Balanced Dynamics in Rotating, Stratified Flows: An IPAM Tutorial Lecture

James C. McWilliams

Atmospheric and Oceanic Sciences, UCLA

tutorial (n) 1. schooling on a topic of common knowledge. 2. opinion offered with incomplete proof or illustration.

These are notes for a lecture intended to introduce a primarily mathematical audience to the idea of balanced dynamics in geophysical fluids, specifically in relation to climate modeling. A more formal presentation of some of this material is in McWilliams (2003).

1 A Fluid Dynamical Hierarchy

The vast majority of the kinetic and available potential energy in the general circulation of the ocean and atmosphere, hence in climate, is constrained by approximately "balanced" forces, *i.e.,* geostrophy in the horizontal plane (Coriolis force *vs.* pressure gradient) and hydrostacy in the vertical direction (gravitational buoyancy force *vs.* pressure gradient), with the various external and phase-change forces, molecular viscosity and diffusion, and parcel accelerations acting as small residuals in the momentum balance. This lecture is about the theoretical framework associated with this type of diagnostic force balance.

Parametric measures of dynamical influence of rotation and stratification are the Rossby and Froude numbers,

$$Ro \; = \; \frac{V}{fL} \quad \text{and} \quad Fr \; = \; \frac{V}{NH} \, . \tag{1}$$

V is a characteristic horizontal velocity magnitude, $f = 2\Omega$ is Coriolis frequency (Ω is Earth's rotation rate), N is buoyancy frequency for the stable density stratification, and $(L, \; H)$ are (horizontal, vertical) length scales. For flows on the planetary scale and mesoscale, Ro and Fr are typically not large. Furthermore, since for these flows the atmosphere and ocean are relatively thin, the aspect ratio,

$$\lambda \; = \; \frac{H}{L} \, , \tag{2}$$

is small.

Consider a hierarchical sequence of fluid dynamical PDE systems for rotating, density stratified flows, motivated by planetary atmospheres and oceans (Fig. 1): compressible, anelastic, Boussinesq (incompressible), Primitive (hydrostatic), gradient-wind balanced, and Quasigeostrophic. They are ordered in an increasing degree of physical simplification and approximately ordered in their relevance to flows with increasing spatial and temporal scales.

This hierarchy involves a succession of physically justifiable approximations that progressively limit the range of solution behaviors and increase the equation solvability, hence theoretician's utility. (*N.b.,* it's all turbulence and their are no general solutions, not even proofs of regularity, existence, uniqueness). There is also advantage in computational efficiency to advancing in this hierarchy, although the present practice of realistic climate simulations has settled on a mid-sequence model, the Primitive Equations, to be relatively safe in the dynamical approximations.

A HIERARCHY OF FLUID SYSTEMS WITH (f, N)

molecules

compressible

acoustic waves

anelastic

Boussinesq (incompressible)

Primitive (hydrostatic)

inertia−gravity waves

gradient−wind balanced

quasigeostrophic

planets

From a mathematical perspective these steps transform prognostic equations of the type,

$$\frac{dC}{dt} = F(C). \tag{3}$$

into a diagnostic "balanced" approximation:

$$F(C) = 0. \tag{4}$$

This implies a reduction in the available modes of the system when linear analyses are made. For example, consider simple thermodynamics with some state variable a such that density $\rho = \rho[a, p]$, then the compressible system at the top of this hierarchy has 3 momentum equations for $\mathbf{u} = (u, v, w)$, an a equation, and a mass continuity equation, all in the form of (3) above; *i.e.*, it is a 5th order PDE system in time. In contrast, the Quasigeostrophic (QG) system at the bottom of the hierarchy is 1st order in time.

In this hierarchy the PDE types become progressively less hyperbolic and more elliptic, with fewer propagating wave modes and more "balance" relations. For example, in the first step from compressible to anelastic equations, acoustic waves are excluded from the solution space, and certain pressure forces have instantaneous action at a distance rather than evolve into this spatial distribution through acoustic wave propagation. Sound speeds are typically very rapid for most fluid systems (*i.e.*, the Mach number is not large), so the mass-continuity balance approximation excises the fastest wave mode and the largest time derivative in the PDE system. This allows a computational time step to be much larger.

A narrower definition of balance refers to the momentum equations, a.k.a. diagnostic force balance. Thus, the hydrostatic balance approximation drops $\partial_t w$ in the vertical momentum equation — where $\hat{\mathbf{z}}$ is parallel to gravity and the rotation vector — in favor of

$$\partial_z p = -g\rho. \tag{5}$$

Analogously, the geostrophic balance approximation in $\hat{\mathbf{x}}_\perp$ drops (or at least demotes in a multi-scale asymptotic sense) $\partial_t u$ and $\partial_t v$ in favor of

$$fv = \frac{1}{\rho_0}\partial_x p, \qquad fu = -\frac{1}{\rho_0}\partial_y p. \tag{6}$$

In this hierarchy the two most consequential steps in solution behavior are to go beyond compressibility, excising "very fast" acoustic waves, and beyond the Primitive Equations, excising "fast" inertia-gravity waves (IGW) while retaining advective and Rossby-wave modes. The ocean and atmosphere are full of acoustic and inertia-gravity waves. They are mostly generated by fast events like a falling tree, a convective thunderstorm, or the tide, and it has been very difficult to establish that they have an important energy exchange with the general circulations, hence that they need to be a central component of a climate model.

2 Two Realms: Balanced and Unbalanced Flows

The Boussinesq equations are

$$\frac{D\mathbf{u}}{Dt} + f\hat{\mathbf{z}} \times \mathbf{u} = -\frac{1}{\rho_0}\boldsymbol{\nabla}p - g\rho\hat{\mathbf{z}}$$

$$\frac{D\rho}{Dt} = 0$$
$$\boldsymbol{\nabla} \cdot \mathbf{u} = 0. \tag{7}$$

Gravity and rotation axis are aligned with $\hat{\mathbf{z}}$. The equation of state is very simple, $\rho = \rho[a]$. Inviscid and adiabatic assumptions are also made for simplicity, although of course diffusion is needed in solutions in which energy and density variance evolve toward small scales. The Coriolis frequency is locally approximated as $f = f_0 + \beta y$. The advective time derivative is

$$\frac{D}{Dt} = \partial_t + \mathbf{u} \cdot \boldsymbol{\nabla}. \tag{8}$$

With this PDE system we follow an asymptotic path toward Quasigeostrophic theory. First, define a mean stratification and its deviation density field:

$$\rho = \overline{\rho}(z,t) + \rho'(\mathbf{x},t). \tag{9}$$

Next, specify geostrophic, hydrostatic, advective scaling estimates and non-dimensionalization for the model variables: given characteristic values for ρ_0, V, H, L, f_0, β, N_0, then

$$x, y \sim L, \quad z \sim H, \quad t \sim L/V,$$
$$u, v \sim V, \quad w \sim Ro\, VH/L,$$
$$p \sim \rho_0 f_0 VL, \quad \rho' \sim \rho_0 f_0 VL/gH, \quad \overline{\rho} \sim \rho_0 N_0^2 H/g. \tag{10}$$

Assume a small parameter:

$$Ro, \ Fr, \ \beta/Lf_0 = \epsilon \to 0. \tag{11}$$

With geostrophic scaling the non-dimensional Boussinesq equations are

$$\epsilon \frac{Du}{Dt} + (1 + \epsilon y)v = -\partial_x p$$
$$\epsilon \frac{Dv}{Dt} - (1 + \epsilon y)u = -\partial_y p$$
$$\epsilon^2 \lambda^2 \frac{Dw}{Dt} = -\partial_z p - \rho'$$
$$\partial_t \overline{\rho} + \epsilon \left(\frac{D\rho'}{Dt} + w\partial_z \overline{\rho} \right) = 0$$
$$\boldsymbol{\nabla}_\perp \cdot \mathbf{u}_\perp + \epsilon \partial_z w = 0, \tag{12}$$

with

$$\frac{D}{Dt} = \partial_t + \mathbf{u}_\perp \cdot \boldsymbol{\nabla}_\perp + \epsilon w \partial_z. \tag{13}$$

Leading order in ϵ balances are geostrophic, hydrostatic, and with horizontally non-divergent flow, which comprises an under-determined system for (u, v, p, ρ'), with constant mean stratification, $\overline{\rho}(z)$. To close leading the leading order dynamics, eliminate w at $\mathcal{O}(\epsilon)$ between the curl of horizontal momentum, density, and continuity equations and obtain the QG potential vorticity equation, which has the single time derivative in the QG PDE system for (u, v, p, ρ'):

$$\partial_t q + J_\perp[p, q] = 0, \qquad q = \boldsymbol{\nabla}_\perp^2 p + \partial_z \left(\frac{\partial_z p}{\partial_z \overline{\rho}} \right) + y. \tag{14}$$

4

$J_\perp[a, b] = \partial_x a \partial_y b - \partial_y a \partial_x b$ is the horizontal Jacobian operator. This system has solutions with Rossby waves, baroclinic and barotropic instabilities, and geostrophic turbulence with inverse energy cascade and forward enstrophy (q variance) cascade (Charney, 1971). It is the essential theoretical basis for initialization of weather forecasts and for the dynamical interpretation of the general circulation.

An alternative model-derivation path toward inertia-gravity wave (IGW) dynamics can also be made with small ϵ. Non-dimensionalize the Boussinesq equations as above except assume with $t \sim 1/f_0$ and $w \sim VH/L$ and with $\lambda = 1$ (WLOG):

$$\begin{aligned}
\partial_t u + \epsilon(\mathbf{u} \cdot \boldsymbol{\nabla})u + (1 + \epsilon y)v &= -\partial_x p \\
\partial_t v + \epsilon(\mathbf{u} \cdot \boldsymbol{\nabla})v - (1 + \epsilon y)u &= -\partial_y p \\
\partial_t w + \epsilon(\mathbf{u} \cdot \boldsymbol{\nabla})w &= -\partial_z p - \rho' \\
\partial_t \overline{\rho} + \epsilon\left(\partial_t \rho' + \epsilon(\mathbf{u} \cdot \boldsymbol{\nabla})\rho' + w\partial_z\overline{\rho}\right) &= 0 \\
\boldsymbol{\nabla}_\perp \cdot \mathbf{u}_\perp + \partial_z w &= 0 \,.
\end{aligned} \tag{15}$$

The $\epsilon \to 0$ asymptotic system again has $\overline{\rho}(z)$ and linear inertia-gravity wave solutions,

$$(u, v, w, p, \rho') \propto exp[i(kx + \ell y + mz - \omega t)], \qquad \omega = \pm\sqrt{\frac{\partial_z\overline{\rho}(k^2 + \ell^2) + m^2}{k^2 + \ell^2 + m^2}}, \tag{16}$$

in the WKB limit. We see that the lowest possible frequency for an IGW is one, *i.e., f* dimensionally. Since the dimensional evolutionary rate for balanced flows is ϵf, we call the former "fast" and the latter "slow". At finite ϵ nonlinear evolution also occurs for inertia-gravity waves, most dramatically when wave breaking occurs. The cascade of IGW energy is forward to small-scale dissipation.

Smaller scale flows characteristically have large ϵ values, since V does not decrease as rapidly as L and H do. Thus, their dynamics lies largely outside the realms of balanced flows and inertia-gravity waves. In a broad sense this is the realm of Kolmogorov's isotropic, homogeneous turbulence, with forward energy cascade and efficient material mixing. These flows complete the route to dissipation and material mixing energized from larger scales flows with small ϵ.

3   Accuracy of Balance Models

Advective evolutionary rates are slower than IGW rates whenever $Ro, Fr < 1$. The concept of the slow manifold is the subset of all possible solutions to the fundamental fluid equations that evolve only on the slow rates of advection, $V/L \sim \epsilon f$, or potential-vorticity differences (*e.g.,* Rossby waves with $\omega \sim \beta L \sim \epsilon f$), thereby excluding all the faster acoustic wave and IGW solution behaviors. Balanced dynamics denotes the processes controlling the evolution on the slow manifold, and Balance Equations (BE) are a PDE system that manifests balanced dynamics only.

Obviously, QG is a BE, and at a conceptual level it is a very important model because it manifests the paradigmatic properties of the BE family. It is also computationally easy to solve. However, the asymptotic accuracy of QG is only $\mathcal{O}(\epsilon^0)$, and it turns out that in many situations with

small but finite ϵ, QG solutions are not particularly accurate, *e.g.,* compared to Primitive or Boussinesq solutions with adequately balanced initial conditions and forcing. Thus, a necessary condition for BE accuracy is higher-order asymptotic validity. The design of alternative higher-order BE systems has been a very popular activity, with, *e.g.,* notable early contributions from many able scientists (*e.g.,* Lorenz, 1960; Charney, 1962). It turns out that higher-order asymptotic accuracy is a necessary but not sufficient for good solution behavior, and that several differently defined BE systems show good behavior, *i.e.,* their design is not unique. I will skip over the physical and mathematical aesthetics of alternative good balance models.

A common property for the good BE models is a generalization of geostrophic balance for what is called gradient-wind balance (also called nonlinear balance). For a simple axisymmetric flow with azimuthal velocity V(r,z), gradient-wind balance includes a centrifugal force correction to the Coriolis force in the radial momentum equation, *i.e.,*

$$\frac{1}{\rho_0} \frac{\partial p}{\partial r} = fV + \frac{V^2}{r}. \tag{17}$$

More generally, if we make Helmholtz decomposition of the horizontal velocity in an incompressible flow,

$$\mathbf{u}_\perp = \hat{\mathbf{z}} \times \boldsymbol{\nabla}_\perp \psi + \boldsymbol{\nabla}_\perp \chi, \tag{18}$$

then a good BE must encompass the following approximation to the divergence of the horizontal momentum equations:

$$\begin{aligned}
\frac{1}{\rho_0} \nabla_\perp^2 p &= \hat{\mathbf{z}} \cdot \boldsymbol{\nabla}_\perp \times f\mathbf{u}_\perp + 2J_\perp[u,\, v] \\
&\approx \boldsymbol{\nabla}_\perp \cdot (f\boldsymbol{\nabla}_\perp \psi) + 2J_\perp[\partial_x \psi,\, \partial_y \psi]\,.
\end{aligned} \tag{19}$$

ψ is the streamfunction associated with vertical vorticity, and χ is the divergence potential associated with vertical velocity. Balance models typically are based on $\chi \ll \psi$ (*cf.,* the small w scale in (10)). At one higher level of differentiation than (17) the same terms can be identified: pressure-gradient, Coriolis, and a centrifugal effect along curved parcel trajectories. In many examples BE solutions have been shown to be highly accurate even beyond what might be expected from their provable $\mathcal{O}(\epsilon^1)$ asymptotic accuracy. Within the sampling estimation errors in measurements and obfuscation by regularization choices in general circulation models (*e.g.,* eddy viscosity), the true level of inaccuracy of balance dynamics has not been well determined.

Some form of BE is necessary for accurate weather forecasts. A widely-used operational means of accomplishing this is called nonlinear-normal mode initialization (NNMI). Its idea is that the linear normal modes of a model for a stratified resting state provide a complete basis set for a general fluid model can be decomposed into two groups, slow (or Rossby) modes $\mathcal{R}$ and fast (or IGW) modes $\mathcal{G}$. Under the hypothesis that weather evolution occurs on the slow manifold, or at least very close to it, then initial data should be projected onto $\mathcal{R}$ directly and onto $\mathcal{G}$ with some form of balanced constraint, $F]\mathcal{G}, \mathcal{R}] = 0$ (with F a nonlinear differential functional), rather than being allowed to evolve freely as IGW (Daley, 1991). The latter relation is sometimes referred to as "slaving" $\mathcal{G}$ to $\mathcal{R}$. This projection operation is equivalent to defining a dynamically consistent initial condition for a BE system, but then using it for integration of a more general fluid model (the Primitive equations for most weather forecasting). In QG theory the slaving relation is equivalent to the "omega equation" that determines w (hence χ) from the primary fields.

4 Limits to Balance

Finally, in a necessarily cursory way in a tutorial lecture, I briefly mention some of the modern ideas about balanced dynamics. The story is still unfolding ...

- No satisfactory asymptotic theory as a regular expansion in powers of ϵ has ever been established for the coupling (*e.g.,* energy exchange) between $\mathcal{R}$ and $\mathcal{G}$ modes. This supports an approximate view that their dynamics are superimposable: coexistence with mutual independence. However, this is not a defensibly exact view for the reasons described in the rest of this section.

- Is the slow manifold an inertial manifold for finite ϵ? *I.e.,* does an initial state on the putative manifold stay there as it it evolves? An early indication that this might not be true is that operational NNMI procedures were implemented by iterative solution of their defining nonlinear PDE systems, and, while several iterations often sufficed to make the subsequent evolution seem mostly slow, it proved generally impossible to achieve a high degree of iteration convergence.

- Speaking loosely, the slow manifold can be called a fuzzy manifold (Warn, 1996). The idea is that the solution phase space for more fundamental models has regions where the evolution is mostly slow and consistent with good BE systems, but the solution is not exactly slow and balanced. In an ODE system motivated by fluid PDEs, Vanneste and Yavneh (2004) show that the inescapable fast component of the solution as a relative amplitude $\propto \epsilon^{-1/2} exp[-\alpha/\epsilon]$ for $\epsilon \ll 1$, with α a positive constant. Here the fuzz has a very small nap indeed, smaller than any power of ϵ.

- A non-asymptotic mathematical analysis of PDE type was performed for one accurate BE system (McWilliams *et al.*, 1998). It shows that the PDE operator that must be inverted for time integration loses its ellipticity if any of three conditions fail to be satisfied everywhere in the flow domain: stable density stratification ($N > 0$ locally), positive potential vorticity ($q > 0$, where q is Ertel potential vorticity, a generalization of the QG quantity in (14)), and positivity of $A-S$, where $A = f + \hat{\mathbf{z}} \cdot \boldsymbol{\nabla} \times \mathbf{u}$ is the absolute vertical vorticity and S is the strain rate for the horizontal shear. Each of these conditions can be violated only when $\epsilon = \mathcal{O}(1)$. The implication is that, if a good BE system cannot be integrated further in time, then some non-slow, unbalanced behavior must arise. These conditions suggest that the best places to look for balance breakdown events are weak stratification, anticyclonic vorticity, high strain, and small q.

- The energy requirement for achieving climate equilibrium is strongly suggestive of significant energy leakage from the slow manifold. Large-scale forces (*e.g.,* differential planetary rotation, surface wind stress) input energy to the slow manifold, and the resulting mean circulation is unstable primarily to balanced synoptic and mesoscale eddies within the slow manifold. If balanced turbulence leads primarily to inverse energy cascade, then how does its energy cycle connect with viscous dissipation occurring on very small scales, ~ 0.1-1 cm? While several dissipation routes are conceivable (*e.g.,* turbulent boundary layers), an appealing hypothesis is that finite-ϵ breakdown of balance occurs often enough to provide an important route for the general circulation. When and how?

- One means of realizing fuzziness and leakage from the slow manifold is through an ageostrophic or unbalanced instability of a balanced flow. Two of the three BE integrability conditions above are known as critical instability onset conditions: $N < 0$ leads to convection, and $q < 0$ leads to centrifugal instability. The third condition, $A - S < 0$, is not known to be associated with an instability onset, but several examples have been adduced where it seems relevant (*e.g.,* Taylor-Couette flow, where the growth rate has exponential smallness $\propto exp[-\alpha/\epsilon]$ for small ϵ; Yavneh *et al.*, 2001). Other agesotrophic instability examples are Kelvin-Helmholtz instability and Ripa's (1983) second criterion, both of which require finite values of Fr. Once an agoestrophic instability occurs it is plausible that it will evolve to finite-amplitude turbulent behavior and forward energy cascade.

- Another means is the super-exponentially rapid scale contraction that occurs for density fronts and filaments under the influence of the strain field in balanced flows (Hoskins, 1982; McWilliams *et al.*, 2009). Scale contraction implies that the local ϵ value increases with time. Ageostrophic "secondary" flows arise in the cross-frontal plane and accelerate the scale contraction. Furthermore, sharp fronts and filaments are potentially vulnerable to ageostrophic instability.

- Rotating and/or stably stratified flows do manifest a forward energy cascade that has dual inertial ranges in kinetic and available potential energy (Lindborg, 2006; Molemaker *et al.*, 2010). This provides a dynamical route to small-scale turbulence and dissipation when energy leaks off the slow manifold into this type of cascade.

- Balanced flows with finite ϵ sometimes spontaneously emit IGW, *e.g.,* at strong fronts or within a strong dipole vortex (Snyder *et al.*, 2009).

These fragmentary perspectives comprise the scientific frontier for balanced advective dynamics. Apart from turbulence itself, the nature of balance is perhaps the deepest issue in natural fluid dynamics. Its mysteries are manifested most strongly in flows with intermediate scales and intermediate ϵ values, and a full understanding still lies ahead of us.

References

Charney, J., 1962: Integration of the Primitive and Balance Equations. *Proc. Int. Symp. Numerical Weather Prediction*, Meteor. Soc. Japan, Tokyo, p131-152.

Charney, J., 1971: Geostrophic Turbulence. *J. Atmos. Sci.* **28**, 1087-1095.

Daley, R., 1991: *Atmospheric Data Analysis*, Cambridge University Press.

Hoskins, B.J. , 1982: The mathematical theory of frontogenesis. *Ann. Rev. Fluid Mech.* **1**, 131-151.

Lindborg, E., 2006: The energy cascade in a strongly stratified fluid. *J. Fluid Mech.* **550**, 207-242.

Lorenz, E., 1960: Energy and Numerical Weather Prediction. *Tellus* **12**, 364-373.

McWilliams, J.C., I. Yavneh, M.J.P. Cullen, & P.R. Gent, 1998: The breakdown of large-scale flows in rotating, stratified fluids. *Phys. Fluids* **10**, 3178-3184.

McWilliams, J.C., 2003: Diagnostic force balance and its limits. In: *Nonlinear Processes in Geophysical Fluid Dynamics*, O.U. Velasco Fuentes, J. Sheinbaum, J. Ochoa, eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, 287-304.

McWilliams, J.C., F. Colas, and M.J. Molemaker, 2009: Cold filamentary intensification and oceanic surface convergence lines. *Geophys. Res. Lett.* **36**, L18602.

Molemaker, M.J., J.C. McWilliams, and X. Capet, 2010: Balanced and unbalanced routes to dissipation in an equilibrated Eady flow. *J. Fluid Mech.*, in press.

Ripa, P., 1983: General Stability Conditions for Zonal Flows in a One-layer Model on the Beta-plane or the Sphere. *J. Fluid Mech.* **126**, 463-487.

Snyder, C., R. Plougonven, and D. Muraki, 2009: Mechanisms for spontaneous gravity wave generation within a dipole vortex. *J. Atmos. Sci.* **66**, 3464-3478.

Vanesste, J., and I. Yavneh, 2004: Exponentially small inertia-gravity waves and the breakdown of quasigeostrophic balance. *J. Atmos. Sci.* **61**, 211-223.

Warn, T., 1997: Nonlinear balance and quasigeostrophic sets. *Atmosphere-Ocean* **32**, 135-145.

Yavneh, I., J.C. McWilliams, and M.J. Molemaker, 2001: Non-axisymmetric instability of centrifugally stable, stratified Taylor-Couette flow. *J. Fluid. Mech.* **448**, 1-21.