coefficients as

$$\frac{E(t)}{R^2/\tau^2} = \frac{\tau^2}{R^2} \sum_{m,\ell \neq 0} E_{m,\ell}(t) = \frac{1}{2} \sum_{m,\ell \neq 0} \frac{|\tau \hat{\zeta}_{m,\ell}(t)|^2}{R^2 \ell(\ell+1)}. \quad (3.15)$$

After the initial relaxation phase, when the energy injection balances the energy dissipation, the total energy in the system fluctuates around a statistical mean. The active $m = 0$ modes carry most of the total energy, implying that the bulk dynamics are dominated by these few modes. This also explains the bandedness of the flow patterns since spherical harmonics with $m = 0$ do not vary with the azimuthal angle ϕ ; see figure 3-1 (panels j-l). Time-averaged energy spectra of the statistically steady states are plotted in figure 3-4. The energy shows a clear peak within the active bandwidth further suggesting that the active modes carry most of the total energy.

Strikingly, the statistically stationary states exhibit Rossby waves. For high rotation rates, the Rossby number Ro defined by (3.6) is $\ll 1$. Thus, we can directly compare the linear phase speed given by (3.11) with the slopes of the linear least squares fits to the phase evolution of the coefficients of vorticity $\hat{\zeta}_{m,\ell}(t)$ from the simulations. The normalized phase speed of the modes for different values of R/Λ and $\Omega\tau = 500$ is shown in figure 3-5. The plotted modes have ℓ in the active bandwidth, corresponding to the grey regions in figure 3-3 (panels c, g, k). The phase speed of the modes from the simulations are close to 1 when normalized by the analytical prediction (figure 3-5), implying that the linearized theory captures the main characteristics of the nonlinear dynamics.

We also analyze the phase speed of the waves as a function of latitude. According to the dispersion relation in (3.14), the Rossby-wave phase speed c_p^x depends on the co-latitude θ_0 through $\beta = 2\Omega \sin \theta_0/R$. To check this prediction, we examine the local dynamics at a number of discrete latitudes $(\pi/2 - \theta_0)$ shown in panels (a-d) of figure 3-6 for $R/\Lambda = 8$ and $\Omega\tau = 500$ (the β -plane approximation is valid for these parameters). We show the 2D discrete Fourier transform in time and spatial coordinate $x = R(\sin \theta_0)\phi$ in panels (e-l) for the same northern and southern latitudes. The unstable modes lie within the forcing bandwidth or when $k_- < |\mathbf{k}| < k_+$. Hence, we plot the expected wave dispersion (3.14) making the approximations $|\mathbf{k}| \simeq k_+$ and $|\mathbf{k}| \simeq k_-$. This produces two analytical curves for $\sigma(k_x)$ which are linear in k_x for $k_x < k_+$. These curves capture the spread of the spectral power in the nonlinear dynamics at every latitude; see the white lines in figures 3-6(e-l). We subsequently infer that the

nonlinear, statistically stationary states contain modes with phase speeds matching those of linear β -plane Rossby waves.

3.5 Conclusions

This chapter has addressed the broader question of how internally driven phenomenological models can be utilized to minimally model other chaotic pattern-forming systems. Specifically, motivated by pattern formation in rotating planetary atmospheres, we have presented analytical and numerical solutions of generalized Navier-Stokes (GNS) equations on a 2D rotating sphere. This phenomenological model generalizes the widely studied barotropic vorticity equation by adding an internal forcing that injects energy within a fixed spectral bandwidth. We derived a family of exact time-dependent solutions to the GNS equations on the rotating sphere as well as in the local β -plane. These solutions correspond to a superposition of zonal jets and westward-propagating Rossby waves. Simulations at high rotation rates confirm that the statistically stationary states are close to these exact solutions. We further showed that the phase speeds of waves in the simulations agree with those predicted for linear Rossby waves.

Given the generic nature of the GNS approach, it is possible to modify the functional form of the spectral forcing to incorporate more than one dominant length scale. Our results suggest that the GNS framework can serve as a useful minimal model for providing analytical insight into complex flows on rotating spheres, such as planetary atmospheres. Additionally, they also indicate that phenomenological modeling of collective dynamics in active fluids can help enhance our understanding of chaotic pattern-forming systems in general.

Chapter 4

Learning hydrodynamic equations for active matter from particle simulations and experiments

The contents of this chapter are available in the preprint: R. Supekar, B. Song, A. Hastewell, A. Mietke, and J. Dunkel, *arXiv:2101.06568*, 2021 [134].

4.1 Introduction

Natural and engineered active matter, from cells [135], tissues [136] and organisms [137] to self-propelled particle suspensions [9, 6] and autonomous robots [15, 138, 139], exhibits complex dynamics across a wide range of length and time scales. Predicting the collective self-organization and emergent behaviors of such systems requires extensions of traditional theories that go beyond conventional physical descriptions of non-living matter [40, 11, 12]. Due to the inherent complexity of active matter interactions in multi-cellular communities [46, 140] and organisms [47], or even non-equilibrium chemical [141] or colloidal [9, 6, 142] systems, it becomes increasingly difficult and inefficient for humans to formulate and quantitatively validate continuum theories from first principles. A key question is therefore whether one can utilize computing machines [49] to identify interpretable systems of equations that elucidate

the mechanisms underlying collective active matter dynamics.

Enabled by recent major advances in microscopic imaging [143, 144, 46, 47] and agent-based computational modeling [48], active matter systems can now be observed and analyzed at unprecedented spatiotemporal [145, 146, 147] resolution. To infer interpretable predictive theories, the high-dimensional data recorded in experiments or simulations have to be compressed and translated into low-dimensional models. Such learned models must faithfully capture the macroscale dynamics of the relevant collective properties. Macroscale properties can be efficiently encoded through hydrodynamic variables, continuous fields that are linked to the symmetries and conservation laws of the underlying microscopic system [11, 12]. Although much theoretical progress has been made in the field of dynamical systems learning over the last two decades [148, 58, 149, 59, 56, 57, 65], the inference of hydrodynamic models from particle data has remained largely unsuccessful in practice, not least due to severe complications arising from measurement noise, inherent fluctuations and self-organized scale-selection in active systems. Yet, extrapolating the current experimental revolution [143, 144, 9, 6, 46, 140], data-driven equation learning will become increasingly more important as simultaneous observations of physical, biological, and chemical properties of individual cells and other active units will become available in the near future [150, 151].

Learning algorithms for ordinary differential equations (ODEs) and partial differential equations (PDEs) have been proposed and demonstrated based on least-squares fitting [148, 58], symbolic regression [149, 59], and sparse regression [56, 57] combined with weak formulations [62, 63], artificial neural networks [60, 61, 54], and stability selection [64, 65]. These groundbreaking studies, however, focused primarily on data that describes *a priori* known continuum models, and recent coarse-graining applications have remained limited to ODEs [152] or one-dimensional PDEs [153, 154]. By contrast, it is still an open challenge to infer higher-dimensional hydrodynamic PDE models directly from microscopic active matter simulations or experiments.

In this chapter, we present a comprehensive learning framework that takes microscopic particle data as input and generates sparse interpretable predictive hy-

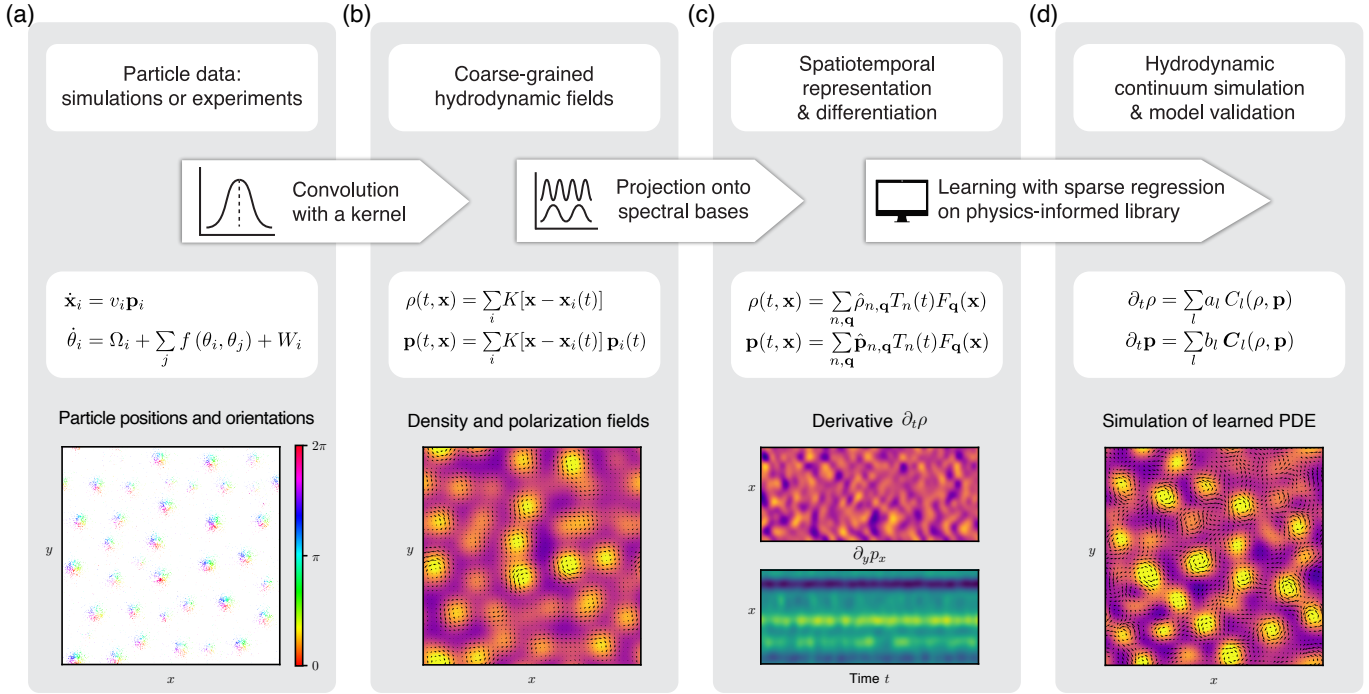


Figure 4-1: Learning hydrodynamic models from particle simulations and experiments. (a) Inputs are time-series data for particle positions $\mathbf{x}_i(t)$, particle orientations $\mathbf{p}_i(t) = (\cos \theta_i, \sin \theta_i)^\top$, etc., measured in simulations or experiments with microscale resolution (§ 4.2.1). (b) Spatial kernel coarse-graining of the discrete microscopic variables provides continuous hydrodynamic fields, such as the density $\rho(t, \mathbf{x})$ or the polarization density $\mathbf{p}(t, \mathbf{x})$ (§ 4.2.2). (c) Coarse-grained fields are sampled on a spatiotemporal grid and projected onto suitable spectral basis functions. Systematic spectral filtering (compression) ensures smoothly interpolated hydrodynamic fields, enabling efficient and accurate computation of spatiotemporal derivatives (§ 4.2.3). (d) Using these derivatives, a library of candidate terms $C_l(\rho, \mathbf{p})$ and $\mathbf{C}_l(\rho, \mathbf{p})$ consistent with prior knowledge about conservation laws and broken symmetries is constructed. A sparse regression algorithm determines subsets of relevant phenomenological coefficients a_l and b_l (§ 4.2.4). The resulting hydrodynamic models are sparse and interpretable, and their predictions can be directly validated against analytic coarse-graining results (§ 4.2.5) or experiments (§ 4.3). Bottom: Snapshots illustrating the workflow for microscopic data generated from simulations of chiral active Brownian particles [equation (4.1)].

hydrodynamic models as output (figure 4-1). We demonstrate its practical potential in applications to active particle data from simulations and recent experiments [9]. In both cases, we find that the learned hydrodynamic models predict the emergent collective dynamics not only qualitatively but also quantitatively. Conceptually, this advance is made possible by leveraging spectral basis representations [155] for systematic denoising and robust numerical differentiation, and by explicitly incorporating physical constraints into the learning framework. Our analysis achieves the goal of this thesis to show how insights from analytic coarse-graining calculations and prior knowledge of conservation laws and broken symmetries can enhance the robustness of automated equation discovery from microscopic data.

4.2 Learning framework

Our model learning approach combines recent advances in sparse PDE recovery [57, 64] with spectral filtering and compression [156, 157, 129]. We first demonstrate the key steps of the general framework (figure 4-1) for an experimentally motivated chiral active particle model, for which the hydrodynamic continuum equations were not known previously. Later on, we will apply the same methodology to infer a quantitative hydrodynamic model directly from video data recorded in recent colloidal micro-roller experiments [9] (§ 4.3).

4.2.1 Active particle simulations ¹

To generate challenging test data for the learning algorithm, we simulated a 2D system of interacting self-propelled chiral particles [158, 37, 38, 159, 160]. Microscopic models of this type are known to capture essential aspects of the experimentally observed self-organization of protein filaments [161, 162], bacterial swarms [145, 163, 39] and cell monolayers [164]. In the simulations, a particle i with orientation $\mathbf{p}_i = (\cos \theta_i, \sin \theta_i)^\top$

¹These simulations were performed by Boya Song, Department of Mathematics, MIT.

moved and changed orientation according to the Brownian dynamics

$$\frac{d\mathbf{x}_i}{dt} = v_i \mathbf{P}_i, \quad (4.1a)$$

$$\frac{d\theta_i}{dt} = \Omega_i + g \sum_{j \in \mathcal{N}_i} \sin(\theta_j - \theta_i) + \sqrt{2D_r} \eta_i, \quad (4.1b)$$

where $\eta_i(t)$ denotes orientational Gaussian white noise, with $\langle \eta_i(t) \eta_j(t') \rangle = \delta_{ij} \delta(t - t')$ and zero mean, modulated by the rotational diffusion constant D_r . The parameter $g > 0$ determines the alignment interaction strength between particles i and j within a neighborhood \mathcal{N}_i of radius R . The self-propulsion speed $v_i \geq 0$ and orientational rotation frequency $\Omega_i \geq 0$ were drawn from a joint distribution $p(v_i, \Omega_i)$ (appendix C.1.1). This heuristic distribution was chosen such that long-lived vortex states, similar to those observed in swimming sperm cell suspensions [4], formed spontaneously from arbitrary random initial conditions (figure 4-2a). Emerging vortices are left-handed for $\Omega_i \geq 0$, and their typical size is $\sim \langle v_i \rangle_p / \langle \Omega_i \rangle_p$, where $\langle \cdot \rangle_p$ denotes an average over the parameter distribution $p(v_i, \Omega_i)$. We simulated equation (4.1) in non-dimensionalized form, choosing the interaction radius R as reference length and $R / \langle v_i \rangle_p$ as time scale. Accordingly, we set $R = 1$ and $\langle v_i \rangle_p = 1$ from now on. Simulations were performed for $N = 12,000$ particles on a periodic domain of size 100×100 (figure 4-2a).

From a learning perspective, this model poses many of the typical challenges that one encounters when attempting to infer hydrodynamic equations from active matter experiments: spontaneous symmetry breaking and meso-scale pattern formation, microscopic parameter variability, noisy dynamics, anisotropic interactions, and so on. Indeed, similar to many experimental systems, it is not even clear *a priori* whether or not equations (4.1) permit a quantitative description in terms of a sparse hydrodynamic continuum model.

4.2.2 Hydrodynamic fields

Given particle-resolved data, hydrodynamic fields are obtained by coarse-graining. A popular coarse-graining approach is based on convolution kernels [165, 166], weight

functions that translate discrete fine-grained particle densities into continuous fields, analogous to the point spread function of a microscope. For example, given the particle positions $\mathbf{x}_i(t)$ and orientations $\mathbf{p}_i(t)$, an associated particle number density field $\rho(t, \mathbf{x})$ and polarization density field $\mathbf{p}(t, \mathbf{x})$ can be defined by

$$\rho(t, \mathbf{x}) = \sum_i K[\mathbf{x} - \mathbf{x}_i(t)], \quad (4.2a)$$

$$\mathbf{p}(t, \mathbf{x}) = \sum_i K[\mathbf{x} - \mathbf{x}_i(t)] \mathbf{p}_i(t). \quad (4.2b)$$

The symmetric kernel $K(\mathbf{x})$ is centered at $\mathbf{x} = 0$ and normalized, $\int d^2\mathbf{x} K(\mathbf{x}) = 1$, so that the total number of particles is recovered from $\int d^2\mathbf{x} \rho(t, \mathbf{x}) = N$. Equations (4.2) generalize to higher tensorial density fields in a straightforward manner, and can be readily adapted to accommodate different boundary conditions (appendix C.1.2).

We found that, in the context of hydrodynamic model learning, the coarse-graining with a Gaussian kernel $K(\mathbf{x}) \propto \exp[-|\mathbf{x}|^2/(2\sigma^2)]$ in (4.2) presents a useful preprocessing step that simplifies the use of fast transforms at later stages. In practice, the coarse-graining scale σ has to be chosen larger than the particle’s mean-free path, but small enough to resolve the relevant collective structures. For the microscopic test data from equations (4.1), $\sigma = 5$ provides a sufficient resolution of the emerging vortex patterns (figure 4-2a, bottom panel).

4.2.3 Spatiotemporal representation and differentiation

A central challenge in PDE learning is the computation of spatial and temporal derivatives of the coarse-grained fields. Our framework exploits that hydrodynamic models aim to capture the long-wavelength dynamics of the slow collective modes [11]. This fact allows us to project the coarse-grained fields on suitable basis functions that additionally enable sparse representations (high compression), fast transforms

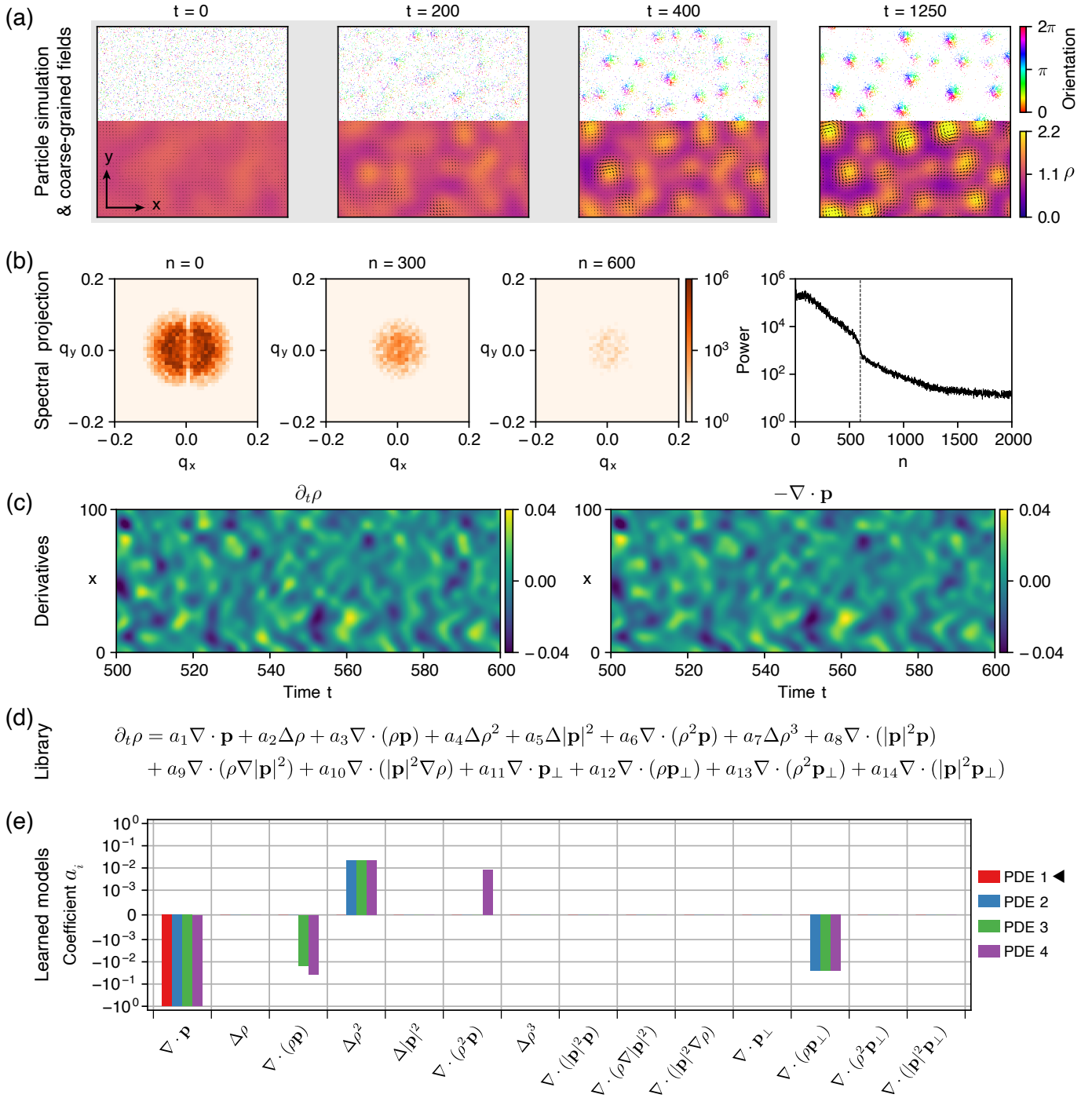


Figure 4-2: Learning mass conservation dynamics. (a) Top: Time evolution of positions and orientations of 12,000 particles following the dynamics in equations (4.1). Bottom: Coarse-grained density ρ (color code) and polarization field \mathbf{p} (arrows). Starting from random initial conditions ($t = 0$), a long-lived vortex pattern with well-defined handedness emerges ($t = 1250$). Training data were randomly sampled from the time window $t \in [40, 400]$, contained within the gray box. Domain size: 100×100 . (b) Slices through the power spectrum $S_{x;n,\mathbf{q}} = |\mathbf{e}_x \cdot \hat{\mathbf{p}}_{n,\mathbf{q}}|^2$ for different values of the Chebyshev polynomial order $n \in \{0, 300, 600\}$, corresponding to modes with increasing temporal frequencies. The rightmost panel depicts the total spectral power $\sum_{\mathbf{q}} S_{x;n,\mathbf{q}}$ [see equation (4.3b)] of each Chebyshev mode n . The slowly decaying long tail of fast modes indicates a regime in which fluctuations dominate over a smooth signal. The cut-off $n_0 = 600$ removes these modes, in line with the goal to learn a hydrodynamic model for the slow long-wavelength modes. (c) Kymographs of the spectral derivatives $\partial_t \rho$ and $-\nabla \cdot \mathbf{p}$ at $y = 50$, obtained from the spectrally truncated data. (d) Mass conservation in the microscopic system restricts the physics-informed candidate library to terms that can be written as divergence of a vector field. (e) Learned phenomenological coefficients a_i of PDEs with increasing complexity (decreasing sparsity) (appendix C.3). PDE 1 (\blacktriangleleft) is given by $\partial_t \rho = a_1 \nabla \cdot \mathbf{p}$ with $a_1 = -0.99$. As PDE 1 is the sparsest PDE that agrees well with analytic coarse-graining results (table 4.1), it is selected for the hydrodynamic model.

and efficient differentiation. Here, we work with representations of the form

$$\rho(t, \mathbf{x}) = \sum_{n, \mathbf{q}} \hat{\rho}_{n, \mathbf{q}} T_n(t) F_{\mathbf{q}}(\mathbf{x}), \quad (4.3a)$$

$$\mathbf{p}(t, \mathbf{x}) = \sum_{n, \mathbf{q}} \hat{\mathbf{p}}_{n, \mathbf{q}} T_n(t) F_{\mathbf{q}}(\mathbf{x}), \quad (4.3b)$$

where $T_n(t)$ denotes a degree- n Chebyshev polynomial of the first kind [167, 168], $F_{\mathbf{q}}(\mathbf{x}) = \exp(2\pi i \mathbf{q} \cdot \mathbf{x})$ is a Fourier mode with wave vector $\mathbf{q} = (q_x, q_y)^\top$, and $\hat{\rho}_{n, \mathbf{q}}$ and $\hat{\mathbf{p}}_{n, \mathbf{q}}$ are complex mode coefficients (figure 4-2b and appendix C.1.3). Generally, the choice of the basis functions should be adapted to the spatiotemporal boundary conditions of the microscopic data (§ 4.3).

The spectral representation (4.3) enables the efficient and accurate computation of space and time derivatives [169]. Preprocessing via spatial coarse-graining (§ 4.2.2) ensures that the mode coefficients $\hat{\rho}_{n, \mathbf{q}}$ and $\hat{\mathbf{p}}_{n, \mathbf{q}}$ decay fast for $|\mathbf{q}| \gg 1/(2\pi\sigma)$ (figure 4-2b, left). Furthermore, focusing on the slow hydrodynamic modes, we can filter out the fast modes with $n > n_0$ by keeping only the dominant Chebyshev terms in equation (4.3). The cut-off value n_0 can usually be directly inferred from a characteristic steep drop-off in the power spectrum of the data, which signals the transition to hydrodynamically irrelevant fast fluctuations [170] (figure 4-2b, right). Choosing n_0 according to this criterion yields accurate, spatiotemporally consistent derivatives as illustrated for the kymographs of the derivative fields $\partial_t \rho$ and $-\nabla \cdot \mathbf{p}$, which are essential to capture mass conservation. More generally, combining kernel-based and spectral coarse-graining also mitigates measurement noise, enabling a direct application to experimental data (§ 4.3).

4.2.4 Inference of hydrodynamic equations

To infer hydrodynamic models that are consistent with the coarse-grained projected fields (4.3), we build on a recently proposed sparse regression framework [56, 57]. The specific aim is to determine sparse PDEs for the density and polarization dynamics

of the form

$$\partial_t \rho = \sum_l a_l C_l(\rho, \mathbf{p}), \quad (4.4a)$$

$$\partial_t \mathbf{p} = \sum_l b_l \mathbf{C}_l(\rho, \mathbf{p}). \quad (4.4b)$$

The candidate library terms $\{C_l(\rho, \mathbf{p})\}$ and $\{\mathbf{C}_l(\rho, \mathbf{p})\}$ are functions of the fields and their derivatives, which can be directly evaluated at various sample points using the spectral representation (4.3). Equations (4.4) thus define a linear system for the phenomenological coefficients a_l and b_l , and the objective is to find sparse solutions such that the resulting hydrodynamic model recapitulates the collective particle dynamics.

Learned hydrodynamic models must respect the symmetries of the underlying microscopic dynamics. Prior knowledge of such symmetries can greatly accelerate the inference process by placing constraints on the model parameters a_l and b_l . The learning ansatz (4.4b) already assumes global rotational invariance by using identical coefficients b_l for the x and y components of the polarization field equations. Generally, coordinate-independence of hydrodynamic models demands that the dynamical fields and the library functions C_l , \mathbf{C}_l , etc. have the correct scalar, vectorial or tensorial transformation properties. This fact imposes stringent constraints on permissible libraries, as do microscopic conservation laws.

Symmetries and conservation laws: Generating a physics-informed candidate library

Whenever prior knowledge about (broken) symmetries and conservation laws is available, it should inform the candidate library construction to ensure that the PDE learning is performed within a properly constrained model space. For example, the active particle model in equation (4.1) describes a chiral dynamical system with intrinsic microscopic rotation rates $\Omega_i \geq 0$. The space of valid hydrodynamic models therefore includes PDEs in which the mirror symmetry is explicitly broken. Formally, this implies the Levi-Civita symbol ϵ_{ij} can be used to generate a pseudo-vector $\mathbf{p}_\perp := \boldsymbol{\epsilon}^\top \cdot \mathbf{p} = (-p_y, p_x)^\top$ that has to be included in the construction of the candidate

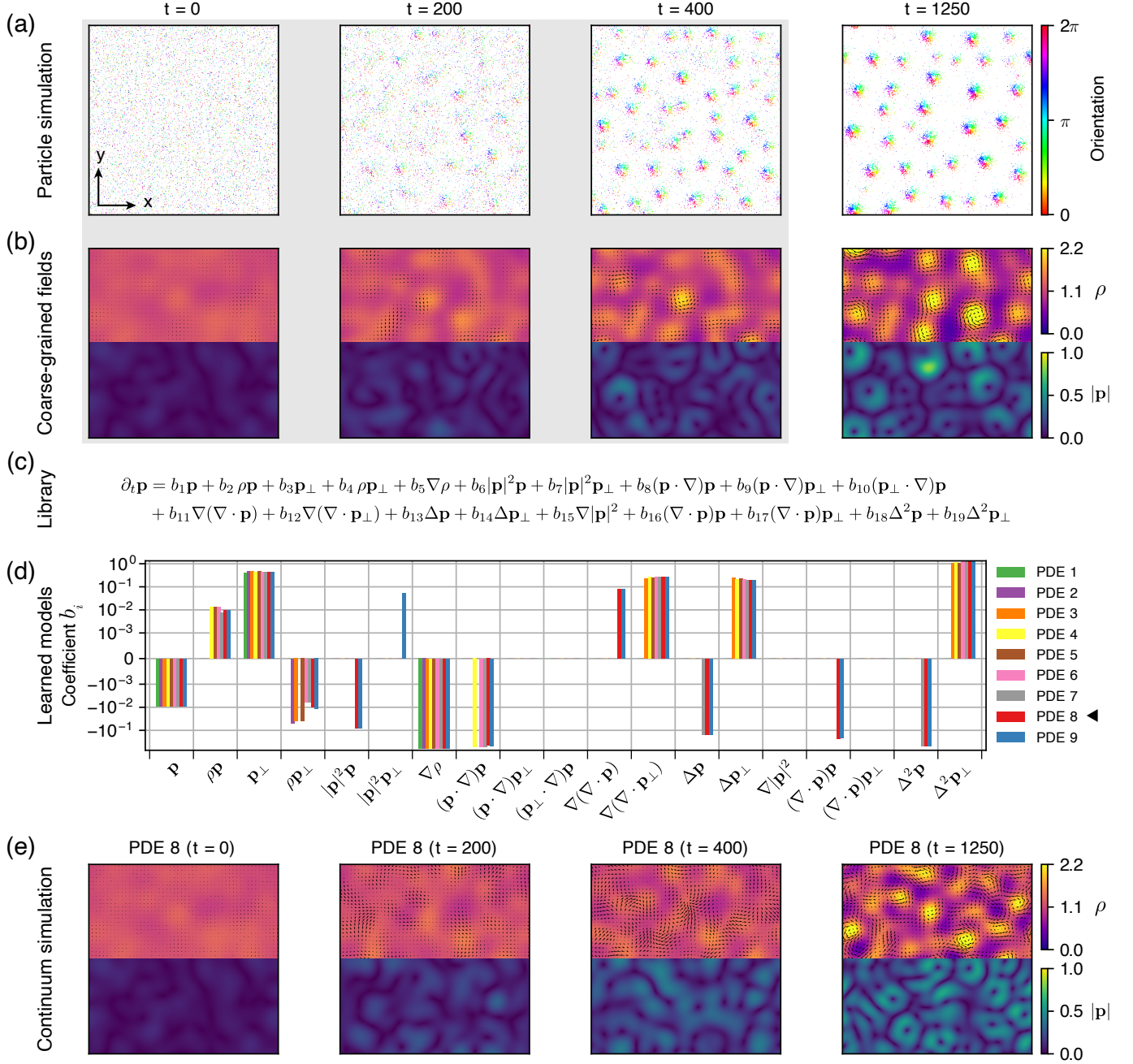


Figure 4-3: Learning polarization dynamics. (a) Same particle dynamics as in figure 4-2a for visual reference. (b) Top: Coarse-grained density and polarization field as in figure 4-2a. Bottom: Magnitude $|\mathbf{p}|$ of the coarse-grained polarization field. Emerging vortices ($t = 400, 1250$) appear as ring-like patterns in $|\mathbf{p}|$. Training data were randomly sampled from the time window $t \in [40, 400]$, enclosed within the gray box. **c**: Physics-informed candidate library (with $b_1 = -D_r$) including terms constructed from $\mathbf{p}_\perp = (-p_y, p_x)^\top$, which are allowed due to the chirality of the microscopic system. **d**: Learned phenomenological coefficients b_i of PDEs with increasing complexity (appendix C.3). For all PDEs, learned coefficients of the linear terms \mathbf{p}_\perp and $\nabla \rho$ compare well with analytic predictions (table 4.1, appendix C.2.2). **e**: Simulation of the final hydrodynamic model (PDE 8 for the polarization dynamics and PDE 1 in figure 4-2e for the density dynamics). Starting from random initial conditions ($t = 0$), long-lived vortex states emerge on a similar time scale, with similar spatial patterns, and with comparable density and polarization amplitudes as in the coarse-grained microscopic model data (b). Hydrodynamic models with PDEs sparser than PDE 8 do not form stable vortex patterns.

libraries $\{C_l(\rho, \mathbf{p})\}$ and $\{\mathbf{C}_l(\rho, \mathbf{p})\}$.

Additional constraints on the scalar library terms $\{C_l(\rho, \mathbf{p})\}$ in equation (4.4a) arise from particle number conservation, which is known to hold for many experimental active matter systems as well as for microscopic models such as equation (4.1). In this case, it suffices to consider functions C_l that can be written as the divergence of a vector field that represents some flux contribution. In the application below, we included fluxes up to first order in derivatives and third order in the fields (figure 4-2d); if needed, this scalar library can be easily extended.

The vectorial library $\{\mathbf{C}_l(\rho, \mathbf{p})\}$ for the chiral polarization dynamics, equation (4.4b), cannot be constrained further by symmetries or conservation laws. Mechanical substrate interactions with the environment as invoked by the microscopic model (4.1) and present in many active matter experiments explicitly break Galilean invariance, leading to external forces and torques whose form is not known *a priori*. We therefore included in equation (4.4b) also vector fields that cannot be written as a divergence, such as \mathbf{p}_\perp , $\rho\mathbf{p}$ or $(\mathbf{p} \cdot \nabla)\mathbf{p}$, in our candidate library $\{\mathbf{C}_l(\rho, \mathbf{p})\}$.

In general, higher order terms and terms including derivatives can be systematically constructed from the basic set of available fields and operators $\mathcal{B} = \{\rho, \mathbf{p}, \mathbf{p}_\perp, \nabla\}$. We illustrate the general procedure for an example library containing up to cubic terms from \mathcal{B} . The first step is to write the list of distinct rank-2 tensors

$$\mathcal{S} = \{s\mathbb{I}, \mathbf{p}\mathbf{p}, \mathbf{p}\mathbf{p}_\perp, \mathbf{p}_\perp\mathbf{p}_\perp, \nabla\mathbf{p}, \nabla\mathbf{p}_\perp\},$$

where $s \in \{1, \rho, \nabla \cdot \mathbf{p}, \nabla \cdot \mathbf{p}_\perp\}$ represents one of the linearly independent scalars that can be formed from elements in \mathcal{B} . From any tensor $\Sigma \in \mathcal{S}$ and its transpose, we can then generate vectorial terms \mathbf{C}_l by forming scalar products with the elements in \mathcal{B} . In particular, terms $\nabla \cdot \Sigma$ yield possible contributions from internal stresses and torques due to alignment interactions, while $\Sigma \cdot \mathbf{p}$ and $\Sigma \cdot \mathbf{p}_\perp$ correspond to substrate-dependent interactions.

For pattern-forming systems with emergent length scale selection, the library should be extended to include Swift-Hohenberg-type [171] terms $\Delta^2\mathbf{p}$, $\Delta^2\mathbf{p}_\perp$, etc. [3, 19].

Such terms can stabilize small-wavelength modes and, combined with $\Delta \mathbf{p}$ and $\Delta \mathbf{p}_\perp$, can give rise to patterns of well-defined length [171]. The final 19-term library with linearly independent terms (appendix C.1.5) used to learn the polarization dynamics for the chiral particle model from equation (4.1) is summarized in figure 4-3c.

Sparse model learning

To determine the hydrodynamic parameters a_l and b_l in equations (4.4), we randomly sampled the coarse-grained fields $\rho(t, \mathbf{x})$ and $\mathbf{p}(t, \mathbf{x})$ and their derivatives at $\sim 10^6$ space-time points within a predetermined learning interval (appendix C.1). Generally, the success or failure of hydrodynamic model learning depends crucially on the choice of an appropriate space-time sampling interval. As a guiding principle, learning should be performed during the relaxation stage, when both time and space derivatives show the most substantial variation.

Evaluating equations (4.4a) and (4.4b) at all sample points yields linear systems of the form $\mathbf{U}_t = \Theta \boldsymbol{\xi}$, where the vector \mathbf{U}_t contains the time derivatives (appendix C.1.4). The columns of the matrix Θ hold the numerical values of the library terms $C_l(\rho, \mathbf{p})$ and $\mathcal{C}_l(\rho, \mathbf{p})$ computed from the spectral representations (4.3). The aim is to infer a parsimonious model so that the vector $\boldsymbol{\xi}$ containing the hydrodynamic parameters a_l or b_l is sparse. In this case, the corresponding PDE only contains a subset of the library terms, and we refer to the total number of terms in a PDE as its *complexity*.

To estimate sparse parameters $\boldsymbol{\xi}$, we used the previously proposed sequentially thresholded least-squares (STLSQ) algorithm from SINDy [56]. STLSQ first finds the least-squares estimate $\hat{\boldsymbol{\xi}} = \arg \min_{\boldsymbol{\xi}} \|\mathbf{U}_t - \Theta \boldsymbol{\xi}\|_2^2$. Subsequently, sparsity of $\hat{\boldsymbol{\xi}}$ is imposed by iteratively setting coefficients below a thresholding hyperparameter τ to zero. Adopting a stability selection approach [172, 173, 64, 65] in which τ is systematically varied over a regularization path $[\tau_{\max}, \tau_{\min}]$ (appendix C.1.4), we obtain candidate PDEs of increasing complexity (Figs. 4-2e and 4-3d) whose predictions need to be validated against the phenomenology of the input data.

Performance improvements and pitfalls

Sparse regression-based learning becomes more efficient and robust if known symmetries or other available information can be used to reduce the number of undetermined parameters a_l and b_l in equations (4.4). Equally helpful and important is prior knowledge of the relevant time and length scales. The coarse-grained field data need to be sampled across spatiotemporal scales that contain sufficient dynamical information; over-sampling in a steady-state typically prevents algorithms from learning terms relevant to the relaxation dynamics. Systems exhibiting slow diffusion time-scales can pose additional challenges. For example, generic analytic coarse-graining (appendix C.2.1) shows that additive rotational noise as in equation (4.1b) implies the linear term $-D_r \mathbf{p}$ in the polarization dynamics in (4.4b). If the diffusive time scale $1/D_r$ approaches or exceeds the duration of the sampling time interval, then the learned PDEs may not properly capture the relaxation dynamics of the polarization field. From a practical perspective, this is not a prohibitive obstacle, as the rotational diffusion coefficient D_r can be often measured independently from isolated single-particle trajectories [174]. In this case, fixing $-D_r \mathbf{p}$ in equation (4.4b) and performing the regression over the remaining parameters produced satisfactory learning results (see figure 4-3, where $1/D_r \sim 100$ is comparable to the length of the learning interval $t \in [40, 400]$).

4.2.5 Validation of learned models

The STLSQ algorithm with stability selection proposes PDEs of increasing complexity – the final learning step is to identify the sparsest acceptable hydrodynamic model among these (figure 4-1). This can be achieved by simulating all the candidate PDEs (appendix C.1.6) and comparing their predictions against the original data and, if available, against analytic coarse-graining results (appendix C.2).

For the microscopic particle model from equation (4.1), the sparsest learned PDE for the particle number density is $\partial_t \rho = a_1 \nabla \cdot \mathbf{p}$ (figure 4-2e); this mass conservation equation is also predicted by analytic coarse-graining (appendix C.2). The learned co-

Table 4.1: Parameters of the hydrodynamic model learned for the microscopic dynamics in equation (4.1) and values predicted by analytic coarse-graining (appendix C.2.2). $\langle \cdot \rangle_p$ denotes averages over the distribution $p(v_i, \Omega_i)$ of particle velocities v_i and rotation rates Ω_i (appendix C.1.1).

Term	Learned value	Analytic coarse-graining
Density dynamics		
$a_1 \nabla \cdot \mathbf{p}$	$a_1 = -0.99$	$-\langle v_i \rangle_p = -1.00$
Polarization dynamics		
$b_3 \mathbf{p}_\perp$	$b_3 = 0.44$	$\langle v_i \Omega_i \rangle_p / \langle v_i \rangle_p = 0.50$
$b_5 \nabla \rho$	$b_5 = -0.60$	$-\frac{1}{2} \langle v_i^2 \rangle_p / \langle v_i \rangle_p = -0.57$

efficient $a_1 = -0.99$ implies an effective number density flux $-a_1 \mathbf{p} \approx \mathbf{p}$, which agrees very well with the analytic prediction $\langle v_i \rangle_p \mathbf{p} = \mathbf{p}$. Additional coefficients appearing in more complex models proposed by the algorithm are at least one order of magnitude smaller than a_1 (figure 4-2e). Hence, as part of a hydrodynamic description of the microscopic system equation (4.1), we adopt the minimal density dynamics $\partial_t \rho = a_1 \nabla \cdot \mathbf{p}$ from now on.

The sparsest learned PDE for the dynamics of the polarization field \mathbf{p} only contains three terms. However, together with the density dynamics, the resulting hydrodynamic models are either unstable or do not lead to the formation of vortex patterns. Our simulations showed that a certain level of complexity is required to reproduce the dynamics observed in the test data. In particular, there exists a unique sparsest model (PDE 8 in figure 4-3d) for which long-lived vortex states emerge from random initial conditions. The resulting hydrodynamic model exhibits density and polarization patterns quantitatively similar to those observed in the original particle system (figure 4-3a,b,e), which also form on a similar time scale. Furthermore, the learned coefficients of the linear terms $\sim \mathbf{p}_\perp$ and $\sim \nabla \rho$ agree well with the analytic predictions (table 4.1, appendix C.2.2). As the learning only used coarse-grained field data in the time interval $t \in [40, 400]$, simulation results for $t > 400$ represent predictions of the learned hydrodynamic model (figure 4-3e). The close agreement between original data and the model simulations (figure 4-3b,e) show that the inference framework has succeeded in learning a previously unknown hydrodynamic description for a chiral polar active particle system with broadly distributed microscopic parameters.

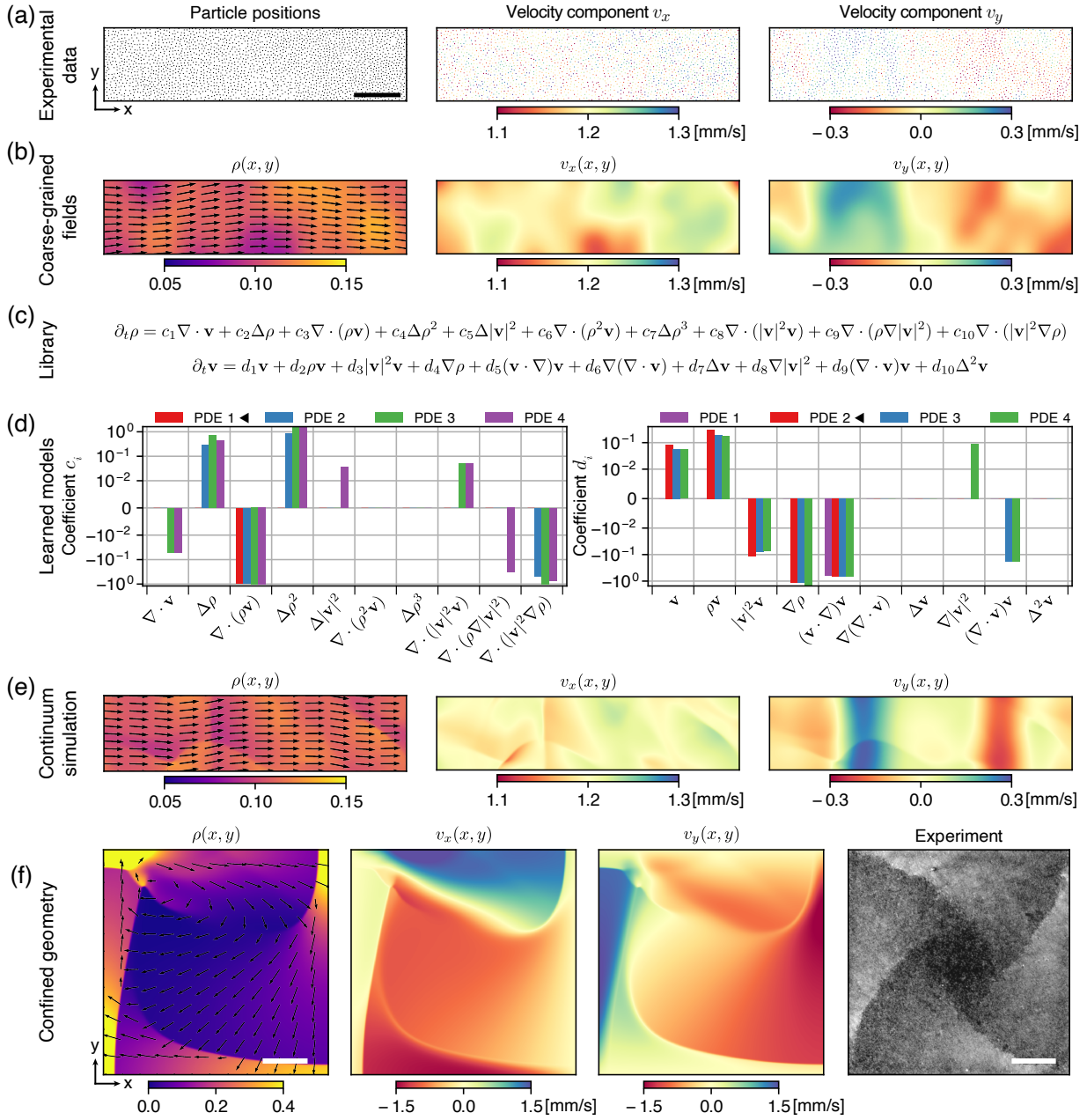


Figure 4-4: Learning from active polar particle experiments. (a) Snapshot of particle positions and velocity components of $\sim 2,200$ spontaneously moving Quincke rollers in a microfluidic channel [9]. Scale bar, $200 \mu\text{m}$. (b) Coarse-grained density field $\rho(t, \mathbf{x})$, expressed as the fraction of area occupied by the rollers with diameter $D_c = 4.8 \mu\text{m}$, and components $v_x(t, \mathbf{x})$ and $v_y(t, \mathbf{x})$ of the coarse-grained velocity field ($\sigma = 45 \mu\text{m}$). 5×10^5 randomly sampled data points from ~ 580 such snapshots over a time duration of 1.4s were used for the learning algorithm. (c) Physics-informed candidate libraries for the density and velocity dynamics, $\{\bar{C}_l(\rho, \mathbf{v})\}$ and $\{\bar{D}_l(\rho, \mathbf{v})\}$, respectively [equation (4.5)]. These are the same libraries as shown in Figs. 4-2e and 4-3d, but without the chiral terms and replacing $\mathbf{p} \rightarrow \mathbf{v}$. (d) Learned phenomenological coefficients c_l and d_l of the four sparsest PDEs for the density (left) and velocity (right) dynamics. The coefficients are non-dimensionalized with length scale σ and time scale σ/v_0 , where $v_0 = 1.2\text{mm s}^{-1}$ is the average roller speed. PDE 1 for density dynamics corresponds to $\partial_t \rho = c_3 \nabla \cdot (\rho \mathbf{v})$ with $c_3 \simeq -0.95$. PDE 2 for the velocity dynamics is shown in equation (4.6b). Learned coefficients compare well with the values reported in Ref. [9] (table 4.2). (e) Simulation snapshot at $t = 1.8\text{s}$ of the learned hydrodynamic model (PDEs marked by \blacktriangleleft in (a)) in a doubly periodic domain. Spontaneous flow emerges from random initial conditions, and exhibits density and velocity fluctuations that show similar spatial patterns and amplitudes as seen in the experiments (a). (f) Simulation snapshots at $t = 18.5\text{s}$ of the same hydrodynamic model as in (e) on a square domain with reflective boundary conditions. The model predicts the emergence of a vortex-like flow permeated by density shock waves. This prediction agrees qualitatively with experimental observations (rightmost panel) of Quincke rollers in a $5\text{mm} \times 5\text{mm}$ confinement with average density $\rho_0 \approx 0.1$ (Image credits: Alexandre Morin, Delphine Geyer, and Denis Bartolo). Scale bars, $200 \mu\text{m}$ (simulation) and 1mm (experiment).

4.3 Learning from experimental data

The inference framework can be readily applied to experimental data. We illustrate this by learning a quantitative hydrodynamic model directly from a video recorded in a recent study [9] of driven colloidal suspensions (figure 4-4a). In these experiments, an electro-hydrodynamic instability enables micron-sized particles to self-propel with speeds up to a few millimeters per second across a surface. The rich collective dynamics of these so-called Quincke rollers [5, 9] provides a striking experimental realization of self-organization in active polar particle systems [41, 175, 11].

4.3.1 Coarse-graining and spectral representation of experimental data

To gather dynamic particle data from experiments, we extracted particle positions $\mathbf{x}_i(t)$ from the Supplementary Movie S2 in Ref. [9], with particle velocities $\mathbf{v}_i(t) = \frac{d}{dt}\mathbf{x}_i$ replacing the particle orientations $\mathbf{p}_i(t)$ from before. This data set captures a weakly compressible suspension of Quincke rollers in a part of a racetrack-shaped channel (figure 4-4a). We then applied the kernel coarse-graining [equations (4.2), $\sigma = 45\mu\text{m}$] to obtain the density field ρ and the velocity field $\mathbf{v} = \mathbf{p}/\rho$. Accounting for the non-periodicity of the data, ρ and \mathbf{v} were projected on a Chebyshev polynomial basis [equation (4.3)] in time and space (figure 4-4b). Filtering out non-hydrodynamic fast modes with temporal mode numbers $n > n_0$, we found that the final learning results were robust for a large range of cut-off modes n_0 (appendix C.3).

4.3.2 Physics-informed library

The goal is to learn a hydrodynamic model of the form

$$\partial_t \rho = \sum_l c_l \bar{C}_l(\rho, \mathbf{v}), \quad (4.5a)$$

$$\partial_t \mathbf{v} = \sum_l d_l \bar{C}_l(\rho, \mathbf{v}), \quad (4.5b)$$

Table 4.2: Parameters of the learned hydrodynamic model for the Quincke roller system are close to values expected from analytic coarse-graining (*) and reported in Ref. [9] for experiments performed at mean area fraction $\rho_0 \approx 0.11$.

Term	Learned values	Ref. [9]
Density dynamics		
$c_3 \nabla \cdot (\rho \mathbf{v})$	$c_3 = -0.95$	-1.0^*
Velocity dynamics		
$(d_1 + d_2 \rho) \mathbf{v}$	$\sqrt{\frac{d_1 + d_2 \rho_0}{-d_3}} = 1.21 \text{ mm/s}$	1.20 mm/s
$+ d_3 \mathbf{v} ^2 \mathbf{v}$		
$d_4 \nabla \rho$	$d_4 = -1.62 \text{ mm}^2/\text{s}^2$	$-5.0 \pm 2.0 \text{ mm}^2/\text{s}^2$
$d_5 (\mathbf{v} \cdot \nabla) \mathbf{v}$	$d_5 = -0.67$	-0.7 ± 0.1

where $\bar{C}_l(\rho, \mathbf{v})$ and $\bar{\mathcal{C}}_l(\rho, \mathbf{v})$ denote library terms with coefficients c_l and d_l , respectively. The experimental Quincke roller system shares several key features with the particle model in equation (4.1), so the construction of the candidate libraries $\{\bar{C}_l(\rho, \mathbf{v})\}$ and $\{\bar{\mathcal{C}}_l(\rho, \mathbf{v})\}$ follows similar principles (figure 4-4c). Conservation of particle number implies that \bar{C}_l can be written as divergences of vector fields. However, rollers do not explicitly break mirror symmetry, so chiral terms can be dropped from the $\{\bar{\mathcal{C}}_l(\rho, \mathbf{v})\}$ library, leaving the candidate terms shown in figure 4-4c.

4.3.3 Learned hydrodynamic equations and validation

The sparse regression algorithm proposed a hierarchy of hydrodynamic models with increasing complexity (figure 4-4d). The sparsest learned model that recapitulates the experimental observations is given by

$$\partial_t \rho = c_3 \nabla \cdot (\rho \mathbf{v}), \quad (4.6a)$$

$$\partial_t \mathbf{v} = d_1 \mathbf{v} + d_2 \rho \mathbf{v} + d_3 |\mathbf{v}|^2 \mathbf{v} + d_4 \nabla \rho + d_5 (\mathbf{v} \cdot \nabla) \mathbf{v}. \quad (4.6b)$$

Notably, equations (4.6) contain all the relevant terms to describe the propagation of underdamped sound waves, a counter-intuitive, but characteristic feature of overdamped active polar particle systems [9].

Although the finite experimental observation window and imperfect particle tracking was expected to limit the accuracy of the learned models, the learned coeffi-

cient values agree well with corresponding parameters estimated in Ref. [9] by fitting a linearized Toner-Tu model to the experimental data (table 4.2). The coefficient $c_3 \simeq -0.95$ in the mass conservation equation is close to the theoretically expected value -1 . The learned coefficient d_4 in the velocity equation (4.6b) is of similar magnitude but slightly less negative than the dispersion-based estimate in Ref. [9]. The learned coefficients d_1 , d_2 , and d_3 (table C.4), to our knowledge, had not been determined previously. Despite being inferred from a single video, these parameters yield a remarkably accurate prediction $v_0(\rho_0) = \sqrt{-(d_1 + d_2\rho_0)/d_3}$ for the typical roller speed as a function of the area fraction ρ_0 (Supplementary figure 4 in Ref. [9] and figure 4-5). Similarly, the learned coefficient d_5 of the nonlinear advective term $\sim (\mathbf{v} \cdot \nabla)\mathbf{v}$, is in excellent agreement with the value reported in Ref [9]. Interestingly, $d_5 \neq -1$ reveals the broken Galilean invariance [40, 11] due to fluid-mediated roller-substrate interaction, a key physical aspect of the experimental system that is robustly discovered by the hydrodynamic model learning framework.

To validate the learned hydrodynamic model, we simulated equations (4.6) on a periodic domain comparable to the experimental observation window (see figure 4-4e

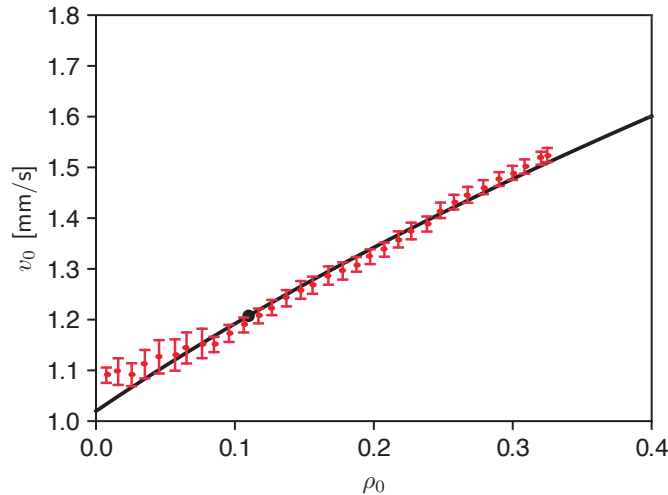


Figure 4-5: The learned model accurately predicts collective Quincke roller speeds v_0 at different average area fractions ρ_0 . Although equation (4.6) was learned from a single experiment (Supplementary Movie S2 in Ref. [9]) at fixed average area fraction $\rho_0 = 0.11$ (filled black circle), the model prediction $v_0(\rho_0) = \sqrt{-(d_1 + d_2\rho_0)/d_3}$ (solid line) with inferred parameters d_1, d_2, d_3 (table C.4), agrees well the experimentally measured speed values (red symbols) reported in Supplementary figure 4 of Ref. [9].

and appendix C.1.6). Starting from random initial conditions, spontaneously flowing states emerge, even though the spontaneous onset of particle flow is not a part of the experimental data from which the model was learned. The emergent density and flow patterns are quantitatively similar to the experimentally observed ones. In particular, the learned model predicts the formation of transverse velocity bands as seen in the experiments (figure 4-4b,e).

4.3.4 Predicting collective roller dynamics in confinement

Useful models can make predictions for a variety of experimental conditions. At minimum, if a learned hydrodynamic model captures the most relevant physics of an active system, then it should remain valid in different geometries and boundary conditions. To confirm this for the Quincke system, we simulated equations (4.6) on a square domain using no-flux and shear-free boundary conditions (appendix C.1.6). Starting from random initial conditions, our learned model predicts the formation of a vortex-like flow, permeated by four interwoven density shock waves, which arise from reflections at the boundary (figure 4-4f, left). Remarkably, this behavior has indeed been observed in experiments [5], in which Quincke rollers were confined within a square domain (figure 4-4f, right). These results demonstrate the practical potential of automated model learning for complex active matter systems.

4.4 Conclusions

Leveraging spectral representations of field observables, we have presented a PDE learning framework that robustly identifies quantitative hydrodynamic models for the self-organized dynamics of active matter systems from both microscopic simulations and experimental data. This approach complements traditional analytic coarse-graining techniques [36, 37, 11, 38, 39] which require *ad hoc* moment closures to truncate infinite hierarchies of coupled mode equations (appendix C.2). Analytic closures typically neglect correlations and rely on approximations that may not be valid in interacting active matter systems. Automated learning of hydrodynamic equations

yields data-informed closure relations, while simultaneously providing quantitative measurements of phenomenological coefficients (viscosities, elastic moduli, etc.) from video data [176].

Successful model learning requires both good data and a good library. Good data need to sample all dynamically relevant length and time scales [177]. A good library is large enough to include all hydrodynamically relevant terms and small enough to enable robust sparse regression [65]. Since the number of possible terms increases combinatorially with the number of fields and differential operators, library construction should be guided by prior knowledge of global, local, and explicitly broken symmetries. Such physics-informed libraries ensure properly constrained model search spaces, promising more robust and efficient sparse regression. Equally important is the use of suitable spectral field representations – without these an accurate evaluation of the library terms seems nearly impossible even for very-high quality data.

The presented framework complements modern machine learning approaches, from model-free methods [51, 50] to those that leverage *a priori* known model structure to predict complex dynamics [53, 178, 166], infer model parameters [176], or partially replace PDE models with suitably trained neural networks [52, 54]. Our approach provides interpretable hydrodynamic equations and a direct quantitative mapping from possibly heterogeneous microscopic systems to macroscopic hydrodynamic parameters. Due to its generic nature, the underlying framework can be easily extended to a larger number of hydrodynamic variables and fields with higher tensorial symmetry.

Chapter 5

Summary and outlook

5.1 Conclusions

In this thesis, we have taken a three-fold approach to understand and gain insights into the collective dynamics of active fluids. Broadly, this involved – 1) modeling the swimming behavior of individual micro-organisms in complex fluids, 2) investigating continuum phenomenological models for active fluids to gain insights into pattern forming systems, and 3) developing a learning framework that translates microscopic particle data from active particle experiments and simulations into continuum hydrodynamic models.

First, in Chapter 2, we investigated cylindrical squirmers in viscoplastic fluids to study the impact of a yield stress on the swimming characteristics of micro-organisms. To begin, we first addressed the classical problem of a translating cylinder in a viscoplastic fluid, particularly in the high Bingham number limit. When slip is allowed at the cylinder surface, as is commonly observed experimentally, we obtained previously unknown analytical solutions using slipline theory from ideal plasticity and also provided numerical evidence for these solutions. The associated analytical calculations and numerical solvers allowed us to modify the boundary conditions on the cylinder surface to investigate squirmers. Our results showed that the yield stress localizes the flow patterns, sometimes within thin viscoplastic boundary layers; this indicates that viscoplasticity can directly impact inter-particle interactions as well as

the nutrient uptake of organisms. We also found that the swimming mode with a normal surface velocity, for example via secretion and absorption of mucus [87], is energetically more efficient rather than ciliary propulsion in viscoplastic fluids.

Next, in Chapter 3, we investigated a continuum phenomenological model that couples active energy injection through a generic linear instability with advective non-linearity. Specifically, we considered the generalized Navier-Stokes (GNS) equations; such models have shown promising results in modeling active turbulence [17, 18]. With a view to explore if such phenomenological models could be relevant in other pattern-forming systems, we analyzed linearly driven flows on rotating spheres by extending a recent study of the GNS model on a stationary sphere [66]. At high rotation rates, direct numerical simulations show the emergence of alternating zonal jets, similar to those observed in planetary atmospheres. We also derived exact solutions which are a combination of alternating zonal jets and westward-propagating Rossby waves. Analyzing the phase speed of the waves in the simulations showed that they are close to the analytical solutions, further highlighting the usefulness of such minimal phenomenological models.

Finally, in Chapter 4, we successfully demonstrated how more specific and previously unknown phenomenological models could be discovered from active particle simulation or experimental data. To achieve this, we presented a learning framework that leverages recent sparse regression methods and spectral basis representations to translate noisy particle data into continuum hydrodynamic models. This framework was successfully applied to data from a chiral active particle model mimicking swimming cells as well as from experiments of self-propelled colloidal rollers. The learned models quantitatively reproduced the self-organized phenomenology in the microscopic data and were validated against coarse-grained equations.

To conclude, existing phenomenological models and domain knowledge of active matter provided us key insights about conservation laws, global and explicitly broken symmetries, and dynamical principles behind pattern formation. These insights were crucial for the successful application of equation discovery methods to the non-trivial active systems we have considered. Using this framework, we anticipate

that many previously intractable physical and biological systems can soon find interpretable quantitative continuum descriptions that may reveal novel ordering and self-organization principles.

5.2 Future extensions

Our work raises several questions that present avenues for future research.

Viscoplastic squirmers

Beyond the cylindrical squirmers we have considered, more realistic spherical models might further enhance our understanding of micro-organism swimming in complex fluids. Such spherical squirmers have been investigated in the context of Newtonian or viscoelastic fluids [82, 29, 86, 33], but not viscoplastic fluids. Furthermore, in our study, we have ignored the impact of viscoplasticity on the cilia, which could themselves become immobile due to the yield stress. Therefore, further work needs to be pursued to quantify the energetics of ciliary motion in yield stress fluids, perhaps by building on viscoplastic slender body theory [35].

Rotating GNS models

The GNS model we investigated parameterizes the dominant length scales that arise from sub-meso-scale processes through higher-order derivatives acting on the velocity field. In Fourier space, these terms correspond to a growth rate dependent on the wavenumber through a polynomial function with a single positive maximum (figure 3-2). This spectral forcing can be easily modified, for example, to take into account multiple dominant length scales by modifying the polynomial or replacing it with an arbitrary function. Through direct numerical simulations of such a model, it would be interesting to explore how the dominant modes corresponding to different length scales interact. Beyond direct numerical simulations, the nonlinear dynamics of the rotating GNS system could also be investigated using triad truncation [179] as was done recently for non-rotating GNS flows [18]. The flat-plane triad dynamics based

on Fourier modes can be adapted to the β -plane equations in (3.7b) to explore the effect of rotation through the β parameter. Additionally, the triad analysis can also be done using spherical harmonics, extending recent results for the barotropic vorticity equation [(3.3b) without the GNS forcing] by Lynch [125]. Both these analyses would be of interest in the context of geophysical flows.

Learning framework for hydrodynamic models

The field of equation discovery and machine learning for dynamical systems is rapidly evolving [50]. Hence, there is vast potential for future enhancements in the presented learning framework by adopting recent advances in sparse regression algorithms such as Iterative Hard Thresholding [64] and Sparse Relaxed Regularized Regression (SR3) [180], and weak formulations [62, 63]. Furthermore, model selection strategies such as cross-validation [181] and information criteria [182] can also be integrated. In so doing, error metrics between the data and the simulations must be defined in a statistical and spectral sense so as to quantify the phenomenological similarities. Additionally, since the library of terms is crucial for the successful application of sparse regression, it would be essential in the future to have automated ways of generating libraries that respect the relevant vectorial and tensorial symmetries up to desired orders in the fields and derivatives.

In view of the successful applications presented in this thesis, we expect that the underlying computational framework can be directly applied to a wide variety of passive and active matter systems. Some examples include controlled experimental data for bacterial swarms [145] or fish schools [183], whose hydrodynamic parameters are largely unknown so far.

Appendix A

This appendix is associated with Chapter 2.

A.1 Viscoplastic boundary layers for a no-slip cylinder

In this appendix, we outline a boundary layer theory for a translating cylinder with a no-slip surface. To set the scene, figure A-1 shows a magnification of the boundary-layer structure below the upper rotating plug.

A.1.1 Beyond the rotating plug

Outside the region directly underneath the plug, the main balance of forces expected for the boundary layer against the surface of the cylinder is given by

$$\partial_r p \sim 0 \quad \text{and} \quad \partial_\theta p \sim \partial_r(r^2 \tau_{r\theta}), \quad \text{where} \quad \tau_{r\theta} \sim \partial_r v + \text{Bi} \operatorname{sgn}(\partial_r v) \quad (\text{A.1})$$

(see Balmforth et al. [89]). The solution must match to the nearly perfectly plastic flow outside the boundary layer, where the pressure is given by (2.18), the velocity is directed along the α -lines and the shear rates are much weaker. The latter two conditions translate to $v(r_b) \sim 0$ and $\partial_r v(r_b) \sim 0$, where r_b is the edge of the boundary layer. Hence, after incorporating the no-slip condition $v(1) = \sin \theta$, we find

$$v = -2\text{Bi}(r_b - r)^2 \operatorname{sgn}(\sin \theta), \quad r_b = 1 + \text{Bi}^{-1/2} \sqrt{\frac{1}{2} |\sin \theta|}, \quad (\text{A.2})$$

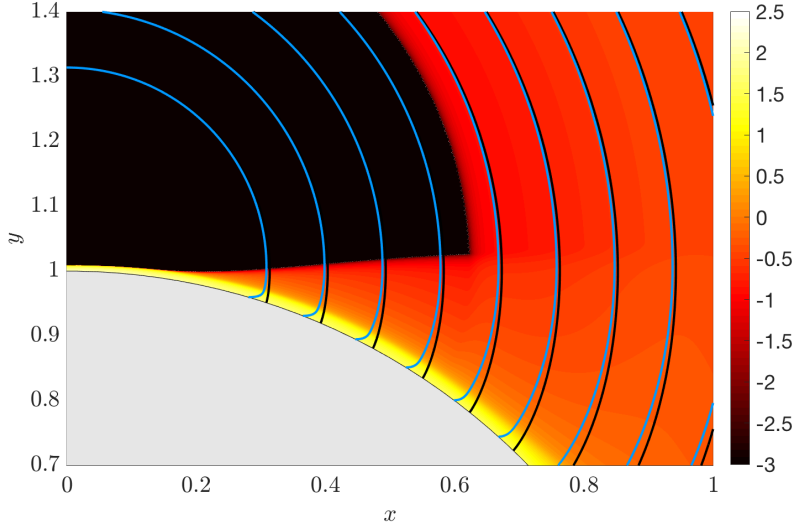


Figure A-1: A magnification of the region underneath the upper rotating plug for a no-slip cylinder with $\text{Bi} = 2^{12}$. The color shading shows $\log_{10} \dot{\gamma}$, and the blue lines are streamlines. The α -lines (black) of the slipline solution are selected to coincide with the streamlines at $\theta = \frac{1}{2}\pi$ and they match with the streamlines except near the boundary layer attached to the cylinder.

indicating that the thickness of the boundary layer is $\mathcal{O}(\text{Bi}^{-1/2})$. Such parabolic velocity profiles have been noted previously by Tokpavi et al. [75].

A.1.2 Underneath the rotating plug

Directly underneath the rotating plug where the pressure jump is smoothed, the angular scale becomes smaller and we rescale the variables to reflect this whilst maintaining the main balances demanded by force balance and the continuity equation:

$$(r, \theta) = \left(1 + \text{Bi}^{-a}\xi, \frac{1}{2}\pi - \text{Bi}^{-b}\Theta\right), \quad [u, v] \sim [\text{Bi}^{b-a}U(\xi, \Theta), V(\xi, \Theta)], \quad (\text{A.3})$$

$$p \sim \text{Bi}P(\xi, \Theta) \quad \text{and} \quad \tau_{\xi\Theta} \sim \text{Bi}^a \partial_{\xi} V - \text{Bi}, \quad (\text{A.4})$$

where η and Θ are $\mathcal{O}(1)$, and the exponents $a > \frac{1}{2}$ and $b > 0$ satisfy $2a = 1 + b$. The force balance in (A.1) then becomes

$$\partial_{\xi} P \sim 0 \quad \& \quad \partial_{\Theta} P \sim \partial_{\xi}^2 V. \quad (\text{A.5})$$

The no-slip condition is now $V(0) \sim 1$, whereas matching again demands that $(V, \partial_\xi V) \rightarrow 0$ at the edge of the boundary layer $\xi = \Xi(\Theta)$. Hence,

$$V \sim - \left(1 - \frac{\xi}{\Xi}\right)^2 \text{sgn}(y) \quad \text{and} \quad \partial_\Theta P = -\frac{2}{\Xi^2} \text{sgn}(y). \quad (\text{A.6})$$

At this stage, unlike in (A.2), we cannot match P to the pressure of the slipline solution to determine the boundary-layer profile $\Xi(\Theta)$ because of the intervention of the rotating plug. Instead, we proceed by dividing up the boundary layer into the part surrounding $\theta = \frac{1}{2}\pi$ that is directly attached to the rotating plug, and the part beyond where the boundary layer detaches from the plug and another nearly perfectly plastic flow separates the two. For the first part, the rigid rotation of the plug implies a velocity field of $(u, v) = \omega(\cos \theta, r - \sin \theta)$, where the rotation rate is observed from the numerical computations to be $\omega \sim 1 - \text{Bi}^{-c}\Omega$, with $\Omega > 0$; see figure 2-3a. If we now integrate the leading-order continuity equation, $\partial_\Xi U - \partial_\Theta V \sim 0$, over the boundary layer from $\xi = 0$ to $\xi = \Xi$, we find

$$\left[\int_0^\Xi \left(1 - \frac{\xi}{\Xi}\right)^2 d\xi \right]_\Theta = \frac{1}{3} \partial_\Theta \Xi = U(0, \Theta) - U(\Xi, \Theta) \quad (\text{A.7})$$

$$\equiv \text{Bi}^{a-b}(1 - \omega) \cos \theta \sim \text{Bi}^{a-2b-c}\Omega\Theta. \quad (\text{A.8})$$

Hence $2b + c = a$ and

$$\Xi(\Theta) = \Xi(0) + \frac{3}{2}\Omega\Theta^2, \quad (\text{A.9})$$

which thickens away from the centre of the fan at A , unlike the profile in (A.2).

For the part of the smoothing region where the plug has detached from the boundary layer, there is an intervening window of purely plastic deformation in which the velocity field is adjusted away from rigid rotation. Because this window is relatively small, the α -lines remain close to the involutes of the perfectly plastic slipline solution, which begin at the angular location $\theta = \frac{\pi}{2} - \text{Bi}^{-b}\Theta$ and reach the base of the plug for $y \sim 1$ and $x = \text{Bi}^{-b}\Theta$. This proximity indicates that $u \sim \text{Bi}^{-b}\omega\Theta - \text{Bi}^{b-a}\varpi(\Theta)$

at the edge of the boundary layer, where the correction $\text{Bi}^{b-a}\varpi(\Theta)$ represents the velocity adjustment incurred by the modification to the slipline. Thence,

$$\frac{1}{3}\partial_{\Theta}\Xi \sim U(0, \Theta) - U(\Xi, \theta) \sim \varpi(\Theta) + \Omega\Theta - \frac{1}{6}\text{Bi}^{a-4b}\Theta^3, \quad (\text{A.10})$$

which indicates that $a = 4b$, since all the terms must come in at the same order of Bi because the boundary layer remains continuous across the point of detachment. The combined results for the scalings now indicate that

$$a = \frac{4}{7}, \quad b = \frac{1}{7} \quad \text{and} \quad c = \frac{2}{7}; \quad (\text{A.11})$$

i.e. the boundary layer thickness scales as $\text{Bi}^{-4/7}$ underneath the rotating plug, the angular width of the smoothing region scales as $\text{Bi}^{-1/7}$ and the rotation rate of the plug approaches 1 with a scaling of $\text{Bi}^{-2/7}$ (figure 2-3b).

Beyond the detachment of the plug, (A.10) implies that the boundary layer profile becomes modified to

$$\Xi(\Theta) \sim \Xi_* + \int_{\Theta_*}^{\Theta} \varpi d\Theta - \frac{1}{8}(\Theta^4 - \Theta_*^4) + \frac{3}{2}\Omega\Theta^2, \quad (\text{A.12})$$

where Θ_* denotes the angle of detachment and the corresponding boundary layer thickness is Ξ_* . However, the term $\frac{1}{8}(\Theta^4 - \Theta_*^4)$ is problematic in view of its sign: as Θ increases, this correction opposes the thickening of the boundary layer. Yet Ξ must continue to thicken to become $O(\text{Bi}^{-1/2})$ in order to meet the boundary layer beyond the rotating plug described earlier. The correction $\varpi(\Theta)$ must therefore be chosen to eliminate the offending quartic term, suggesting that the profile remains close to (A.9) throughout the smoothing region. If we assume this to be the case, then a final estimate can be derived from the known pressure jump of $2\pi\text{Bi}$ across the smoothing region (see (2.18)). This jump implies that

$$\int_{-\infty}^{+\infty} \partial_{\Theta} P d\Theta = \int_{-\infty}^{+\infty} \frac{2}{[\Xi(0) + \frac{3}{2}\Omega\Theta^2]^2} d\Theta = 2\pi, \quad \text{or} \quad \Xi(0) = \left(\frac{1}{6\Omega}\right)^{1/3} \quad (\text{A.13})$$

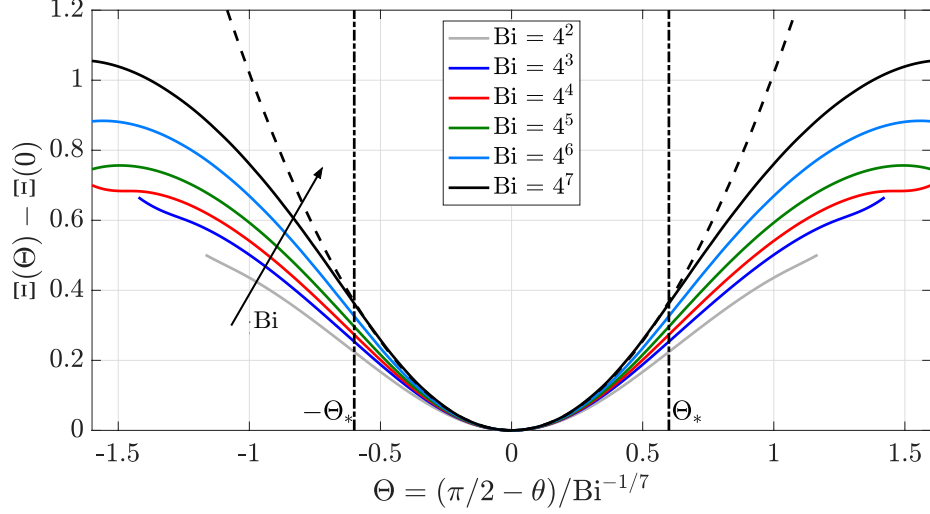


Figure A-2: Rescaled boundary layer profiles ($\Xi(\Theta) - \Xi(0)$) with the boundary-layer theory prediction from (A.9)[dashed]. The vertical dashed-dotted line marks the angular location Θ_* where the rotating plug separates from the boundary layer for $\text{Bi} = 4^7$. Within those limits, the numerical profiles match well with the boundary-layer theory predictions.

From the computations, $\Omega \approx 0.68$, and so $\Xi(0) \approx 0.63$. The numerical solutions, for which the boundary-layer thickness can be determined by fitting the quadratic velocity profile in (A.6) to $v(r, \theta)$ suggest that $\Xi(0)$ is closer to 0.5. The predicted boundary layer profile in (A.9) is plotted in figure A-2 and compared with results extracted from the numerical computations. The departure from the quadratic profile at the location that the plug detaches from the boundary layer is evident.

The assumption that the boundary-layer profile remains close to (A.9) even beyond the detachment of the plug also allows us to estimate the radius of the rotating plug: in order that this boundary layer meet the $O(\text{Bi}^{-1/2})$ -thick profile outside the smoothing region, we must have that $\text{Bi}^{-a}(\Xi_* + \frac{3}{2}\Omega\Theta^2) \sim \text{Bi}^{-1/2}$. That is, $\Theta = \mathcal{O}(\text{Bi}^{(a-1/2)/2}) = \mathcal{O}(\text{Bi}^{-3/28})$. Thus, the radius of the plug is $\sim (\pi/2 - \theta) \sim \mathcal{O}(\text{Bi}^{-3/28})$, which comfortably captures the scaling observed in the numerical computations (figure 2-3a).

A.2 Slipline results for a partially rough cylinder

A.2.1 The drag force and lower bound

For the bounds, it suffices to consider the top right half of the slipline solution in view of its symmetries about $x = 0$ and $y = 0$ (see figure 2-5 for reference). We find the net horizontal force by summing up the contributions on the curves EI and EB , and line BC since the pressure and stresses can be determined along these yielded regions. In the polar coordinates (r_P, θ_P) centered at P, the force per unit length on the circular arc EI is

$$\begin{bmatrix} -p & -\text{Bi} \\ -\text{Bi} & -p \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -p \\ -\text{Bi} \end{bmatrix} \quad (\text{A.14})$$

where $p = 2\text{Bi}(\pi - \vartheta) \equiv \text{Bi}(\pi - 2\theta_P)$. Integrating the horizontal force per unit length (given by $-p \cos \theta_P + \text{Bi} \sin \theta_P$) along arc EI , the net horizontal force is therefore

$$\frac{F_{EI}}{\text{Bi}} = -r_{EI} [2(\pi - \beta_2) \cos \beta_2 + \sin \beta_2] \quad (\text{A.15})$$

with $r_{EI} = \lambda \cot \beta_2 + \sqrt{1 - \lambda^2}$ (the radius of the circular arc).

On curve EB , the local slipline angle is $\vartheta = \theta + \frac{1}{4}\pi - \frac{1}{2}\Delta$ and so the pressure is $p = 2\text{Bi}(\pi - \vartheta) = \text{Bi}(\frac{3}{2}\pi - 2\theta + \Delta)$. In the $x - y$ coordinate system, the force per unit length on this curve is

$$\begin{bmatrix} -p - \text{Bi} \sin 2\vartheta & \text{Bi} \cos 2\vartheta \\ \text{Bi} \cos 2\vartheta & -p + \text{Bi} \sin 2\vartheta \end{bmatrix} \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}, \quad (\text{A.16})$$

where $\frac{1}{2}\Delta < \theta < \beta_2 - \frac{1}{4}\pi + \frac{1}{2}\Delta$. The horizontal force per unit length is $-p \cos \theta - \text{Bi} \cos(\Delta - \theta)$, which can be integrated along EB to get the net horizontal force to be

$$\begin{aligned} \frac{F_{EB}}{\text{Bi}} = & 2(\beta_2 - \pi) \sin\left(\beta_2 - \frac{1}{4}\pi + \frac{1}{2}\Delta\right) + \left(\frac{3}{2}\pi - 1\right) \sin \frac{1}{2}\Delta - \sin\left(\beta_2 - \frac{1}{4}\pi - \frac{1}{2}\Delta\right) + \\ & 2 \cos\left(\beta_2 - \frac{1}{4}\pi + \frac{1}{2}\Delta\right) - 2 \cos \frac{1}{2}\Delta. \quad (\text{A.17}) \end{aligned}$$

Finally, on the surface BC , the slipline angle is $\vartheta = \frac{1}{4}\pi$ and the pressure is $\frac{3}{2}\pi\text{Bi}$.

The horizontal force per unit length, $-(1/\sqrt{2})\text{Bi} - (1/\sqrt{2})p$, is then multiplied by the length of BC to obtain

$$\frac{F_{\text{BC}}}{\text{Bi}} = -\left(\frac{2+3\pi}{2\sqrt{2}}\right)(\lambda - \sqrt{1-\lambda^2}). \quad (\text{A.18})$$

Combining (A.15), (A.17) and (A.18), we may compute $F_x = 4(F_{\text{EI}} + F_{\text{EB}} + F_{\text{BC}})$.

A.2.2 Angular momentum balance

About any arbitrary origin, the two arcs of the rigidly rotating crescent $AEIE'$ exert moments that must cancel in order to balance the net angular momentum of that plug. The cross product of momentum equations $\mathbf{0} = \nabla \cdot \boldsymbol{\sigma}$, with the position vector \mathbf{x} from that origin, followed by the integral over the crescent, implies

$$0 = \int \int_{AEIE'} \nabla \cdot (\mathbf{x} \times \boldsymbol{\sigma}) dx dy = \int_{EAE'} \mathbf{x} \times \boldsymbol{\sigma} \cdot \mathbf{n} d\ell + \int_{EIE'} \mathbf{x} \times \boldsymbol{\sigma} \cdot \mathbf{n} d\ell \quad (\text{A.19})$$

where \mathbf{n} is the outward normal to EAE' and EIE' , and $d\ell$ is the line element proceeding around the arcs in an anti-clockwise sense with respect to the interior. That is,

$$0 = \int_{EIE'} (\mathbf{x} - \mathbf{x}_P) \times \boldsymbol{\sigma} \cdot \mathbf{n} d\ell + \int_{EAE'} (\mathbf{x} - \mathbf{x}_O) \times \boldsymbol{\sigma} \cdot \mathbf{n} d\ell \\ + \mathbf{x}_P \times \int_{EIE'} \boldsymbol{\sigma} \cdot \mathbf{n} d\ell + \mathbf{x}_O \times \int_{EAE'} \boldsymbol{\sigma} \cdot \mathbf{n} d\ell \quad (\text{A.20})$$

where \mathbf{x}_P and \mathbf{x}_O denote the positions of the centres of the circular arcs. But $\mathbf{x}_P - \mathbf{x}_O = (\lambda/\sin\beta_2)\hat{\mathbf{y}}$ and

$$-\int_{EAE'} \boldsymbol{\sigma} \cdot \mathbf{n} d\ell = \int_{EIE'} \boldsymbol{\sigma} \cdot \mathbf{n} d\ell \equiv 2F_{\text{EI}}\hat{\mathbf{x}}, \quad (\text{A.21})$$

given that the crescent must be in net force balance (which prescribes the net force on EAE' even though the full stress tensor is not known there). The angular momentum

balance therefore implies

$$0 = T_{EIE'} - T_{EAE'} - 2 \frac{\lambda F_{EI}}{\sin \beta_2}, \quad (\text{A.22})$$

where $T_{EIE'}$ and $T_{EAE'}$ denote the torques on the arcs about their respective centres; *i. e.*

$$T_{EIE'} = \hat{\mathbf{z}} \cdot \int_{EIE'} (\mathbf{x} - \mathbf{x}_P) \times \boldsymbol{\sigma} \cdot \mathbf{n} d\ell = 2r_{EI}^2 \int_{\beta_2 - \pi/2}^{\pi/2} (-Bi) d\theta_p \quad (\text{A.23})$$

and

$$T_{EAE'} = \hat{\mathbf{z}} \cdot \int_{EAE'} (\mathbf{x} - \mathbf{x}_O) \times \boldsymbol{\sigma} \cdot \hat{\mathbf{r}} d\ell = 2 \int_{\theta_E}^{\pi/2} \varrho Bi d\theta, \quad (\text{A.24})$$

where $\hat{\mathbf{r}}$ is the radial unit vector and $\theta_E = \beta_2 - \frac{1}{4}\pi + \frac{1}{2}\Delta$. Altogether (and after removing a factor of $2Bi$),

$$0 = r_{EI}^2 (\pi - \beta_2) + \varrho \left(\frac{1}{2}\pi - \theta_E \right) + \frac{\lambda F_{EI}}{Bi \sin \beta_2}. \quad (\text{A.25})$$

Geometry of the slipline solution

Referring to figure 2-5, over the region of involutes, the geometry of the α -lines is given by

$$\begin{aligned} x &= x(\vartheta, \Theta) = \lambda [\cos \vartheta + (\vartheta - \Theta) \sin \vartheta], \\ y &= y(\vartheta, \Theta) = \lambda [\sin \vartheta - (\vartheta - \Theta) \cos \vartheta], \end{aligned} \quad (\text{A.26})$$

where Θ is the polar angle at the intersection with the circle of radius λ . The length of a slipline over the section $\vartheta_1 < \vartheta < \vartheta_2$ is

$$\ell = \frac{1}{2} \lambda [(\vartheta - \Theta)^2]_{\vartheta_1}^{\vartheta_2}. \quad (\text{A.27})$$

The lines have a radius of curvature,

$$\varpi = \lambda(\vartheta - \Theta), \quad (\text{A.28})$$

and the area element can be written as

$$dx dy \rightarrow \lambda \varpi d\vartheta d\Theta. \quad (\text{A.29})$$

Along the involutes, the velocity v_α remains constant. The α -lines emerging from BC have $v_\alpha = 1/\sqrt{2}$, whereas those from EB have $v_\alpha = \cos \theta / \cos \Psi$, where θ is the polar angle of the cylinder surface point and

$$\Psi = \frac{1}{4}\pi - \frac{1}{2}\Delta, \quad (\text{A.30})$$

which corresponds to the angle that the characteristic makes with the normal to the cylinder ($\vartheta - \theta = \Psi$ on EB). The angles two polar angles are related by

$$\Theta = \theta - \tan \Psi + \Psi. \quad (\text{A.31})$$

The three key sliplines have

$$\Theta = \frac{1}{4}\pi - 1 \text{ for } CDG, \quad \theta = \theta_B = \frac{1}{2}\Delta \text{ for } BFH, \quad \theta = \theta_E = \beta_2 - \Psi \text{ for } EI. \quad (\text{A.32})$$

The portions of these sliplines that are given by circular arcs have the radii,

$$r_{DG} = \lambda \cot \beta_2 + \lambda(\beta_2 + 1 - \frac{1}{4}\pi), \quad (\text{A.33})$$

$$r_{FH} = \lambda \cot \beta_2 + \lambda(\beta_2 - \frac{1}{4}\pi + \tan \Psi), \quad (\text{A.34})$$

$$r_{EI} = \lambda \cot \beta_2 + \sqrt{1 - \lambda^2}. \quad (\text{A.35})$$

Finally, the rotating plug has a rotation rate of $\lambda^{-1} \sin \beta_2$.

A.2.3 The net dissipation rate and upper bound

The contributions to the dissipation rate arise from either spatially extended regions of plastic shear or from velocity jumps. For an extended region the contribution is

$$\frac{1}{2} \int \int \boldsymbol{\tau} : \dot{\boldsymbol{\gamma}} \, dx dy \rightarrow \text{Bi} \int \int \dot{\gamma} dx dy, \quad (\text{A.36})$$

where, because the velocity is directed only along the α -lines, the shear rate can be calculated from

$$\dot{\gamma}_{\varpi\vartheta} = \left. \frac{\partial v_\alpha}{\partial \varpi} \right|_{\vartheta} - \frac{v_\alpha}{\varpi}, \quad (\text{A.37})$$

if ϖ now denotes the local radius of curvature. A tangential velocity jump of ΔV across an α -line provides the dissipation rate

$$\int \tau_{ns} \Delta V(s) ds, \quad (\text{A.38})$$

where s is the arc length and τ_{ns} is the local shear stress for Cartesian coordinates aligned with that curve. The various contributions to the net dissipation rate are summarized below.

Region $BCDF$

Over the region $BCDF$ the characteristic angle ϑ increases from $\frac{1}{4}\pi$ to β_2 . The dissipation rate is

$$\dot{\gamma} = |\dot{\gamma}_{\varpi\vartheta}| = \frac{v_\alpha}{\varpi} = \frac{1}{\varpi\sqrt{2}}, \quad (\text{A.39})$$

and so

$$\int \int_{BCDF} \dot{\gamma} dx dy = \lambda \int \int_{BCDF} \frac{1}{\sqrt{2}} d\Theta d\vartheta \equiv \int \int_{BCDF} \frac{1}{\sqrt{2}} d\vartheta d\varpi \quad (\text{A.40})$$

$$= \frac{1}{\sqrt{2}}(\beta_2 - \frac{1}{4}\pi)(\lambda - \sqrt{1 - \lambda^2}), \quad (\text{A.41})$$

given that the length of BC is $\lambda - \sqrt{1 - \lambda^2}$, which is the length of all the β -lines across this region, or the change in the radius of curvature at each value of ϑ .

Region EBF

Here, the ranges of the angles are $\Psi + \theta < \vartheta < \beta_2$ and $\theta_B < \theta < \theta_E$, and the dissipation rate becomes

$$\int \int_{EBF} \dot{\gamma} dx dy = \lambda \int \int_{EBF} (\vartheta - \Theta) \left| \frac{\partial v_\alpha}{\partial \theta} + \frac{v_\alpha}{\vartheta - \Theta} \right| d\vartheta d\theta \quad (\text{A.42})$$

$$= \int_{\theta_B}^{\theta_E} \int_{\theta + \pi/4 - \Delta/2}^{\beta_2} |\cos \theta - (\vartheta - \Theta) \sin \theta| d\vartheta d\theta \quad (\text{A.43})$$

(since $\lambda \equiv \cos \Psi$). The absolute value here makes further progress difficult. Moreover, $\text{sgn}(\dot{\gamma}_{\varpi\vartheta}) < 0$ if the velocity field is to be consistent with the slipline solution, for which $\tau_{\vartheta\varpi} = -\text{Bi}$. This consistency condition fails for $\beta_2 > \frac{1}{2}\pi$, which is equivalent to the flaw in Randolph & Houlsby's original construction (which is recovered here when $\beta_2 \rightarrow \frac{1}{2}\pi \cos^{-1} \lambda$).

Region $EFDGHI$

This region has circular α -lines centred at P . The radius r_p of each is given by

$$r_p = \lambda(\beta_2 - \Theta + \cot \beta_2), \quad r_{\text{EI}} < r_p < r_{\text{DG}}. \quad (\text{A.44})$$

The local dissipation rate is

$$\dot{\gamma} = \left| \frac{\partial v_\alpha}{\partial r_p} - \frac{v_\alpha}{r_p} \right|. \quad (\text{A.45})$$

Since there is no angle dependence, the net dissipation rate is given by

$$\int \int_{EFDGHI} \dot{\gamma} dx dy = (\pi - \beta_2) \int_{r_{EI}}^{r_{DG}} \dot{\gamma} r_p dr_p \quad (\text{A.46})$$

$$= \frac{1}{\sqrt{2}} (\pi - \beta_2) (r_{DG} - r_{FH}) + (\pi - \beta_2) \int_{\theta_B}^{\theta_E} |(\beta_2 - \Theta + \cot \beta_2) \sin \theta - \cos \theta| d\theta. \quad (\text{A.47})$$

Like in region EBF , the shear rate $\dot{\gamma}_{\vartheta\varpi}$ again becomes inconsistent with $\tau_{\vartheta\varpi} = -\text{Bi}$ for $\beta_2 > \frac{1}{2}\pi$.

Velocity jumps

Along CDG , there is a tangential velocity jump of $-1/\sqrt{2}$ and a shear stress of $-\text{Bi}$, leading to a net dissipation rate of

$$\frac{1}{\sqrt{2}} \text{Bi} [(\pi - \beta_2) r_{DG} + \ell_{FG}], \quad (\text{A.48})$$

where

$$\ell_{CD} = \frac{1}{2} \lambda (\beta_2 - \frac{1}{4}\pi) (\beta_2 - \frac{1}{4}\pi + 2) \quad (\text{A.49})$$

is the length of the involute CD . Along BFH , the velocity jump of $1/\sqrt{2} - \lambda^{-1} \cos \theta_B$ and shear stress of $-\text{Bi}$, give a contribution

$$\left(\frac{1}{\lambda} \cos \theta_B - \frac{1}{\sqrt{2}} \right) \ell_{BFH} \text{Bi} = \left(\frac{1}{\lambda} \cos \theta_B - \frac{1}{\sqrt{2}} \right) [\ell_{BF} + r_{FH} (\pi - \beta_2)] \text{Bi}, \quad (\text{A.50})$$

with

$$\ell_{BF} = \frac{1}{2} \lambda (\beta_2 - \frac{1}{4}\pi) (\beta_2 + 2 \tan \Psi - \frac{1}{4}\pi). \quad (\text{A.51})$$

Next, along BC , the velocity jumps by $-1/\sqrt{2}$ and the shear stress is $-\text{Bi}$, so there is a contribution

$$\frac{1}{\sqrt{2}}\ell_{\text{BC}}\text{Bi} = \frac{1}{\sqrt{2}}\text{Bi}(\lambda - \sqrt{1 - \lambda^2}). \quad (\text{A.52})$$

Finally, there are the partially slipping boundary layers along EB and AE . Here, $\tau_{r\theta} = \rho\text{Bi}$ and the angular velocity jump is $\tan \Psi \cos \theta + \sin \theta = \lambda^{-1} \sin(\theta + \Psi)$ for EB and $\lambda^{-1} \sin \beta_2$ for AE . The the net dissipation rates are given by

$$\frac{\rho}{\lambda}\text{Bi} \int_{EB} \sin(\theta + \Psi)d\theta = \frac{\rho}{\lambda}\text{Bi} \left(\frac{1}{\sqrt{2}} - \cos \beta_2 \right) \quad (\text{A.53})$$

and

$$\frac{\rho}{\lambda}\text{Bi} \left(\frac{1}{2}\pi - \theta_{\text{E}} \right) \sin \beta_2. \quad (\text{A.54})$$

Appendix B

This appendix is associated with Chapter 3.

B.1 Formulas for Λ , τ and κ

For a sphere of radius R , (Λ, τ, κ) are related to $(\Gamma_0, \Gamma_2, \Gamma_4)$ as follows [66]:

$$\Lambda = \frac{2\pi R}{2\sqrt{\frac{17}{4} - \frac{\Gamma_2}{2\Gamma_4}R^2 - 1}}, \quad (\text{B.1a})$$

$$\tau = \left[\left(\frac{\Gamma_2}{2\Gamma_4} - \frac{2}{R^2} \right) \left(\Gamma_0 - \frac{\Gamma_2^2}{4\Gamma_4} \right) \right]^{-1}, \quad (\text{B.1b})$$

$$\kappa = \left(\frac{17}{2R^2} - \frac{\Gamma_2}{\Gamma_4} - 2\sqrt{\frac{17^2}{16R^4} - \frac{17}{4R^2}\frac{\Gamma_2}{\Gamma_4} + \frac{\Gamma_0}{\Gamma_4}} \right)^{1/2}. \quad (\text{B.1c})$$

Letting $R \rightarrow \infty$ in (B.1), one obtains for the planar case [18]

$$\Lambda = \pi\sqrt{\frac{2\Gamma_4}{-\Gamma_2}}, \quad \tau = \left[\frac{\Gamma_2}{2\Gamma_4} \left(\Gamma_0 - \frac{\Gamma_2^2}{4\Gamma_4} \right) \right]^{-1}, \quad \kappa = \left(-\frac{\Gamma_2}{\Gamma_4} - 2\sqrt{\frac{\Gamma_0}{\Gamma_4}} \right)^{1/2}. \quad (\text{B.2})$$

B.2 Total angular momentum

Taking the surface mass density to be 1, the total angular momentum is given by

$$\begin{aligned} M(t) &= \int_0^\pi \int_0^{2\pi} dA (R \sin \theta v_\phi) = \int_0^\pi d\theta \int_0^{2\pi} d\phi (R^3 \sin^2 \theta v_\phi) \\ &= \int_0^\pi d\theta \int_0^{2\pi} d\phi [R^3 \sin^2 \theta (-\partial_\theta \psi / R)] = -R^2 \int_0^\pi d\theta \int_0^{2\pi} d\phi (\sin^2 \theta \partial_\theta \psi). \end{aligned} \quad (\text{B.3})$$

Applying integration by parts for the θ -integral,

$$\begin{aligned}
M(t) &= -R^2 \int_0^{2\pi} d\phi \left\{ \int_0^\pi d\theta \sin^2 \theta \partial_\theta \psi \right\} \\
&= -R^2 \int_0^{2\pi} d\phi \left\{ [\sin^2 \theta \psi]_0^\pi - \int_0^\pi d\theta (2\psi \sin \theta \cos \theta) \right\} \\
&= 2R^2 \int_0^{2\pi} d\phi \int_0^\pi d\theta (\psi \sin \theta \cos \theta). \tag{B.4}
\end{aligned}$$

We may expand the stream function as $\psi(\theta, \phi) = \sum_{\ell, m} \hat{\psi}_\ell^m(t) Y_\ell^m(\theta, \phi)$. The ϕ -integral survives only for $m = 0$ and $Y_\ell^0(\theta, \phi) = P_\ell^0(\cos \theta)$ where P represents the associated Legendre polynomials. Also realizing that $P_1^0(\cos \theta) = \cos \theta$, we get

$$M(t) = 4\pi R^2 \sum_\ell \hat{\psi}_\ell^0(t) \left\{ \int_0^\pi d\theta P_1^0(\cos \theta) P_\ell^0(\cos \theta) \sin \theta \right\}. \tag{B.5}$$

Finally, using the orthogonality relation

$$\int_0^\pi d\theta P_k^m(\cos \theta) P_\ell^m(\cos \theta) \sin \theta = \frac{2(\ell + m)!}{(2\ell + 1)(\ell - m)!} \delta_{k, \ell}, \tag{B.6}$$

we obtain

$$M(t) = \frac{8\pi}{3} R^2 \hat{\psi}_1^0(t). \tag{B.7}$$

Following the results by Lynch [184], it can be shown that there is no contribution to $\hat{\psi}_1^0(t)$ from any triad interactions that result from the nonlinear dynamics. To see this, we let $\hat{\psi}_{l_\gamma}^{m_\gamma}(t) = \hat{\psi}_1^0(t)$ which interacts with coefficients $\hat{\psi}_{l_\alpha}^{m_\alpha}$ and $\hat{\psi}_{l_\beta}^{m_\beta}$. For a non-vanishing triad interaction, the necessary conditions, $|\ell_\alpha - \ell_\beta| < \ell_\gamma = 1$ and $\ell_\alpha \neq \ell_\beta$, cannot be simultaneously satisfied for $\ell_\gamma = 1$. Additionally, equation (3.5) shows that the GNS forcing is zero for $\ell = 1$. Thus, $\hat{\psi}_1^0(t)$ remains constant and the total angular momentum is conserved.

Appendix C

This appendix is associated with Chapter 4.

C.1 Methods

C.1.1 Particle simulations

The microscopic model in Eqs. (4.1) has been previously studied for fixed particle velocities $v_i = v_0$ and rotation frequencies $\Omega_i = \Omega_0$. In this scenario, particles form small clusters of aligned particles and each cluster orbits on a circle of radius $\sim v_0/\Omega_0$ [159]. To generate the microscopic test data used in Sec. 4.2.4, we considered instead a heuristic distribution $p(v_i, \Omega_i)$ for which particles spontaneously organize into proper vortices (Fig. 4-2a, top). It is convenient to define and draw from this distribution using propagation speeds v_i and the curvature radii $R_i = v_i/\Omega_i$ of a particle's noise-free trajectory as independent variables. In particular, we considered $\tilde{p}(v_i, R_i) \sim g(v_i; \mu_v, \sigma_v)g(R_i; \mu_R, \sigma_R)$, where $g(x; \mu_x, \sigma_x)$ represents a Gaussian normal distribution with mean μ_x and standard deviation σ_x . $\tilde{p}(v_i, R_i)$ then defines $p(v_i, \Omega_i)$ implicitly through the relation $v_i = \Omega_i R_i$. In units of the characteristic scales – mean velocity $\langle v_i \rangle_p$ and interaction radius R – the particle properties v_i and $\Omega_i = v_i/R_i$ used for simulating Eqs. (4.1) (Fig. C-1) were drawn from $\tilde{p}(v_i, R_i)$ with $\mu_v = 1$ ($\langle v_i \rangle_{\tilde{p}} = \langle v_i \rangle_p = 1$), $\sigma_v = 0.4$, $\mu_R = 2.2$ and $\sigma_R = 1.7$. From these samples, we finally removed all particles with $\Omega_i > 1.4$.

For simulations of the microscopic model in Eqs. (4.1), we set $g \simeq 0.35$ and $D_r \simeq 0.009$ ($\Rightarrow D_r \ll \langle \Omega_i \rangle_p$, $D_r \ll \langle v_i \rangle_p$) and initially placed particles randomly distributed and oriented on a domain of size 100×100 (in units of the interaction radius). Equations (4.1) were then numerically integrated using the Euler-Maruyama

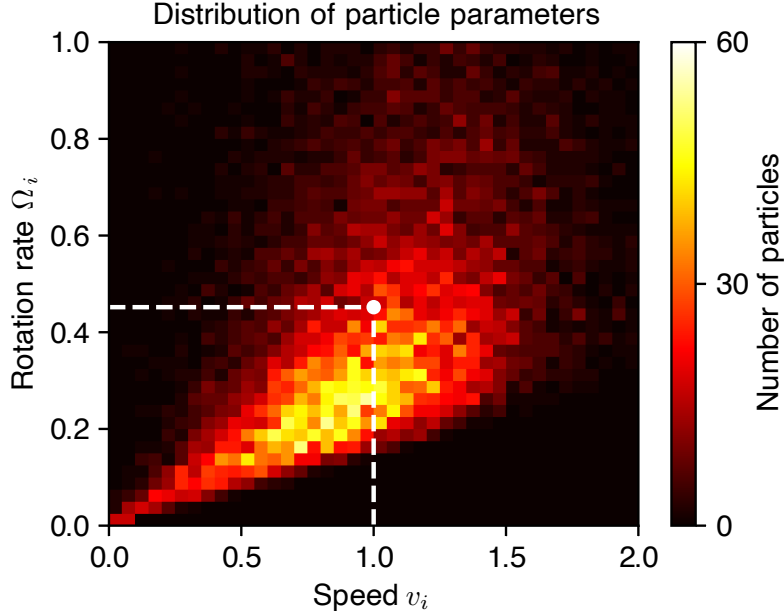


Figure C-1: Distribution of particle speeds v_i and rotation rates Ω_i used to simulate Eqs. (4.1). Those parameters were drawn from a heuristic distribution $p(v_i, \Omega_i)$ that is explained in more detail in Sec. C.1.1. The white marker and dashed lines depict the mean velocity $\langle v_i \rangle_p \simeq 1$ and $\langle \Omega_i \rangle_p \simeq 0.45$.

method with a time step of $dt \simeq 0.0176$. For the subsequent coarse-graining, the data were saved at time intervals of $\Delta t \simeq 0.44$.

C.1.2 Kernel coarse-graining with periodic and non-periodic boundaries

To coarse-grain the discrete microscopic data through Eqs. (4.2), we used a 2D Gaussian kernel

$$K(\mathbf{x}) = (2\pi\sigma^2)^{-1} \exp(-|\mathbf{x}|^2/2\sigma^2). \quad (\text{C.1})$$

Periodicity of the coarse-grained fields for the microscopic test data (Sec. 4.2.2) was ensured by placing ghost particles periodically around the domain.

Coarse-graining in non-periodic domains (Sec. 4.3) was performed by truncating and renormalizing the kernel. This was achieved by defining the integral over the non-periodic domain \mathcal{X} to be $N(\mathbf{x}) = \int_{\mathcal{X}} d^2\mathbf{x}' K(\mathbf{x}' - \mathbf{x})$, and then replacing $K[\mathbf{x} - \mathbf{x}_i(t)]$ with $K[\mathbf{x} - \mathbf{x}_i(t)]/N(\mathbf{x}_i(t))$ in Eqs. (4.2). This renormalization ensured that the

coarse-grained density $\rho(t, \mathbf{x})$ integrated to the total particle number and strongly reduced artefacts near the boundary.

C.1.3 Spectral representation

The coarse-grained hydrodynamic fields [Eq. (4.2)] were evaluated at $[N_t, N_x, N_y]$ uniformly spaced grid points in the respective directions. The resulting discrete data were projected onto the spectral basis functions [Eq. (4.3)] using multidimensional discrete cosine and Fourier transforms provided by the FFTW library [185], with an efficient time complexity of $O(N \log(N))$, where $N = N_t N_x N_y$. For the Chebyshev transforms, the data were interpolated onto the required Chebyshev extrema grid using spline functions of degree 5.

C.1.4 Sparse regression

To perform sparse regression using the sequentially thresholded least squares (STLSQ) algorithm [56], we used the same parameters when working with data from the test microscopic model (Sec. 4.2.1) as well as the Quincke roller experiments (Sec. 4.3). The details of various steps in the learning framework are provided below.

Construction of linear systems: To construct the linear system $\mathbf{U}_t = \Theta \boldsymbol{\xi}$, we randomly sampled the coarse-grained fields at $N_d = 5 \times 10^5$ time-space points. The

explicit form of the linear systems constructed for Eqs. (4.4) was given by

$$\underbrace{\begin{bmatrix} \vdots \\ \partial_t \rho \\ \vdots \end{bmatrix}}_{\mathbf{U}_t(N_d \times 1)} = \underbrace{\begin{bmatrix} \vdots & & \vdots \\ \nabla \cdot \mathbf{p} & \cdots & \nabla \cdot (\rho \mathbf{p}_\perp) \\ \vdots & & \vdots \end{bmatrix}}_{\Theta(N_d \times r)} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{bmatrix}}_{\xi(r \times 1)}, \quad (\text{C.2a})$$

$$\underbrace{\begin{bmatrix} \partial_t p_x \\ \vdots \\ \hline \partial_t p_y \\ \vdots \end{bmatrix}}_{\mathbf{U}_t(2N_d \times 1)} = \underbrace{\begin{bmatrix} (\nabla \rho)_x & \cdots & ((\mathbf{p} \cdot \nabla) \mathbf{p})_x \\ \vdots & & \vdots \\ \hline (\nabla \rho)_y & \cdots & ((\mathbf{p} \cdot \nabla) \mathbf{p})_y \\ \vdots & & \vdots \end{bmatrix}}_{\Theta(2N_d \times m)} \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}}_{\xi(m \times 1)}. \quad (\text{C.2b})$$

Here, the subscripts denote components of the vectors, and r, m are the total number of library terms in each equation. The vertical dots denote the respective terms evaluated at different time-space (t, \mathbf{x}) locations. The linear system in Eq. (C.2b) was generated by stacking data for the x - and y -components of the time-derivatives and the library terms. Such a construction enforced the same coefficients for both the components of the polarization equation, ensuring rotational invariance (coordinate-independence) of the learned PDE.

Pre-processing: Since the thresholding hyperparameter τ in STLSQ is agnostic to the scales of the library terms, as a pre-processing step, we performed transformations so that columns of the data matrix Θ had zero mean and unit variance, and the time-derivative vector \mathbf{U}_t had zero mean.

Stability selection [64]: With equal spacing on a \log_{10} scale, we chose 40 values for τ over the regularization path $[\tau_{\max}, \epsilon \tau_{\max}]$. The value of τ_{\max} was chosen so that all the terms get thresholded out and ϵ was set to 10^{-2} . For every τ , the data were split into 200 sub-samples each with 50% randomly selected data points. Every library term was assigned an importance score as the fraction of sub-samples in which it was learned by STLSQ; in general, this importance score was larger for smaller values of τ .

Along the regularization path, unique combinations of terms that had an importance score larger than 0.6 were considered and their coefficients were refitted to the full data without normalization. This procedure resulted in a small number of PDEs of increasing complexity (Figs. 4-2e, 4-3d, and 4-4d).

C.1.5 Linear dependencies of the library terms

The procedure outlined in Sec. 4.2.4 leads to a few library terms for the polarization dynamics [Eq. (4.4b)] that are linearly dependent on each other. For completeness, we provide here a list of identities that can be used to eliminate these dependencies. Terms with ∇_{\perp} can be replaced with those involving ∇ by using the following relations:

$$\nabla_{\perp} \cdot \mathbf{p} = -\nabla \cdot \mathbf{p}_{\perp}, \quad (\text{C.3a})$$

$$\nabla_{\perp} \cdot \mathbf{p}_{\perp} = \nabla \cdot \mathbf{p}, \quad (\text{C.3b})$$

$$\mathbf{p} \cdot \nabla_{\perp} = -\mathbf{p}_{\perp} \cdot \nabla, \quad (\text{C.3c})$$

$$\mathbf{p}_{\perp} \cdot \nabla_{\perp} = \mathbf{p} \cdot \nabla, \quad (\text{C.3d})$$

$$\frac{1}{2} \nabla_{\perp} |\mathbf{p}|^2 = (\nabla \cdot \mathbf{p}) \mathbf{p}_{\perp} - (\mathbf{p}_{\perp} \cdot \nabla) \mathbf{p}, \quad (\text{C.3e})$$

$$\nabla_{\perp} (\nabla \cdot \mathbf{p}) = \Delta \mathbf{p}_{\perp} - \nabla (\nabla \cdot \mathbf{p}_{\perp}), \quad (\text{C.3f})$$

$$\nabla_{\perp} \mathbf{p}_{\perp} = (\nabla \cdot \mathbf{p}) \mathbb{I} - (\nabla \mathbf{p})^{\top}. \quad (\text{C.3g})$$

One may set $\mathbf{p} \rightarrow \mathbf{p}_\perp$ and $\mathbf{p}_\perp \rightarrow -\mathbf{p}$ in Eqs. (C.3f) and (C.3g) to obtain two more relations. Additional linear dependencies are given by the following identities:

$$\frac{1}{2}\nabla|\mathbf{p}|^2 = (\nabla\mathbf{p}) \cdot \mathbf{p}, \quad (\text{C.4a})$$

$$= (\mathbf{p} \cdot \nabla)\mathbf{p} + (\nabla \cdot \mathbf{p}_\perp)\mathbf{p}_\perp, \quad (\text{C.4b})$$

$$\nabla \cdot (\mathbf{p}\mathbf{p}) = (\mathbf{p} \cdot \nabla)\mathbf{p} + (\nabla \cdot \mathbf{p})\mathbf{p}, \quad (\text{C.4c})$$

$$\nabla \cdot (\mathbf{p}\mathbf{p}_\perp) = (\mathbf{p} \cdot \nabla)\mathbf{p}_\perp + (\nabla \cdot \mathbf{p})\mathbf{p}_\perp, \quad (\text{C.4d})$$

$$(\nabla \cdot \mathbf{p})\mathbf{p}_\perp = (\nabla\mathbf{p}) \cdot \mathbf{p}_\perp + (\mathbf{p} \cdot \nabla)\mathbf{p}_\perp, \quad (\text{C.4e})$$

$$\begin{aligned} &(\mathbf{p} \cdot \nabla)\mathbf{p}_\perp + (\mathbf{p}_\perp \cdot \nabla)\mathbf{p} \\ &= (\nabla \cdot \mathbf{p})\mathbf{p}_\perp + (\nabla \cdot \mathbf{p}_\perp)\mathbf{p}, \end{aligned} \quad (\text{C.4f})$$

$$(\nabla\mathbf{p}_\perp) \cdot \mathbf{p} = -(\nabla\mathbf{p}) \cdot \mathbf{p}_\perp. \quad (\text{C.4g})$$

In Eq. (C.4e), we follow the convention, $[(\nabla\mathbf{a}) \cdot \mathbf{b}]_i = b_j \partial_i a_j$, with $i = x, y$ and repeated indices indicating summation. One may set $\mathbf{p} \rightarrow \mathbf{p}_\perp$ and $\mathbf{p}_\perp \rightarrow -\mathbf{p}$ in Eqs. (C.4a) and (C.4e) to obtain two additional identities.

To ensure in general that no linear dependencies remain after a library of terms has been constructed, it should be checked that the columns the data matrix Θ are linearly independent, that is, Θ is full rank. This maybe achieved by, for example, a singular value decomposition of Θ .

C.1.6 Continuum simulations

Continuum simulations were performed using the spectral PDE solver Dedalus [129] with four-step Runge-Kutta time stepping scheme RK443. For simulation of the PDEs learned from the microscopic test data (Fig. 4-3e) we used 256×256 Fourier modes in a doubly periodic domain with time step 4×10^{-3} . To facilitate a comparison, simulations shown in Fig. 4-3e were initialized using the initial density and polarization field of the coarse-grained particle data. It was verified that similar vortex patterns also form from fully random initial conditions.

For the doubly periodic simulation of Eqs. (4.6) (Fig. 4-4e), we used $1024 \times$

1024 Fourier modes and time step 10^{-4} s. The initial conditions were random with mean density 0.11, and mean horizontal and vertical velocities, $\langle v_x \rangle = 0.1 \text{ mm s}^{-1}$ and $\langle v_y \rangle = 0$, respectively.

The simulation presented in Fig. 4-4f was performed on a confined square domain using the Sine/Cosine basis functions with 1024×1024 modes and time step 10^{-4} s. The basis combinations in the (x, y) directions were chosen to be (\cos, \cos) for density ρ , (\sin, \cos) for v_x and (\cos, \sin) for v_y . These imply that normal density gradients, normal velocities and all remaining shear rates $\partial_x v_y$ and $\partial_y v_x$ vanish at the domain boundaries. The simulations were initialized with random perturbations around a mean density of 0.11.

Since the learned model in Eqs. (4.6) generates density shock waves, we added numerical diffusivities of $10^{-4} \text{ mm}^2/\text{s}$ (Fig. 4-4e) and $10^{-3} \text{ mm}^2/\text{s}$ (Fig. 4-4f) in both density and velocity equation to avoid Gibbs ringing.

C.2 Analytic coarse-graining of the particle model

We describe two approaches to analytically determine mean-field approximations of the microscopic model Eqs. (4.1). These approaches *(i)* provide guidance for developing a physics-informed learning library, *(ii)* allow discussing our PDE learning framework as a tool to effectively infer moment closure relations, and *(iii)* predict the dependency of certain PDE coefficients on distributions of microscopic parameters, which can be used to validate learned hydrodynamic models.

C.2.1 Dynamic equation of the one-particle probability density

A commonly used approach to determine mean-field description of models such as Eqs. (4.1) is to find an approximate dynamic equation for the one-particle probability

density [186, 36, 37, 11, 38, 159]

$$f(\theta, \mathbf{x}, t) = \sum_{i=1}^N \langle \delta(\theta - \theta_i(t)) \delta(\mathbf{x} - \mathbf{x}_i) \rangle, \quad (\text{C.5})$$

where $\langle \cdot \rangle$ denotes a Gaussian white noise average. Neglecting multiplicative noise and factorizing pair-correlations gives rise to a nonlinear integro-differential equation [37, 159] that can be transformed into an infinite hierarchy of coupled PDEs for the angular moments $f_n(\mathbf{x}, t)$ defined by

$$f_n(\mathbf{x}, t) = \int_0^{2\pi} d\theta f(\theta, \mathbf{x}, t) e^{in\theta}. \quad (\text{C.6})$$

For the microscopic model Eqs. (4.1) and equal swimming parameters $v_i = v_0$ and $\Omega_i = \Omega_0$ for all particle, this procedure leads to [37, 159]

$$\begin{aligned} \partial_t f_n + \frac{v_0}{2} [\partial_x (f_{n+1} + f_{n-1}) - i\partial_y (f_{n+1} - f_{n-1})] \\ = n(i\Omega_0 - D_r n) f_n + \frac{gn\pi}{2} (f_{n-1} f_1 - f_{n+1} f_{-1}). \end{aligned} \quad (\text{C.7})$$

Each complex angular moment f_n can be identified as a mean-field variable that represents different orientational order parameters encoded by the probability density $f(\theta, \mathbf{x}, t)$ [11]. In particular, f_0 represents the particle number density ρ and $f_1 =: p_x + ip_y$ represents the polarization density $\mathbf{p} = (p_x, p_y)^\top$. These fields correspond to the coarse-graining information in Eqs. (4.2) that our learning framework extracts explicitly from given microscopic data. For $n = 0$ and $n = 1$, we can therefore write Eqs. (C.7) as

$$\partial_t \rho + v_0 \nabla \cdot \mathbf{p} = 0, \quad (\text{C.8a})$$

$$\partial_t \mathbf{p} + \frac{v_0}{2} (\nabla \rho + \nabla \cdot \mathbf{Q}) = \Omega_0 \mathbf{p}_\perp - D_r \mathbf{p} + \frac{g\pi}{2} (\rho \mathbb{I} - \mathbf{Q}) \cdot \mathbf{p}, \quad (\text{C.8b})$$

which also shows the coupling to the next higher mode $f_2 =: Q_{xx} + iQ_{xy}$, corresponding to the independent degrees of freedom of a nematic tensor. The chiral term

$\sim \Omega_0 \mathbf{p}_\perp$ with $\mathbf{p}_\perp = (-p_y, p_x)^\top$ breaks the mirror symmetry. Terms constructed from \mathbf{p}_\perp are therefore generally allowed in chiral systems and consequently included into the library in Eqs. (4.4).

The final step that is key to analytically close the infinite hierarchy of Eqs. (C.7) requires the introduction of moment closure assumptions [36, 37, 9]. Depending on the structure of the mode coupling, the resulting closure relation allows to express the nearest coupled modes with $|n| = k$ in terms of modes $|n| < k$ and neglects the remaining modes. For example, in the case of Eq. (C.8) a moment closure assumption must provide an expression $\mathbf{Q}(\rho, \mathbf{p})$ [37, 9]. Our framework does not require explicit closure assumptions but instead takes a different route by inferring – directly from the data – an effective closure relation that best explains the observed systems dynamics.

C.2.2 Dynamic equations from conventional kernel coarse-graining

While the previous approach provides a clear coarse-graining strategy to find a closed set of PDEs from a system of stochastic ODEs with homogeneous microscopic parameters, it is more challenging to understand how the phenomenological coefficients will depend on the distribution $p(v_i, \Omega_i)$ of microscopic kinetic parameters described in Sec. C.1.1. We therefore consider an alternative strategy, for which we write Eqs. (4.1) as

$$\frac{d\mathbf{x}_i}{dt} = v_i \mathbf{p}_i, \quad (\text{C.9a})$$

$$\frac{d\mathbf{p}_i}{dt} = \Omega_i \boldsymbol{\epsilon} \cdot \mathbf{p}_i + \mathbf{F}_i, \quad (\text{C.9b})$$

where $\boldsymbol{\epsilon} \cdot \mathbf{p}_i = \mathbf{p}_{i,\perp} = (-\sin \theta_i, \cos \theta_i)^\top$, and \mathbf{F}_i contains forces from interactions and rotational diffusion. Taking directly the time derivative of the coarse-graining prescription in Eq. (4.2a) and using Eq. (C.9a), we find

$$\partial_t \rho(t, \mathbf{x}) + \nabla \cdot \mathbf{J}(t, \mathbf{x}) = 0, \quad (\text{C.10})$$

where we have defined a flux

$$\mathbf{J}(t, \mathbf{x}) = \sum_i K [\mathbf{x} - \mathbf{x}_i(t)] v_i \mathbf{p}_i(t). \quad (\text{C.11})$$

Using this definition and Eq. (C.9b), we find a dynamic equation for \mathbf{J} of the form

$$\partial_t \mathbf{J}(t, \mathbf{x}) + \nabla \cdot \boldsymbol{\sigma}(t, \mathbf{x}) = \mathbf{T}(t, \mathbf{x}) + \boldsymbol{\Phi}(t, \mathbf{x}). \quad (\text{C.12})$$

Here, we have defined the tensor and vector fields

$$\boldsymbol{\sigma}(t, \mathbf{x}) = \sum_i K [\mathbf{x} - \mathbf{x}_i(t)] v_i^2 \mathbf{p}_i(t) \mathbf{p}_i(t), \quad (\text{C.13a})$$

$$\mathbf{T}(t, \mathbf{x}) = \boldsymbol{\epsilon} \cdot \sum_i K [\mathbf{x} - \mathbf{x}_i(t)] v_i \Omega_i \mathbf{p}_i(t), \quad (\text{C.13b})$$

$$\boldsymbol{\Phi}(t, \mathbf{x}) = \sum_i K [\mathbf{x} - \mathbf{x}_i(t)] v_i \mathbf{F}_i(t). \quad (\text{C.13c})$$

Averaging the fields in Eq. (C.11) and Eqs. (C.13) over the particle parameter distribution $p(v_i, \Omega_i)$ yields

$$\langle \mathbf{J}(t, \mathbf{x}) \rangle_p = \left\langle \sum_i K [\mathbf{x} - \mathbf{x}_i(t)] v_i \mathbf{p}_i(t) \right\rangle_p, \quad (\text{C.14a})$$

$$\langle \boldsymbol{\sigma}(t, \mathbf{x}) \rangle_p = \left\langle \sum_i K [\mathbf{x} - \mathbf{x}_i(t)] v_i^2 \mathbf{p}_i(t) \mathbf{p}_i(t) \right\rangle_p, \quad (\text{C.14b})$$

$$\langle \mathbf{T}(t, \mathbf{x}) \rangle_p = \left\langle \boldsymbol{\epsilon} \cdot \sum_i K [\mathbf{x} - \mathbf{x}_i(t)] v_i \Omega_i \mathbf{p}_i(t) \right\rangle_p, \quad (\text{C.14c})$$

$$\langle \boldsymbol{\Phi}(t, \mathbf{x}) \rangle_p = \left\langle \sum_i K [\mathbf{x} - \mathbf{x}_i(t)] v_i \mathbf{F}_i(t) \right\rangle_p. \quad (\text{C.14d})$$

We then adopt a moment factorization approximation

$$\langle \mathbf{J}(t, \mathbf{x}) \rangle_p \simeq \langle v_i \rangle_p \mathbf{p}, \quad (\text{C.15a})$$

$$\langle \boldsymbol{\sigma}(t, \mathbf{x}) \rangle_p \simeq \frac{1}{2} \langle v_i^2 \rangle_p (\rho \mathbb{I} + \mathbf{Q}), \quad (\text{C.15b})$$

$$\langle \mathbf{T}(t, \mathbf{x}) \rangle_p \simeq \langle v_i \Omega_i \rangle_p \mathbf{p}_\perp, \quad (\text{C.15c})$$

where we used the definition of the particle number density in Eq. (4.2a), the polarization density in Eq. (4.2b), and $|\mathbf{p}_i|^2 = 1$. Additionally, we have defined in Eq. (C.15b) a nematic moment of the form

$$\mathbf{Q} = \sum_i K[\mathbf{x} - \mathbf{x}_i(t)] [2\mathbf{p}_i(t)\mathbf{p}_i(t) - \mathbb{I}]. \quad (\text{C.16})$$

Finally, averaging Eqs. (C.10) and (C.12) over the microscopic parameter distributions and using Eqs. (C.15), we obtain

$$\partial_t \rho + \langle v_i \rangle_p \nabla \cdot \mathbf{p} = 0, \quad (\text{C.17a})$$

$$\partial_t \mathbf{p} + \frac{\langle v_i^2 \rangle_p}{2\langle v_i \rangle_p} (\nabla \rho + \nabla \cdot \mathbf{Q}) = \frac{\langle v_i \Omega_i \rangle_p}{\langle v_i \rangle_p} \mathbf{p}_\perp + \langle v_i \rangle_p^{-1} \langle \Phi \rangle_p. \quad (\text{C.17b})$$

From this, we can read off predictions about the coefficients we expect to find by the learning framework for the terms $\nabla \cdot \mathbf{p}$, $\nabla \rho$ and \mathbf{p}_\perp (Tab. 4.1).

C.3 Parameters of learned models

The parameters of the PDEs learned from simulations of the active polar particle model in Eq. (4.1) are summarized in Tab. C.1 (density dynamics) and Tab. C.2 (polarization dynamics). For the experimental Quincke roller system [9], the learned hydrodynamic model parameters are given in Tab. C.3 (density dynamics) and Tab. C.4 (velocity dynamics).

Table C.1: Parameters a_l of the density dynamics PDE (Fig. 4-2d) learned from simulations the microscopic active particle system in Eq. (4.1). The sparsest model (\blacktriangleleft) agrees well with the analytic coarse-graining prediction (Tab. 4.1).

Term	PDE 1 \blacktriangleleft	PDE 2	PDE 3	PDE 4
$a_1 \nabla \cdot \mathbf{p}$	-0.991	-0.991	-0.972	-0.957
$a_2 \Delta \rho$	–	–	–	–
$a_3 \nabla \cdot (\rho \mathbf{p})$	–	–	-0.015	-0.037
$a_4 \Delta \rho^2$	–	0.022	0.022	0.022
$a_5 \Delta \mathbf{p} ^2$	–	–	–	–
$a_6 \nabla \cdot (\rho^2 \mathbf{p})$	–	–	–	0.008
$a_7 \Delta \rho^3$	–	–	–	–
$a_8 \nabla \cdot (\mathbf{p} ^2 \mathbf{p})$	–	–	–	–
$a_9 \nabla \cdot (\rho \nabla \mathbf{p} ^2)$	–	–	–	–
$a_{10} \nabla \cdot (\mathbf{p} ^2 \nabla \rho)$	–	–	–	–
$a_{11} \nabla \cdot \mathbf{p}_\perp$	–	–	–	–
$a_{12} \nabla \cdot (\rho \mathbf{p}_\perp)$	–	-0.026	-0.026	-0.026
$a_{13} \nabla \cdot (\rho^2 \mathbf{p}_\perp)$	–	–	–	–
$a_{14} \nabla \cdot (\mathbf{p} ^2 \mathbf{p}_\perp)$	–	–	–	–

Table C.2: Parameters b_l of the nine sparsest PDEs for the polarization dynamics (Fig. 4-3c), learned from simulations of the microscopic system in Eq. (4.1). PDE 8 (\blacktriangleleft) reproduces the characteristic vortex dynamics as in the microscopic simulations (Fig. 4-3a,b,e) and the coefficients of the linear terms compare well with analytic coarse-graining predictions (Tab. 4.1).

Term	PDE 1	PDE 2	PDE 3	PDE 4	PDE 5	PDE 6	PDE 7	PDE 8 \blacktriangleleft	PDE 9
$b_1 \mathbf{p}$	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009	-0.009
$b_2 \rho \mathbf{p}$	–	–	–	0.013	0.013	0.013	0.007	0.009	0.009
$b_3 \mathbf{p}_\perp$	0.414	0.476	0.477	0.428	0.478	0.436	0.436	0.440	0.441
$b_4 \rho \mathbf{p}_\perp$	–	-0.050	-0.040	–	-0.040	-0.006	-0.006	-0.010	-0.012
$b_5 \mathbf{p} ^2 \mathbf{p}$	–	–	–	–	–	–	–	-0.080	-0.080
$b_6 \mathbf{p} ^2 \mathbf{p}_\perp$	–	–	–	–	–	–	–	–	0.054
$b_7 \nabla \rho$	-0.638	-0.637	-0.600	-0.595	-0.601	-0.596	-0.596	-0.595	-0.595
$b_8 (\mathbf{p} \cdot \nabla) \mathbf{p}$	–	–	–	-0.536	–	-0.510	-0.510	-0.463	-0.479
$b_9 (\mathbf{p} \cdot \nabla) \mathbf{p}_\perp$	–	–	–	–	–	–	–	–	–
$b_{10} (\mathbf{p}_\perp \cdot \nabla) \mathbf{p}$	–	–	–	–	–	–	–	–	–
$b_{11} \nabla (\nabla \cdot \mathbf{p})$	–	–	–	–	–	–	–	0.078	0.077
$b_{12} \nabla (\nabla \cdot \mathbf{p}_\perp)$	–	–	0.225	0.265	0.248	0.265	0.270	0.277	0.277
$b_{13} \Delta \mathbf{p}$	–	–	–	–	–	–	-0.151	-0.155	-0.156
$b_{14} \Delta \mathbf{p}_\perp$	–	–	0.252	0.202	0.222	0.203	0.198	0.196	0.197
$b_{15} \nabla \mathbf{p} ^2$	–	–	–	–	–	–	–	–	–
$b_{16} (\nabla \cdot \mathbf{p}) \mathbf{p}$	–	–	–	–	–	–	–	-0.225	-0.213
$b_{17} (\nabla \cdot \mathbf{p}) \mathbf{p}_\perp$	–	–	–	–	–	–	–	–	–
$b_{18} \Delta^2 \mathbf{p}$	–	–	–	–	–	–	-0.475	-0.483	-0.484
$b_{19} \Delta^2 \mathbf{p}_\perp$	–	–	1.100	1.197	1.085	1.212	1.215	1.235	1.243

Table C.3: Parameters c_i of the PDE for the density dynamics (Fig. 4-4c) learned from experimental data for self-propelled Quincke rollers (Supplementary Movie S2 of Ref. [9]). The dimensions of the coefficients are such that $[\mathbf{v}] = \text{mm/s}$ and $[\rho] = 1$, where the density ρ represents the area fraction of rollers of diameter $D_c = 4.8 \mu\text{m}$. The four sparsest PDEs are shown corresponding to the cut-off $n_0 \in \{50, 100\}$ above which the temporal Chebyshev modes in Eq. (4.3) are set to zero to ignore high frequencies. The sparsest PDEs (\blacktriangleleft) have coefficients close to each other and agree well with the mass conservation equation obtained from analytic coarse-graining (Ref. [9]).

Term	Unit	$n_0 = 50$				$n_0 = 100$			
		PDE 1 \blacktriangleleft	PDE 2	PDE 3	PDE 4	PDE 1 \blacktriangleleft	PDE 2	PDE 3	PDE 4
$c_1 \nabla \cdot \mathbf{v}$	–	–	–	-0.052	-0.052	–	-0.051	-0.051	-0.055
$c_2 \Delta \rho$	$\text{mm}^2 \text{s}^{-1}$	–	0.016	0.040	0.023	–	0.055	0.040	0.041
$c_3 \nabla \cdot (\rho \mathbf{v})$	–	-0.950	-0.950	-1.068	-1.067	-0.945	-1.057	-1.054	-0.985
$c_4 \Delta \rho^2$	$\text{mm}^2 \text{s}^{-1}$	–	0.047	0.080	0.081	–	-0.076	-0.051	-0.062
$c_5 \Delta \mathbf{v} ^2$	s	–	–	–	0.001	–	–	0.001	0.001
$c_6 \nabla \cdot (\rho^2 \mathbf{v})$	–	–	–	–	–	–	–	–	-0.313
$c_7 \Delta \rho^3$	$\text{mm}^2 \text{s}^{-1}$	–	–	–	–	–	0.427	0.341	0.366
$c_8 \nabla \cdot (\mathbf{v} ^2 \mathbf{v})$	$\text{mm}^{-2} \text{s}^2$	–	–	0.035	0.035	–	0.034	0.034	0.034
$c_9 \nabla \cdot (\rho \nabla \mathbf{v} ^2)$	s	–	–	–	-0.013	–	–	-0.008	-0.007
$c_{10} \nabla \cdot (\mathbf{v} ^2 \nabla \rho)$	s	–	-0.018	-0.039	-0.028	–	-0.036	-0.028	-0.027

Table C.4: Parameters d_i of the PDE for the velocity dynamics (Fig. 4-4c) learned from experimental data for self-propelled Quincke rollers (Supplementary Movie S2 of Ref. [9]). The dimensions of the coefficients are such that $[\mathbf{v}] = \text{mm/s}$ and $[\rho] = 1$, where the density ρ represents the area fraction of rollers of diameter $D_c = 4.8 \mu\text{m}$. The four sparsest PDEs are shown corresponding to the cut-off $n_0 \in \{50, 100\}$ above which the temporal Chebyshev modes in Eq. (4.3) are set to zero to ignore high frequencies. The sparsest PDEs which reproduce the experimental observations (\blacktriangleleft) have coefficients that are close to each other for different values of n_0 , and they agree well with corresponding values reported in Ref. [9] (Tab. 4.2).

Term	Unit	$n_0 = 50$				$n_0 = 100$			
		PDE 1	PDE 2 \blacktriangleleft	PDE 3	PDE 4	PDE 1	PDE 2 \blacktriangleleft	PDE 3	PDE 4
$d_1 \mathbf{v}$	s^{-1}	–	2.281	1.524	1.491	–	1.825	1.252	1.122
$d_2 \rho \mathbf{v}$	s^{-1}	–	8.356	5.156	4.745	–	6.135	3.143	3.083
$d_3 \mathbf{v} ^2 \mathbf{v}$	$\text{mm}^{-2} \text{s}$	–	-2.194	-1.436	-1.382	–	-1.710	-1.095	-0.999
$d_4 \nabla \rho$	$\text{mm}^2 \text{s}^{-2}$	–	-1.620	-1.711	-2.074	–	-1.689	-2.438	-2.430
$d_5 (\mathbf{v} \cdot \nabla) \mathbf{v}$	–	-0.639	-0.674	-0.678	-0.679	-0.662	-0.696	-0.702	-0.702
$d_6 \nabla (\nabla \cdot \mathbf{v})$	$\text{mm}^2 \text{s}^{-1}$	–	–	–	–	–	–	–	–
$d_7 \Delta \mathbf{v}$	$\text{mm}^2 \text{s}^{-1}$	–	–	–	–	–	–	–	0.002
$d_8 \nabla (\mathbf{v} ^2)$	–	–	–	–	0.090	–	–	0.169	0.168
$d_9 (\nabla \cdot \mathbf{v}) \mathbf{v}$	–	–	–	-0.189	-0.190	–	–	-0.178	-0.179
$d_{10} \Delta^2 \mathbf{v}$	$\text{mm}^4 \text{s}^{-1}$	–	–	–	–	–	–	–	0.000

Bibliography

- [1] Geophysical Fluid Dynamics Program. URL <https://gfd.whoi.edu>.
- [2] A. Reuther, J. Kepner, C. Byun, S. Samsi, W. Arcand, D. Bestor, B. Bergeron, V. Gadepally, M. Houle, M. Hubbell, et al. Interactive supercomputing on 40,000 cores for machine learning and data analysis. In *2018 IEEE High Performance extreme Computing Conference (HPEC)*, pages 1–6. IEEE, 2018.
- [3] H. H. Wensink, J. Dunkel, S. Heidenreich, K. Drescher, R. E. Goldstein, H. Löwen, and J. M. Yeomans. Meso-scale turbulence in living fluids. *Proc. Natl. Acad. Sci. U.S.A.*, 109(36):14308–14313, 2012.
- [4] I. H. Riedel, K. Kruse, and J. Howard. A self-organized vortex array of hydrodynamically entrained sperm cells. *Science*, 309(5732):300–303, 2005.
- [5] A. Bricard, J. B. Caussin, N. Desreumaux, O. Dauchot, and D. Bartolo. Emergence of macroscopic directed motion in populations of motile colloids. *Nature*, 503:95–98, 2013.
- [6] V. Soni, E. S. Bililign, S. Magkiriadou, S. Sacanna, D. Bartolo, M. J. Shelley, and W. T. M. Irvine. The odd free surface flows of a colloidal chiral fluid. *Nat. Phys.*, 15(11):1188–1194, 2019.
- [7] M. F. Randolph and G. T. Houlsby. The limiting pressure on a circular pile loaded laterally in cohesive soil. *Géotechnique*, 34:613–623, 1984.
- [8] C. M. Martin and M. F. Randolph. Upper-bound analysis of lateral pile capacity in cohesive soil. *Geotechnique*, 56(2):141–145, 2006.
- [9] D. Geyer, A. Morin, and D. Bartolo. Sounds and hydrodynamics of polar active fluids. *Nat. Mater.*, 17(9):789–793, 2018.
- [10] S. Ramaswamy. The Mechanics and Statistics of Active Matter. *Annu. Rev. Condens. Matter Phys.*, 1(1):323–345, 2010.
- [11] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha. Hydrodynamics of soft active matter. *Rev. Mod. Phys.*, 85:1143–1189, 2013.

- [12] F. Jülicher, S. W. Grill, and G. Salbreux. Hydrodynamic theory of active matter. *Rep. Prog. Phys.*, 81(7):076601, 2018.
- [13] A. Cavagna and I. Giardina. Bird flocks as condensed matter. *Annu. Rev. Condens. Matter Phys.*, 5(1):183–207, 2014.
- [14] I. Ashraf, H. Bradshaw, Thanh T. Ha, J. Halloy, R. Godoy-Diana, and B. Thiria. Simple phalanx pattern leads to energy saving in cohesive fish schooling. *Proc. Natl. Acad. Sci. U. S. A.*, 114(36):9599–9604, 2017.
- [15] M. Rubenstein, A. Cornejo, and R. Nagpal. Programmable self-assembly in a thousand-robot swarm. *Science*, 345(6198):795–800, 2014.
- [16] J. Toner and Y. Tu. Flocks, herds, and schools: A quantitative theory of flocking. *Phys. Rev. E*, 58(4):4828–4858, 1998.
- [17] J. Dunkel, S. Heidenreich, K. Drescher, H. H. Wensink, M. Bär, and R. E. Goldstein. Fluid dynamics of bacterial turbulence. *Phys. Rev. Lett.*, 110(22):228102, 2013.
- [18] J. Słomka and J. Dunkel. Spontaneous mirror-symmetry breaking induces inverse energy cascade in 3D active fluid. *Proc. Natl. Acad. Sci. U.S.A.*, 114(9):2119–2124, 2017.
- [19] M. James, W. J. T. Bos, and M. Wilczek. Turbulence and turbulent pattern formation in a minimal model for active fluids. *Phys. Rev. Fluids*, 3:061101, 2018.
- [20] S. Ramaswamy. Active fluids. *Nat. Rev. Phys.*, 1(11):640–642, 2019.
- [21] Y. Hatwalne, S. Ramaswamy, M. Rao, and R. A. Simha. Rheology of Active-Particle Suspensions. *Phys. Rev. Lett.*, 92(11):118101, 2004.
- [22] D. Saintillan. Rheology of Active Fluids. *Annu. Rev. Fluid Mech.*, 50:563–592, 2018.
- [23] E. Lauga and T. R. Powers. The hydrodynamics of swimming microorganisms. *Rep. Progr. Phys.*, 72(9):096601, 2009.
- [24] B. M. Friedrich and F. Jülicher. Chemotaxis of sperm cells. *Proc. Natl. Acad. Sci. U.S.A.*, 104(33):13256–13261, 2007.
- [25] K. Drescher, R. E. Goldstein, N. Michel, M. Polin, and I. Tuval. Direct measurement of the flow field around swimming microorganisms. *Phys. Rev. Lett.*, 105(16):1–4, 2010.
- [26] K. Drescher, K. C. Leptos, I. Tuval, T. Ishikawa, T. J. Pedley, and R. E. Goldstein. Dancing *Volvox*: Hydrodynamic bound states of swimming algae. *Phys. Rev. Lett.*, 102(16):1–4, 2009.

- [27] J. R. Blake. Infinite models for ciliary propulsion. *J. Fluid Mech.*, 49(02): 209–222, 1971.
- [28] Sir J. Lighthill. *Mathematical Biofluidynamics*. SIAM, 1975.
- [29] T. J. Pedley. Spherical squirmers: models for swimming micro-organisms. *IMA J. Appl. Math.*, 81(3):488–521, 2016.
- [30] J. Patarin, É. Ghiringhelli, G. Darsy, M. Obamba, P. Bochu, B. Camara, S. Qué- tant, J. Cracowski, C. Cracowski, and M. Robert de Saint Vincent. Rheological analysis of sputum from patients with chronic bronchial diseases. *Sci. Rep.*, 10: 15685, 2020.
- [31] D. Cepeda, F. Pardos, and N. Sánchez. From biggest to smallest mud dragons: size-latitude trends in a group of meiobenthic animals worldwide. *Org. Divers. Evol.*, 21:43–58, 2021.
- [32] D. G. Crowdy and Y. Or. Two-dimensional point singularity model of a low-reynolds-number swimmer near a wall. *Phys. Rev. E*, 81(3):036313, 2010.
- [33] Z. Ouyang, J. Lin, and X. Ku. The hydrodynamic behavior of a squirmer swimming in power-law fluid. *Phys. Fluids*, 30(8):083301, 2018.
- [34] D. R. Hewitt and N. J. Balmforth. Taylor’s swimming sheet in a yield-stress fluid. *J. Fluid Mech.*, 828:33–56, 2017.
- [35] D. R. Hewitt and N. J. Balmforth. Viscoplastic slender body theory. *J. Fluid Mech.*, 856:870–897, 2018.
- [36] E. Bertin, M. Droz, and G. Grégoire. Hydrodynamic equations for self-propelled particles: microscopic derivation and stability analysis. *J. Phys. A*, 42(44): 445001, 2009.
- [37] F. D. C. Farrell, M. C. Marchetti, D. Marenduzzo, and J. Tailleur. Pattern formation in self-propelled particles with density-dependent motility. *Phys. Rev. Lett.*, 108:248101, 2012.
- [38] B. Liebchen, M. E. Cates, and D. Marenduzzo. Pattern formation in chemically interacting active rotors with self-propulsion. *Soft Matter*, 12:7259–7264, 2016.
- [39] H. Chaté. Dry aligning dilute active matter. *Annu. Rev. Condens. Matter Phys.*, 11(1):189–212, 2020.
- [40] J. Toner and Y. Tu. Long-range order in a two-dimensional dynamical XY model: How birds fly together. *Phys. Rev. Lett.*, 75(23):4326–4329, 1995.
- [41] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.*, 75(6): 1226–1229, 1995.

- [42] J. Słomka and J. Dunkel. Generalized Navier-Stokes equations for active suspensions. *Eur. Phys. J. Spec. Top.*, 224(7):1349–1358, 2015.
- [43] J. Słomka and J. Dunkel. Geometry-dependent viscosity reduction in sheared active fluids. *Phys. Rev. Fluids*, 2(4):9–12, 2017.
- [44] J. B. Marston. Planetary Atmospheres as Nonequilibrium Condensed Matter. *Annu. Rev. Condens. Matter Phys.*, 3(1):285–310, 2012.
- [45] P. Delplace, J. B. Marston, and A. Venaille. Topological origin of equatorial waves. *Science*, 358:1075–1077, 2017.
- [46] R. Hartmann, P. K. Singh, P. Pearce, R. Mok, B. Song, F. Díaz-Pascual, J. Dunkel, and K. Drescher. Emergence of three-dimensional order and structure in growing biofilms. *Nat. Phys.*, 15:251–256, 2019.
- [47] G. Shah, K. Thierbach, B. Schmid, J. Waschke, A. Reade, M. Hlawitschka, I. Roeder, N. Scherf, and J. Huisken. Multi-scale imaging and analysis identify pan-embryo cell dynamics of germlayer formation in zebrafish. *Nat. Commun.*, 10:5753, 2019.
- [48] M. R. Shaebani, A. Wysocki, R. G. Winkler, G. Gompper, and H. Rieger. Computational models for active matter. *Nat. Rev. Phys.*, 2(4):181–199, 2020.
- [49] F. Cichos, K. Gustavsson, B. Mehlig, and G. Volpe. Machine learning for active matter. *Nat. Mach. Intell.*, 2(2):94–103, 2020.
- [50] S. L. Brunton, B. R. Noack, and P. Koumoutsakos. Machine Learning for Fluid Mechanics. *Annu. Rev. Fluid Mech.*, 52(1):477–508, 2020.
- [51] J. Pathak, B. Hunt, M. Girvan, Z. Lu, and E. Ott. Model-Free Prediction of Large Spatiotemporally Chaotic Systems from Data: A Reservoir Computing Approach. *Phys. Rev. Lett.*, 120(2):24102, 2018.
- [52] Y. Bar-Sinai, S. Hoyer, J. Hickey, and M. P. Brenner. Learning data-driven discretizations for partial differential equations. *Proc. Natl. Acad. Sci. U. S. A.*, 116(31):15344–15349, 2019.
- [53] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.*, 378:686–707, 2019.
- [54] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, A. Ramadhan, and A. Edelman. Universal differential equations for scientific machine learning, 2020. arXiv:2001.04385.
- [55] M. Raissi, A. Yazdani, and G. E. Karniadakis. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science*, 367(6481):1026–1030, 2020.

- [56] S. L. Brunton, J. L. Proctor, and J. N. Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl. Acad. Sci. U.S.A.*, 113(15):3932–3937, 2016.
- [57] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz. Data-driven discovery of partial differential equations. *Sci. Adv*, 3(4):e1602614, 2017.
- [58] M. Bär, R. Hegger, and H. Kantz. Fitting partial differential equations to space-time dynamics. *Phys. Rev. E*, 59(1):337–342, 1999.
- [59] M. Schmidt and H. Lipson. Distilling Free-Form Natural Laws. *Science*, 324: 81–86, 2009.
- [60] K. Champion, B. Lusch, J. N. Kutz, and S. L. Brunton. Data-driven discovery of coordinates and governing equations. *Proc. Natl. Acad. Sci. U. S. A.*, 116 (45):22445–22451, 2019.
- [61] G.-J. Both, S. Choudhury, P. Sens, and R. Kusters. Deepmod: Deep learning for model discovery in noisy data. *J. Comput. Phys*, page 109985, 2020.
- [62] P. A. K. Reinbold, D. R. Gurevich, and R. O. Grigoriev. Using noisy or incomplete data to discover models of spatiotemporal dynamics. *Phys. Rev. E*, 101 (1):010203(R), 2020.
- [63] P. A. K. Reinbold, L. M. Kageorge, M. F. Schatz, and R. O. Grigoriev. Robust learning from noisy, incomplete, high-dimensional experimental data via physically constrained symbolic regression. *Nat. Commun.*, (12):3219, 2021.
- [64] S. Maddu, B. L. Cheeseman, I. F. Sbalzarini, and C. L. Müller. Stability selection enables robust learning of partial differential equations from limited noisy data. 2019. arXiv:1907.07810.
- [65] S. Maddu, B. L. Cheeseman, C. L. Müller, and I. F. Sbalzarini. Learning physically consistent differential equation models from data using group sparsity. *Phys. Rev. E*, 103(4):042310, 2021.
- [66] O. Mickelin, J. Słomka, K. J. Burns, D. Lecoanet, G. M. Vasil, L. M. Faria, and J. Dunkel. Anomalous Chained Turbulence in Actively Driven Flows on Spheres. *Phys. Rev. Lett.*, 120(16):164503, 2018.
- [67] R. Supekar, D. R. Hewitt, and N. J. Balmforth. Translating and squirming cylinders in a viscoplastic fluid. *J. Fluid Mech.*, 882:A11, 2020.
- [68] J. S. Ultman and M. M. Denn. Slow viscoelastic flow past submerged objects. *Chem. Eng. J.*, 2(2):81–89, 1971.
- [69] O. G. Harlen. The negative wake behind a sphere sedimenting through a viscoelastic fluid. *J. Non-Newtonian Fluid Mech.*, 108(1-3):411–430, 2002.

- [70] G. F. Brookes and R. L. Whitmore. Drag forces in bingham plastics. *Rheol. Acta*, 8(4):472–480, 1969.
- [71] K. Adachi and N. Yoshioka. On creeping flow of a visco-plastic fluid past a circular cylinder. *Chem. Eng. Sci.*, 28(1):215–226, 1973.
- [72] Y. Ding, N. Gravish, and D. I. Goldman. Drag induced lift in granular media. *Phys. Rev. Lett.*, 106(2):028001, 2011.
- [73] A. E. Hosoi and D. I. Goldman. Beneath our feet: strategies for locomotion in granular media. *Ann. Rev. Fluid Mech.*, 47:431–453, 2015.
- [74] N. Roquet and P. Saramito. An adaptive finite element method for Bingham fluid flows around a cylinder. *Comput. Meth. Appl. Mech. Eng.*, 192:3317–3341, 2003.
- [75] D. L. Tokpavi, A. Magnin, and P. Jay. Very slow flow of bingham viscoplastic fluid around a circular cylinder. *J. Non-Newtonian Fluid Mech.*, 154:65–76, 2008.
- [76] H. Ozogul, P. Jay, and A. Magnin. Slipping of a viscoplastic fluid flowing on a circular cylinder. *J. Fluids Eng.*, 137(7):071201, 2015.
- [77] E. Chaparian and I. A. Frigaard. Yield limit analysis of particle motion in a yield-stress fluid. *J. Fluid Mech.*, 819:311–351, 2017.
- [78] D. L. Tokpavi, A. Magnin, P. Jay, and L. Jossic. Experimental study of the very slow flow of a yield stress fluid around a circular cylinder. *J. Non-Newtonian Fluid Mech.*, 164:35–44, 2009.
- [79] N. J. Balmforth, I. A. Frigaard, and G. Ovarlez. Yielding to stress: recent developments in viscoplastic fluid mechanics. *Ann. Rev. Fluid Mech.*, 46:121–146, 2014.
- [80] J. D. Murff, D. A. Wagner, and M. F. Randolph. Pipe penetration in cohesive soil. *Géotechnique*, 39(2):213–229, 1989.
- [81] H. A. Barnes. A review of the slip (wall depletion) of polymer solutions, emulsions and particle suspensions in viscometers: its cause, character, and cure. *J. Non-Newtonian Fluid Mech.*, 56(3):221–251, 1995.
- [82] J. R. Blake. A spherical envelope approach to ciliary propulsion. *J. Fluid Mech.*, 46(1):199–208, 1971.
- [83] J. R. Blake. Self propulsion due to oscillations on the surface of a cylinder at low Reynolds number. *Bull. Aust. Math. Soc.*, 5(2):255–264, 1971.
- [84] M. J. Lighthill. On the squirring motion of nearly spherical deformable bodies through liquids at very small reynolds numbers. *Commun. Pure Appl. Maths*, 5(2):109–118, 1952.

- [85] R. J. Clarke, M. D. Finn, and M. MacDonald. Hydrodynamic persistence within very dilute two-dimensional suspensions of squirmers. *Proc. R. Soc. A*, 470 (2167):20130508, 2014.
- [86] S. Yazdi, A. M. Ardekani, and A. Borhan. Locomotion of microorganisms near a no-slip boundary in a viscoelastic fluid. *Phys. Rev. E*, 90(4):1–11, 2014.
- [87] S. E. Spagnolie and E. Lauga. Jet propulsion without inertia. *Phys. Fluids*, 22 (8):369, 2010.
- [88] W. Prager and P. G. Hodge. *Theory of perfectly plastic solids*. Wiley, 1951.
- [89] N. J. Balmforth, R. V. Craster, D. R. Hewitt, S. Hormozi, and A. Maleki. Viscoplastic boundary layers. *J. Fluid Mech.*, 813:929–954, 2017.
- [90] E. J. Hinch. *Perturbation methods*. Cambridge University Press, 1991.
- [91] R. Supekar, V. Heinonen, K. J. Burns, and J. Dunkel. Linearly forced fluid flow on a rotating sphere. *J. Fluid Mech.*, 892:A30, 2020.
- [92] G. Falkovich and K. R. Sreenivasan. Lessons from hydrodynamic turbulence. *Phys. Today*, 59(4):43–49, 2006.
- [93] U. Frisch. *Turbulence: The Legacy of A. N. Kolmogorov*. Cambridge University Press, 1995.
- [94] G. Boffetta and R. E. Ecke. Two-dimensional turbulence. *Annu. Rev. Fluid Mech*, 44(1):427–451, 2012.
- [95] S. Grossmann, D. Lohse, and C. Sun. High-reynolds number taylor-couette turbulence. *Annu. Rev. Fluid Mech*, 48(1):53–80, 2016.
- [96] D. Lucas and R. Kerswell. Spatiotemporal dynamics in two-dimensional Kolmogorov flow over large domains. *J. Fluid Mech*, 750:518–554, 2014.
- [97] D. H. Rothman. Negative-viscosity lattice gases. *J. Stat. Phys*, 56(3-4):517–524, 1989.
- [98] M. I. Tribelsky and K. Tsuboi. New scenario for transition to turbulence? *Phys. Rev. Lett.*, 76(10):1631–1634, 1996.
- [99] S. Sukoriansky, B. Galperin, and A. Chekhlov. Large scale drag representation in simulations of two-dimensional turbulence. *Phys. Fluids*, 11(10):3043–3053, 1999.
- [100] C. Rosales and C. Meneveau. Linear forcing in numerical simulations of isotropic turbulence: Physical space Implementations and convergence properties. *Phys. Fluids*, 17(9):095106, 2005.

- [101] J. Słomka, P. Suwara, and J. Dunkel. The nature of triad interactions in active turbulence. *J. Fluid Mech.*, 841:702–731, 2018.
- [102] M. Linkmann, G. Boffetta, M. C. Marchetti, and B. Eckhardt. Phase Transition to Large Scale Coherent Structures in Two-Dimensional Active Matter Turbulence. *Phys. Rev. Lett*, 122(21):214503, 2019.
- [103] V. I. Arnold. Kolmogorov’s hydrodynamic attractors. *Proc. R. Soc. Lond. A*, 434:19–22, 1991.
- [104] E. N. Lorenz. Deterministic Nonperiodic Flow. *J. Atmos. Sci*, 20:130–141, 1963.
- [105] I. A. Beresnev and V. N. Nikolaevskiy. A model for nonlinear seismic waves in a medium with instability. *Physica D*, 66:1–6, 1993.
- [106] G. M. Vasil. On the magnetorotational instability and elastic buckling. *Proc. R. Soc. Lond A*, 471:20140699, 2015.
- [107] M. Heimpel, J. Aurnou, and J. Wicht. Simulation of equatorial and high-latitude jets on Jupiter in a deep convection model. *Nature*, 438(7065):193–196, 2005.
- [108] T. Schneider and J. Liu. Formation of Jets and Equatorial Superrotation on Jupiter. *J. Atmos. Sci*, 66(3):579–601, 2008.
- [109] T. Sanchez, D. T. N. Chen, S. J. Decamp, M. Heymann, and Z. Dogic. Spontaneous motion in hierarchically assembled active matter. *Nature*, 491(7424):431–434, 2012.
- [110] R. Zhang, Y. Zhou, M. Rahimi, and J. J. De Pablo. Dynamic structure of active nematic shells. *Nat. Commun*, 7:13483, 2016.
- [111] S. Henkes, M. C. Marchetti, and R. Sknepnek. Dynamical patterns in nematic active matter on a sphere. *Phys. Rev. E*, 97(4):042605, 2018.
- [112] I. Nitschke, S. Reuther, and A. Voigt. Hydrodynamic interactions in polar liquid crystals on evolving surfaces. *Phys. Rev. Fluids*, 4:044002, 2019.
- [113] H. Löwen. Active particles in noninertial frames: How to self-propel on a carousel. *Phys. Rev. E*, 99(6):1–12, 2019.
- [114] J. Słomka, A. Townsend, and J. Dunkel. Stokes’ second problem and reduction of inertia in active fluids. *Phys. Rev. Fluids*, 3:103304, 2018.
- [115] J. G. Charney, R. Fjörtoft, and J. V. Neumann. Numerical Integration of the Barotropic Vorticity Equation. *Tellus*, 2(4):237–254, 1950.
- [116] P. B. Rhines. Waves and turbulence on a beta-plane. *J. Fluid Mech.*, 69(3):417–443, 2006.

- [117] G. K. Vallis and M. E. Maltrud. Generation of Mean Flows and Jets on a Beta Plane and over Topography. *J. Phys. Oceanogr.*, 23:1346–1362, 1993.
- [118] K. Srinivasan and W. R. Young. Zonostrophic Instability. *J. Atmos. Sci.*, 69(5):1633–1656, 2012.
- [119] S. M. Tobias and J. B. Marston. Direct statistical simulation of out-of-equilibrium jets. *Phys. Rev. Lett.*, 110:104502, 2013.
- [120] B. Galperin, S. Sukoriansky, and N. Dikovskaya. Geophysical flows with anisotropic turbulence and dispersive waves: flows with a β -effect. *Ocean Dyn.*, 60(2):427–441, 2010.
- [121] N. A. Bakas, N. C. Constantinou, and P. J. Ioannou. Statistical State Dynamics of Weak Jets in Barotropic Beta-Plane Turbulence. *J. Atmos. Sci.*, 76(3):919–945, 2019.
- [122] F. Waleffe. Exact coherent structures in channel flow. *J. Fluid Mech*, 435:93–102, 2001.
- [123] H. Wedin and R. Kerswell. Exact coherent structures in pipe flow: Travelling wave solutions. *J. Fluid Mech*, 508(508):333–371, 2004.
- [124] B. J. Hoskins. Stability of the Rossby-Haurwitz wave. *Q. J. R. Meteorol. Soc.*, 99:723–745, 1973.
- [125] P. Lynch. On resonant Rossby-Haurwitz triads. *Tellus*, 61A(3):438–445, 2009.
- [126] R. A. Madden. How I Learned to Love Normal-Mode Rossby–Haurwitz Waves. *Bull. Am. Meteorol. Soc*, 100(3):503–511, 2018.
- [127] P. D. Thompson. A generalized class of exact time-dependent solutions of the vorticity equation for nondivergent barotropic flow. *Mon. Weather Rev.*, 110(9):1321–1324, 1982.
- [128] J. Pedlosky. *Waves in the Ocean and Atmosphere: Introduction to Wave Dynamics*. Springer, 2003.
- [129] K. J. Burns, G. M. Vasil, J. S. Oishi, D. Lecoanet, and B. P. Brown. Dedalus: A flexible framework for numerical simulations with spectral methods. *Phys. Rev. Res.*, 2(2):023068, 2020.
- [130] D. Lecoanet, G. M. Vasil, K. J. Burns, B. P. Brown, and J. S. Oishi. Tensor calculus in spherical coordinates using Jacobi polynomials. Part-II: Implementation and examples. *J. Comp. Phys.: X*, 3:100012, 2019.
- [131] G. M. Vasil, D. Lecoanet, K. J. Burns, J. S. Oishi, and B. P. Brown. Tensor calculus in spherical coordinates using Jacobi polynomials. Part-I: Mathematical analysis and derivations. *J. Comp. Phys.: X*, 1:100013, 2019.

- [132] J. B. Parker and J. A. Krommes. Zonal flow as pattern formation. *Phys. Plasmas*, 20(10), 2013.
- [133] B. Galperin and P. L. Read. *Zonal Jets: Phenomenology, Genesis, and Physics*. Cambridge University Press, 2019.
- [134] R. Supekar, B. Song, A. Hastewell, A. Mietke, and J. Dunkel. Learning hydrodynamic equations for active matter from particle simulations and experiments, 2021. arXiv:2101.06568.
- [135] D. T. Tambe, C. Corey Hardin, T. E. Angelini, K. Rajendran, C. Y. Park, X. Serra-Picamal, E. H. Zhou, M. H. Zaman, J. P. Butler, D. A. Weitz, J. J. Fredberg, and X. Trepat. Collective cell guidance by cooperative intercellular forces. *Nat. Mater.*, 10(6):469–475, 2011.
- [136] C.-P. Heisenberg and Y. Bellaïche. Forces in Tissue Morphogenesis and Patterning. *Cell*, 153(5):948–962, 2013.
- [137] M. Tennenbaum, Z. Liu, D. Hu, and A. Fernandez-Nieves. Mechanics of fire ant aggregations. *Nat. Mater.*, 15:54–59, 2016.
- [138] L. M. Nash, D. Kleckner, A. Read, V. Vitelli, A. M. Turner, and W. T. M. Irvine. Topological mechanics of gyroscopic metamaterials. *Proc. Natl. Acad. Sci. U. S. A.*, 112(47):14495–14500, 2015.
- [139] W. Savoie, T. A. Berrueta, Z. Jackson, A. Pervan, R. Warkentin, S. Li, T. D. Murphey, K. Wiesenfeld, and D. I. Goldman. A robot made of robots: Emergent transport and control of a smarticle ensemble. *Sci. Robot.*, 4(34):eaax4316, 2019.
- [140] Y. Li, M. Chen, J. Hu, R. Sheng, Q. Lin, X. He, and M. Guo. Volumetric compression induces intracellular crowding to control intestinal organoid growth via wnt/ β -catenin signaling. *Cell Stem Cell*, 2020.
- [141] N. J. Cira, A. Benusiglio, and M. Prakash. Vapour-mediated sensing and motility in two-component droplets. *Nature*, 519:446–450, 2015.
- [142] W. B. Rogers, W. M. Shih, and V. N. Manoharan. Using DNA to program the self-assembly of colloidal nanoparticles and microparticles. *Nat. Rev. Mater.*, 1:16008, 2016.
- [143] E. H. K. Stelzer. Light-sheet fluorescence microscopy for quantitative biology. *Nat. Meth.*, 12(1):23–26, 2015.
- [144] R. M. Power and J. Huisken. A guide to light-sheet fluorescence microscopy for multiscale imaging. *Nat. Meth.*, 14(4):360–373, 2017.
- [145] H. Jeckel, E. Jelli, R. Hartmann, P. K. Singh, R. Mok, J. F. Tetzlaff, L. Vidakovic, B. Eckhardt, J. Dunkel, and K. Drescher. Learning the space-time phase diagram of bacterial swarm expansion. *Proc. Natl. Acad. Sci. U. S. A.*, 116(5):1489–1494, 2019.

- [146] B. Qin, C. Fei, A. A. Bridges, A. A. Mashruwala, H. A. Stone, N. S. Wingreen, and B. L. Bassler. Cell position fates and collective fountain flow in bacterial biofilms revealed by light-sheet microscopy. *Science*, 369(6499):71–77, 2020.
- [147] R. Hartmann, H. Jeckel, E. Jelli, P. K. Singh, S. Vaidya, M. Bayer, D. K. H. Rode, L. Vidakovic, F. Díaz-Pascual, J. C. N. Fong, A. Dragoš, O. Lamprecht, J. G. Thöming, N. Netter, S. Häussler, C. D. Nadell, V. Sourjik, Á. T. Kovács, F. H. Yildiz, and K. Drescher. Quantitative image analysis of microbial communities with BiofilmQ. *Nat. Microbiol.*, 6:151–156, 2021.
- [148] D. P. Vallette, G. Jacobs, and J. P. Gollub. Oscillations and spatiotemporal chaos of one-dimensional fluid fronts. *Phys. Rev. E*, 55(4):4274–4287, 1997.
- [149] J. Bongard and H. Lipson. Automated reverse engineering of nonlinear dynamical systems. *Proc. Natl. Acad. Sci. U. S. A.*, 104(24):9943–9948, 2007.
- [150] C. Linghu, S. L. Johnson, P. A. Valdes, O. A. Shemesh, W. M. Park, D. Park, K. D. Piatkevich, A. T. Wassie, Y. Liu, B. An, S. A. Barnes, O. T. Celiker, C. Yao, C. Yu, R. Wang, K. P. Adamala, M. F. Bear, A. E. Keating, and E. S. Boyden. Spatial multiplexing of fluorescent reporters for imaging signaling network dynamics. *Cell*, 183(6):1682 – 1698.e24, 2020.
- [151] N. Cermak, S. K. Yu, R. Clark, Y. Huang, S. N. Baskoylu, and S. W. Flavell. Whole-organism behavioral profiling reveals a role for dopamine in state-dependent motor program coupling in *C. elegans*. *eLife*, 9:e57093, 2020.
- [152] J. T. Nardini, R. E. Baker, M. J. Simpson, and K. B. Flores. Learning differential equation models from stochastic agent-based model simulations. *J. R. Soc. Interface*, 18(176):20200987, 2021.
- [153] L. Felsberger and P. Koutsourelakis. Physics-constrained, data-driven discovery of coarse-grained dynamics. *Comput. Phys. Commun.*, 25(5):1259–1301, 2019.
- [154] Joseph Bakarji and Daniel M. Tartakovsky. Data-driven discovery of coarse-grained equations. *J. Comput. Phys.*, 434:110219, 2021.
- [155] H. Schaeffer. Learning partial differential equations via data discovery and sparse optimization. *Proc. R. Soc. A*, 473:20160446, 2017.
- [156] S. Smith. *Digital signal processing: a practical guide for engineers and scientists*. Elsevier, 2013.
- [157] T. A Driscoll, N. Hale, and L. N. Trefethen. *Chebfun Guide*. Pafnuty Publications, 2014. URL <http://www.chebfun.org/docs/guide/>.
- [158] F. Peruani, A. Deutsch, and M. Bär. A mean-field theory for self-propelled particles interacting by velocity alignment mechanisms. *Eur. Phys. J. Spec. Top.*, 157(1):111–122, 2008.

- [159] B. Liebchen and D. Levis. Collective behavior of chiral active matter: Pattern formation and enhanced flocking. *Phys. Rev. Lett.*, 119:058002, 2017.
- [160] N. Kruk, J. A. Carrillo, and H. Koeppl. Traveling bands, clouds, and vortices of chiral active matter. *Phys. Rev. E*, 102(2):22604, 2020.
- [161] Y. Sumino, K. H. Nagai, Y. Shitaka, D. Tanaka, K. Yoshikawa, H. Chaté, and K. Oiwa. Large-scale vortex lattice emerging from collectively moving microtubules. *Nature*, 483(7390):448–452, 2012.
- [162] L. Huber, R. Suzuki, T. Krüger, E. Frey, and A. R. Bausch. Emergence of coexisting ordered states in active matter systems. *Science*, 361(6399):255–258, 2018.
- [163] H. Li, X.-Q. Shi, M. Huang, X. Chen, M. Xiao, C. Liu, H. Chaté, and H. P. Zhang. Data-driven quantitative modeling of bacterial active nematics. *Proc. Natl. Acad. Sci. U.S.A.*, 116(3):777–785, 2019.
- [164] F. Giavazzi, C. Malinverno, S. Corallino, F. Ginelli, G. Scita, and R. Cerbino. Giant fluctuations and structural effects in a flocking epithelium. *J. Phys. D: Appl. Phys.*, 50(38):384003, 2017.
- [165] A. P. Solon, J. Stenhammar, M. E. Cates, Y. Kafri, and J. Tailleur. Generalized thermodynamics of phase equilibria in scalar active matter. *Phys. Rev. E*, 97:020602, 2018.
- [166] E. Wallin and M. Servin. Data-driven model order reduction for granular media, 2020. arXiv:2004.03349.
- [167] J. P. Boyd. *Chebyshev and Fourier spectral methods*. Courier Corporation, 2001.
- [168] J. C. Mason and D. C. Handscomb. *Chebyshev polynomials*. CRC press, 2002.
- [169] O. Bruno and D. Hoch. Numerical differentiation of approximated functions with limited order-of-accuracy deterioration. *SIAM J. Numer. Anal.*, 50(3):1581–1603, 2012.
- [170] J. L. Aurentz and L. N. Trefethen. Chopping a Chebyshev Series. *ACM Trans. Math. Softw.*, 43(4), 2017.
- [171] M. Cross and H. Greenside. *Pattern Formation and Dynamics in Nonequilibrium Systems*. Cambridge University Press, Cambridge, 2009.
- [172] N. Meinshausen and P. Bühlmann. Stability selection. *J. R. Statist. Soc. B*, 72(4):417–473, 2010.
- [173] R. D. Shah and R. J. Samworth. Variable selection with error control: Another look at stability selection. *J. R. Statist. Soc. B*, 75(1):55–80, 2013.

- [174] K. V. Edmond, M. T. Elsesser, G. L. Hunter, D. J. Pine, and E. R. Weeks. Decoupling of rotational and translational diffusion in supercooled colloidal fluids. *Proc. Natl. Acad. Sci. U. S. A.*, 109(44):17891–17896, 2012.
- [175] J. Toner, Y. Tu, and S. Ramaswamy. Hydrodynamics and phases of flocks. *Ann. Phys. (N. Y.)*, 318:170–244, 2005.
- [176] J. Colen, M. Han, R. Zhang, S. A. Redford, L. M. Lemma, and L. Morgan. Machine learning active-nematic hydrodynamics. *Proc. Natl. Acad. Sci. U. S. A.*, 118(10):e2016708118, 2021.
- [177] K. P. Champion, S. L. Brunton, and J. N. Kutz. Discovery of nonlinear multi-scale systems: Sampling strategies and embeddings. *SIAM J. Appl. Dyn. Syst.*, 18(1):312–333, 2019.
- [178] D. Zhang, L. Guo, and G. E. Karniadakis. Learning in modal space: Solving time-dependent stochastic pdes using physics-informed neural networks. *SIAM J. Sci. Comput.*, 42(2):A639–A665, 2020.
- [179] H. K. Moffatt. Note on the triad interactions of homogeneous turbulence. *J. Fluid Mech.*, 741:1–11, 2014.
- [180] P. Zheng, T. Askham, S. L. Brunton, J. N. Kutz, and A. Y. Aravkin. A Unified Framework for Sparse Relaxed Regularized Regression: SR3. *IEEE Access*, 7:1404–1423, 2019.
- [181] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer New York Inc., 2001.
- [182] N. M. Mangan, J. N. Kutz, S. L. Brunton, and J. L. Proctor. Model selection for dynamical systems via sparse regression and information criteria. *Proc. R. Soc. London A*, 473:20170009, 2017.
- [183] T. Walter and I. D. Couzin. Trex, a fast multi-animal tracking system with markerless identification, and 2d estimation of posture and visual fields. *eLife*, 10:e64000, 2021.
- [184] P. Lynch. Resonant Rossby wave triads and the swinging spring. *Bull. Am. Meteorol. Soc.*, 84(5):605–616, 2003.
- [185] M. Frigo and S. Johnson. The design and implementation of FFTW3. *Proceedings of the IEEE*, 93(2):216–231, 2005.
- [186] D. S. Dean. Langevin equation for the density of a system of interacting Langevin processes. *J. Phys. A*, 29(24):L613–L617, 1996.