Chapter 3

Accounting scales of motion

Fluids are fascinating because they contain a lot of disparate actions at the same time, like the potentialities of mind. Actions of different sizes are especially intriguing, like vigorous little scrambles of transient cloud structure (a dragon! no, a peacock!) caught up in whole-sky-sized swirls, all drifting by on continent or planet sized currents. Profound questions of the predictability and persistence of meaning are here to explore, but we need bookkeeping tools. In chapter 1 we budgeted stuff in space. Now we wish to budget action in the abstract domain of *size*, whose vastness requires us to use logarithmic *scale*. The crudeness of available tools will disappoint, pulling us into the entity descriptions of Part II, but these accounting fundamentals offer lessons worth their study.

3.1 One size cut: molecular vs. macroscopic

Borders are crossed by individual careening molecules, and also by great masses of them moving together. These cases should be distinguished, if only because the continuum limit of Calculus only applies easily to the latter. Randomly hurtling molecules carry specific "stuff" (like their momentum and energy) from conditions at their place of 'origin' (last contact with other matter) to the place of their next collision, where that stuff is shared out into a new *neighborhood*, a term by *spatial averaging*.

Pure logic (not detailed here; think) tells us that the net effect of numerous molecules sharing stuff both ways across a border in this way is a *flux down the gradient, proportional to the strength of the gradient,* which is called a *diffusive* flux. Treating that proportionality coefficient as constant, *diffusion* (the convergence of diffusive flux) is proportional to the *Laplacian operator*, $\boldsymbol{\nabla}^2 = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}$, which we also met in the pressure equation (problem 1.5.2 and its solution). This separation of molecular flux (viscosity) from fluid flux is enshrined in the fundamental Navier-Stokes equations of fluid dynamics, with $\nu\boldsymbol{\nabla}^2 u,\ \nu\boldsymbol{\nabla}^2 v,\ \nu\boldsymbol{\nabla}^2 w$ added to the RHS of the b,c,d equations in sets (2.10) and (2.14).

The Laplacian operator (double derivatives) measures *curvature* of a line or surface, so we can use our special brainfeel for the word "curvaceous" to appreciate the nature of the 3D Laplacian. Inverting the Laplacian requires constants of integration (boundary conditions), because a curved bowl can be held at different angles and different heights above the floor. We will explore Laplacian inversion for the pressure field (from problem 1.4.3) in particular geometries in Chapter 4.

3.2 Another cut: large-scale flow vs. small "eddies"

Do fluid blobs *smaller than the flow we care about* (the "eddies" in whatever stream we are pondering) act like molecules? That is, as they cross a border we care about (both ways, because of mass continuity), can their effect be represented as a *diffusive flux* of specific stuff? If so, hooray! We simply reuse the argument above, enhancing molecular viscosity ν with an *eddy diffusivity,* in this case *eddy viscosity* K when momentum is the stuff being transported. Eddy flux is orders of magnitude larger than molecular diffusion for Earthly scales, so ν would just be replaced with a K. But then, the Laplacian's curvaceousness feeler must have a broader reach, spanning the size of those flow features we care about.

Specifically, the words *mean conditions* and *neighborhood* from the section above can be redefined to denote a broader averaging operation (a *spatial smoothing filter* denoted by overbars $\bar{u}, \bar{v}, \bar{w}$) that just preserves (with acceptable blurring) the features we care about. We then (following custom) write eddy quantities, representing all the smaller scales removed by the filter with a prime[a]:

$$u' \equiv u - \bar{u} \qquad (3.1)$$

Such a definition for all variables can be substituted into our equations and the eddy terms moved to the right, perhaps to be treated as diffusion as just discussed. For u, Eq. (2.14b) can be elaborated in this way as

$$\begin{aligned}\bar{u}_t = -\bar{u}\,\overline{u_x} - \bar{v}\,\overline{u_y} - \bar{w}\,\overline{u_z} - \overline{\pi_x} + \overline{fv} \\ +\overline{(u'u')_x} \; + \overline{(v'u')_y} \; + \overline{(w'u')_z} \qquad (3.2)\end{aligned}$$

by transforming to flux form, averaging, and then using the continuity equation for the large-scale flow (since that powerful Law applies to all scales) to transform back to the *apparent advection* form on the LHS (problem 3.10.1)[b]. What has been gained from this exercise? Nothing yet, until terms on the RHS are treated somehow, but we have made a rigorous scale cut. It's a bookkeeping technique, and the messy eddy flux terms are the notational cost of just one size-cut to isolate the larger scales. Only an intrepid few feel inspired to write the bookkeping for *higher order* expansions of RHS terms (e.g. André et al. 1976), or make another size cut with another averaging operation (e.g. Tan et al. 2018).

For those interested in planetary-scale flows, a typical horizontal filter scale is Earth's size divided by about 1000 (for our visual screens) or 100 (fitting our brain's social hardware[c]) or 10 (for counting on our fingers). Such a filter means that all of the white parts of several or many convective clouds are implied in the definition of the "eddies", from which we want to know only a *net effect* in fluxing stuff across scales we

---

[a] The prime was also used in chapter 2 where we only *cared about* the whole-fluid horizontal mean denoted by the special averaging operators $\bar{\bar{\rho}}$ and [KE] in problem 2.4.1.

[b] Averaged prime terms like $\overline{\bar{u}u'}$ vanish only if the averaging is done over *fixed limits of integration*, not for a sliding boxcar or convolution smoother that leaves $\bar{u}$ a continuous function. For this reason, x,y,z differentiation subscripts are kept under the bar.

[c] https://en.wikipedia.org/wiki/Dunbars_number.

care about. This is the realm of *convection parameterization* (confronted in Part III), whose engineering-like challenges also serve as a genuinely scientific test of whether our appreciation of convection is *usefully* faithful to nature (Arakawa 2004). Although computers can now track much finer spatial distinctions, appreciation happens in human brains: high-resolution data is as much a burden as a blessing without a screen-viewable region of interest to focus on, or a right-sized rigorous framing like (3.2) to average it into.

3.3 On anomalies, deviations, perturbations, eddies, etc.

Primes and overbars like (3.2) are sprinkled all over the fluids literature, with various meanings. Sometimes different notations are helpfully invoked: In the atmospheric general circulation, square brackets and asterisks $[v^*u^*]$ represent *zonal* (east-west around the globe) spatial averages and *zonal eddy* deviations therefrom, because overbars and primes are used for *time* averages and *anomalies* or *fluctuations.* Fortunately all these words were chosen carefully, so let's keep science clearer by using them that way. Sometimes primes are called *perturbations,* which makes the most sense if someone or something perturbs something else, like in an experiment minus a control. The Reynolds interpretation of fluid scale interactions (where bars and primes are an *ensemble* mean plus individual *member* deviations) is increasingly practical, as computation's power is used to explicitly compute realizations and tally them. All of these are measures of *deviations* from some baseline, a very agnostic term for a simple mathematical difference. In the practical realm of *time series data analysis*, primes and bars are sometimes substituted for space or for realizations, with caveats like *stationarity* and its rabbit-hole cousin *ergodicity*. Suffice it to say that confusion is possible, so readers should allocate mental capacity for conceptual complications and vocabulary care at least as great as the messiness of the notation forces upon us.

3.4 Fourier decomposition and logarithmic *scale*

To make more than one or two awkward size distinctions, we can leap to another kind of size-accounting, using sines and cosines (*Fourier harmonics*). These are solutions to wave equations like (2.13), where curvature is negatively proportional to displacement, so they fit best for actual waves, but any continuous function can be expressed in terms of this complete, orthogonal (or mutually exclusive) *basis set*, a particular case of *Galerkin decomposition*.

The almost household physical-realm words *wavelength* (and *period* in time) need to be rethought as spectral-realm words for 2π times their inverses, *wavenumber* (and *frequency*). You should also know the careful wave-word *amplitude*. The abstractness of wavenumber is reduced somewhat by accounting it in units of *number of cycles over a given distance* (or frequency *over a given time interval*). This invocation of our brain's counting faculty is useful, because when we partition conserved stuff like energy (or variance, squared amplitude) over a discrete set of bins in wavenumber or frequency space[d], the amount of that stuff falling in each bin of this inverse-size domain depends on the widths (that is, the spacing) of the bins.

Larger-sized (long-wavelength or long-period) energy is emphasized by the inverse implied in wavenumber or frequency counting. For instance, transforming the fluctuations in a 4km discrete spatial grid boxes around the Earth's equator (10,000 points, recalling section 1.1), the energy or variance (squared amplitude) "stuff" falling in the wavelength range from 4000-40,000 km is divided into just 10 wavenumber bins, while energy in the wavelength range 40-400 km is divided over 1000 bins. Since our goal in scale separations is often to focus on larger scales, this tool is a useful lens, but its inherent emphasis on redness[e] should be noted. The Fourier Transform and its inverse are linear and reversible, allowing us to quickly jump back and forth between the *physical domain* (space and

---

[d] We can do this meaningfully: https://en.wikipedia.org/wiki/Parsevals_theorem

[e] Named based on light's spectrum: longer wavelengths are what we call red.

time) and *spectral domain* (wavenumber and frequency) without any loss of information, all with a few pencil strokes or keystrokes. This is an essential power every science student should learn to appreciate and apply, if not fully understand to its depths[f]. Computer exercises in section 3.10 illustrate some fun puzzles of Fourier data analysis.

To tame the vastness of the size (and inverse-size or wavenumber) domain, use of the log function suggests itself. In fact, logarithms are fundamentally what we mean by the word *scale* as opposed to *size*. We speak of *meter-scale eddies* whose size is 1-10m but not 100m, while *kilometer-scale* clouds are 1-10km across but not 100km. Scale is measured in *decades* (factors of 10, based on our fingers), or sometimes more finely in *octaves* (factors of 2, but named for the mysteriously pleasing brainfeel of tones 1/8 of an octave apart in the ancient art of music on stringed instruments). The physics of ears and neurology are involved, specialized head hardware that we should leverage but not project too carelessly onto nature.

Any spatial or temporal series can be *decomposed into* a sum of Fourier components. But does this mean that nature is truly, deeply, secretly *composed of* those components? The *spectr-* in spectral analysis means *ghostly*, for good reason: a superstitious age was spooked to see white light become rainbow-colored on passage through a prism. Consider a short wave (with large wavenumber K), $\psi = A\,sin(Kt)$, whose amplitude A is *modulated by* a long wave with small wavenmber k: $A = C\,sin(kt)$ so that $\psi = C\,sin(kt)\,sin(Kt)$. This modulation process might be quite physical and real, but Fourier analysis decomposes it into an *interference pattern* or *beating pattern* between two high wavenumbers near K, $\psi = C/2\,[cos(kt - Kt) - cos(kt + Kt)]$, with no long-wavelength energy involved at all. Which is the truer picture: the large-scale and small-scale interaction of modulation, or mere additive superposition (not an inter-action at all in the terms of chapter 7)? AM

[f] The world-changing Fast Fourier Transform (FFT) is a masterpiece of mid-20th century matrix factorization, letting NlogN computations give the answer to an N^2 sized question, where N is the number of points in a space or time series resolving N/2 frequencies.

and FM radio show us that both are *useful* descriptions. Bookkeeping systems must not be mistaken for natural truths.

3.5 A downscale energy transfer: shear instability

Whatever the interpretation of Fourier spectra, Parseval's theorem does allow us to decompose energy into bins which are a monotonic (if non-uniformly spaced) measure of size in the form of *scale*, the negative log of its inverse (wavenumber). Consider *shear instability* (the roll-up of sheets or filaments of vorticity into ball-like vortices). This process is a transfer of kinetic energy from larger to smaller scales, with no trace in the global kinetic energy budget's source term the *buoyancy flux* [$w'b'$] (Problem 2.4.1), at least for horizontal roll-ups.

Shear instability of filaments and sheets of vorticity is the essence of turbulence. The *cascade* theory (chapter 7) envisions energy flux across scale as *local in the scale domain*; in other words, it postulates that only almost-big eddies can feed on a big eddy. This process is envisioned to end when cascaded energy reaches the *Kolmogorov* scale of molecular mean free paths where the kinetic energy of flow (groups of molecules acting coherently) gets scattered into the random motion of the individual molecules, raising the "temperature" by which we measure motions below our first size-cut, a process (or is it just a category redefinition?) called *dissipation.*

3.6 Upscale energy flux and convection-LS interaction

What if eddies carry momentum *up its mean gradient (up-shear)*, acting as a negative viscosity (Starr 1968)? From a kinetic energy perspective, such eddies pass energy from their own scale to the shear. In a bounded fluid, that local shear emebedding the eddy is part of some larger-scale flow feature, but that could be a much-larger (planetary-scale) jet stream, not simply a swirl one decade larger in scale. In other words, such up-shear momentum flux also need not be local in the scale domain. This situation is actually quite common, since shear's effect is to tilt eddies in such a way that the v'u' or w'u' terms in (3.2) act upshear. As long as there is some ongoing energy source for the eddies (like buoyant convection), the process has no exotic impossibility.

3.7 Spectral energetics and the cascade fallacy

The energy equations of problem 2.4.1 can be decomposed into orthogonal basis functions such as Fourier sinusoids, for instance in the vertical (the z domain). Analytic work in the pre-computer era showed by scale analysis that *if* a turbulent flow were driven at a large scale (like stirring by a huge spoon), and *if* the only dissipation were by molecular diffusion at a Kolmogorov microscale, and *if* the energy transfer in the intervening "inertial subrange" of scales were restricted to being local (only eddies of comparable size exchanging energy), *then* by purely dimensional analysis, a logarithmic graph of the kinetic energy spectrum must have a slope of -5/3. This paradigm is called a *cascade*[g] because of its local-in-scale energy transfer assumption.

Paradigms are sticky in the brain, and many people don't appreciate that a cascade is a very particular and distinctive type of waterfall. Is the long-standing local-in-scale assumption proven because a gross statistic (a single number) from observations happens to agrees with the mechanistically hazy prediction (from dimensional considerations alone) given by that paradigm? We have all seen tiny roll-ups on much-larger currents like a smoke plume. Although this is not an instance of fully-developed turbulence, it shows that locality is not a necessary condition for energy exchange across scales. Convection is an entirely different example: for instance, the kinetic energy source [wb] due to a local warm updraft is spectrally white, driving all scales of motion, yet a 2D cloud model in the x-z plane also exhibits a -5/3 slope (Mapes et al. 2008). Matching a low-order statistic (like a single number) is at best weak evidence that one specific but nonunique paradigm's core assumption, the whole basis for its name, is correct.

---

[g] A *cascading* waterfall is one with many short drops from one pool to the next.

3.8 Multiscale information, DOFs, and macro-entropy

There are ten times more Fourier wavenumber bins in the 10s of km scale range than the 100s of km scale range. Each is a *degree of freedom* (DOF), a container for *information* (the drawing of distinctions) and for *energy*. Such a limitlessly growing multiplicity of DOFs with smallness of scale is a big problem for sensible thinking about macroscopic phenomena.

In early 20th century physics, a prediction of infinite energy in the smallest scales (the "UV catastrophe", Wikipedia) was solved only by invoking an irreducible *quantum*, a low but finite floor to the bottomless pit of spectral microstructure. This idea led to Einstein's 1921 Nobel prize, and gave physical reality to classical philosophy's postulate of Atomism. Still, above that lowest floor that props up "physicalism" (Wikipedia), there remains a crisis in multiscale (micro to macro) reasoning. A strict reductionism insists that the most fundamental causality is at the smallest scale, with all larger-scales being mere "epiphenomena" which are "supervenient" (Wikipedia) on the smallest scale. This view leaves no meaningful role for macroscopic identity or agency, from convective cells to biological cells, neurons to behavior. While rather abstract, supervenience is a philosophical abyss under discourse in all the macro-sciences, privileging the hardest of hard sciences (physics) which hands off its causality to nihilistic quantum indeterminacy. There is no bottom, it is turtles all the way down! This is a merely crisis of description: nature operates sensibly at macroscales.

Information theory may offer a mathematical paradigm (and toolkit) for building firmer catwalks over this abyss, and then platforms for causally satisfying treatments for phenomena of multiple macroscales (like convection). This revolution is led by neuroscience, where biology's complex structures are obviously important, yet information is also a basic currency of discourse (Hoel 2018). Others bring information-entropy ideas to still larger systems like ecological succession (Annila 2019), which might seem a good analogue for convection developing and organizing over land on a summer day.

But biology may strain the analogy of informational entropy too far, because its outcomes (organismal structure) are too complicated to constrain via simple unlikeliness arguments (measured and expressed in information terms). The long memory of inheritance makes biomorphology evolve slowly, with low maintenance cost for complexity once evolved. Nature is a scrap-heap of strategies that may be geared to long-dead ancestors thriving in long-past environments. Biological evolution is also subject to an interdependent web of selection pressures that may favor aspects other than slight gains by 'unlikely' complex structure development in the efficiency of exploitation of a simple univariate energy resource. While convergent evolution on different continents may suggest some universality, the nonuniqueness of large life is even clearer (humans vs. dinosaurs for instance).

Meanwhile *thermodynamic* entropy, a drift of unspecified speed toward equilibrium ($dS \geq 0$ is all the Second Law offers), to maximize the redundancies of zillions of identical molecules, is a rather dull use for information theory. It was overkill as our measure of "heat stuff" (section 2.1), for which static energy is clearer. Might the organization of turbulent convection be the Goldilocks ("just right") application of entropy (in information and probability terms) as a guiding principle? Entropy production optimization principles have been explored before in atmospheric dynamics (e.g. Paltridge 1978), but always with the objective of *microscopic or thermodynamic* entropy. Those works have a fringe feel, and have not unlocked compelling predictive or even explanatory power as great as their initial intellectual promise. Might progress be made by setting a smallest-macroscopic-scale floor (the Kolmogorov scale or mean free path of molecules, about 1 cm) under the enumeration of flow possibilities in P(*situation*), to be quantified as information and optimized against energy efficiency in some type of Free Energy Principle (Wikipedia)? This might prevent macrostructure information from being diluted beyond reach in the essentially infinite multiplicity of molecular situations within each such cubic cm of air.

We will return to these ideas in Part III, but it is useful to set this marker now: perhaps information theory may offer a relevant new facet to the foundations of convection science, alongside the accounting of physical quantities like momentum and energy. At the very least, it can offer a well-defined currency in which to state interesting physical postulates about macroscopic entities and structures in convective cloud fields. But before framing those big questions, we need to *define* some entities, and consider their basic mechanisms of both efficacy at doing their teleological job (gravity's work) and interacting with each other both competitively and cooperatively.

3.9 Problems and solutions

3.9.1 *Scale separation (large scale vs. eddy)*

Derive (3.2) from (3.1), noting the footnote there.

3.9.2 *Spectra of spatial data (your photograph)*

View the Jupyter notebook Spectra.ipynb from this Github page: (*URL*). Using Jupyter-Python (easily installed as explained at *Unidata's python page URL*), operate the notebook to replicate its figures. Replace the photograph with your own, and adjust code there to explore how Fourier analysis in 1 or 2 horizontal dimension works to decompose an energy-like quantity (*variance of brightness* in an image domain).

3.9.3 *Modulation vs. beating*

a. In a Jupyter notebook or other coding and exposition environment, illustrate the modulation vs. beating (interference) interpretations at the end of section 3.4.

b. Read about and explain in your own words the difference between AM and FM radio.

3.9.4 *Multiscale solutions to fluid equations*

Show and explain the truth of this statement from Palmer (2019): "Although the Navier–Stokes equations cannot be solved directly, they have certain symmetry properties ... One of these is a scaling symmetry: if u(x,t) is the velocity field and p(x,t) is the pressure field associated with a solution to the Navier–Stokes equations, then so are:

$$u_\tau(x,t) = \tau^{-\frac{1}{2}} u\left(\frac{x}{\tau^{1/2}}, \frac{t}{\tau}\right) \tag{1}$$

and

$$p_\tau(x,t) = \tau^{-1} p\left(\frac{x}{\tau^{1/2}}, \frac{t}{\tau}\right) \tag{2}$$

See also section 2.2 of Lovejoy and Schertzer 2014.

Palmer 2019:
https://www.nature.com/articles/s42254-019-0062-2.epdf?shared_access_token=NTUP2o3zPRmrbYc3n_ct4tRgN0jAjWel9jnR3ZoTv0OYY1utT_3Qf75BoayYXiqirMhLELF4e50ASuCZROGu80BLY6qOdXk0Gy77vjtfrw9iEyH7M8ITOQIeCyejUHgxXL2H9cSURpdMQt55uX2lMA%3D%3D