Appendix A

Some useful classical results

This appendix collects the statements of the classical, general-purpose theorems that are invoked – but not proved – at various points in this book. None of these results is specific to kinetic theory or transport equations; they belong to the common toolbox of analysis, complex analysis, ordinary differential equations, linear algebra, probability, and numerical analysis. We state each result precisely, indicate where it is used in the text, and give a reference to a standard textbook where a complete proof can be found.

A.1 Measure theory and integration

Lemma A.1.1 (Riemann–Lebesgue lemma)**.** *Let $g \in L^1(\mathbb{R}^n)$. Then its Fourier transform*

$$\hat{g}(\xi) \;=\; \int_{\mathbb{R}^n} g(x)\, e^{-i\xi\cdot x}\,\mathrm{d}x$$

satisfies $\hat{g}(\xi) \to 0$ as $|\xi| \to +\infty$.

This lemma is used in the proof of Theorem 2.1.2(i) (Chapter 2) to obtain the qualitative decay of the Fourier coefficients $\hat{m}_{\varphi,k}(t)$ of macroscopic quantities under free transport. We refer to Folland [15], Chapter 8, for a proof.

Theorem A.1.2 (Dominated convergence theorem)**.** *Let $(f_n)_{n\geq 1}$ be a sequence of measurable functions on a measure space (X, μ) converging pointwise (or almost everywhere) to a function f. If there exists $g \in L^1(X, \mu)$ such that $|f_n| \leq g$ almost everywhere for every n, then $f \in L^1(X, \mu)$ and*

$$\int_X f_n\,\mathrm{d}\mu \;\longrightarrow\; \int_X f\,\mathrm{d}\mu.$$

This is an important theorem often used in these lecture notes. For instance, in the proof of Theorem 2.1.2(ii) (Chapter 2), it is applied to justify passing to the limit $\varepsilon \to 0$ when shrinking the rectangular contour R_T^{ε} onto the boundary of the strip of holomorphy. We refer to Folland [15], Chapter 2, for a proof.

Definition A.1.3 (Radon measure)**.** *A* (positive) Radon measure *on $\mathbb{R}^d$ is a Borel measure μ that is* locally finite *(every point has a neighbourhood of finite measure, equivalently $\mu(K) < \infty$*

for every compact $K \subset \mathbb{R}^d$) and regular *(for every Borel set A, $\mu(A) = \inf\{\mu(U) : A \subset U,\ U \text{ open}\} = \sup\{\mu(K) : K \subset A,\ K \text{ compact}\}$). On $\mathbb{R}^d$, every locally finite Borel measure is automatically regular, so a Radon measure is simply a Borel measure that is finite on compact sets; this includes all finite measures, all measures with a locally integrable density, and all sums of Dirac masses such as the empirical measures of Chapter 2 (Section 3.2 of Chapter 3) and Chapter 5.*

This is the class of measures μ^{in} for which the push-forward formula of Theorem 2.3.3 (Chapter 2) is stated. We refer to Folland [15], Chapter 7, for the general theory of Radon measures on locally compact Hausdorff spaces.

Theorem A.1.4 (Fubini's theorem)**.** *Let (X, μ) and (Y, ν) be σ-*finite *measure spaces (i.e. X, resp. Y, is a countable union of measurable sets of finite measure) and let $f \in L^1(X \times Y, \mu \otimes \nu)$. Then, for μ-almost every $x \in X$, the function $y \mapsto f(x, y)$ is ν-integrable, $x \mapsto \int_Y f(x, y)\,\mathrm{d}\nu(y)$ is μ-integrable, and*

$$\int_{X\times Y} f\,\mathrm{d}(\mu\otimes\nu) \;=\; \int_X \left(\int_Y f(x,y)\,\mathrm{d}\nu(y)\right)\mathrm{d}\mu(x) \;=\; \int_Y \left(\int_X f(x,y)\,\mathrm{d}\mu(x)\right)\mathrm{d}\nu(y).$$

Used after Lemma 4.3.7 (Chapter 4) to extend the one-dimensional B-spline interpolation error estimate to several dimensions by integrating successively in each coordinate direction. We refer to Folland [15], Chapter 2, for a proof.

A.2 Complex analysis

Theorem A.2.1 (Cauchy's integral theorem)**.** *Let $\Omega \subset \mathbb{C}$ be an open set and let $g : \Omega \to \mathbb{C}$ be holomorphic. If $R \subset \Omega$ is a closed rectangle (with sides parallel to the axes) entirely contained in Ω, and $\Gamma = \partial R$ is its boundary, positively oriented, then*

$$\oint_\Gamma g(z)\,\mathrm{d}z \;=\; 0.$$

(The same conclusion holds, more generally, for any closed contour that can be continuously shrunk to a point within Ω.)

Remark A.2.2. *Used in the proof of Theorem 2.1.2(ii) (Chapter 2) to deform the contour of integration from the real axis to the shifted line $\Im z = -\sigma$, converting a purely oscillatory integral into one carrying an explicit exponential decay factor (see Figure 2.1). We refer to Ahlfors [1], Chapter 4, for a proof.*

A.3 Ordinary differential equations

Theorem A.3.1 (Cauchy–Lipschitz theorem)**.** *Let $I \subset \mathbb{R}$ be an interval, $D \subset \mathbb{R}^d$ an open set, and $V : I \times D \to \mathbb{R}^d$ continuous and locally Lipschitz in its second argument, uniformly on compact subsets of I. Then, for every $(t_0, z_0) \in I \times D$, the Cauchy problem*

$$\dot\gamma(s) = V(s, \gamma(s)), \qquad \gamma(t_0) = z_0,$$

has a unique maximal solution, defined on an open subinterval of I containing t_0; moreover, if V has linear growth in z, uniformly in $t \in I$, this maximal solution is global (defined on all of I).

This is the classical existence-uniqueness theorem invoked in the proof of Theorem 2.2.2 (Chapter 2) to construct the flow map of the characteristic ODE; the remainder of that proof (global existence under the linear-growth assumption (H3), and smooth dependence on parameters) is then given in full. We refer to Hartman [19], Chapter II, for a proof.

Lemma A.3.2 (Grönwall's lemma)**.** *Let $u, \phi : [t_0, T] \to \mathbb{R}_+$ be continuous, with $\phi \geq 0$, and let $C \geq 0$. If*

$$u(t) \;\leq\; C + \int_{t_0}^{t} \phi(s)\, u(s)\, \mathrm{d}s \qquad \forall\, t \in [t_0, T],$$

then

$$u(t) \;\leq\; C\, \exp\!\Big(\int_{t_0}^{t} \phi(s)\, \mathrm{d}s\Big) \qquad \forall\, t \in [t_0, T].$$

Equivalently, in differential form: if u is differentiable and $u'(t) \leq \phi(t)\, u(t) + \psi(t)$ for some $\psi \geq 0$, then

$$u(t) \leq \big(u(t_0) + \int_{t_0}^{t} \psi(s)\, e^{-\int_{t_0}^{s} \phi(\tau) \mathrm{d}\tau} \mathrm{d}s\big)\, e^{\int_{t_0}^{t} \phi(\tau) \mathrm{d}\tau}.$$

This is by far the most frequently used classical tool in the book: it underlies the global-existence bound in the proof of Theorem 2.2.2 (Chapter 2), the well-posedness of the mean-field characteristic flow in Theorem 3.3.1 (Chapter 3), the convergence estimates for particle and semi-Lagrangian methods in Theorem 4.2.2 and Theorem 4.4.1 (Chapter 4), and the flocking estimates of Theorem 5.2.2 (Chapter 5), among many other places. We refer to Evans [14], Appendix B.2, for a proof.

Theorem A.3.3 (Banach fixed-point theorem)**.** *Let (X, d) be a non-empty complete metric space and let $\Phi : X \to X$ be a contraction, i.e. there exists $k \in [0, 1)$ such that $d(\Phi(x), \Phi(y)) \leq k\, d(x, y)$ for all $x, y \in X$. Then Φ has a unique fixed point $x^* \in X$, and for every $x_0 \in X$ the sequence of iterates $x_{n+1} := \Phi(x_n)$ converges to x^*.*

Remark A.3.4. *Used in the proof of Theorem 3.3.1 (Chapter 3) to construct the mean-field characteristic flow as the fixed point of a Picard map on a weighted space of continuous functions with linear growth. We refer to Brezis [6], Chapter 2, for a proof.*

A.4 Linear algebra

Lemma A.4.1 (Jacobi's formula)**.** *Let $A \in \mathrm{GL}_d(\mathbb{R})$. The differential of the determinant map at A is given by*

$$(D \det)(A) \cdot B \;=\; \det(A)\, \mathrm{tr}(A^{-1} B), \qquad \forall\, B \in \mathcal{M}_d(\mathbb{R}).$$

Remark A.4.2. *Used in the proof of Theorem 2.2.3 (Chapter 2) to derive the evolution equation $\partial_s J = (\nabla_z \cdot V)(s, Z)\, J$ satisfied by the Jacobian $J(s, t, z) = \det(D_z Z(s, t, z))$ of the characteristic flow. We refer to Horn and Johnson [20], Chapter 0, for a proof (there stated in the equivalent form $\det(A + \varepsilon B) = \det(A)(1 + \varepsilon\, \mathrm{tr}(A^{-1} B) + o(\varepsilon))$).*

A.5 Probability

Theorem A.5.1 (Strong law of large numbers)**.** *Let $(X_i)_{i\geq 1}$ be independent, identically distributed random variables with $\mathbb{E}[|X_1|] < \infty$. Then, almost surely,*

$$\frac{1}{N}\sum_{i=1}^{N} X_i \longrightarrow \mathbb{E}[X_1] \qquad \textit{as } N \to +\infty.$$

Remark A.5.2. *Invoked informally in Chapter 3 (Section 3.2) to motivate the mean-field limit of the empirical average, and again in Chapter 4 (Section 4.2) to justify the $\mathcal{O}(N^{-1/2})$ convergence, in Wasserstein distance, of a Monte–Carlo initial particle distribution towards the target density μ^{in}. We refer to Durrett [13], Chapter 2, for a proof (together with the companion central limit theorem, which explains the $N^{-1/2}$ rate).*

A.6 Functional analysis and weak compactness

We collect here the basic facts about L^p spaces and weak/weak-* compactness that underlie the various compactness arguments used (sometimes only implicitly) in Chapters 3–5: the mean-field and hydrodynamic limits, and the existence theory for the kinetic flocking equation, all rely at some point on extracting a convergent subsequence from a family of approximate solutions, and this is the classical machinery that makes such extractions rigorous.

Definition A.6.1 (Separable and reflexive Banach spaces)**.** *Let E be a Banach space with topological dual E' and bidual $E'' = (E')'$.*

- *E is* separable *if it contains a countable dense subset (equivalently, every element of E can be approximated, in norm, by a sequence drawn from a fixed countable set).*
- *The* canonical embedding *$J : E \to E''$ is defined by $\langle J(x), \phi\rangle_{E'',E'} := \langle \phi, x\rangle_{E',E}$ for $\phi \in E'$; it is always a linear isometry. E is* reflexive *if J is onto, i.e. $J(E) = E''$, so that E may be identified with its own bidual.*

Proposition A.6.2 (L^p spaces)**.** *Let (X, μ) be a measure space and $1 \leq p \leq \infty$.*

(i) (Riesz–Fischer) *$L^p(X, \mu)$, equipped with the norm*

$$\|f\|_{L^p} = \Big(\int_X |f|^p \,\mathrm{d}\mu\Big)^{1/p},$$

the essential supremum when $p = \infty$), is a Banach space; it is separable, in the sense of Definition A.6.1, whenever X is (e.g. $X = \mathbb{R}^d$ or an open subset thereof) and $p < \infty$.

(ii) (Duality) *For $1 \leq p < \infty$ with conjugate exponent $p' \in (1, \infty]$ (defined by $1/p + 1/p' = 1$), the topological dual of $L^p(X, \mu)$ is isometrically isomorphic to $L^{p'}(X, \mu)$, via $g \mapsto (f \mapsto \int_X f g \,\mathrm{d}\mu)$, and Hölder's inequality*

$$\left|\int_X f g \,\mathrm{d}\mu\right| \leq \|f\|_{L^p} \|g\|_{L^{p'}}$$

holds.

(iii) (Reflexivity) *For $1 < p < \infty$, $L^p(X,\mu)$ is reflexive, in the sense of Definition A.6.1 (this follows from (ii): the bidual of L^p is identified with $L^{(p')'} = L^p$ itself). $L^1(X,\mu)$ and $L^\infty(X,\mu)$ are not reflexive, in general (unless X is purely atomic with finitely many atoms).*

Remark A.6.3. *These structural facts are used implicitly throughout the L^p estimates of Chapter 4 (Section 4.3) and the a priori bounds (5.11)–(5.12) of Chapter 5. We refer to Brezis [6], Chapter 4, for a proof.*

Definition A.6.4 (Weak and weak-* convergence)**.** *Let E be a Banach space with topological dual E'. A sequence $(x_n)_{n\geq 1} \subset E$ converges* weakly *to $x \in E$, written $x_n \rightharpoonup x$, if $\langle \phi, x_n\rangle \to \langle \phi, x\rangle$ for every $\phi \in E'$.*
A sequence $(\phi_n)_{n\geq 1} \subset E'$ converges weak-* *to $\phi \in E'$, written $\phi_n \overset{*}{\rightharpoonup} \phi$, if $\langle \phi_n, x\rangle \to \langle \phi, x\rangle$ for every $x \in E$. When $E = L^p(X,\mu)$ and $E' = L^{p'}(X,\mu)$ as in Proposition A.6.2(ii), this reads*

$$\int f_n\, g \,\mathrm{d}\mu \;\to\; \int f\, g\,\mathrm{d}\mu$$

for every g in the relevant dual space.

Theorem A.6.5 (Banach–Alaoglu theorem)**.** *Let E be a Banach space with topological dual E'. The closed unit ball of E' is compact for the weak-* topology $\sigma(E', E)$. Equivalently, if E is separable (Definition A.6.1), every bounded sequence in E' admits a weak-* convergent subsequence.*

Remark A.6.6. *Since $L^1(\mathbb{R}^{2d})$ is separable with dual $L^\infty(\mathbb{R}^{2d})$, this is the tool that turns the L^∞ a priori bound obtained from the hypothesis $f_0 \in L^\infty$ in Theorem 5.3.1 (Chapter 5) into the existence of a weak-* limit point for a sequence of approximate solutions f^n – the "compactness" referred to, but not detailed, in Step 2 of that proof. We refer to Brezis [6], Chapter 3, for a proof.*

Corollary A.6.7 (Weak compactness in reflexive spaces)**.** *If E is a reflexive Banach space (Definition A.6.1), every bounded sequence in E has a weakly convergent subsequence. In particular, by Proposition A.6.2(iii), every sequence bounded in $L^p(X,\mu)$ for some $1 < p < \infty$ admits a weakly convergent subsequence in $L^p(X,\mu)$.*

Remark A.6.8. *Combined with the L^p bound (5.11) (Chapter 5), this is what allows one to extract a weak limit from a bounded sequence of approximate solutions before applying the averaging lemma (Proposition 5.3.3) to upgrade this weak convergence into strong convergence of the macroscopic moments ρ and J. We refer to Brezis [6], Chapter 3 (Kakutani's and the Eberlein–Šmulian theorems), for a proof.*

Theorem A.6.9 (Dunford–Pettis theorem)**.** *Let (X,μ) be a finite measure space and let $\mathcal{F} \subset L^1(X,\mu)$ be bounded in L^1-norm. Then $\mathcal{F}$ is relatively weakly compact in $L^1(X,\mu)$ if and only if it is uniformly integrable, i.e.*

$$\lim_{R\to+\infty} \sup_{f\in\mathcal{F}} \int_{\{|f|>R\}} |f|\,\mathrm{d}\mu \;=\; 0.$$

Remark A.6.10. *Since L^1 is not reflexive, Corollary A.6.7 does not apply directly to the mass bound $\|f\|_{L^1} \le \|f_0\|_{L^1}$ used throughout Chapter 5; the Dunford–Pettis theorem is the substitute that yields weak L^1 compactness once uniform integrability has been secured, typically from the higher-moment bound* (5.12). *We refer to Brezis [6], Chapter 4, for a proof.*

Definition A.6.11 (Polish space)**.** *A topological space X is* Polish *if it is separable (it contains a countable dense subset) and* completely metrizable *(there exists a metric d, inducing the topology of X, for which (X, d) is a complete metric space). Every finite- or countable-dimensional Euclidean space $\mathbb{R}^d$, every closed subset of such a space, and every separable Banach space (in particular every separable $L^p(X, \mu)$ space, $1 \le p < \infty$, by Proposition A.6.2(i)) is a Polish space; this covers every space on which probability measures appear in this book.*

Theorem A.6.12 (Prokhorov's theorem)**.** *Let X be a Polish space and let $(\mu_n)_{n\ge1}$ be a* tight *sequence of probability measures on X, i.e. for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset X$ such that $\mu_n(K_\varepsilon) \ge 1 - \varepsilon$ for every n. Then $(\mu_n)_{n\ge1}$ admits a subsequence converging* narrowly *(i.e. weakly against bounded continuous test functions) to a probability measure μ on X.*

Remark A.6.13. *This is the classical compactness tool behind statements of the type "$\mu_N^{in} \to \mu^{in}$ weakly" for empirical measures, as used informally in Chapter 3 (Section 3.2) and Chapter 4 (Section 4.2, Monte–Carlo initialisation): on $X = \mathbb{R}^d$, tightness of the family of empirical measures follows from a uniform bound on a moment $\int |z|\, \mu_N^{in}(\mathrm{d}z)$, and Prokhorov's theorem then extracts a narrowly convergent subsequence, independently of the quantitative Wasserstein rate obtained separately from the law of large numbers (Theorem A.5.1). We refer to Billingsley [4], Chapter 1, for a proof.*

A.7 Classical inequalities

Proposition A.7.1 (Young's convolution inequality)**.** *Let ν be a finite signed Borel measure on $\mathbb{R}^d$, with* total variation norm

$$\|\nu\|_{TV} := \sup \sum_i |\nu(A_i)|,$$

the supremum being over all finite measurable partitions $(A_i)_i$ of $\mathbb{R}^d$.
For a measure with density $h \in L^1(\mathbb{R}^d)$, this is simply

$$\|\nu\|_{TV} = \|h\|_{L^1},$$

and let $\psi \in L^p(\mathbb{R}^d)$, $1 \le p \le \infty$. Then $\nu \star \psi \in L^p(\mathbb{R}^d)$ and

$$\|\nu \star \psi\|_{L^p(\mathbb{R}^d)} \;\le\; \|\nu\|_{TV}\, \|\psi\|_{L^p(\mathbb{R}^d)}.$$

Remark A.7.2. *Used in the proof of Lemma 4.3.6 (Chapter 4) to bound the Sobolev norm of a B-spline convolution kernel applied to a test function. We refer to Brezis [6], Chapter 4 (stated there for $\nu \in L^1(\mathbb{R}^d)$; the extension to a finite signed measure ν follows from the same duality argument, approximating ν by L^1 densities), for a proof.*

Proposition A.7.3 (Poincaré inequality on an interval). *Let $e \in H^2(0,1)$ satisfy $e(0) = e(1) = 0$. Then*

$$\|e\|_{L^p(0,1)} \le \frac{1}{8}\,\|e''\|_{L^p(0,1)}, \qquad 1 \le p \le \infty.$$

Remark A.7.4. *Used in the proof of Lemma 4.3.7 (Chapter 4) to bound the error of the cubic B-spline interpolant on a single mesh cell. This is a quantitative, elementary form of the general Poincaré–Sobolev inequalities discussed in Evans [14], Chapter 5; the explicit constant $1/8$ follows from solving the two-point boundary value problem $-e'' = -\varphi''$, $e(0) = e(1) = 0$, by the Green's function of $-\mathrm{d}^2/\mathrm{d}x^2$ on $(0,1)$.*

Theorem A.7.5 (Minkowski's integral inequality). *Let (X,μ) and (Y,ν) be σ-finite measure spaces, let $1 \le p < \infty$, and let $F : X \times Y \to \mathbb{R}$ be measurable. Then*

$$\left\| \int_Y F(\cdot,y)\,\mathrm{d}\nu(y) \right\|_{L^p(X,\mu)} \le \int_Y \|F(\cdot,y)\|_{L^p(X,\mu)}\,\mathrm{d}\nu(y).$$

Remark A.7.6. *Used in the proof leading to (4.9) (Chapter 4) to control the L^p convolution error of the smoothed-particle approximation, both in the convolution variable y and in the Taylor-remainder parameter θ. We refer to Folland [15], Chapter 6 (Minkowski's inequality for integrals), for a proof.*

A.8 Numerical analysis

Theorem A.8.1 (Baker–Campbell–Hausdorff formula). *Let $\mathcal{A}$ and $\mathcal{B}$ be two (possibly unbounded) operators generating flows $e^{s\mathcal{A}}$ and $e^{s\mathcal{B}}$. Then, formally,*

$$e^{s\mathcal{A}}e^{s\mathcal{B}} = e^{s(\mathcal{A}+\mathcal{B})+\frac{s^2}{2}[\mathcal{A},\mathcal{B}]+\mathcal{O}(s^3)}, \qquad e^{\frac{s}{2}\mathcal{A}}e^{s\mathcal{B}}e^{\frac{s}{2}\mathcal{A}} = e^{s(\mathcal{A}+\mathcal{B})+\mathcal{O}(s^3)},$$

where $[\mathcal{A},\mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}$. In particular, the symmetric (Strang) composition on the right is accurate to one order higher than the naive (Lie) composition on the left, because the leading commutator term cancels by symmetry.

Remark A.8.2. *Used in Section 4.4 (Chapter 4) to justify the local error $\mathcal{O}(\Delta t^3)$ – hence global error $\mathcal{O}(\Delta t^2)$ – of the Strang splitting scheme for the Vlasov–Poisson system, as compared to the $\mathcal{O}(\Delta t^2)$ local error of the (unsymmetrized) Lie splitting. We refer to Hairer, Lubich and Wanner [18] for a complete treatment of the Baker–Campbell–Hausdorff formula and its use in the error analysis of splitting methods.*