Tellus (1994), 46A, 671–680
Printed in Belgium – all rights reserved

Copyright © Munksgaard, 1994
TELLUS
ISSN 0280–6495

A simple model for the three-dimensional, thermally and wind-driven ocean circulation

By LEO R. M. MAAS, *Centre for Earth and Ocean Research, University of Victoria*, Victoria, BC, Canada*

(Manuscript received 6 August 1993; in final form 2 December 1993)

ABSTRACT

As a generalization to box models of the large-scale, thermally and wind-driven ocean circulation, nonlinear equations, describing the evolution of two vectors characterizing the state of the ocean, are derived for a rectangular ocean on an f-plane. These state vectors represent the basin-averaged density gradient and the overall angular momentum vector of the ocean. Neglecting rotation, the Howard-Malkus loop oscillation is retrieved, governed by the Lorenz equations. This has the equations employed in box models, in the restricted sense where no distinction is made between the restoring time scales of the temperature and salinity fields, as a special case. In another approximation, with rotation included, the equations are equivalent to a set E. N. Lorenz introduced to describe the "simplest possible atmospheric general circulation model". Although the atmospheric circulation may be chaotic, parameter values in the ocean are such that the circulation is steady or, at most, exhibits a self-sustained oscillation. For a purely thermally forced ocean, this is always a unique state. Addition of a wind-induced horizontal circulation allows for multiple equilibria, despite the neglect of the salinity field.

1. Introduction

Following Stommel (1961), the use of box models to describe aspects of the large-scale circulation in the oceans has seen a revival over the last few years (Welander, 1986; Weaver and Hughes, 1992). Stommel's work was concerned with the possible existence of multiple equilibria in a two-box model (crudely representing the equatorial and polar regions) of a single hemisphere ocean due to a difference in restoring time scales of the temperature and salinity fields. The equilibria are characterized by a different sense of the meridional circulation. Multiple equilibria are also traceable in many of the more complicated two- and three-dimensional numerical models (e.g., Bryan, 1986; Manabe and Stouffer, 1988; Marotzke et al., 1988; Weaver et al., 1993). Much more elaborate box models have been proposed since (Huang and Stommel, 1992; Thual and McWilliams, 1992), but these retain the ad hoc nature of the simpler box-models. Moreover, these models, as well as other analytical models (Källén and Huang, 1987; Cessi and Young, 1992), seem incapable of incorporating the Coriolis force in their description. It is the purpose of this paper to show how both of these difficulties can be overcome by generalizing box models, not by adding more of them, but by deriving the governing equations directly from the equations of motion.

To see how this can be achieved it is useful to formulate an even simpler version of Stommel's (1961) model, by assuming that the temperature and salinity fields have the same restoring coefficients (or neglecting salinity altogether). His model of the thermohaline circulation would then read:

$$\frac{\mathrm{d}\,\Delta\rho^{(y)}}{\mathrm{d}t} = F - c\,\Delta\rho^{(y)} - |q|\,\Delta\rho^{(y)}, \tag{1a}$$

$$q = \lambda\,\Delta\rho^{(y)}, \tag{1b}$$

* On leave from The Netherlands Institute for Sea Research, PO Box 59, 1790 AB Texel, Netherlands.

where $\Delta\rho^{(y)}$ signifies the north–south (y) density difference in the meridional ($y-z$) plane, F denotes the difference in density (buoyancy) flux between both boxes and c is a "restoring coefficient" (related to diffusive processes). The last term describes the advective exchange, with q denoting the (scaled) flow rate between the boxes (and λ a Rayleigh damping term). Stommel reasoned that the absolute value of this quantity should appear, as it does not matter to the actual exchange which sense the circulation has. The second equation constitutes the "dynamics" in Stommel's model, which may, in fact, be recognized as a relic of more general vorticity equations:

$$\cdots + g\frac{\partial \rho}{\partial y} = -r\omega_1 + \cdots,$$
$$\vdots$$

since the flow q is, due to its confinement, always of a differential nature and therefore, when scaled with a proper length scale, equivalent to the x-component of the vorticity vector, ω_1. Here, g denotes the acceleration of gravity and r is a friction coefficient Conspicuously absent in this simple model, however, is the Coriolis term, which should couple the dynamics in the meridional plane to that in the zonal plane, i.e., to ω_2. At the same time it seems desirable to allow the density field to develop an east–west and top–bottom contrast, i.e., to consider evolution equations, like (1a), for $\Delta\rho^{(x)}$ and $\Delta\rho^{(z)}$, which will be changed by "differential advection". However, efforts to formulate an extension of Stommel's original (1961) model exactly along these lines are tedious and somewhat superficial. An alternative and perhaps more direct formulation of the problem can be cast in terms of the "lowest (non-vanishing) moments" of the momentum and density fields.

2. Evolution of basin-averaged angular momentum and density gradient

2.1. *Evolution of angular momentum*

Consider an idealized, rectangular ocean basin, of sizes L, B and H in the x, y and z-directions (oriented along the east, north and upward direction) respectively. The basin is assumed to have a rigid lid, any free surface displacements being translated into corresponding pressure variations at the lid. Consequently, the volume and any other geometric factors characterizing the basin are fixed quantities. Choose the origin of the coordinate system to be in the geometric centre of the basin. Let the motion of the fluid be described on the f-plane by the Navier-Stokes equations in Boussinesq approximation, where f denotes the Coriolis parameter. The lowest moment of the momentum field is, in an enclosed ocean, given by the basin-averaged angular momentum vector, which, in Boussinesq approximation (and per unit mass), is defined with respect to Cartesian axes as

$$\boldsymbol{L} = \frac{1}{V}\int \boldsymbol{x}\times\boldsymbol{u}\,\mathrm{d}V.$$

Here V denotes the volume of the basin (and $\mathrm{d}V$ a volume increment) and $\boldsymbol{u}$ the velocity vector. Note that the smallness of the magnitude of the vertical velocity in a small aspect-ratio basin like the ocean is offset by the larger arm over which it extends and contributes as much to the angular momentum as does the horizontal flow.

Since the boundaries are fixed, time-differentiation and integration over the basin can be exchanged. Hence, taking the cross-product of $\boldsymbol{x}$ with the momentum equations and subsequently integrating over the entire basin yields:

$$\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} + \frac{f}{2}\boldsymbol{k}\times\boldsymbol{L} = -\frac{1}{V}\int \nabla\times(\boldsymbol{x}P)\,\mathrm{d}V$$
$$+\frac{1}{V}\int g'z\int_{H/2}^{z}[-\rho_y\boldsymbol{i}+\rho_x\boldsymbol{j}]\,\mathrm{d}z'\,\mathrm{d}V$$
$$+\frac{1}{V}\int \boldsymbol{x}\times\nabla\cdot(\boldsymbol{A}\,\nabla\boldsymbol{u})\,\mathrm{d}V. \qquad (2)$$

Here,

$$P = p + \int_{H/2}^{z}\rho g'\,\mathrm{d}z', \qquad (3)$$

describes the non-hydrostatic and free-surface contributions to the pressure field, with perturbation pressure and density defined as $p=(p_*-p_0)/p_0$ and $\rho=(\rho_*-\rho_0)/\delta\rho$, in which the starred quantities denote the original, dimensional quantities and ρ_0 is the constant reference density with which

a pressure p_0 is hydrostatically related. The scale of the density perturbations, $\delta\rho$ ($\ll\rho_0$), which will be made more explicit later, has been used to render ρ an O(1) quantity and has consequently been combined with g into a reduced gravity $g' = g\,\delta\rho/\rho_0$. Also, $\boldsymbol{i}, \boldsymbol{j}$ and $\boldsymbol{k}$ denote unit vectors in the x, y and z-directions and $\boldsymbol{A}$ the eddy viscosity tensor. The basin-averaged angular momentum is thus changed by the Coriolis torque, the pressure torque, the buoyancy torque and the Reynolds stress torque respectively. Note that nonlinear advective terms cancel identically for the basin-averaged angular momentum vector.

The pressure torque will be assumed to vanish. The integral containing this term can be rewritten in terms of boundary integrals, but for the symmetric (f-plane) model adopted here, is not likely to contribute a net torque if only the response to buoyancy forcing is considered. In case a wind stress is applied a net pressure torque, due to divergences in Ekman transport, is likely to be set up. However, it seems to require more complicated wind patterns than the uniform, or simply sheared wind-stress fields considered in the following. Torques due to the Reynolds stress can again be partially integrated and, at the surface, are related to torques of the wind stress, τ:

$$\boldsymbol{T} = \frac{1}{V}\int\int\left(-\frac{H}{2}\tau^{(y)}, \frac{H}{2}\tau^{(x)}, x\tau^{(y)} - y\tau^{(x)}\right)\mathrm{d}x\,\mathrm{d}y. \tag{4}$$

The other terms contain torques due to friction at the rigid walls and internal friction. Evaluation of the former usually requires some knowledge of the internal dynamics of the ocean as the shear terms at the boundaries are determined by the boundary layer physics (Wright and Stocker, 1991; Wright and Vreugdenhil, 1994), but the aggregate of all frictional terms is here interpreted as yielding a general Rayleigh damping term, $-(r_h L_1, r_h L_2, r_v L_3)$, with $r_{h,v}$ friction coefficients.

The buoyancy torque can be evaluated by employing a "global" Taylor expansion:

$$\rho = x\bar{\rho}_x(t) + y\bar{\rho}_y(t) + z\bar{\rho}_z(t) + (x^2 - L^2/3)\,\bar{\rho}_{xx}(t) + \cdots, \tag{5}$$

where the basin-averaged density gradients, or moments of the density field, are defined consistently as

$$\bar{\rho}_x(t) = \frac{\int x\rho\,\mathrm{d}V}{\int x^2\,\mathrm{d}V}, \qquad \bar{\rho}_y(t) = \frac{\int y\rho\,\mathrm{d}V}{\int y^2\,\mathrm{d}V}, \quad \text{etc.,} \tag{6}$$

which are clearly related to the coordinates of the centre of mass. Note that the expansion is chosen such that higher order moments do not contribute to the lower order moments, so that the evolution equations for the latter may be safely truncated in any desired order. Presently we will neglect terms beyond the first order and hence approximate the isopycnal field by a series of parallel planes. With representation (5) the buoyancy torque in eq. (2) evaluates to

$$g'/V\int z^2\,\mathrm{d}V[-\bar{\rho}_y(t)\,\boldsymbol{i} + \bar{\rho}_x(t)\,\boldsymbol{j}], \tag{7}$$

where the remaining integral is just a constant geometrical factor.

With these considerations, the evolution equation of the angular momentum vector is given by

$$\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} + \frac{f}{2}\boldsymbol{k}\times\boldsymbol{L} = g'/V\int z^2\,\mathrm{d}V[-\bar{\rho}_y(t)\,\boldsymbol{i} + \bar{\rho}_x(t)\,\boldsymbol{j}] - (r_h L_1, r_h L_2, r_v L_3) + \boldsymbol{T}. \tag{8}$$

2.2. *Evolution of density gradients*

An evaluation of the buoyancy torque (7) obviously requires a consideration of the evolution equations of the basin-averaged density gradients. These are obtained by multiplying the equation which describes the evolution of the density field with x and averaging over the basin. Assuming that the eddy diffusivities are constants, albeit of different magnitude in the horizontal (K_h) and vertical (K_v) direction, this equation, after a partial integration and employing the no-flux condition at the solid boundaries, reduces with (5) to

$$\frac{1}{V}\int x^2\,\mathrm{d}V\frac{\mathrm{d}\bar{\rho}_x}{\mathrm{d}t} + \frac{1}{2}L_3\bar{\rho}_y - \frac{1}{2}L_2\bar{\rho}_z = -K_h\bar{\rho}_x + F^{(x)}, \tag{9a}$$

$$\frac{1}{V}\int y^2\,\mathrm{d}V\frac{\mathrm{d}\bar{\rho}_y}{\mathrm{d}t}+\frac{1}{2}L_1\bar{\rho}_z-\frac{1}{2}L_3\bar{\rho}_x$$
$$=-K_\mathrm{h}\bar{\rho}_y+F^{(y)}, \tag{9b}$$

$$\frac{1}{V}\int z^2\,\mathrm{d}V\frac{\mathrm{d}\bar{\rho}_z}{\mathrm{d}t}+\frac{1}{2}L_2\bar{\rho}_x-\frac{1}{2}L_1\bar{\rho}_y$$
$$=-K_\mathrm{v}\bar{\rho}_z+F^{(z)}, \tag{9c}$$

where the moments of the buoyancy flux, Q, assumed to enter the basin only through the surface, appear in

$$\boldsymbol{F}=(F^{(x)}, F^{(y)}, F^{(z)})$$
$$=(1/V)\int\int(x, y, H/2)\,Q\,\mathrm{d}x\,\mathrm{d}y. \tag{10}$$

In a similar way account may be given for buoyancy fluxes entering the basin through one of the lateral boundaries.

Together with (8) these constitute a closed system governing the evolution of two state vectors, $\boldsymbol{L}(t)$ and $\nabla\bar{\rho}(t)$, describing the circulation and density field in a rigid lid, f-plane ocean, driven by differential momentum (4) and buoyancy (10) fluxes.

2.3. Scaling

It is convenient to assume that the horizontal scales B and L are identical and keep, in an application to a "thin" ocean, only the distinction in horizontal and vertical scales explicit and employ the following scaling (denoted by square brackets):

$$[x, y, z, t]=[L, L, H, L^2/K_\mathrm{h}],$$

in which the time scale is associated with horizontal diffusion of heat. Accordingly

$$[u, v, w, \delta\rho]$$
$$=[K_\mathrm{h}/L, K_\mathrm{h}/L, K_\mathrm{h}H/L^2, 12\rho_0 r_\mathrm{h}K_\mathrm{h}/gH],$$

suggested by continuity, while the (internal) scale of the density variations (that had been absorbed in g' up to now) has been chosen such that the buoyancy torque is (for O(1) density variations) of the same order of magnitude as the frictional torque. The external density and momentum fluxes (per unit mass) are scaled as

$$[Q, \tau^{(x,y)}]=[\delta\rho_\mathrm{e}HK_\mathrm{h}/\delta\rho L^2, u_*^2],$$

which really defines related scales of the external density contrast, $\delta\rho_\mathrm{e}$ and frictional velocity, u_*. The numerical factors in eqs. (8) and (9) may be absorbed by rescaling time with a factor 1/12 and $\delta\rho$ and $\boldsymbol{u}$ with a factor 2. This produces the numerical quantities in the nondimensional parameters below.

2.4. Zero-dimensional model of three-dimensional ocean circulation

With these scalings, the following simplified model of the three-dimensional ocean circulation can be obtained:

$$\mathrm{Pr}^{-1}\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t}+f'\boldsymbol{k}\times\boldsymbol{L}=-\bar{\rho}_y(t)\,\boldsymbol{i}+\bar{\rho}_x(t)\,\boldsymbol{j}$$
$$-(L_1, L_2, rL_3)+\hat{T}\boldsymbol{T}, \tag{11a}$$

$$\frac{\mathrm{d}\nabla\bar{\rho}}{\mathrm{d}t}+\nabla\bar{\rho}\times\boldsymbol{L}$$
$$=-(\bar{\rho}_x, \bar{\rho}_y, \mu\bar{\rho}_z)+\mathrm{Ra}\boldsymbol{F}. \tag{11b}$$

Here $\mathrm{Pr}=r_\mathrm{h}L^2/12K_\mathrm{h}$ denotes the Prandtl number (interpreting the Rayleigh damping term as proportional to an eddy diffusion term), $f'=f/2r_\mathrm{h}$ the scaled Coriolis parameter, $r=r_\mathrm{v}/r_\mathrm{h}$ the ratio of the frictional time scales of horizontal and vertical angular momentum, and $\hat{T}=u_*^2L/2r_\mathrm{h}K_\mathrm{h}H$ is the dimensionless magnitude of the torque exerted by the wind stress. In (11b), the ratio of the vertical to horizontal diffusion time scales is denoted as $\mu=K_\mathrm{v}L^2/K_\mathrm{h}H^2$, while the forcing strength is determined by the Rayleigh number $\mathrm{Ra}=g'_\mathrm{e}H/2r_\mathrm{h}K_\mathrm{h}$, with g'_e a reduced gravity based on the externally imposed density difference, $\delta\rho_\mathrm{e}$. The dimensionless momentum and density fluxes, $\boldsymbol{T}$ and $\boldsymbol{F}$, are similar to the expressions in (4) and (10) but with H replaced by 1, and $F^{(y)}=1$ and $T^{(z)}=-1$, as their magnitudes have been employed in the definition of Ra and $\hat{T}$, respectively, while a clockwise torque is negative.

Taking ocean-like parameter values: $L=5\times10^6$ m, $H=5\times10^3$ m, $g=10\ \mathrm{m\,s^{-2}}$, $f=10^{-4}\ \mathrm{s^{-1}}$, $K_\mathrm{h}=10^2\ \mathrm{m^2\,s^{-1}}$, $K_\mathrm{v}=10^{-4}\ \mathrm{m^2\,s^{-1}}$, $u_*=10^{-2}\ \mathrm{m\,s^{-1}}$, $r_\mathrm{h}=r_\mathrm{v}=10^{-6}\ \mathrm{s^{-1}}$, $\delta\rho_\mathrm{e}/\rho_0=4\times10^{-3}$, typical magnitudes of the nondimensional numbers are: $\mathrm{Pr}=2\times10^4$, $f'=50$, $r=1$, $\mu=1$, $\mathrm{Ra}=10^6$ and $\hat{T}=500$, while the dimensional time scale, $L^2/12K_\mathrm{h}$, is of order 500 years.

3. Special cases

3.1. Circulation in the meridional plane: $f' = 0$

With $f' = 0$ in (11), the circulation in the zonal and meridional plane decouples and the latter is described by

$$\mathrm{Pr}^{-1}\frac{dL_1}{dt} + \bar{\rho}_y = -L_1, \tag{12a}$$

$$\frac{d\bar{\rho}_y}{dt} + L_1\bar{\rho}_z = -\bar{\rho}_y + \mathrm{Ra}, \tag{12b}$$

$$\frac{d\bar{\rho}_z}{dt} - L_1\bar{\rho}_y = -\mu\bar{\rho}_z + \mathrm{Ra}\,F^{(z)}. \tag{12c}$$

Assuming that there is no inertia in the angular momentum equation ($\mathrm{Pr} \to \infty$) we retrieve Stommel's dynamics

$$\bar{\rho}_y = -L_1, \tag{13a}$$

see (1b), in which the buoyancy torque is exactly balanced by the frictional torque. However, in contrast to his model, eqs. (12b, c) still carry information about the vertical density structure: any change in the meridional density gradient can only occur via an intermediate change in the vertical density gradient. This process can, at best, happen instantaneously, as when the first term in (12c) is much smaller than the other two (assuming that the net heating term, $F^{(z)}$, vanishes, such as seems likely for a large-scale basin). Formally this is obtained by rescaling $\bar{\rho}_y$, L_1 and Ra with $\mu^{1/2}$ and subsequently taking the limit $\mu \to \infty$. The remaining balance can be used to eliminate $\bar{\rho}_z$ from (12b) to yield the following equation, recast in the original variables, for the evolution of the meridional density gradient:

$$\frac{d\bar{\rho}_y}{dt} = \mathrm{Ra} - \bar{\rho}_y - \frac{L_1^2}{\mu}\bar{\rho}_y. \tag{13b}$$

Comparison with (1a) shows that the "insensitivity" to the sense of the circulation is accomplished by the squared magnitude of the "advective flow field", rather than by the modulus of it. This "stripped" version of Stommel's original equations, has only one, stable (thermal, $L_1 < 0$) steady state.

Without these last assumptions, the circulation in the meridional plane described by (12) is interesting in itself, as it is equivalent to that in a Howard-Malkus loop (Malkus, 1972). This is a thermal oscillator in which fluid in a ring, placed in the vertical, is differentially heated vertically. The buoyancy torque creates angular momentum, which advects the buoyancy field. The process is checked by thermal and frictional damping. Welander (1991) also included asymmetric forcing and pointed out that eqs. (12) are of Lorenz-type (Lorenz, 1963). However, they become identical to Lorenz' equations, except for a trivial transformation of variable, only if the lateral differential heating (the forcing term in (12b)) vanishes; convection in this case being driven purely by differential heating in the vertical. (A limit formally obtained by taking simultaneously $\mathrm{Ra} \to 0$ and $F^{(z)} \to \infty$ while retaining a finite value of their product.) If lateral differential heating is added, Welander (1991) showed, that chaos may still be prevalent. For the oceanic case, however, the other extreme, viz. the vanishing of the differential heating in the vertical, is more likely to occur. In that case it can be shown that the equations acquire a single, steady state, which is stable for all possible values of the parameters.

3.2. Three-dimensional circulation for $f' \neq 0$ and $Pr \to \infty$

It seems worthwhile to consider the effect of rotation in what seems to be the realistic limit $\mathrm{Pr} \to \infty$ (Weaver and Sarachik, 1991; Wright and Stocker, 1991; Zhang et al., 1992). In an application to a large-scale ocean it also seems realistic to assume that there is no direct wind forcing of the zonal and meridional circulation: $T^{(x)} = T^{(y)} = 0$. Eqs. (11a) then yield:

$$L_1 = \frac{1}{1+f'^2}(-\bar{\rho}_y + f'\bar{\rho}_x), \tag{14a}$$

$$L_2 = \frac{1}{1+f'^2}(\bar{\rho}_x + f'\bar{\rho}_y), \tag{14b}$$

$$L_3 = -\frac{\hat{T}}{r}. \tag{14c}$$

For the moment we will also drop the remaining wind-stress torque (hence $L_3 = 0$), which will be considered in Section 5, as well as differential heating in the east–west direction ($F^{(x)} = 0$). The latter

assumption is not really restrictive because of the isotropy of the f-plane. This means that if this term was non-zero, we might rotate our coordinate system in such a way that the new y-direction is aligned with the direction of the net differential horizontal heating and the same equations as (15) below would result. If we redefine the basin-averaged density gradients and Rayleigh number in (11b) by dividing them by $1+f'^2$, this factor is eliminated from the denominator in (14a, b) such that, after replacing L_1 and L_2 in (11b) by these expressions, the evolution of the average density gradients are given by

$$\frac{d\bar{\rho}_x}{dt} = (\bar{\rho}_x + f'\bar{\rho}_y)\,\bar{\rho}_z - \bar{\rho}_x, \tag{15a}$$

$$\frac{d\bar{\rho}_y}{dt} = (\bar{\rho}_y - f'\bar{\rho}_x)\,\bar{\rho}_z - \bar{\rho}_y + \mathrm{Ra}, \tag{15b}$$

$$\frac{d\bar{\rho}_z}{dt} = -\mu\bar{\rho}_z - (\bar{\rho}_x^2 + \bar{\rho}_y^2) + \mathrm{Ra}\,F^{(z)}. \tag{15c}$$

Remarkably, these equations are equivalent to a model that Lorenz (1984, 1990) proposed as "possibly the simplest model describing the atmospheric general circulation". The correspondence is

$$\bar{\rho}_x \leftrightarrow Z, \qquad \bar{\rho}_y \leftrightarrow Y, \qquad \bar{\rho}_z \leftrightarrow X,$$

where X, Y, Z are the dependent variables in Lorenz' model denoting the strength of the globe-encircling jet stream, and the cosine and sine components of a "chain of superposed large-scale eddies" driven by thermal contrast between continents and oceans. Lorenz (1984, 1990) remarks that this model can be derived from a severe truncation of a geostrophic baroclinic model in spectral form. Because X is geostrophically related to the large-scale pole-ward temperature gradient, it is the analog of $F^{(z)}$ which, in his case, is the dominant driving term. He notes that if this term is sufficiently strong (and positive) this system too may exhibit chaos. However in those circumstances it has positive values for X, which would correspond, in the present context, to statically unstable states, driven by vigorous surface cooling or evaporation. Such processes may perhaps be locally of dominating influence and contribute to deep-water formation like in the Labrador Sea (Weaver and Sarachik, 1991; Weaver et al. 1994), or some areas of the Mediterranean, but is unlikely to dominate an entire ocean basin. Therefore the thermally driven, general circulation in the ocean is probably more realistically described by (15) with $F^{(z)}=0$.

4. Steady states and stability of the thermally driven ocean circulation

The steady state of eq. (15) is obtained by setting the right-hand sides equal to zero. In this way we obtain:

$$(\bar{\rho}_x, \bar{\rho}_y) = \frac{\mathrm{Ra}(f'\bar{\rho}_z,\, 1-\bar{\rho}_z)}{[(1-\bar{\rho}_z)^2 + f'^2\bar{\rho}_z^2]}, \tag{16}$$

from the first two equations. Inserting this into the third (with $F^{(z)}=0$) gives

$$\bar{\rho}_z[(1-\bar{\rho}_z)^2 + f'^2\bar{\rho}_z^2] + \mathrm{Ra}^2/\mu = 0. \tag{17}$$

This cubic in $\bar{\rho}_z$ has, for all (positive) values of the parameters, only one real root, which is negative and thus corresponds to a statically stable, physical state. From (16) it follows that this corresponds to a positive meridional and a negative zonal average density gradient. The negative correlation between these two in the thermohaline context (and in a spherical coordinate frame) was first suggested by Wright and Stocker (1991) and further substantiated by Wright and Vreugdenhil (1994) and has been used to bring three-dimensional aspects in a two-dimensional (zonally averaged) numerical model of the thermohaline circulation (see Hughes and Weaver, 1994). From (15a, b) it follows that the steady state has $L_1<0$ and $L_2>0$, or in other words, has a circulation with a northeastward-directed surface flow.

There may be just one steady state, but, interestingly, it can turn unstable for some parameter values. The stability of the steady state in phase space is determined by the eigenvalues, ν, of (15), linearized around the steady state described by (16) and (17). The eigenvalue equation reads

$$(1+\nu)^2(\mu+\nu) - 2\bar{\rho}_z(1+\nu)(2\mu+\nu) + \bar{\rho}_z^2(1+f'^2)(3\mu+\nu) = 0, \tag{18}$$

where $\bar{\rho}_z$ in (17) and in this equation pertain to its value at steady state. The steady state, which for

small values of the parameters is stable, i.e., for which (18) has roots with negative real parts, becomes unstable through a Hopf-bifurcation. At the transition the cubic in v can be written as

$$(v+a)(v+ib)(v-ib),$$

and by equating this with the cubic in (18) expressions for a and b can be obtained. From the remaining consistency equation we may obtain f' as a function of μ and the value of $\bar{\rho}_z$ at steady state. Viewing the latter in this expression and in (17) as a parameter a parametric relation of Ra versus f', defining the stability boundary for a given value of μ, is obtained (Fig. 1). Points in parameter space below and to the left of the curves are stable. In the unstable case a simple periodic limit cycle, a self-sustained oscillation, occurs. An example is given in Fig. 2. The meridional circulation remains negative (thermal) but varies in strength; the zonal circulation however, changes its sense of direction and may become negative when the meridional density contrast vanishes. For the parameter values in Fig. 2 the circulations are nearly geostrophically coupled to the density gradients in the meridional and zonal planes respectively. The oscillatory cycle is best described starting from the moment that the meridional density gradient is slightly negative. In that case, the east-west density difference relaxes due to diffusive and frictional processes and hence the geostrophically coupled thermal circulation weakens. This allows the ever-present thermal forcing to generate a positive pole-equator density contrast, which, however, promotes the geostrophically coupled

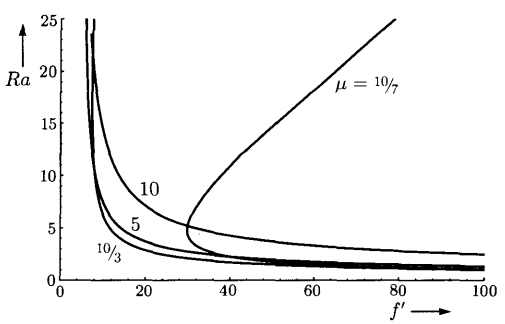

Fig. 1. Stability diagram of the steady state of eqs. (15) (with $F^{(z)}=0$). Curves have been drawn for different values of μ above and to the right of which the single steady state is unstable and gives rise to a limit cycle.

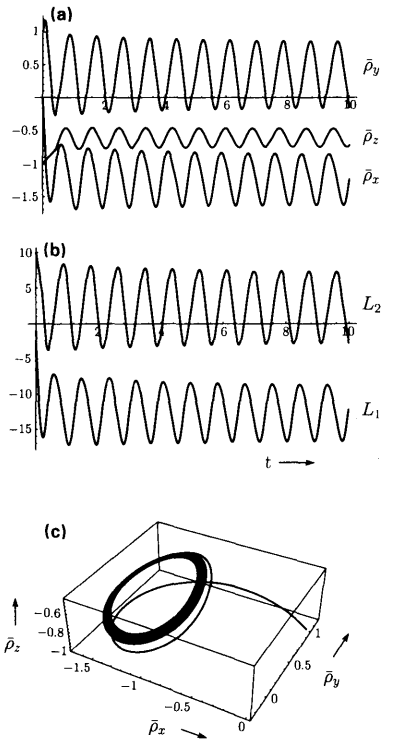

Fig. 2. Example of a limit cycle in eqs. (15) (with $F^{(z)}=0$), for $\mu=3$, $f'=10$ and Ra $=5$. The curves give $\bar{\rho}_y$, $\bar{\rho}_z$, $\bar{\rho}_x$ (a), L_2 and L_1 (b) from above to below, respectively, as a function of time. The three-dimensional diagram (c) gives its behaviour in phase space.

zonal circulation. This, in turn, amplifies the remaining east–west density difference and, with it, thermal overturning which again removes the pole-equator density difference and associated zonal overturning, which closes the cycle.

Similar, purely thermally-driven limit cycles occur in the Labrador Sea region of a numerical model of the North Atlantic Ocean (Weaver et al., 1994). In that case, however, it is likely that the net buoyancy flux over this area ($F^{(z)}\neq 0$) plays a role.

5. The thermally and wind-driven ocean circulation

In addition to the meridional thermal forcing we now consider the case where there is a nonzero torque of the wind stress generating vertical angular momentum. Under the assumption

Tellus A/46/5-9

$\mathrm{Pr} \to \infty$, this remains a constant (given by (14c)). Eqs. (15a, b, c), extended with terms related to the L_3 component in (11b) (see eqs. (22) below), however, are isomorphic with eqs. (15) themselves. This result is obtained by a transformation in which the new variables (capped below) are related to the old by:

$$(\hat{\rho}_x, \hat{\rho}_y, \hat{\rho}_z, \hat{\mu}) = \frac{(\bar{\rho}_x, \bar{\rho}_y, \bar{\rho}_z - L_3/f', \mu)}{(1 - L_3/f'),} \tag{19a}$$

$$\hat{t} = t(1 - L_3/f'), \tag{19b}$$

$$\widehat{\mathrm{Ra}} = \mathrm{Ra}/(1 - L_3/f')^2. \tag{19c}$$

Due to the transformation a non-zero forcing term now appears at the right-hand of the transformed version of (15c). This appears exactly in Lorenz's (1984) format as the product of the damping coefficient $\hat{\mu}$, and a function

$$F = -L_3/(f' - L_3). \tag{20}$$

The steady states of the thus extended eqs. (15) are determined by

$$\widehat{\mathrm{Ra}}^2 = \hat{\mu}(F - \hat{\rho}_z)[(1 - \hat{\rho}_z)^2 + f'^2\hat{\rho}_z^2], \tag{21}$$

which, as Lorenz notes, can be directly plotted if the $\widehat{\mathrm{Ra}}$ number is viewed here as a function of $\hat{\rho}_z$. Values of $\hat{\rho}_z$ are restricted by the requirement that the right-hand side of (21) is positive and hence $\hat{\rho}_z < F$. With (19a) and (20), this implies that each steady state is statically stable, $\bar{\rho}_z < 0$, but does not indicate that they are necessarily dynamically stable, i.e., the steady state may have eigenvalues with positive real parts. In fact, the system may be chaotic, as Lorenz (1984, 1990) showed with a set of particular parameter values, but the chaos which ensues is different from that obtained in his celebrated (1963) model. Chaos in the latter results by the simultaneous instability of two steady states (representing convection with opposite senses of direction) when the model acquires enough inertia (i.e., for a particular, small enough value of the Prandtl number in (12)). In contrast, in his (1984) model, the particular Rayleigh number used ($\mathrm{Ra} = 1$) is below the critical value above which multiple steady states exist and hence there is just one steady state. This state, however, is unstable and states in phase space are attracted to a region (approximately given by $-2 < \hat{\rho}_z < 2$) where a new (stable) steady state will come into existence for values exceeding the critical Rayleigh number. However, as this steady state does not yet exist, it defies a direct perturbation analysis by which the classical (1963) model can be attacked, and the extent of it, in parameter space, is consequently poorly understood. Lorenz (1984), for the particular values of the other parameters, concludes that chaos requires "large enough values of F", well above a value of one. Realistic estimates of the torque due to the wind stress will likely yield negative values of L_3, and (20) shows that our F is below one and, again, is not likely to induce chaotic motion in the large-scale ocean circulation. Moreover, values of F greater than one are obtained only when $L_3 > f'$, in which case the transformation (19) has some peculiarities in that the direction of time changes. A direct integration of the original equations, (11b) and (14), is, in that case, warranted.

Without employing transformation (19), but retaining the assumption that $F^{(z)} = 0$, eqs. (15), extended with terms related to the vertical angular momentum, read:

$$\frac{\mathrm{d}\bar{\rho}_x}{\mathrm{d}t} = (\bar{\rho}_z - 1)\,\bar{\rho}_x + (f'\bar{\rho}_z - L_3)\,\bar{\rho}_y, \tag{22a}$$

$$\frac{\mathrm{d}\bar{\rho}_y}{\mathrm{d}t} = -(f'\bar{\rho}_z - L_3)\,\bar{\rho}_x + (\bar{\rho}_z - 1)\,\bar{\rho}_y + \mathrm{Ra}, \tag{22b}$$

$$\frac{\mathrm{d}\bar{\rho}_z}{\mathrm{d}t} = -\mu\bar{\rho}_z - (\bar{\rho}_x^2 + \bar{\rho}_y^2). \tag{22c}$$

The steady states are determined, as before, by the zero's of a cubic:

$$h(\bar{\rho}_z) \equiv \bar{\rho}_z[(1 - \bar{\rho}_z)^2 + (f'\bar{\rho}_z - L_3)^2]$$
$$= -\frac{\mathrm{Ra}^2}{\mu} \equiv -l^2 < 0. \tag{23}$$

The existence of multiple equilibria is determined by the occurrence of extrema in $h(\zeta)$, which in turn exist if the discriminant of the roots of $\mathrm{d}h/\mathrm{d}\zeta = 0$ is real and positive. Note that $h(\zeta)$ may have local extrema for positive values of ζ, but these are unattainable for positive values of Ra^2/μ. Therefore physically realisable multiple equilibria will occur for $f' > f'_{\min} \equiv (\sqrt{3}\,L_3 - 1)/(L_3 + \sqrt{3})$. Evaluating $h(\zeta)$ at the positions of the extrema determines a range of l^2 values in which multiple

equilibria can be found. These boundaries are given by

$$l^2_{\max,\min} = -\frac{2}{27(1+f'^2)^2} \times [(1+f'L_3)[9(f'-L_3)^2+(1+f'L_3)^2] \pm [(1+f'L_3)^2-3(f'-L_3)^2]^{3/2}], \quad (24)$$

see Fig. 3. These boundaries coincide when the argument of the 3/2 power vanishes, which occurs at the above defined $f'_{\min}$. Alternatively, this may be expressed as a condition on L_3, which, when reintroduced in the non-vanishing part of $l^2_{\max,\min}$, gives the dashed curve in Fig. 3. It is defined by

$$l^2(f') = \frac{8(1+f'^2)}{3^{3/2}(f'-\sqrt{3})^3},$$

above and to the right of which multiple steady states may exist in some range of Ra^2/μ-values for certain L_3-values.

There may thus either be one or two steady states that may be steady or (one of which may be) of limit cycle type. A complete understanding of these equations and (15) is lacking at present and requires further work in a more complete parameter space.

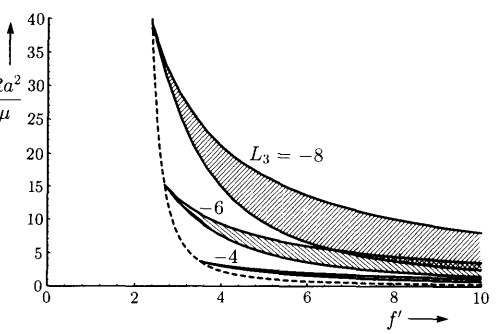

Fig. 3. Diagram showing where multiple steady states reside in parameter space. For points above and to the right of the dashed line there exists a value of L_3 having a range of Ra^2/μ-values for which multiple steady states are obtained. Specific examples of these ranges are given for three values of L_3.

6. Conclusions and discussion

It is shown that the equations describing the circulation in a simple box model are, in a certain limit, related to those describing the evolution of the lowest moments of the momentum and density fields in a rectangular, f-plane ocean. In contrast to box-models, however, Coriolis effects and wind forcing can be readily incorporated. For certain parameter values the model formally reduces to some well-known systems of equations (Lorenz, 1963 and Lorenz, 1984). Parameter values in the ocean avoid the chaos which these equations may exhibit. The thermally driven ocean circulation, however, may feature self-sustained oscillations with a typical time scale of 500 years, though their occurrence for large Ra-numbers, as Fig. 1 shows, strongly depends on the values of μ, the ratio of diffusivities scaled with the square of the aspect ratio. Notwithstanding the neglect of salinity effects the present f-plane model still exhibits multiple equilibria when wind-forcing is included, see Fig. 3. It is evident from (24) that for the oceanic values discussed at the end of Subsection (2.4) (taking into account the rescaling referred to above eqs. (15) which yields $\mathrm{Ra} \approx 400$), the ocean is in a state where multiple steady states might occur.

7. Acknowledgements

This work was initiated by a course on the atmosphere-ocean climate system given at the University of Victoria by Andrew Weaver. Discussions with him, Chris Garrett, Amit Tandon, Hide Yamazaki, Ming Li, Michael Winton, Ed Sarachik, Mitsuhiro Kawase and suggestions by Huib de Swart, Kees Vreugdenhil and Dan Wright have been very beneficial. Thanks are due to Rosalie Rutka who helped with the preparation of this manuscript. A NATO-Science Fellowship, granted by the Netherlands Organization for Scientific Research (NWO), and the hospitality of the Centre for Earth and Ocean Research at the University of Victoria enabled the author to perform the present study during a sabbatical leave. Their support is greatly appreciated.

REFERENCES

Bryan, F. 1986. High-latitude salinity effects and interhemispheric thermohaline circulations. *Nature* **323**, 301–304.

Cessi, P. and Young, W. R. 1992. Multiple equilibria in two-dimensional thermohaline circulation. *J. Fluid Mech.* **241**, 291–309.

Huang, R. X. and Stommel, H. M. 1992. Convective flow patterns in an eight-box cube driven by combined wind stress, thermal and saline forcing. *J. Geophys. Res.* **C97**, 2347–2364.

Hughes, T. M. C. and Weaver, A. J. 1994. Multiple equilibria of an asymmetric two-basin ocean model. *J. Phys. Ocean.* **24**, in press.

Källén, E. and Huang, X.-Y. 1987. A simple model for large-scale thermohaline convection. *Dyn. Atm. Oceans* 11, 153–173.

Lorenz, E. N. 1963. Deterministic non-periodic flow. *J. Atm. Sci.* **20**, 130–141.

Lorenz, E. N. 1984. Irregularity: a fundamental property of the atmosphere. *Tellus* **36A**, 98–110.

Lorenz, E. N. 1990. Can chaos and intransitivity lead to interannual variability? *Tellus* **42A**, 378–389.

Malkus, V. W. R. 1972. Non-periodic convection at high and low Prandtl number. *Mem. Soc. Roy. Sci. Liège*, **X6e** Serie, Tome IV, 125–128.

Manabe, S. and Stouffer, R. J. 1988. Two stable equilibria of a coupled ocean-atmosphere model. *J. Climate* **1**, 841–866.

Marotzke, J., Welander, P. and Willebrand, J. 1988. Instability and multiple steady states in a meridional plane model of the thermohaline circulation. *Tellus* **40A**, 162–172.

Stommel, H. 1961. Thermohaline convection with two stable regimes of flow. *Tellus* **13**, 224–230.

Thual, O. and McWilliams, J. C. 1992. The catastrophe structure of thermohaline convection in a two-dimensional fluid model and a comparison with low-order box models. *Geophysical and Astrophysical Fluid Dynamics* **64**, 67–95.

Weaver, A. J. and Sarachik, E. S. 1991. Evidence for decadal variability in an ocean general circulation model. *Atm.-Ocean* **29**, 197–231.

Weaver, A. J. and Hughes, T. M. C. 1992. Stability and variability of the thermohaline circulation and its link to climate. *Trends in physical oceanography*, Council of Scientific Research Integration, Research Trends Series 1, Trivandrum, India.

Weaver, A. J., Aura, S. M. and Myers, P. G. 1994. Interdecadal variability in an idealized model of the North Atlantic. *J. Geophys. Res.*, in press.

Weaver, A. J., Marotzke, J., Cummins, P. F. and Sarachik, E. S. 1993. Stability and variability of the thermohaline circulation. *J. Phys. Ocean.* **23**, 39–60.

Welander, P. 1986. Thermohaline effects in the ocean circulation and related simple models. In: *Large-scale transport processes in oceans and atmosphere*, eds. Willebrand, J. and Anderson, D. L. T., 163–200. D. Reidel Publ. Co.

Welander, P. 1991. On the ocean heat engine, stiffness of chaotic systems and climate prediction. *Tellus* **43A/B**, 116–120.

Wright, D. G. and Stocker, T. F. 1991. A zonally averaged model for the thermohaline circulation. Part I: model development and flow dynamics. *J. Phys. Ocean.* **21**, 1713–1724.

Wright, D. G. and Vreugdenhil, C. B. 1994. Vorticity dynamics and zonally averaged ocean circulation models. *J. Phys. Ocean.*, resubmitted.

Zhang, S., Lin, C. A. and Greatbatch, R. J. 1992. A thermocline model for ocean-climate studies. *J. Mar. Res.* **50**, 99–124.