A derivation of the QG equations

J.A. Biello

November 4, 2012

I pursued my own derivation of the QG equation because I found it hard to sort through Pedlosky - and I needed to do this in order to understand the matching to, say MEWTG. I begin with the hydrostatic primitive equations which have already been non-dimensionalized. In fact, I non-dimensionalize them as we would in the tropical IPESD scaling. (A lot of this falls into my philosophy about connecting tropics through midlatitudes by carefully matching the order of the vertical velocity in the two regions. But I digress)

$$\begin{aligned} \frac{D}{Dt}u - \frac{\sin(\epsilon y)\, v + p_x}{\epsilon} &= 0 \\ \frac{D}{Dt}v + \frac{\sin(\epsilon y)\, u + p_y}{\epsilon} &= 0 \\ \frac{D}{Dt}\theta + \frac{w}{\epsilon} &= S_\theta \\ p_z &= \theta \\ u_x + v_y + w_z &= 0 \end{aligned} \tag{1}$$

where horizontal length scales are measured in units of 500 km, vertical in 5 km, time in 1 day. Horizontal velocities are measured in units of 5 m/s, vertical in units of 5 cm/s. Again, this is a scaling appropriate to the tropics, in the midlatitudes we will see that the vertical velocity enters at higher order. The temperature is measured in units of 3 Kelvin (which is order ϵ times the vertical scale times the vertical lapse rate of temperature) so that the heating rate is measured in units of 3 Kelvin/day. Of course the pressure (which is really the Emden function) comes along for the ride, but the units are 250 m^2/s^2. In this scaling

$$\epsilon = \frac{1}{4\pi} = \frac{500 \text{ km}}{R_{\text{Earth}}} \tag{2}$$

and the advective derivative is the full 3-dimensional advection. Again, I emphasize that I used this structure because I wanted to match to the tropics. However, for QG we will be at latitudes where

$$\sin(\epsilon y) \sim O(\epsilon^0) \tag{3}$$

and thereby we define

$$f = \sin(\epsilon y) \tag{4}$$

which depends weakly on y in the midlatitudes

$$\frac{d}{dy}f = \epsilon f' \equiv \epsilon\beta(\epsilon y) \tag{5}$$

where β is everywhere order one (except near the poles).

So as to make all the terms have nice ordering, we will clear a power of ϵ from the equations and rewrite them in the classical way, linear operator on the left hand side, nonlinearities on the right

$$\begin{aligned}
-f\,v + p_x &= -\,\epsilon\frac{D}{Dt}u \\
f\,u + p_y &= -\,\epsilon\frac{D}{Dt}v \\
p_z - \theta &= 0 \\
u_x + v_y + w_z &= 0 \\
w &= -\,\epsilon\left[\frac{D}{Dt}\theta - S_\theta\right]
\end{aligned} \tag{6}$$

where I've also eliminated θ in terms of p_z using the hydrostatic constraint and re-ordered the equations (for a reason!). Since we will be successively approximating a linear operator, we should look at that linear operator

$$\mathbf{L}V \equiv \begin{bmatrix} 0 & -f & 0 & \partial_x & 0 \\ f & 0 & 0 & \partial_y & 0 \\ 0 & 0 & 0 & \partial_z & -1 \\ \partial_x & \partial_y & \partial_z & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ p \\ \theta \end{bmatrix}. \tag{7}$$

It is clear that $\mathbf{L}$ is skew-self adjoint with respect to the standard L^2 norm in $\mathbf{R^3}$. Therefore, any eigenvectors are also eigenvectors of the adjoint, $\mathbf{L}^\dagger$. Also, we can see that the equations in (6) can be written formally as

$$\mathbf{L}\,V = \epsilon N(V), \tag{8}$$

where $N(V)$ represents the nonlinear terms on the right; this is my favorite way to write things!

We solve this system of equations in (6) with a regular asymptotic expansion

$$\theta \rightarrow \theta^\epsilon = \theta_0 + \epsilon\theta_1 + ... \tag{9}$$

etc. for each of the variables.

$\underline{O(\epsilon^0)}$: Formally the equations are

$$\mathbf{L}V_0 = 0 \tag{10}$$

which simply means the classical geostrophic constraint

$$V_0 = \begin{bmatrix} u_0 \\ v_0 \\ w_0 \\ p_0 \\ \theta_0 \end{bmatrix} = \begin{bmatrix} -f^{-1}\partial_y \Pi \\ f^{-1}\partial_x \Pi \\ 0 \\ \Pi \\ \partial_z \Pi \end{bmatrix} \tag{11}$$

for any $\Pi(x, y, z, t)$. It is also crucial to recognize that the adjoint eigenfunction is

$$V^\dagger = \begin{bmatrix} -f^{-1}\partial_y \Phi \\ f^{-1}\partial_x \Phi \\ 0 \\ \Phi \\ \partial_z \Phi \end{bmatrix} \tag{12}$$

for any function $\Phi(x, y, z, t)$. (You definitely need some L^2 integrability here and some derivatives, but I'm not going to worry about that for now.)

An important point to realize here is that the divergence of the horizontal velocity is not exactly zero, but rather $O(\epsilon)$

$$\partial_x u_0 + \partial_y v_0 = -\epsilon \frac{\beta}{f} \frac{\partial_y \Pi}{f} = -\epsilon \frac{\beta}{f} v_0. \tag{13}$$

$\underline{O(\epsilon^1)}$: At this order we will see how the time evolution arises. We have

$$\begin{aligned} -f\, v_1 + \partial_x p_1 &= -\frac{\bar{D}}{\bar{D}t} u_0 \\ f\, u_1 + \partial_y p_1 &= -\frac{\bar{D}}{\bar{D}t} v_0 \\ \partial_z p_1 - \theta_1 &= 0 \\ \partial_x u_1 + \partial_y v_1 + \partial_z w_1 &= \frac{\beta}{f} v_0 \\ w_1 &= -\left[\frac{\bar{D}}{\bar{D}t} \theta_0 - S_\theta \right] \end{aligned} \tag{14}$$

which can be rewritten

$$\mathbf{L} V_1 = \tilde{N}(V_0) \tag{15}$$

in which we need to solve for V_1. Note that the overbar time derivative denotes advection with respect to the lowest order flow, (u_0, v_0). I have used $\tilde{N}$ to distinguish it from N because of the source of velocity divergence at this order. Notice, this is well accepted, i.e. that there is a source of divergence to the higher order velocity in QG theory arising from the lower order velocity.

The solvability condition for (14) or (15) is that the projection of the right hand side on the adjoint eigenvector from (12) must be zero. We find, after integrating by parts

$$\int_{\mathbf{R}^3} \Phi \left\{ \partial_x \left[f^{-1} \frac{\bar{D}}{\bar{D}t} v_0 \right] - \partial_y \left[f^{-1} \frac{\bar{D}}{\bar{D}t} u_0 \right] + \frac{\beta}{f} v_0 + \partial_z \left[\frac{\bar{D}}{\bar{D}t} \theta_0 - S_\theta \right] \right\} d^3x = 0 \tag{16}$$

for all functions $\Phi(x, y, z, t)$. Therefore the quantity in parentheses must be zero $\forall(x, y, z, t)$ and so

$$\partial_x \left[f^{-1} \frac{\bar{D}}{\bar{D}t} \frac{\partial_x \Pi}{f} \right] + \partial_y \left[f^{-1} \frac{\bar{D}}{\bar{D}t} \frac{\partial_y \Pi}{f} \right] + \partial_z \left[\frac{\bar{D}}{\bar{D}t} \partial_z \Pi \right] + \frac{\beta}{f} \frac{\partial_x \Pi}{f} = \partial_z S_\theta \tag{17}$$

So now we need only simplify this equation. It is important to remember that all derivatives of f or β with respect to y are $O(\epsilon)$ and must therefore be neglected in this asymptotic expansion. Therefore, we can pass f through all the derivatives on the left hand side and multiply out by f^2. We can simplify this expression to read

$$\begin{aligned} &\partial_t \partial^2_{xx} \Pi + \frac{1}{f} \partial_x J(\Pi, \partial_x \Pi) + \\ &\partial_t \partial^2_{yy} \Pi + \frac{1}{f} \partial_y J(\Pi, \partial_y \Pi) + \\ &\partial_t f^2 \partial^2_{zz} \Pi + \frac{f^2}{f} \partial_z J(\Pi, \partial_z \Pi) + \beta \partial_x \Pi = f^2 \partial_z S_\theta + O(\epsilon) \end{aligned} \tag{18}$$

where $J(A, B) = \partial_x A \ \partial_y B - \partial_y A \ \partial_x B$, the Jacobian. The Leibniz rule applied to the Jacobian yields

$$\partial_x J(A, B) = J(\partial_x A, B) + J(A, \partial_x B), \tag{19}$$

and, of course,

$$J(A, A) = 0 \quad \forall A. \tag{20}$$

So, equation (18) becomes

$$\frac{\overline{D}}{\overline{\overline{D}t}} q = f^2 \frac{\partial S_\theta}{\partial z} \tag{21}$$

where

$$\frac{\overline{D}}{\overline{\overline{D}t}} (\cdot) = \frac{\partial}{\partial t} (\cdot) + \frac{1}{f} J(\Pi, \cdot) \tag{22}$$

and

$$q = \partial^2_{xx} \Pi + \partial^2_{yy} \Pi + f^2 \partial^2_{zz} \Pi + \frac{f}{\epsilon}. \tag{23}$$

This is just the conservation of potential vorticity equation. Though there seems to be an ϵ^{-1} term, this is not really the case. The QG PV equation (21) can be rewritten

$$\frac{\overline{D}}{\overline{\overline{D}t}} q' + \beta \frac{\partial \Pi}{\partial x} = f^2 \frac{\partial S_\theta}{\partial z} \tag{24}$$

where

$$q' = \partial^2_{xx}\Pi + \partial^2_{yy}\Pi + f^2\partial^2_{zz}\Pi \tag{25}$$

the relative potential vorticity. This shows that all the terms in the QG PV equation are really $O(\epsilon^0)$. Usually we see equation (23) written from a reference latitude, y_0 and the constant, $O(\epsilon^{-1})$ term is subtracted from the definition.

Solving for the $O(\epsilon^1)$ terms

So after the solvability condition is applied, the solution to equation (15) is

$$V_1 = \mathbf{L}^{-1}P\left[\tilde{N}(V_0)\right] + V_H \tag{26}$$

where $P[\cdot]$ denotes the projection of the right hand side of (15) which removes the secular terms. V_H is the homogeneous solution, i.e. anything in the kernel of $\mathbf{L}$. Since this is arbitrary, the way we address this term is to say that it is initially set to zero - i.e. any initial data which fall in the kernel of $\mathbf{L}$ can always be incorporated in the lowest order term, V_0. With this prescription, V_1 is uniquely define in terms of V_0.

Looking at equations (14) we see that all the first order fields are uniquely defined. In particular, it is clear that w_1 is uniquely defined

$$w_1 = -\left[\frac{\overline{D}}{\overline{Dt}}\Pi_z - S_\theta\right]. \tag{27}$$

That being said, I am no convinced that there is no ambiguity in the vertical velocity. The lowest order term is zero, the next order term is given by equation (14). Now, I'm not sure how to reconcile this with either (a) Majda/Embid's derivation from Ertel's theorem or (b) Khouider's derivation.

A diagnostic equation for w_1

Though equation (27) might seem like a prognostic equation, it isn't. Nonetheless, atmospheric scientists like to solve Laplacians and there is a way to re-write it in terms of Laplacians. Let's define the modified Laplacian operator

$$\nabla^2_f \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + f^2\frac{\partial^2}{\partial z^2} \equiv \nabla^2_H + f^2\frac{\partial^2}{\partial z^2} \tag{28}$$

and act on equation (27) with it. We find

$$\nabla^2_f w_1 = -\left[\frac{\partial}{\partial z}\frac{\partial}{\partial t}\nabla^2_f\Pi + \nabla^2_f\ J(\Pi,\Pi_z) - \nabla^2_f S_\theta\right] \tag{29}$$

and using (24) this becomes

$$\begin{aligned}\nabla_f^2 w_1 &= -\left[f^2\frac{\partial^2 S_\theta}{\partial z^2} - \frac{\partial}{\partial z}\, J(\Pi, \nabla_f^2\Pi) - \beta\frac{\partial^2}{\partial z\,\partial x}\Pi + \nabla_f^2\, J(\Pi, \Pi_z) - \nabla_f^2 S_\theta\right] \\ &= -\left[\nabla_f^2\, J(\Pi, \Pi_z) - \frac{\partial}{\partial z}\, J(\Pi, \nabla_f^2\Pi) - \beta\,\Pi_{xz} - \nabla_H^2 S_\theta\right].\end{aligned} \tag{30}$$

Now we've got to do a double Leibniz rule for the modified Laplacian

$$\nabla_f^2\, J(A, B) = J(\nabla^2 A, B) + 2J(\nabla A, \nabla B) + J(A, \nabla^2 B) \tag{31}$$

where the middle term is just summation over the like indices of the gradient. So (30) becomes

$$\begin{aligned}\nabla_f^2 w_1 = &- \left[J(\nabla_f^2\Pi, \Pi_z) + 2J(\nabla_f\Pi, \nabla_f\Pi_z) + J(\Pi, \nabla_f^2\Pi_z)\right. \\ &\left.- J(\Pi_z, \nabla_f^2\Pi) - J(\Pi, \nabla_f^2\Pi_z) - \beta\,\Pi_{xz} - \nabla_H^2 S_\theta\right] \\ = &\left[2J(\Pi_z, \nabla_f^2\Pi) + 2J(\nabla_f\Pi_z, \nabla_f\Pi) + \beta\,\Pi_{xz} + \nabla_H^2 S_\theta\right]\end{aligned} \tag{32}$$

which is the time derivative independent way to get w_1.

There's actually a nicer way to do this. Just go back to (30) and write

$$\begin{aligned}\nabla_f^2\left[w_1 + J(\Pi, \Pi_z)\right] =&\beta\Pi_{xz} + \frac{\partial}{\partial z}\, J(\Pi, \nabla_f^2\,\Pi) + \nabla_H^2 S_\theta \\ =&\frac{\partial}{\partial z}\, J\left(\Pi, \nabla_f^2\,\Pi + \frac{f}{\epsilon}\right) + \nabla_H^2 S_\theta \\ \nabla_f^2\left[w_1 + J(\Pi, \Pi_z)\right] =&\frac{\partial}{\partial z}\, J(\Pi, q) + \nabla_H^2 S_\theta\end{aligned} \tag{33}$$

which looks like the nicest way to describe it.

Comments to Boualem:

Yes, the flow is a-geostrophic, however it is not a gravity wave, per se. What you and I would mean by a gravity wave is that it (a) oscillates and (b) disperses. The first order flow does neither. It can't oscillate because it is the particular solution to equation (26) in my writeup. Since the linear operator L has no time derivatives, then the time dynamics of V_1 are completely parametrically determined by V_0. Since V_0 has no "fast, oscillating" component nor, then, does V_1.

I think the best way to think of V_1 is it solves the forced geostrophic balance equations.... either a forced version of your equation (6) (which is what your equation (7) is, or you can think of a forced version of my equation (10) (which is what my equation (15), and its solution (26) are).

Now, if you have an arbitrary such forcing, then the solution, V_1 is secular - and that will mean secular in space - so it grows as $|x| \to \infty$.

But you don't have arbitrary forcing. Rather, you have removed the part of the force that would cause the spatial secularity. Furthermore, there is no time derivative to invert to solve this equation(15). Therefore the solutions, V_1 remain "pinned" to the spatial location of the forcing.

Basically, there are no waves in V_1. HOWEVER, if you were to look at the the linearized equations - say the straight up linearization of my equation (1), then there would be 3 eigenfunctions for this problem problem (say for each k,l,m, wavenumbers). Of course you have 3 eigenvalues, fast right, fast left, zero, for all k,l,m and we call these right gravity waves, left gravity waves, geostrophic balance.

The spatial structure of V_1 must project on to the eigenstructure of all of these modes, gravity and geostrophic balance - actually, for all I know it may not project onto geostrophic balance, but I bet it will project onto the gravity eigenfunction.

However, this does not mean that the time dynamics of this mode are that of a gravity wave. Its just that the spatial eigenstructure must project onto the gravity wave (it is divergent after all!).

How the hell does this happen? Well, we're actually solving a non-linear problem, albeit with a succession of linear approximations.