

## The shape of the main thermocline, revisited

by Rick Salmon<sup>1</sup>

### ABSTRACT

Using the Monte Carlo method of statistical physics, we compute the equilibrium statistical mechanics of the shallow water equations, considered as a reduced-gravity model of the ocean's upper layer in a square ocean that spans the equator. The ensemble-averaged flow comprises a westward drift at low latitudes, associated with the poleward deepening of the main thermocline, and a more intense compensating eastward flow near the latitudes at which the layer depth vanishes. Inviscid numerical simulations with a model that exactly conserves mass, energy, and potential enstrophy support the theoretical prediction.

### 1. Introduction

Imagine the ocean to be in a state of rest, with a stable density stratification depending only on the vertical coordinate  $z$ . Suppose that a random stirring force—a giant paddle—sets the fluid into vigorous adiabatic motion in every direction. After a time, the random stirring force abruptly stops, but the fluid motion continues—we suppose—*undamped* by friction. What will be the state of this hypothetical ocean after a very long time? Equilibrium statistical mechanics offers a precise quantitative answer to this question under a single, additional, but very strong assumption—that the fluid motion be confined to a finite number of degrees of freedom. All of the foregoing conditions are satisfied by a numerical ocean model, subject to random initial conditions, in which all the heating, driving, and damping forces have been switched off. Thus the hypothetical question posed above could be stated more succinctly as follows: What is the eventual state attained by a numerical ocean model, under random initial conditions, with all the forcing and viscosity turned off?

A much earlier paper (Salmon 1982a, hereafter S82) tried to calculate the statistical-mechanical equilibrium state associated with a particularly simple ocean model, namely, the 'reduced gravity'—also sometimes called 'one-and-one-half-layer'—model, and to show that model simulations support the theoretical prediction. In the reduced-gravity model, the governing equations are the shallow-water equations (4.1) with  $h$  the depth of an upper moving layer of homogeneous fluid; beneath this layer, the ocean is assumed to be at rest. The

1. Scripps Institution of Oceanography, University of California, La Jolla, California, 92093-0213, U.S.A.  
*email: rsalmon@ucsd.edu*

general conclusion of S82 was that the final average flow closely resembles the corresponding flow in the one-layer quasigeostrophic case: quasi-uniform westward flow throughout the interior of the moving layer with a compensating eastward flow along its poleward flanks. The westward interior flow corresponds to a poleward increase in the thermocline depth  $h$  between a narrow equatorial region of nearly uniform  $h$  and the even narrower regions of eastward flow in which  $h$  abruptly vanishes. These ‘outcrop’ regions correspond to the inertial boundary layers in the corresponding quasigeostrophic calculation. Thus, principal observed features of upper-ocean circulation are predicted to arise automatically by the tendency for initially random, rotating flow to induce rectified currents.

The equilibrium statistical mechanics of quasigeostrophic models was first considered by Salmon *et al.* (1976). In a closed beta-plane box, one-layer quasigeostrophic equilibrium flow is quasi-steady, and is essentially identical to the steady inviscid flow studied by Fofonoff (1954), which we hereafter call *Fofonoff flow*. The tendency for unforced, inviscid, quasigeostrophic models to attain Fofonoff flow has been confirmed in numerical experiments by Griffa and Salmon (1989), Wang and Vallis (1994), Dukowicz and Greatbatch (1999) and a number of others.

Many of these papers also investigate the degree to which the models attain Fofonoff-like states in the presence of forcing and viscosity. For example, Griffa and Salmon (1989) investigated the dependence of the statistical equilibrium state on the direction of the wind forcing in model solutions with viscosity and bottom drag. In the case of easterly wind forcing—for which the wind does positive work on Fofonoff flow—they found that the equilibrium mean flow is strong, quasi-steady, and closely resembles Fofonoff flow. In the case of westerly wind forcing—in which the wind opposes Fofonoff flow—the equilibrium mean flow is very weak, and most of the kinetic energy resides in time-dependent, turbulent motions. Thus, in the general forced/dissipated case, Fofonoff flow represents a kind of stable attractor; the system approaches the attractor as closely as the forcing and dissipation will allow. In this paper we are solely concerned with the calculation of ideal, inviscid equilibrium states like Fofonoff flow. However, we come back to the issue of forcing and dissipation in Section 5.

The ideal, inviscid quasigeostrophic equilibrium state is easy to calculate because the primary conserved quantities—the energy and potential enstrophy—are *quadratic* functionals of dynamic variables for which the motion in phase space is nondivergent. (We define all these terms precisely in the following section.) That is, the quasigeostrophic energy (2.3) and potential enstrophy (2.6) depend quadratically on the streamfunction  $\psi(x, y, t)$ . For the same reason, it is very easy to construct quasigeostrophic numerical models that conserve energy and potential enstrophy. In contrast, the shallow-water energy (4.4) is a cubic functional of the variables  $u$ ,  $v$  and  $h$ , and the potential enstrophy (4.7) involves a quotient. This makes it impossible to calculate the shallow-water equilibrium states exactly, and it makes it difficult to construct numerical models of the shallow-water equations that exactly conserve energy and potential enstrophy. In fact, S82 did neither, instead piling one assumption on top of another until the equilibrium calculation yielded the expected result. The numerical

experiments of S82 were likewise unconvincing. Simulated times were short, resolution was poor, and neither conservation law was respected; the emphasis in S82 was on handling the outcrop at which the layer depth vanishes.

In this paper we revisit the problem considered by S82 using better computer technology and superior methods of calculation. We compute the shallow-water equilibrium state using the Monte Carlo method of statistical physics, and we compare our calculation to the results of numerical simulations with a recently developed shallow-water model that exactly conserves its energy and potential enstrophy.

The plan of the paper is as follows. Section 2 reviews the equilibrium statistical mechanics of the quasigeostrophic model. Section 3 introduces the Monte Carlo method and illustrates its use by application to the quasigeostrophic system. We present two quasigeostrophic calculations. The first calculation corresponds to the standard, analytic case in which energy and potential enstrophy are the only recognized invariants; the second calculation also recognizes the fourth power of the potential vorticity. The similarity between the mean flows computed in these two cases supports the view that energy and potential enstrophy are of primary importance.

Section 4 addresses the shallow water system. We use the Monte Carlo method to compute the ensemble-average flow corresponding to the three shallow-water invariants of mass, energy and potential enstrophy. In the case of the full shallow-water system, and in a case intermediate between quasigeostrophy and full shallow-water, we find that the mean flow resembles that hypothesized in S82. Finally, we compute the inviscid, time-average flow in a numerical model that exactly conserves its mass, energy and potential enstrophy. Although the model cannot accommodate vanishing layer depth, the results resemble our Monte Carlo calculations. Section 5 concludes with an assessment of the method and suggestions for future work.

Some results of the statistical mechanical theory may also be understood, qualitatively but more simply, by using ideas from ‘wave-mean interaction theory.’ In this theory, recently summarized in the beautiful book by Bühler (2009), the concept of pseudomomentum plays a central role. From the standpoint of wave-mean theory, Fofonoff flow arises because breaking Rossby waves deposit their westward pseudomomentum into the mean flow. We briefly mention the connection between pseudomomentum theory and equilibrium statistical mechanics in Section 2.

However, from an even simpler viewpoint, the westward Fofonoff flow arises from the tendency for potential-vorticity conservation on fluid particles—combined with weak dissipation or irreversible mixing of potential vorticity—to accumulate negative relative vorticity in the northern part of the ocean and positive relative vorticity in the southern part. This tendency and its ability to explain the emergence of Fofonoff-like mean gyres was recognized in a pioneering paper by Veronis (1970), who considered quasigeostrophic flow governed by (2.1-2.2) with weak Rayleigh friction and random wind forcing terms appended. Veronis (1970) gave what might be the first physical explanation for the emergence of Fofonoff flow:

The local vorticities ...are 'sorted out' by the  $\beta$ -effect ...so that negative vorticity accumulates to the north and positive vorticity to the south. ...The result is that a time average of the streamfunction shows a nonzero mean circulation consisting of a gyre of negative (positive) vorticity in the northern (southern) half basin.

It is a pleasure to dedicate this paper to George Veronis, whose ideas, mentorship and friendship have been so important to its author.

## 2. Review of previous work

This section offers a bare-bones review of some basic principles of equilibrium statistical mechanics and their application to single-layer quasigeostrophic turbulence in a beta-plane ocean. For much more extensive introductions to these subjects, see Salmon (1982b, 1998) and Holloway (1986). The governing quasigeostrophic equation is

$$\frac{\partial q}{\partial t} + J(\psi, q) = 0 \quad (2.1)$$

where

$$q = \nabla^2 \psi + f \quad (2.2)$$

is the potential vorticity;  $\psi(x, y, t)$  is the streamfunction at location  $(x, y)$  and time  $t$ ;  $f = \beta y$  is the Coriolis parameter with  $\beta$  a constant;  $\nabla A \equiv (A_x, A_y)$ ; and  $J(A, B) \equiv A_x B_y - B_x A_y$ . Subscripts denote differentiation. The boundary condition is  $\psi = 0$ . The horizontal velocity in the (east, north) direction is  $\mathbf{u} = (u, v) = (-\psi_y, \psi_x)$ . Hence (2.1) expresses the conservation of potential vorticity on fluid particles. The dynamics (2.1–2.2) conserves the energy<sup>2</sup>

$$E = \iint d\mathbf{x} (\nabla \psi \cdot \nabla \psi) \quad (2.3)$$

and every quantity of the form

$$\iint d\mathbf{x} F(q) \quad (2.4)$$

where  $F(\cdot)$  is an arbitrary function. Particular importance attaches to the circulation

$$C = \iint d\mathbf{x} q \quad (2.5)$$

and the potential enstrophy

$$Z = \iint d\mathbf{x} q^2. \quad (2.6)$$

2. The development is cleanest if we here omit the conventional factor of 1/2 in the definition of energy.

Equilibrium statistical mechanics applies to systems of coupled ordinary differential equations in the form

$$\frac{dy_i}{dt} = f_i(y_1, y_2, y_3, \dots, y_n), \quad i = 1, \dots, n \quad (2.7)$$

where  $n$ , the number of degrees of freedom, is finite. Every state of the system with evolution equation (2.7) corresponds to a point  $\mathbf{y} \equiv (y_1, y_2, y_3, \dots, y_n)$  in phase space. If

$$\sum_{i=1}^n \frac{\partial f_i}{\partial y_i} = 0, \quad (2.8)$$

then the motion in phase space is nondivergent. If motion governed by (2.7) satisfying (2.8) conserves the  $m$  quantities

$$E_1(\mathbf{y}), E_2(\mathbf{y}), \dots, E_m(\mathbf{y}), \quad (2.9)$$

then equilibrium statistical mechanics predicts that the probability distribution in phase space approaches the canonical distribution<sup>3</sup>

$$P(\mathbf{y}) = \alpha_0 \exp(-\alpha_1 E_1 - \alpha_2 E_2 - \dots - \alpha_m E_m) \quad (2.10)$$

where the  $m + 1$  constants  $\alpha_i$  are determined by the normalization requirement

$$\iint \dots \int d\mathbf{y} P(\mathbf{y}) = 1 \quad (2.11)$$

and by the  $m$  requirements

$$\langle E_i \rangle \equiv \iint \dots \int d\mathbf{y} E_i(\mathbf{y}) P(\mathbf{y}) = E_i^0 \quad (2.12)$$

that the average value of each  $E_i$  equals its prescribed value  $E_i^0$ . From the viewpoint of information theory, one obtains (2.10) by maximizing the entropy

$$S = - \iint \dots \int d\mathbf{y} P \ln P \quad (2.13)$$

subject to the constraints (2.11) and (2.12). The  $\alpha_i$  are the Lagrange multipliers corresponding to these constraints. One can also regard  $\alpha_1, \alpha_2, \dots, \alpha_m$  as ‘inverse temperatures.’

The quasigeostrophic system (2.1–2.2) fits the form (2.7) if the streamfunction  $\psi(x, y, t)$  is truncated to a finite number of degrees of freedom, as is always necessary in constructing a

3. If the microcanonical distribution is desired instead, then additional constraints limiting the variance of the  $E_i$  may be attached to the maximization of entropy; see, e.g., Salmon (1998, pp. 251–256).

numerical model. In the conceptually cleanest method, we imagine  $\psi(x, y, t)$  to be expanded as the sum

$$\psi(x, y, t) = \sum_{i=1}^n y_i(t) \phi_i(x, y) \quad (2.14)$$

of orthonormal modes obeying  $\nabla^2 \phi_i = -k_i^2 \phi_i$  and the boundary condition  $\phi_i = 0$ . Then (2.7) corresponds to a spectral model of (2.1). The largest eigenvalue  $k_{max}$  corresponds to the smallest spatial scale  $k_{max}^{-1}$  resolved by the model. However, one could instead regard  $y_i(t)$  as the value of  $\psi$  at the  $i$ -th grid point in a finite-difference model of (2.1). Salmon and Talley (1989) show how to construct finite-difference models of (2.1–2.2) that conserve finite-difference analogs of (2.3, 2.5–2.6). In either the spectral or finite-difference case, the discrete evolution equations (2.7) obey the requirement (2.8). Hence statistical mechanics predicts the equilibrium probability distribution

$$P = \alpha_0 \exp(-\alpha_E E - \alpha_C C - \alpha_Z Z) \quad (2.15)$$

where  $E$ ,  $C$  and  $Z$  are the conserved spectral or finite-difference analogs of (2.3, 2.5–2.6). The four  $\alpha$ 's are determined by the normalization requirement (2.11) and by the requirements that  $\langle E \rangle$ ,  $\langle C \rangle$  and  $\langle Z \rangle$  equal their prescribed initial values. By a standard theorem of statistical mechanics, these requirements uniquely determine the  $\alpha$ -values.

The equilibrium state (2.15) corresponding to (2.1–2.2) was thoroughly analyzed by Carnevale and Frederiksen (1987). They considered the asymptotic limit of perfect model resolution,  $k_{max} \rightarrow \infty$ , with  $\langle E \rangle$ ,  $\langle C \rangle$  and  $\langle Z \rangle$  held fixed. In this limit, the mean streamfunction  $\langle \psi \rangle$  obeys

$$\nabla^2 \langle \psi \rangle + \beta y = \frac{\alpha_E}{\alpha_Z} \langle \psi \rangle + \frac{\alpha_C}{\alpha_Z} \quad (2.16)$$

and the energy in the fluctuating flow represented by  $\psi' = \psi - \langle \psi \rangle$  is concentrated at the smallest resolved scale  $k_{max}^{-1}$ .<sup>4</sup> Carnevale and Frederiksen noted that, for all *realizable* values of the  $\alpha$ 's—that is, for all  $\alpha$ 's corresponding to values of (2.3, 2.5–2.6) that could actually be attained—the mean flow defined by (2.16) is a stable flow in the sense of Arnol'd (1965, 1969). Unless the energy is unrealistically large,  $\alpha_E \gg k_{min}^2 \alpha_Z > 0$  as  $k_{max} \rightarrow \infty$ , and the solution of (2.16) consists of a westward interior flow with speed  $U_I \equiv \beta \alpha_Z / \alpha_E$  closed by boundary layers of width  $\sqrt{U_I / \beta} = \sqrt{\alpha_Z / \alpha_E}$ .

All of these results are *analytic*. Although computer simulations of (2.1–2.2) have been used to test the predictions of equilibrium statistical mechanics, the theory itself is a pencil-and-paper theory. However, this is true only because the conserved quantities (2.3, 2.5–2.6) are either linear or quadratic. Then (2.15) is multivariate Gaussian, and integrals like those needed to establish (2.16) can be performed.

4. In contrast, Fofonoff (1954) *hypothesized* (2.16), noting that its solution is a steady solution of (2.1).

But suppose we ask the statistical mechanical theory to incorporate the *quartic* invariant

$$Q = \iint d\mathbf{x} q^4. \quad (2.17)$$

Then the equilibrium probability distribution acquires an additional term in its exponent, and the average streamfunction is given by

$$\langle \psi \rangle = \iint \cdots \int d\mathbf{y} \psi \alpha_0 \exp(-\alpha_E E - \alpha_C C - \alpha_Z Z - \alpha_Q Q). \quad (2.18)$$

Now, however, there is nothing analogous to the theory of quadratic forms that would allow us to factor (2.18) into a product of simpler integrals. The expression (2.18) is well-defined, but for a model with (say)  $100^2$  grid points, it corresponds to an irreducible integral over a  $100^2$ -dimensional space. The question of how the additional, nonquadratic invariant (2.17) modifies the prediction (2.16) is an interesting one, and in the next section we show how to answer it using a method designed for calculating very high-dimensional integrals. However, the more essential point is this: In the case of shallow-water dynamics, even the energy and the potential enstrophy are nonquadratic functions of the phase coordinates. Thus, in the shallow-water case, the methods of the following section seem indispensable.

### 3. The Monte Carlo method

Metropolis *et al.* (1953) invented an algorithm for calculating integrals like (2.18). Their algorithm and its many subsequent variants are often called the *Monte Carlo* method. For a thorough explanation of the Monte Carlo method, see for example, Kalos and Whitlock (2008, pp. 64–72). Here we present a brief description of the method as it applies to the calculation of (2.18).

To begin the Metropolis algorithm, we choose values for the Lagrange multipliers  $\alpha_E$ ,  $\alpha_C$ ,  $\alpha_Z$ , and  $\alpha_Q$ . Then we choose an arbitrary state  $\mathbf{y}$  of the quasigeostrophic system as the first member of an ensemble. We calculate its probability of occurrence,  $P(\mathbf{y})$ , according to the canonical distribution. Then we randomly choose a nearby state  $\mathbf{y}'$  and calculate its probability  $P(\mathbf{y}')$ . If the new state  $\mathbf{y}'$  is more probable than the old state  $\mathbf{y}$ , then we accept the new state as the next member of the ensemble, and we continue the process by random nomination of another nearby state. If the new state  $\mathbf{y}'$  is *less* probable than the old state  $\mathbf{y}$ , then we accept it with probability  $\mathcal{P} \equiv P(\mathbf{y}')/P(\mathbf{y})$ . That is, we request a random number  $R$  between 0 and 1. If  $R < \mathcal{P}$  we accept  $\mathbf{y}'$  as the next member of our ensemble. If  $R > \mathcal{P}$  we admit the old state  $\mathbf{y}$  as an ensemble member *for a second time* and proceed by making another nomination for  $\mathbf{y}'$ . It can be shown that this process of accumulating ensemble members eventually results in an ensemble of states in which each state occurs with the same probability as in the canonical distribution. However, convergence is typically very slow. To compute the average of any quantity  $G(\mathbf{y})$  that depends on the system state  $\mathbf{y}$ , we simply add up all the values of  $G$  in our accumulated ensemble and divide by the number of ensemble members.

We consider a square ocean of side  $L = 4000$  km covered by 100 grid points in each direction. At the boundary grid points  $\psi = 0$ . At the interior grid points  $\psi$  may assume any value, hence the dimension of our phase space is  $n = 98^2$ . In each iteration of the Monte Carlo method we perturb the value of  $\psi$  at a single interior grid point. Let  $\psi_{ij}$  be the value of  $\psi$  at grid point  $(i, j)$ . We nominate the new value  $\psi'_{ij}$  by choosing a number, with uniform probability, from the range  $(\psi_{ij} - \Delta\psi, \psi_{ij} + \Delta\psi)$ , where  $\Delta\psi$  is a prescribed constant. Then we compute the probability associated with the perturbed state, and either accept it or reject it according to the rules stated in the previous paragraph. We obtain the average streamfunction field  $\langle\psi_{rs}\rangle$  by averaging over all the members of our ensemble. Each ensemble member consists of the entire field of gridded  $\psi$ -values. The most time-consuming part of the calculation is the computation of the discrete approximations to the energy (2.3) and potential-vorticity moments (2.5, 2.6, 2.17) for each nominated state. Efficiency is achieved by using the fact that perturbations in each  $\psi_{ij}$  affect only a few terms in the expressions for the energy and other invariants; only these terms need to be re-computed to find the change in energy from the previous state.

The parameter  $\Delta\psi$ , which determines the average size of the perturbations, must neither be too large nor too small. If  $\Delta\psi$  is too large, then too many of the nominated states will be rejected, and convergence will be slow. If  $\Delta\psi$  is too small, then the acceptance rate will be large, but convergence will be slow because the new states differ little from the old ones. The fastest convergence is achieved by periodically adjusting  $\Delta\psi$  so that about 50% of the nominations result in transitions to a new state. Because the initial guess is inevitably a poor one, one typically finds that relatively large values of  $\Delta\psi$  are required early in the calculation. However,  $\Delta\psi$  must be reduced as the sequence of states approaches equilibrium.

Let the equator  $y = 0$  lie at midbasin. If the average circulation  $\langle C \rangle$  vanishes, then the equilibrium flow will be symmetric about the equator. We impose this by setting  $\alpha_C = 0$ . We wish to compare the average streamfunction corresponding to

$$P = \alpha_0 \exp(-\alpha_E E - \alpha_Z Z) \quad (3.1)$$

to the average streamfunction corresponding to

$$P = \alpha_0 \exp(-\alpha_E E - \alpha_Z Z - \alpha_Q Q) \quad (3.2)$$

which incorporates the quartic moment (2.17). In the *Fofonoff case* (3.1), the average streamfunction is already known—it is given by the solution of (2.16) with  $\alpha_C = 0$ —but our calculation provides a useful check on the Monte Carlo algorithm. In the *quartic case* (3.2), the average streamfunction is unknown, and the Monte Carlo method seems to be the only way to find it.

To calculate the average streamfunction corresponding to (3.1), we choose fixed values for  $\alpha_E$  and  $\alpha_Z$  and an arbitrary beginning state for our system; we begin in the state of rest. According to theory, there is a unique, one-to-one correspondence between the values of  $\alpha_E$

and  $\alpha_Z$  and the values of  $\langle E \rangle$  and  $\langle Z \rangle$ . That is,  $\alpha_E = \alpha_E(\langle E \rangle, \langle Z \rangle)$  and  $\alpha_Z = \alpha_Z(\langle E \rangle, \langle Z \rangle)$ . However, it is simpler to specify  $\alpha_E$  and  $\alpha_Z$  than to specify  $\langle E \rangle$  and  $\langle Z \rangle$ .<sup>5</sup> Trial-and-error adjustments of the  $\alpha$ 's are sufficient to yield realistic values of the conserved quantities, because the results do not depend sensitively on either set of parameters.

We apply the Metropolis algorithm in *stages*. Each stage begins with the final ensemble member of the previous stage and makes 50,000 passes through the  $n = 98^2$  grid points at which  $\psi$  may vary. Each pass consists of a single  $\psi$ -perturbation at each interior grid point. Thus each stage accumulates  $50,000 \times 98^2$  ensemble members. Once again, each ensemble member is a complete set of gridded  $\psi$ -values. We compute the average  $\psi$ -field corresponding to each stage. The first and many subsequent stages are of no interest, because they merely reflect our arbitrary starting point. We keep computing new stages until the averages obtained from successive stages differ by very small amounts.

Figure 1a shows the average streamfunction corresponding to (3.1) with 'energy temperature'  $T_E \equiv \alpha_E^{-1} = 0.6422 \text{ km}^2/\text{day}^2$  and 'potential enstrophy temperature'  $T_Z \equiv \alpha_Z^{-1} = 1.573 \times 10^{-5} \text{ day}^{-2}$ , as computed by the Monte Carlo method.<sup>6</sup> These  $\alpha$ -choices correspond to Fofonoff flow with a westward interior velocity  $U_I = \beta\alpha_Z/\alpha_E = 80.1 \text{ km/day}$  and a boundary layer thickness  $\delta_I \equiv \sqrt{U_I/\beta} = 202 \text{ km}$  of  $5\Delta$ , where  $\Delta$  is the distance between grid points. The Metropolis algorithm was applied until significant trends had disappeared and until the averages computed by successive stages differed by less than 1%. This required about 30 minutes of computer time on a 2 GHz processor. In the stage corresponding to Figure 1a, the average interior velocity is 79.2 km/day, which is very close to the expected value of 80.1. In fact, Figure 1a closely resembles the familiar boundary-layer solution of (2.16) (e.g. Salmon, 1998, p. 273), which applies when  $\delta_I \ll L$ . Of course, this  $U_I$  is much larger than observed interior ocean current velocities. However, the need to resolve inertial boundary layers on our coarse  $100 \times 100$  grid demands a relatively large  $U_I$ . The use of finer resolution greatly increases the computer time required for the Metropolis algorithm to converge.

In trying to determine the influence of the quartic invariant (2.17) on the equilibrium statistical mechanics, we must recognize that the answer depends not only on the prescribed value of (2.17), but also on the information that will be computed from the resulting equilibrium distribution (3.2). The Fofonoff solution of Figure 1a has a value of the quartic equal to  $\langle Q \rangle_{Foff} = 13.9 \text{ day}^{-4}$ . If, for the same values of energy and potential enstrophy, the *prescribed* value of the average quartic should just happen to equal  $\langle Q \rangle_{Foff}$ , then the Lagrange multiplier  $\alpha_Q = 0$ , and *any statistic* computed from (3.2) would be the same as that computed from (3.1). For nonzero  $\alpha_Q$  the statistics computed from (3.2) will generally differ from those computed from (3.1), but the difference will depend on the particular

5. We need not specify  $\alpha_0$ , because the Monte Carlo algorithm requires only the *ratio* of probabilities.

6. From now on, when we speak of energy, we mean the energy *per unit area*, that is, the total energy (2.3) divided by the area  $L^2$ , and similarly for the potential enstrophy and the quartic. Numerical values reflect this convention.

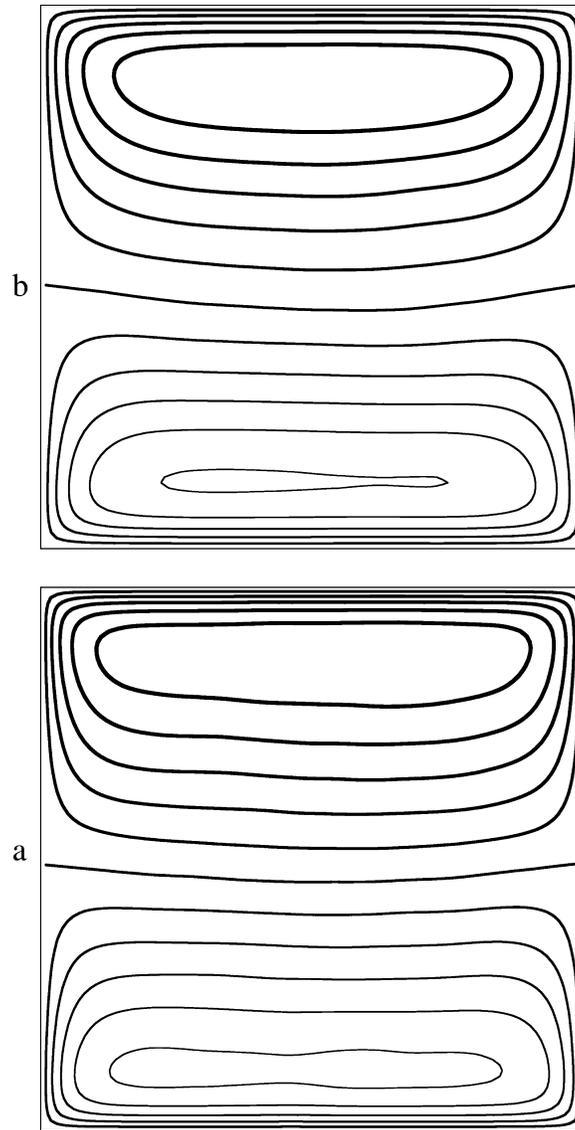


Figure 1. The average streamfunction for quasigeostrophic flow in a beta-plane ocean, as computed by the Monte Carlo method, for (a) the probability density (3.1), which recognizes only the energy and potential enstrophy; and (b) the probability density (3.2), which also recognizes the quartic (2.17).

statistic being compared. For example, the difference in  $\langle Q \rangle$  itself is likely to be significant unless of course  $\langle Q \rangle \approx \langle Q \rangle_{Foff}$ . It seems likely that higher order statistics, such as  $\langle q^8 \rangle$ , would be more sensitive to the prescribed value of the quartic than to the prescribed values of the average energy and potential enstrophy. However, it is also likely that lower order statistics such as  $\langle \psi \rangle$  are not very sensitive to the prescribed value of  $\langle Q \rangle$  or  $\alpha_Q$ . In fact, if this were not the case, then the whole program of statistical mechanics as applied to systems like (2.1–2.2) would be seriously in doubt.

In this section we are only interested in the average streamfunction  $\langle \psi \rangle$ , and we investigate the influence of the quartic by increasing  $\alpha_Q$  from zero, keeping the same values of  $\alpha_E$  and  $\alpha_Z$  as in the Fofonoff case described above.<sup>7</sup> We find no significant changes to the average flow until  $\alpha_Q$  exceeds the value  $\alpha_{Qcrit} \equiv n/\langle Q \rangle_{Foff}$ , where, again,  $n$  is the number of degrees of freedom. Figure 1b shows the streamfunction corresponding to a value of  $\alpha_Q$  equal to *five times* this amount. In the solution of Figure 1b, the average interior flow is 114.5 km/day;  $\langle E \rangle / \langle E \rangle_{Foff} = 1.63$ ;  $\langle Z \rangle / \langle Z \rangle_{Foff} = 0.86$ ; and  $\langle Q \rangle / \langle Q \rangle_{Foff} = 0.67$ . These changes are significant, but the general pattern of the flow is the same as in Figure 1a. This supports the prevailing idea (e.g. Carnevale and Frederiksen, p. 175) that the higher moments of potential vorticity are of secondary importance in determining the equilibrium mean flow.

Before leaving the quasigeostrophic case, we briefly mention a connection between the statistical mechanical explanation for Fofonoff flow and wave-mean theories based upon the concept of pseudomomentum. Suppose that the energy (2.3) were the only conserved quantity. Then equilibrium statistical mechanics predicts equipartition of energy among Fourier modes. Since the number of modes with wavenumbers of size  $k$  increases like  $k$ , most of the energy appears near the highest wavenumber  $k_{max}$  in the truncated system. The enstrophy approaches its maximum value of  $k_{max}^2$  times the energy. But suppose that the potential enstrophy (2.6) is also conserved. Apart from irrelevant additive and multiplicative constants, this implies conservation of

$$Z = \iint d\mathbf{x} \left( \frac{1}{2} \zeta^2 + \beta y \zeta \right). \quad (3.3)$$

If  $\beta = 0$  the conservation of enstrophy  $\iint d\mathbf{x} \zeta^2$  prevents energy from reaching high wavenumbers; the enstrophy constraint traps most of the energy in large spatial scales. The entropy of the system is lower than if energy alone were conserved. However, if  $\beta \neq 0$ , then energy can move to high wavenumbers, increasing the entropy of the system. The resulting increase in enstrophy can be balanced by a negative correlation between  $y$  and  $\zeta$ . The first term in the integrand of (3.3) corresponds to the pseudomomentum;<sup>8</sup> the second term corresponds to the mean momentum, after integrations by parts. A typical argument

7. An alternative procedure would hold  $\langle E \rangle$  and  $\langle Z \rangle$  fixed while varying  $\langle Q \rangle$ , but it would be extremely difficult to determine the corresponding  $\alpha$ 's.

8. In discussing pseudomomentum, it is conventional to divide (3.3) by  $\beta$ .

applies the local, flux-form, conservation law corresponding to (3.3) to an ocean subdomain bounded by a latitude line (e.g. Salmon, 1998, pp. 264–279). Here we focus on the integral (3.3) over the whole domain, which is more closely related to the statistical mechanical approach. The implied negative correlation between vorticity and latitude represents the same sorting mechanism identified by Veronis (1970), and it offers a simple, qualitative explanation for Fofonoff flow. Since quasigeostrophic theory does not distinguish between bottom topography and Coriolis parameter, these same ideas predict anticyclonic flow around islands and seamounts.

Discussions of pseudomomentum often refer to the importance of dissipation or irreversible mixing of potential vorticity, and thus would seem to contradict the basis of equilibrium statistical mechanics. In fact, the two are closely related. Suppose that the inviscid truncated system reaches an equilibrium in which the increase in pseudomomentum  $P \equiv 1/2 \iint d\mathbf{x} \zeta^2/\beta$  balances the decrease in  $U \equiv \iint d\mathbf{x} y\zeta$ . Since  $\zeta + \beta y$  is conserved on fluid particles, the increase in pseudomomentum is associated with the development of thin filaments of vorticity  $\zeta$ . Now imagine that viscous dissipation is suddenly turned on for a short time. It is very plausible that the effect of the dissipation will be to decrease the pseudomomentum  $P$  while hardly affecting the mean flow  $U$  at all. However, this viscous decrease in  $P$  simply clears the way for nonlinear interactions to build  $P$  back up, thereby further decreasing  $U$ , resulting in an even stronger Fofonoff flow. We could achieve much the same result, with no dissipation at all, by simply increasing the truncation wavenumber  $k_{max}$ , that is, by opening up new, high-wavenumber modes, whose excitation would correspond to an increase in  $P$  and a compensating decrease in  $U$ .

#### 4. Statistical mechanics of shallow water

The shallow water equations may be written in the form

$$\frac{\partial u}{\partial t} = qvh - \frac{\partial \Phi}{\partial x} \quad (4.1a)$$

$$\frac{\partial v}{\partial t} = -quh - \frac{\partial \Phi}{\partial y} \quad (4.1b)$$

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(uh) + \frac{\partial}{\partial y}(vh) = 0 \quad (4.1c)$$

where  $(u, v)$  is the horizontal velocity in the  $(x, y)$  direction at time  $t$ ,  $h$  is the fluid depth,

$$q = (\zeta + f)/h \quad (4.2)$$

is the potential vorticity with relative vorticity  $\zeta = v_x - u_y$  and Coriolis parameter  $f$ , and

$$\Phi = \frac{1}{2}(u^2 + v^2) + gh. \quad (4.3)$$

At solid boundaries, both the normal component of velocity and its time derivative must vanish. We regard (4.1–4.3) as a reduced gravity model of the ocean’s upper layer; below this upper layer, the ocean is assumed to be at rest. Hence we take  $g = 0.002 \times 9.8 \text{ m sec}^{-2}$ .

The dynamics (4.1–4.3) conserves the energy

$$E = \iint d\mathbf{x} \left( \frac{1}{2} h \mathbf{u} \cdot \mathbf{u} + \frac{1}{2} g h^2 \right) \quad (4.4)$$

and every quantity of the form  $\iint d\mathbf{x} h q^n$ . Particular importance attaches to the mass

$$M = \iint d\mathbf{x} h, \quad (4.5)$$

the circulation

$$C = \iint d\mathbf{x} h q, \quad (4.6)$$

and the potential enstrophy

$$Z = \iint d\mathbf{x} h q^2. \quad (4.7)$$

Neither (4.4) nor (4.7) is quadratic in the dependent variables. Despite this fact, Salmon (2009, hereafter S09) developed a finite-difference model of the shallow-water equations that—apart from truncation error in the time step—exactly conserves finite-difference analogues of (4.4–4.7) in a bounded domain.

As shown in S09, it is easy to construct shallow-water models that conserve (4.5–4.7) if the fundamental dynamical variables  $u$  and  $v$  are replaced by the vorticity  $\zeta$  and the divergence  $\mu = u_x + v_y$ . In the new independent variables  $\zeta, \mu, h$ , the evolution equations take the form

$$\frac{\partial \zeta}{\partial t} = J(q, \chi) - \nabla \cdot (q \nabla \gamma) \quad (4.8a)$$

$$\frac{\partial \mu}{\partial t} = J(q, \gamma) + \nabla \cdot (q \nabla \chi - \nabla \Phi) \quad (4.8b)$$

$$\frac{\partial h}{\partial t} = -\nabla^2 \gamma. \quad (4.8c)$$

The streamfunction  $\chi$  and the velocity potential  $\gamma$  are defined by

$$h \mathbf{u} = \left( -\frac{\partial \chi}{\partial y} + \frac{\partial \gamma}{\partial x}, \frac{\partial \chi}{\partial x} + \frac{\partial \gamma}{\partial y} \right) \quad (4.9)$$

and are determined from  $\zeta, \mu, h$  by the elliptic equations

$$\nabla \cdot (h^{-1} \nabla \chi) + J(h^{-1}, \gamma) = \zeta \quad (4.10a)$$

$$\nabla \cdot (h^{-1} \nabla \gamma) + J(\chi, h^{-1}) = \mu. \quad (4.10b)$$

We also require

$$\Phi = \frac{1}{2h^2}(\nabla\chi \cdot \nabla\chi + \nabla\gamma \cdot \nabla\gamma + 2J(\chi, \gamma)) + gh. \quad (4.11)$$

The no-normal-flow boundary conditions may be taken as  $\chi = 0$  and  $\partial\gamma/\partial n = 0$ , where  $n$  is the outward normal. S09 derives finite-difference analogues of (4.8) that conserve analogues of (4.5–4.7) for *any* values of the ‘diagnostic variables’  $\chi$ ,  $\gamma$ , and  $\Phi$ . This is possible because (4.5–4.7) take a particularly simple form in terms of the ‘prognostic variables’  $\zeta$ ,  $\mu$ , and  $h$ . However, the conservation of energy

$$E = \iint d\mathbf{x} \left( \frac{1}{2h}(\nabla\chi \cdot \nabla\chi + \nabla\gamma \cdot \nabla\gamma + 2J(\chi, \gamma)) + \frac{1}{2}gh^2 \right) \quad (4.12)$$

depends upon the finite-difference forms of (4.10) and (4.11). More precisely, the finite-difference form of (4.12) determines the finite-difference forms of (4.10) and (4.11). Because the formulation (4.8–4.11) requires the solution of the coupled elliptic equations (4.10) at every time step, the S09 model is much less efficient than numerical models based upon the standard formulation (4.1–4.3). This inefficiency is the price to be paid for maintaining so many conservation laws, especially in the presence of boundaries.<sup>9</sup> For a complete description of the finite-difference model, refer to S09.

In this section we use Monte Carlo to calculate the equilibrium statistical mechanics of the shallow water equations, and we compare our predictions to unforced, inviscid solutions of the S09 model. The domain is similar to that in Section 3—a square ocean with the equator at mid-basin. However, in order to resolve the deformation radii, we take the domain size to be  $L = 2000$  km—half the size of the domain in Section 3. As in the quasigeostrophic case, we impose statistical symmetry about the equator; we assume that the average value of the circulation (4.6) vanishes. The remaining invariants are the mass (4.5), the energy (4.12), and the potential enstrophy (4.7). The Monte Carlo calculation applies the Metropolis algorithm to the probability distribution

$$P = \alpha_0 \exp(-\alpha_M M - \alpha_E E - \alpha_Z Z). \quad (4.13)$$

This requires only the finite-difference expressions for  $M$ ,  $E$  and  $Z$ . Thus, for Monte Carlo calculations, virtually any choice of fundamental variables would suffice. Although the conventional variables  $u$ ,  $v$ ,  $h$  would seem to be a natural choice, there are three reasons to prefer  $\chi$ ,  $\gamma$  and  $h$  as phase-space coordinates that define the state of the shallow water system.

First, as already explained, the shallow-water model derived in S09 is based upon  $\chi$ ,  $\gamma$  and  $h$ . By using these same variables for the Monte Carlo calculation we facilitate inter-comparison.

9. Griffa *et al.* (1996) investigated the inviscid dynamics of a much simpler shallow water model that conserved only the potential enstrophy.

Second, as noted in S82, the phase-space motion is nondivergent if the variables are taken to be  $hu$ ,  $hv$  and  $h$ . That is, uniformly gridded values of these variables obey the requirement (2.8). Since by (4.9) the transformation between  $hu$ ,  $hv$ ,  $h$  and  $\chi$ ,  $\gamma$ ,  $h$  is a linear one, the latter variables also obey (2.8). Said another way, volumes in  $hu$ - $hv$ - $h$  space are proportional to volumes in  $\chi$ - $\gamma$ - $h$  space, and are conserved in both spaces. This means that, in applying the Metropolis algorithm, one nominates new system states by selecting from distributions that are *uniformly* distributed in a local  $\chi$ - $\gamma$ - $h$  neighborhood, in a manner analogous to the nomination of  $\psi$ -states in Section 3.

Third, the choice  $\chi$ - $\gamma$ - $h$  represents a natural partition of phase space into three parts. For our Monte Carlo calculations we use  $100^2$  grid points. Thus, taking the boundary conditions on  $\chi$  and  $\gamma$  into account, we have  $98^2$  values of  $\chi$ ,  $98^2$  values of  $\gamma$ , and  $100^2$  values of  $h$ . The full phase space has  $98^2 \times 98^2 \times 100^2$  dimensions. However, by *imposing* the conditions  $h = H$  (constant) and  $\gamma = 0$ , we can apply the Metropolis algorithm to the  $98^2$ -dimensional *subspace* spanned by the  $\chi$ -values alone. In this case, the mass (4.5) is a trivial invariant, and—apart from irrelevant additive and multiplicative constants—the energy (4.4) and potential enstrophy reduce to the corresponding quasigeostrophic forms with  $\psi = \chi/H$ . Thus, for the  $\chi$ -space truncation, the Monte Carlo calculation reduces to the much-studied quasigeostrophic case reviewed in Sections 2 and 3.

At the next level of complexity, one can consider the equilibrium statistical mechanics of the  $\chi$ - $h$  system, obtained by maintaining the restriction  $\gamma = 0$ , but allowing  $\chi$  to take any value at nonboundary grid points, and  $h$  to take any non-negative value at all grid points. The advantage of the  $\chi$ - $h$  system over the  $\chi$  system—the quasigeostrophic system—is that the thermocline depth  $h$  can vary by an arbitrary amount. In particular,  $h$  can vanish. In the  $\chi$ - $h$  system, the mass takes the form (4.5), the energy takes the form

$$E = \iint d\mathbf{x} \left( \frac{\nabla\chi \cdot \nabla\chi}{2h} + \frac{1}{2}gh^2 \right) \quad (4.14)$$

and the potential enstrophy takes the form

$$Z = \iint d\mathbf{x} h^{-1} (\nabla \cdot (h^{-1} \nabla \chi) + f)^2. \quad (4.15)$$

Since neither (4.14) nor (4.15) is quadratic, the Monte Carlo method is indispensable.

As compared to the  $\chi$  system, the  $\chi$ - $h$  system corresponds to ‘lifting a barrier’ in the sense of thermodynamics. One can lift a second ‘barrier’—the ‘barrier’ that prevents excitation of  $\gamma$ —and consider the equilibrium statistical mechanics of the  $\chi$ - $\gamma$ - $h$  system, corresponding to the full shallow-water dynamics. However, it is not obvious that the full  $\chi$ - $\gamma$ - $h$  system has more to teach us than the  $\chi$ - $h$  system. To understand this point, first realize that the ‘barriers’ mentioned above are not fundamentally different than the ‘barrier’ that confines the energy, potential enstrophy and other invariants to a finite number of spatial degrees of freedom, i.e., the ‘barrier’ between our spatially truncated systems and the many degrees of freedom that reside in wavenumbers greater than the cutoff wavenumber  $k_{max}$ . Suppose that this cutoff

barrier too were lifted. In a real fluid there are still a finite number of degrees of freedom, but these now correspond to the locations and velocities of the molecules comprising the fluid. The equilibrium statistical mechanics of this molecular system is a rather uninteresting one; it is *heat death*, with no motion whatsoever on macroscopic scales. As is generally true in thermodynamics, interesting results are obtained only by considering systems in which barriers prevent global equilibria of this kind. Past studies of spatially truncated, quasigeostrophic systems with their twin barriers confining excitations to  $\chi$ -space on the one hand, and to macroscopic scales of motion on the other, have proved to be valuable: The inviscid, spatially-truncated, quasigeostrophic equilibrium states indicate the general direction towards which nonlinear interactions tend to drive the more realistic, viscous, untruncated system. The precise connection between these two systems—truncated and inviscid versus untruncated and viscous—is unavoidably ambiguous, but it clearly depends on the placement of ‘barriers’ that define the truncation. We come back to these points in Section 5. In the remainder of this section we apply the Monte Carlo method to the  $\chi$ - $h$  system and to the  $\chi$ - $\gamma$ - $h$  system.

To apply the Metropolis algorithm we specify values for  $\alpha_M$ ,  $\alpha_E$  and  $\alpha_Z$ . As in Section 3, we redefine  $M$ ,  $E$  and  $Z$  as the mass, energy and potential enstrophy per unit area of the whole domain. Consider, as a reference state, the state in which the fluid layer is motionless and covers the domain with a uniform depth  $h = H \equiv 250$  m. In this state of rest, the invariants take the values

$$M_0 = H, \quad E_0 = \frac{1}{2}gH^2, \quad Z_0 = \frac{1}{3}\beta^2 y_{max}^2 / H \quad (4.16)$$

where  $y_{max} = 1000$  km. We define the  $\alpha$ 's with respect to these values. We take

$$\alpha_M = \frac{n}{\theta_M M_0}, \quad \alpha_E = \frac{n}{\theta_E E_0}, \quad \alpha_Z = \frac{n}{\theta_Z Z_0} \quad (4.17)$$

where  $n = 100^2$  is the number of grid points, and  $\theta_M$ ,  $\theta_E$ ,  $\theta_Z$  are the ‘nondimensional temperatures’ corresponding to mass, energy and potential enstrophy, respectively. We choose order-one values for  $\theta_M$ ,  $\theta_E$ , and  $\theta_Z$ . Then, starting at the state of rest, we apply the Metropolis algorithms in *stages* that consist of many passes through the  $100^2$  grid points. In each pass, first  $\chi$ , then  $h$ , and then (sometimes)  $\gamma$  is perturbed at each grid point. The perturbations cover a range  $\pm\Delta\chi$ ,  $\pm\Delta h$ ,  $\pm\Delta\gamma$ , which are initially set at  $\Delta\chi = \Delta\gamma = 1$  Sverdrup (Sv),  $\Delta h = 100$  m, and then periodically adjusted to keep the acceptance rate near 50%. Since  $h$  can take only positive values, the Metropolis algorithm must be slightly modified from the form described in Section 3. The Appendix, which describes the needed modification, is an almost self-contained derivation of the Metropolis algorithm.

Table 1 summarizes the Monte Carlo calculations. The calculations differ only in the choice of ‘potential-enstrophy temperature’  $\theta_Z$  and the particular system considered: Experiments E1, E2, E3 and E4 correspond to the  $\chi$ - $h$  system; in these experiments  $\gamma \equiv 0$ . Experiments E5, E6, E7 and E8 correspond to the full  $\chi$ - $\gamma$ - $h$  system. All 8 Monte Carlo

Table 1. Summary of Monte Carlo solutions.

Experiment	E1	E2	E3	E4	E5	E6	E7	E8
Perturbs	$\chi, h$	$\chi, h$	$\chi, h$	$\chi, h$	$\chi, h, \gamma$	$\chi, h, \gamma$	$\chi, h, \gamma$	$\chi, h, \gamma$
$\theta_Z$	1	1/2	1/4	1/8	1	1/2	1/4	1/8
$h_{avg}$ (meters)	301	287	268	242	346	335	315	300
$\sqrt{\langle \mathbf{u}^2 \rangle}$ (km/day)	102	131	138	135	218	235	251	278
$\sqrt{\langle \mathbf{u} \rangle^2}$ (km/day)	94	122	129	127	82	110	127	175
$\langle \chi \rangle_{max}$ (Sv)	130	170	163	144	122	170	175	133
$\zeta_{rms}/f_{max}$	0.61	0.56	0.54	0.54	0.62	0.55	0.53	0.54
$\mu_{rms}/f_{max}$	0.77	1.12	1.65	2.18	10.8	11.1	12.0	14.8

experiments have ‘mass temperature’  $\theta_M = 1$  and ‘energy temperature’  $\theta_E = \pi$ . These values were arbitrarily selected<sup>10</sup> to give an average thermocline depth in the oceanographic range. The statistics reported in Table 1 correspond to the final stage of each calculation—the stage nearest equilibrium—containing 5.8 billion ensemble members. For each calculation, Table 1 gives the ensemble- and area-average depth  $h_{avg}$ , the rms velocity  $\sqrt{\langle \mathbf{u}^2 \rangle}$ , the mean-flow velocity  $\sqrt{\langle \mathbf{u} \rangle^2}$ , the mass transport  $\langle \chi \rangle_{max}$  in either gyre, the rms vorticity  $\zeta_{rms}$ , and the rms divergence  $\mu_{rms}$ . The latter two are normalized by the value of the Coriolis parameter at the northern boundary. Repeated calculations suggest that the numbers presented in Table 1 are accurate to within a few per cent.

If  $\theta_Z = \infty$  ( $\alpha_Z = 0$ ), then there is no constraint on potential vorticity and no equilibrium mean flow. The equilibrium statistical mechanics is that of nonrotating, compressible, two-dimensional turbulence. As  $\theta_Z$  decreases past unity, both the  $\chi$ - $h$  system and the  $\chi$ - $\gamma$ - $h$  system develop strong and very similar mean flows. As measured by the ratio  $U_{rel} \equiv \sqrt{\langle \mathbf{u} \rangle^2} / \sqrt{\langle \mathbf{u}^2 \rangle}$ , the strength of the mean flow increases from 0.92 to 0.94 as  $\theta_Z$  decreases from 1 to 1/8 in the  $\chi$ - $h$  system. In the corresponding  $\chi$ - $\gamma$ - $h$  calculations,  $U_{rel}$  increases from 0.38 to 0.62.

The principal difference between the  $\chi - h$  calculations (E1-4) and the  $\chi - h - \gamma$  calculations (E5-8) is that most of the kinetic energy in the latter resides in small-scale *divergent* motions that make little contribution to the mean flow. That is, as the ‘ $\gamma$ -barrier’ is lifted, the newly available region of phase space rapidly gains energy, and the rms divergence increases by an order of magnitude. In contrast, the rms vorticity hardly changes. The large difference between the rms vorticity and divergence in calculations E5-8 may be traced to the fact that vorticity—but not divergence—appears in the potential vorticity, and thus is constrained by potential enstrophy conservation. The change in the mean flow between E1-4 and E5-8—as measured by the velocity ratio  $U_{rel}$  or by the gyre transport  $\langle \chi \rangle_{max}$ —is relatively slight. Of course, instead of comparing calculations with the same values of  $\theta_M$

10. In the very special case  $\alpha_E = \alpha_Z = 0$  of no constraints on energy or potential enstrophy, the value  $\theta_M = 1$  corresponds to a spatially uniform average depth ( $h$ ) equal to the reference depth  $H$ . Similarly, when  $\alpha_M = \alpha_Z = 0$ , and when the kinetic energy term is omitted from the energy, the value  $\theta_E = \pi$  also corresponds ( $h$ ) =  $H$ . However, in the general case these choices have no special significance.

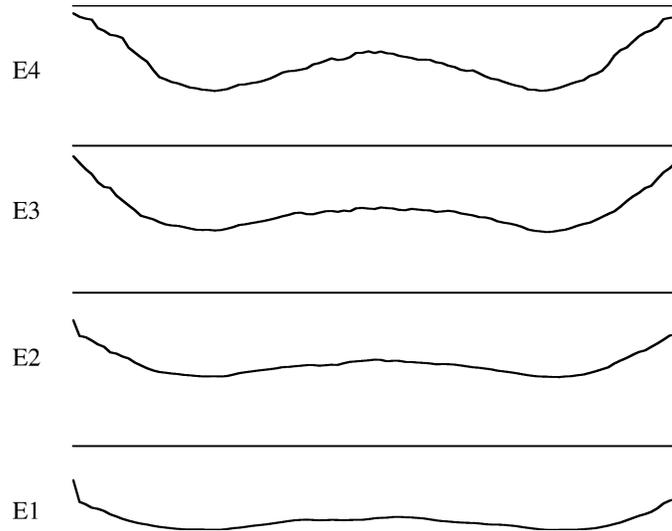


Figure 2. The ensemble- and longitude-averaged thermocline depth  $h$  in Monte Carlo calculations E1-E4, in which  $\chi$  and  $h$  vary freely but the velocity potential  $\gamma$  is constrained to vanish. In each panel, the horizontal line represents the ocean surface, and the curved line represents the main thermocline. The equator lies at the center. The average depth for each calculation is given in Table 1.

and  $\theta_E$ , one might prefer to compare calculations with the same values of mass and energy, but this would require an enormous number of independent Monte Carlo calculations to accurately determine the dependence of  $\theta_M$ ,  $\theta_E$ , and  $\theta_Z$  on  $M$ ,  $E$  and  $Z$ .

Figure 2 shows the ensemble-averaged thermocline depth  $\langle h \rangle$  in the  $\chi - h$  calculations (E1-4) after  $x$ -averaging over the middle half of the ocean basin, to exclude eastern and western boundary layers. In each panel of Figure 2, the horizontal line denotes the ocean surface, and the curve represents the average location of the boundary between isopycnal layers. Figure 3 shows the corresponding information for the  $\chi - h - \gamma$  calculations (E5-8). In every case, as  $\theta_Z$  decreases, the thermocline depth at the equator decreases; the maximum thermocline depth at mid-latitude increases; and the poleward extremities of the layer approach outcrop. In these features, the calculations strongly resemble the ocean's main thermocline. In fact, the general shape of the curves in Figures 2 and 3 strongly resembles the shapes of upper-ocean isopycnals in typical meridional sections throughout the world ocean.

Figures 4-7 show the mean streamfunction  $\langle \chi(x, y) \rangle$  and mean layer depth  $\langle h(x, y) \rangle$  for all 8 Monte Carlo calculations. At the smallest values of  $\theta_Z$ , the flow separates from the northern and southern boundaries in E3 and E4. In this region, the average layer depth is only a few meters.<sup>11</sup> However, the eastern and western boundary currents extend all the

11. These regions of very small layer depth correspond to actual outcrops of the abyssal layer. Average values of zero cannot occur in the Monte Carlo method because the Metropolis algorithm nominates only non-negative  $h$ .

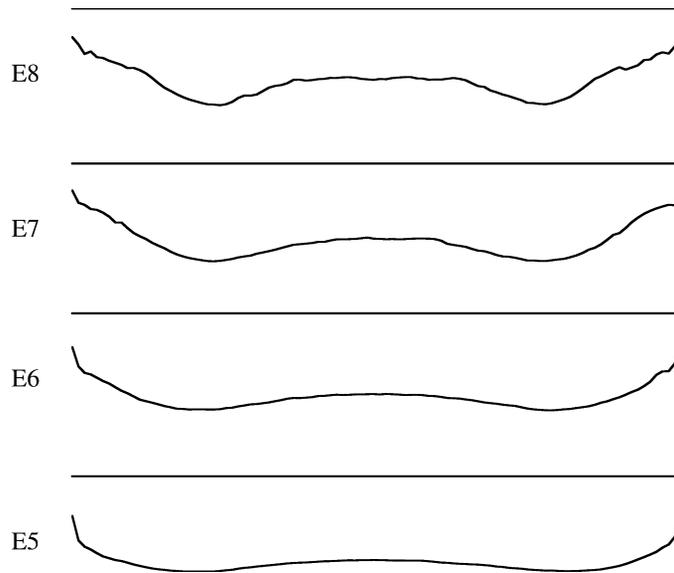


Figure 3. As in Figure 2, but for Monte Carlo calculations E5-8, in which  $\chi$ ,  $h$ , and  $\gamma$  all vary freely.

way to the corners of the domain. This behavior is clearly associated with the extension of the fluid layer into the corners, which is especially notable in E3, E4 (Fig. 6) and E8 (Fig. 7). The noisy behavior of  $\langle h(x, y) \rangle$  in the equatorial region of E8 (Figs. 5 and 7) seems to be caused by incomplete convergence of the Metropolis algorithm in this region. In fact, despite the size of the ensemble, which required several days computing time on a 2.6 GHz processor,  $\langle h(x, y) \rangle$  is everywhere quite noisy, especially along the ‘outcrop’ lines at which the layer depth nearly vanishes.

This noisiness is understandable when one considers the difficulties facing Monte Carlo calculations in which  $h$  nearly vanishes. As  $h \rightarrow 0$  the potential vorticity  $q$  blows up unless  $\zeta$  experiences a compensating change in the numerator of (4.2). In the Metropolis algorithm, this subtle compensating change must occur *by chance*. This at least partly explains why the shallow-water Monte Carlo calculations require so much more computing time, and are still so much noisier, than the quasigeostrophic calculations of Section 3. This agony could be avoided by using the potential vorticity itself as one of the fundamental dynamical variables to be perturbed, and by nominating only *finite* values of the potential vorticity. Unfortunately, changing the potential vorticity at a single gridpoint induces a change in the velocity field throughout the domain. Determining the new velocity would require the solution of an elliptic problem at *every* step of the Monte Carlo calculation. This strategy was briefly tried and found to be impractical. However, modifications of the strategy, in which iterations in the solution of the elliptic equation are somehow interleaved with the Metropolis iterations, may eventually prove successful. We come back to these issues in the following section.

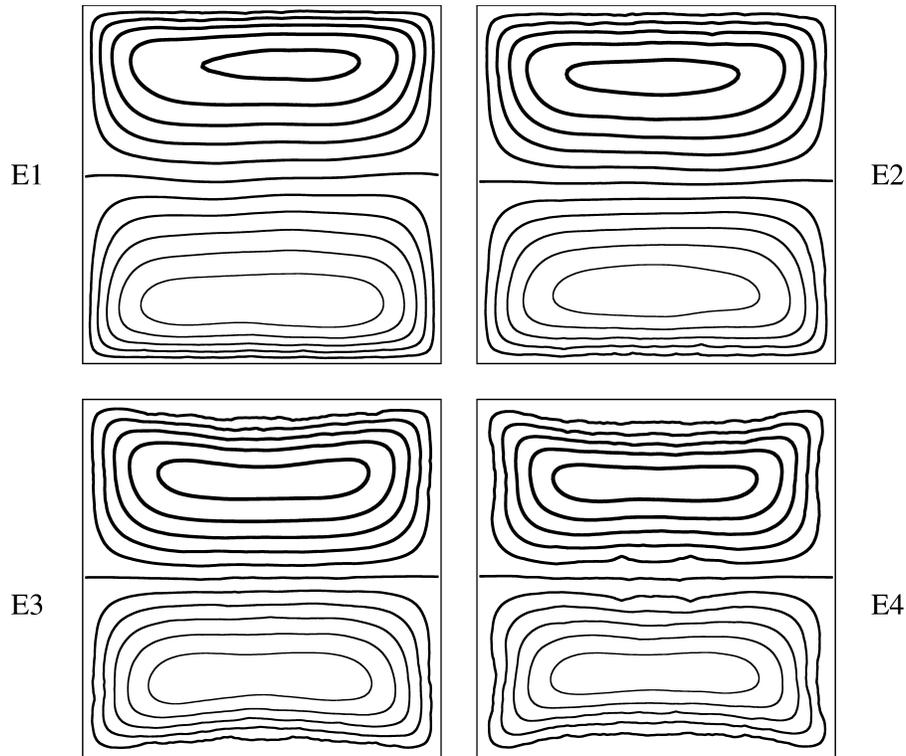


Figure 4. The transport streamfunction  $\chi$  in Monte Carlo calculations E1-4. The maximum transport values are given in Table 1. The corresponding velocity field, which is symmetric about the equator, resembles Fofonoff flow. Darker contours correspond to larger values.

Finally, we examine direct numerical simulations of the shallow water equations using the S09 model. This model comprises a set of coupled ordinary differential equations, analogous to (4.8–11), in which all the space derivatives—but not the time derivatives—are replaced by finite differences. These coupled ordinary differential equations *exactly* conserve finite-difference analogs of (4.5–7) and (4.12). To solve the S09 equations, we use an adaptive fourth-order Runge-Kutta method for the time step, and we use a multi-grid method to solve the analogs of the elliptic equations (4.10). Both of these methods introduce errors that can be made as small as desired. In the experiment to be discussed, the largest errors were caused by incomplete convergence of the multi-grid solver. However, the mass, potential enstrophy and energy were nevertheless conserved to within a few per cent over a time interval comprising 800,000 time steps, corresponding to a period of 3200 simulated days and 19 hours of computer time. Even smaller errors can be achieved by increasing the number of iterations in the multi-grid solver, but the cost in computing time is considerable. For a complete description of the S09 model, refer to S09.

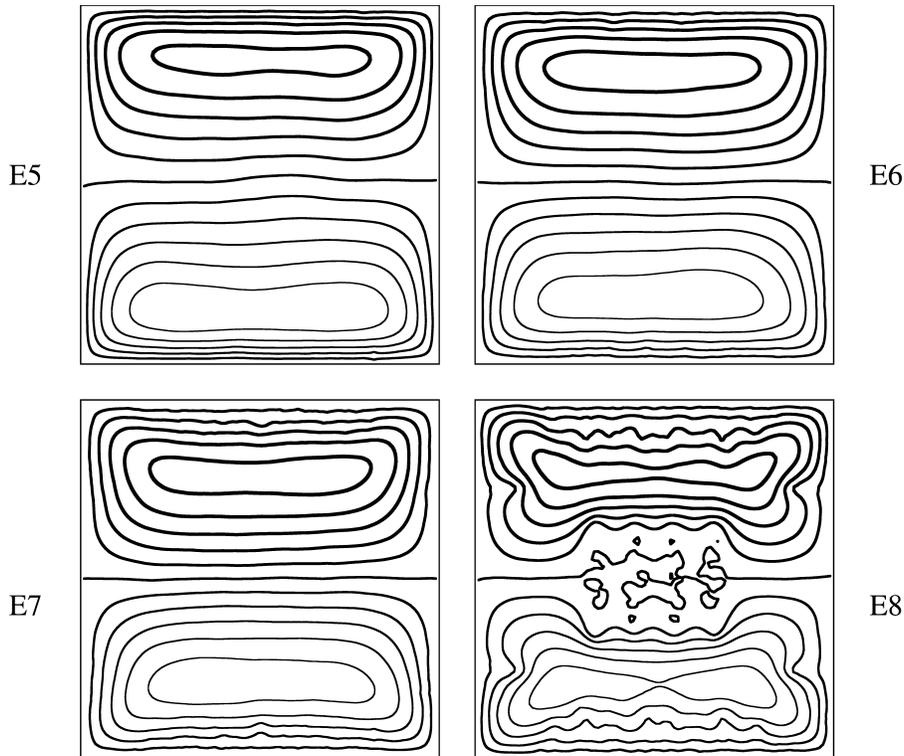


Figure 5. The same as Figure 4, but for Monte Carlo calculations E5-8.

Figures 8 and 9 show the time-average streamfunction  $\chi(x, y)$  and layer depth  $h$  in a shallow-water simulation with  $64^2$  grid points and random initial conditions. These averages correspond to the last 3200 days of a simulation that lasted 4000 days. This simulation is much less energetic than the Monte Carlo cases considered above; its rms velocity of  $20 \text{ km day}^{-1}$  and its average maximum transport of  $42 \text{ Sv}$  are less than half the values given in Table 1. However, because difficulties arise when  $h$  approaches zero at any point within the flow, the direct numerical simulation of Figures 8 and 9 was the most energetic inviscid calculation that could be achieved. We digress to explain the difficulties.

Although the S09 model conserves energy exactly, the model can still blow up, because, if  $h$  becomes even slightly negative, then the kinetic energy in (4.4) becomes negative and the kinetic and potential energy may simultaneously diverge, even as the total energy remains constant. It was hoped to counteract this catastrophe by introducing an additional ‘surface energy’ term into (4.4) and (4.12) that becomes very large when the layer depth falls below a very small positive value. This strategy was used successfully in viscous simulations by Salmon (2002) and Primeau and Newman (2007). In the present inviscid context, the surface-energy strategy offers the essential advantage that it does not affect

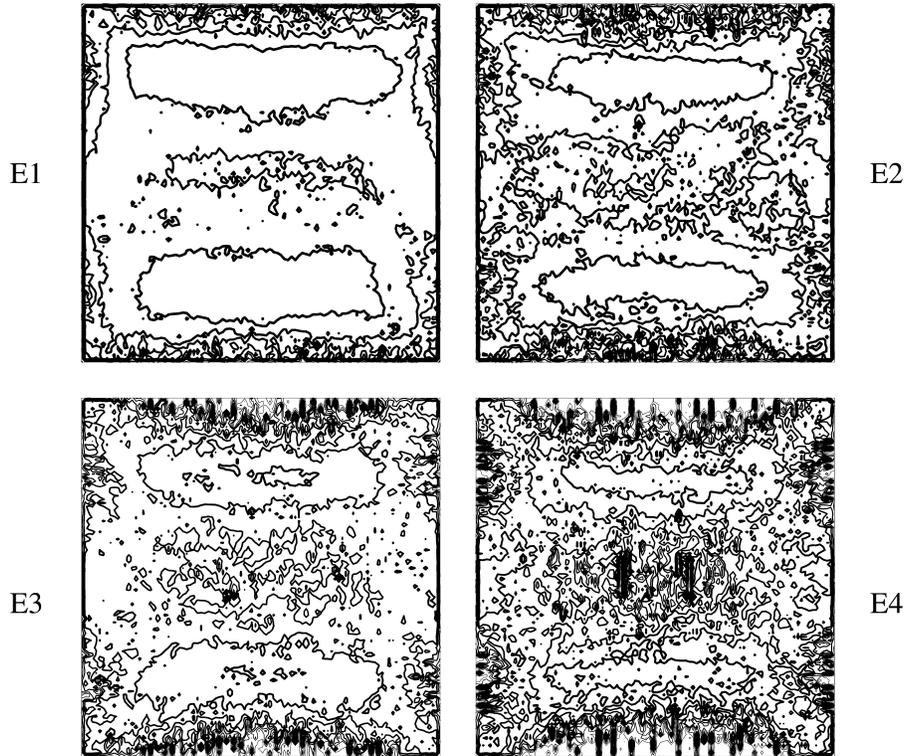


Figure 6. The average thermocline depth  $\langle h(x, y) \rangle$  in Monte Carlo calculations E1-4. Maximum values are, respectively, 353 m, 362 m, 489 m, and 460 m. Darker contours correspond to larger values.

the conservation properties of the S09 model. Inviscid simulations with the surface-energy term did indeed avoid negative  $h$  but at a prohibitively high cost: The system became so stiff, and the adaptive time step so short, that the evolution essentially ground to a halt. In the simulation corresponding to Figures 8 and 9, which again had no surface energy, the instantaneous layer depth was at times perilously close to zero, but the average layer depth in Figure 9 varies only between 452 m and 526 m.

Despite this disappointment, the time-average flow of Figures 8 and 9 supports the general prediction of statistical mechanics and the Monte Carlo calculations that the zeroth-order flow in the ocean's surface layer—a westward drift at low latitudes associated with poleward deepening of the thermocline, and a compensating eastward flow near the latitudes at which the warm layer depth vanishes—requires no *particular* explanation at all. It is simply the average of all possible states with the given average values of mass, energy, and potential enstrophy.

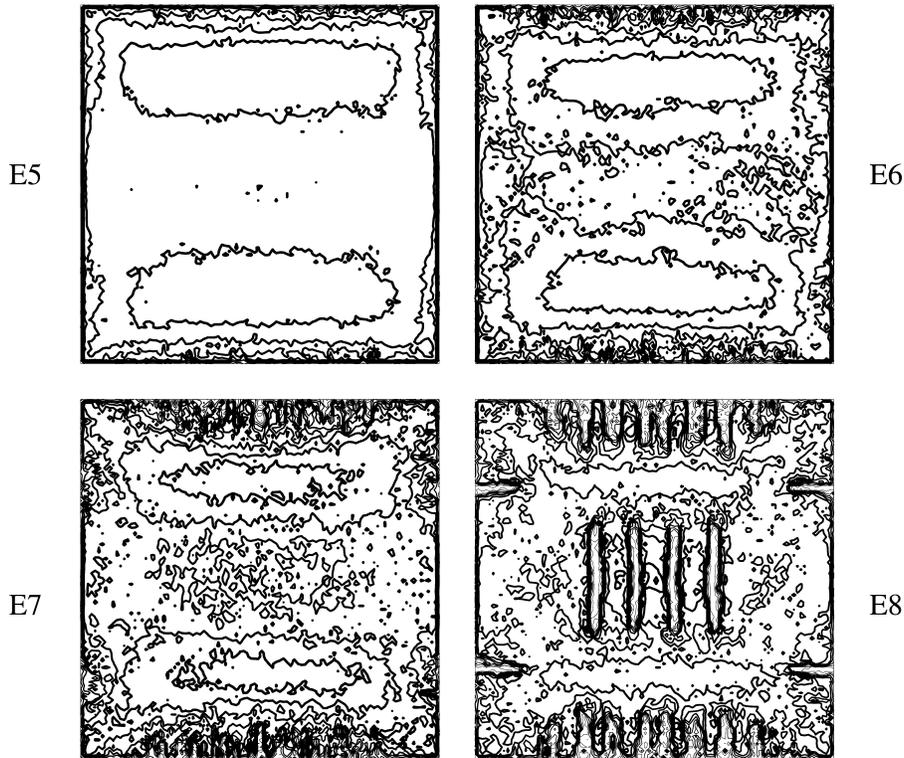


Figure 7. The same as Figure 6, but for Monte Carlo calculations E5-8. The maximum values of average layer depth are, respectively, 403 m, 412 m, 428 m, and 510 m.

## 5. Discussion

In at least one sense these results are surprising. In the quasigeostrophic case, the connection between the streamfunction and the layer depth is built into the model and, most particularly, into the conservation laws (2.3) and (2.6). That is, the quasigeostrophic energy and potential enstrophy *know* about geostrophic balance. Thus, although the quasigeostrophic phase space has only one-third the dimensions of the full shallow-water phase space, its conservation laws incorporate more physics. The statistical mechanics of the shallow-water system depend only on the forms of the mass (4.5), the potential enstrophy (4.7) and the energy (4.12). The fields  $\chi$ ,  $h$  and  $\gamma$  may assume any configuration consistent with the values of these invariants. In particular, there is no imposed connection between the streamfunction  $\chi$  and the fluid depth  $h$ ; the shallow-water Monte Carlo calculation knows nothing at all about geostrophy. At the beginning of this study, it was suspected that this defect might be fatal, and that there might be no connection between the statistical mechanics of the shallow water system and either the statistical mechanics of the quasigeostrophic

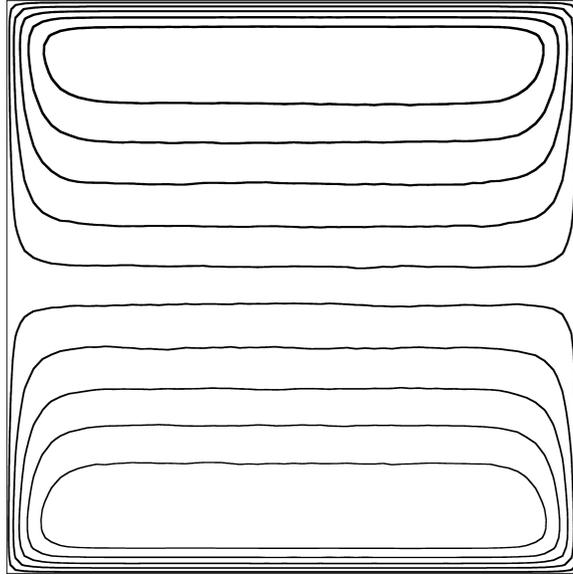


Figure 8. The transport streamfunction  $\chi(x, y)$  averaged over the final 3200 days of a direct numerical simulation of the shallow water equations with random initial conditions. The simulation conserved mass, energy, and potential enstrophy to within a few per cent. The transport of either gyre is 42 Sverdrups. Darker contours correspond to larger values.

system or shallow water inviscid dynamics. The author has still not quite recovered from his surprise that all three yield similar mean flows.

It is still probably worthwhile to investigate the statistical mechanics of a model with fewer independent variables than the shallow water equations, but whose conservation laws incorporate more of the physics. This is the territory of *balanced models* like the  $L_1$  model proposed by Salmon (1985).<sup>12</sup> All balanced models of the shallow water equations incorporate *two* balance conditions to remove the *two* fast inertia-gravity modes from the physics; the remaining physics contains only the one slow Rossby mode, and is most conveniently presented as a single evolution equation for potential vorticity. The balance conditions require the solution of elliptic equations at every time step. The balance conditions in quasigeostrophy and in  $L_1$  are geostrophic balance, but geostrophic balance is not appropriate for flow in the vicinity of the equator. Therefore, the first step in a Monte Carlo calculation based upon a balanced model would be to find physically defensible balance conditions for flow that spans the equator. The assumption  $\gamma = 0$  that the mass transport have no divergence, used for the  $\chi - h$  system in Section 4, is one plausible balance condition, but a second condition relating  $\chi$  and  $h$  is also required. If this second balance

12. Strictly speaking, the quasigeostrophic system is a balanced model, but its key assumption—that the layer depth be nearly uniform—makes it useless for most oceanographic purposes.

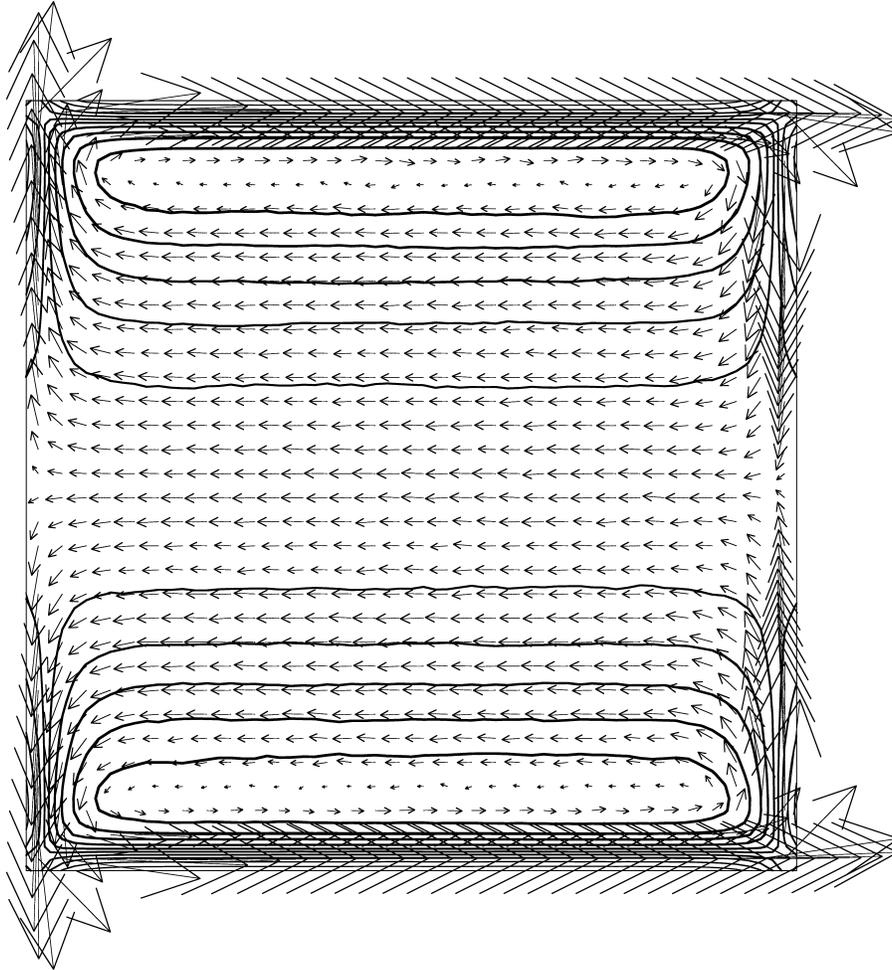


Figure 9. The time-average of the layer depth  $h(x, y)$  in the same simulation of the shallow water equations as depicted in Figure 8. The average depth varies from 452 m to 526 m. Arrows denote the velocity, which has an rms value of 20 km/day and a maximum value of 77 km/day.

condition involves an elliptic equation (as seems likely) then the Monte Carlo method might well become impractical. These matters deserve much further study.

What precisely do our calculations mean? As in every case, the ideal equilibrium states represent target states towards which nonlinear interactions, acting by themselves, would tend to drive the flow. To the extent that flows with realistic forcing and dissipation resemble these ideal states, equilibrium statistical mechanics offers a kind of generic explanation. For example, the resemblance between Figures 2 and 3 and the observed shape of the main thermocline plausibly means that the observed thermocline shape requires no very particular

explanation; it is simply the average shape that results from the system's tendency to explore all the states consistent with its mass, energy and potential enstrophy. Numerical experiments with, say, a very particular forcing might also yield this observed thermocline shape, thus possibly creating the misconception that the thermocline shape depends critically on that very particular forcing. However, the generic explanation offered by equilibrium statistical mechanics suggests otherwise; the system always seeks the equilibrium state.

Of course, many features predicted by equilibrium statistical mechanics are never attained by even weakly dissipative systems. For example, the wavenumber spectra predicted by equilibrium statistical mechanics typically exhibit an 'ultra-violet catastrophe' and bear no resemblance to observed wavenumber spectra. The latter require theories—such as Kolmogorov theory—in which the forcing and dissipation assume a leading-order importance. Unfortunately, the search for a complete theory of turbulence—a *nonequilibrium* statistical mechanics—that would encompass all statistics in all cases of forcing and dissipation has yet proved largely unsuccessful.

In any case, previous applications of equilibrium statistical mechanics to quasigeostrophic models have yielded tantalizing results. These include the emergence of anticyclonic flow around islands and seamounts, and the almost complete 'barotropization' of two-layer quasigeostrophic flow on scales larger than the internal deformation radius. One especially interesting—and possibly unappreciated—result is the theory's prediction that quasigeostrophic turbulence transfers its energy toward the equator *and into high vertical mode*. More precisely, the equilibrium state is one in which the kinetic energy in each vertical mode peaks at the equator. The width of each peak equals the equatorial deformation radius for that mode. All the peaks have equal height, corresponding to energy equipartition among vertical modes at the equator. Thus equilibrium statistical mechanics offers a generic explanation for the existence of the deep equatorial jets. For details, see Salmon (1982b, pp. 55–63). The same features can be understood more simply as resulting from the tendency for energy to flow toward the lowest 'total wavenumber' of the system (Salmon 1982b, p. 47; Salmon, 1998, p. 284).

The present paper is among the first to apply equilibrium statistical mechanics to the shallow water equations. It seems possible that future applications to more complicated, multi-layer, shallow water systems could yield new results as interesting as those previously obtained for quasigeostrophic systems.

*Acknowledgments.* This work was supported by National Science Foundation grant OCE-0542890.

## APPENDIX

### *Metropolis algorithm applied to the fluid depth*

In applying the Metropolis algorithm, we separately nominate new values of  $\chi$ ,  $\gamma$  and  $h$  at each grid point. Apart from boundary conditions, the variables  $\chi$  and  $\gamma$  can assume any values. Hence we nominate new values, with uniform probability, from fixed ranges

centered on the current values of  $\chi$  and  $\gamma$ , as described for  $\psi$  in Section 3. However, the fluid depth  $h$  can assume only positive values. This requires the following modification to the algorithm.

Suppose that we are considering a transition in which the value of  $h$  at a particular grid point is changed from  $h_1$  to  $h_2$ . Let  $P_1$  and  $P_2$  be the canonical probability densities in the system states corresponding to  $h_1$  and  $h_2$ , respectively. Let  $P(1 \rightarrow 2)$  be the probability of transition from state 1 to 2, and let  $P(2 \rightarrow 1)$  be the probability of transition from state 2 to 1. By the principle of detailed balance,

$$P_1 P(1 \rightarrow 2) = P_2 P(2 \rightarrow 1) \quad (\text{A.1})$$

at equilibrium. However,

$$P(1 \rightarrow 2) = N(1 \rightarrow 2) A(1 \rightarrow 2) \quad (\text{A.2})$$

where, *given* that we are in state 1,  $N(1 \rightarrow 2)$  is the probability of *nominating* state 2, and  $A(1 \rightarrow 2)$  is the probability of *accepting* state 2 as the next ensemble member. Similarly,

$$P(2 \rightarrow 1) = N(2 \rightarrow 1) A(2 \rightarrow 1). \quad (\text{A.3})$$

Substituting (A.2) and (A.3) into (A.1), we obtain

$$\frac{A(1 \rightarrow 2)}{A(2 \rightarrow 1)} = \frac{P_2 N(2 \rightarrow 1)}{P_1 N(1 \rightarrow 2)} \equiv \mathcal{P}. \quad (\text{A.4})$$

We satisfy (A.4) by adopting the rules

$$A(1 \rightarrow 2) = 1 \quad \text{if } \mathcal{P} > 1 \quad (\text{A.5a})$$

$$A(1 \rightarrow 2) = \mathcal{P} \quad \text{if } \mathcal{P} < 1. \quad (\text{A.5b})$$

If, as in the case of  $\chi$  and  $\gamma$ , nominations are made from an interval of *fixed* width centered on the current value, then  $N(2 \rightarrow 1) = N(1 \rightarrow 2)$ , and (A.5) reduces to the verbal explanation of the Metropolis algorithm given in Section 3, with  $\mathcal{P} = P_2/P_1$ . However, since  $h$  cannot be negative, we nominate  $h_2$ , uniformly, from the range  $[\max(0, h_1 - \Delta h), h_1 + \Delta h]$  and vice versa. Since the probability of nominating a particular state is inversely proportional to the range from which it is selected, we have

$$\frac{N(1 \rightarrow 2)}{N(2 \rightarrow 1)} = \frac{h_2 + \Delta h - \max(0, h_2 - \Delta h)}{h_1 + \Delta h - \max(0, h_1 - \Delta h)} = \frac{\Delta h + \min(h_2, \Delta h)}{\Delta h + \min(h_1, \Delta h)}. \quad (\text{A.6})$$

If both  $h_1$  and  $h_2$  are greater than  $\Delta h$  then (A.6) is unity. In general, we apply the rules (A.5) with

$$\mathcal{P} = \frac{P_2 (\Delta h + \min(h_1, \Delta h))}{P_1 (\Delta h + \min(h_2, \Delta h))}. \quad (\text{A.7})$$

The foregoing is an almost self-contained derivation of the Metropolis algorithm. For further details, see Kalos and Whitlock (2008, pp. 64–72).

## REFERENCES

- Arnol'd, V. I. 1965. On conditions for nonlinear stability of plane stationary curvilinear flows of an ideal fluid, *Sov. Math. Dokl.*, *6*, 773–777.
- 1969. On an *a priori* estimate in the theory of hydrodynamical stability. *Am. Math. Soc. Trans.*, *79*, 267–269.
- Bühler, O. 2009. *Waves and Mean Flows*, Cambridge Univ. Press, 338 pp.
- Carnevale, G. and J. D. Frederiksen. 1987. Nonlinear stability and statistical mechanics of flow over topography. *J. Fluid Mech.*, *175*, 157–181.
- Dukowicz, J. K. and R. J. Greatbatch. 1999. Evolution of mean-flow Fofonoff gyres in barotropic quasigeostrophic turbulence. *J. Phys. Oceanogr.*, *29*, 1832–1852.
- Fofonoff, N. P. 1954. Steady flow in a frictionless homogeneous ocean. *J. Mar. Res.*, *13*, 254–262.
- Griffa, A. and R. Salmon. 1989. Wind-driven ocean circulation and equilibrium statistical mechanics. *J. Mar. Res.*, *47*, 457–492.
- Griffa, A., E. P. Chassignet, V. Coles, and D. B. Olson. 1996. Inertial gyre solutions from a primitive equation ocean model. *J. Mar. Res.*, *54*, 653–677.
- Holloway, G. 1986. Eddies, waves, circulation and mixing: Statistical geofluid mechanics. *Ann. Rev. Fluid Mech.*, *18*, 91–147.
- Kalos, M. H. and P. A. Whitlock. 2008. *Monte Carlo Methods*, 2nd ed., Wiley-VCH, Weinheim, Germany, 197 pp.
- Metropolis, N., A. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, *21*, 1087–1092.
- Primeau, F. W. and D. Newman. 2007. Bifurcation structure of a wind-driven shallow water model with layer-outcropping. *Ocean Model.*, *16*, 250–263.
- Salmon, R. 1982a. [S82] The shape of the main thermocline. *J. Phys. Oceanogr.*, *12*, 1458–1479.
- 1982b. Geostrophic turbulence, *in* *Topics in Ocean Physics*, Corso LXXX Enrico Fermi School, A. R. Osborne and P. M. Rizzoli, eds., North Holland, Amsterdam.
- 1985. New equations for nearly geostrophic flow. *J. Fluid Mech.*, *153*, 461–477.
- 1998. *Lectures on Geophysical Fluid Dynamics*, Oxford, NY, 362 pp.
- 2002. Numerical solution of the two-layer shallow water equations with bottom topography. *J. Mar. Res.*, *60*, 605–638.
- 2009. [S09] A shallow water model conserving energy and potential enstrophy in the presence of boundaries. *J. Mar. Res.*, *67*, 779–814.
- Salmon, R., G. Holloway and M. C. Hendershott. 1976. The equilibrium statistical mechanics of simple quasigeostrophic models. *J. Fluid Mech.*, *75*, 691–703.
- Salmon, R. and L. D. Talley. 1989. Generalization's of Arakawa's Jacobian. *J. Comput. Phys.*, *83*, 247–259.
- Veronis, G. 1970. Effect of fluctuating winds on ocean circulation. *Deep-Sea Res.*, *17*, 421–434.
- Wang, J. and G. K. Vallis. 1994. Emergence of Fofonoff states in inviscid and viscous ocean circulation. *J. Mar. Res.*, *52*, 83–127.

Received: 10 February, 2010; revised: 6 April, 2010.