J. Fluid Mech. (1997), *vol.* 331, *pp.* 169–198

Symmetry breaking and overturning oscillations in thermohaline-driven flows

By HENK A. DIJKSTRA AND M. JEROEN MOLEMAKER

Institute for Marine and Atmospheric Research Utrecht, Department of Physics and Astronomy, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands

(Received 24 January 1995 and in revised form 9 August 1996)

The bifurcation structure of thermohaline-driven flows is studied within one of the simplest zonally averaged models which captures thermohaline transport: a Boussinesq model of surface-forced thermohaline flow in a two-dimensional rectangular basin. Under mixed boundary conditions, i.e. prescribed surface temperature and fresh-water flux, it is shown that symmetry breaking originates from a codimension-two singularity which arises through the intersection of the paths of two symmetry-breaking pitchfork bifurcations. The physical mechanism of symmetry breaking of both the thermally and salinity dominated symmetric solution is described in detail from the perturbation structures near bifurcation. Limit cycles with an oscillation period in the order of the overturning time scale arise through Hopf bifurcations on the branches of asymmetric steady solutions. The physical mechanism of oscillation is described in terms of the most unstable mode just at the Hopf bifurcation. The occurrence of these oscillations is quite sensitive to the shape of the prescribed fresh-water flux. Symmetry breaking still occurs when, instead of a fixed temperature, a Newtonian cooling condition is prescribed at the surface. There is only quantitative sensitivity, i.e. the positions of the bifurcation points shift with the surface heat transfer coefficient. There are no qualitative changes in the bifurcation diagram except in the limit where both the surface heat flux and fresh-water flux are prescribed. The bifurcation structure at large aspect ratio is shown to converge to that obtained by asymptotic theory. The complete structure of symmetric and asymmetric multiple equilibria is shown to originate from a codimension-three bifurcation, which arises through the intersection of a cusp and the codimension-two singularity responsible for symmetry breaking.

1. Introduction

On the global scale, the ocean circulation is driven by wind forcing and by fluxes of heat and fresh water through the ocean surface. The circulation driven by the latter component is referred to as the thermohaline ocean circulation. In the present Atlantic circulation, the zonally averaged component has a 1-cell structure with a surface flow which is predominantly northward and a compensating southward deep sea circulation. Changes in ocean circulation have an enormous effect on climate, due to a strong modification of the poleward heat transport. There are indications that the Atlantic circulation has been different in the past with corresponding different climate states (Broecker 1993). Also the observed temporal variability on decadal or longer time scales, for instance in temperature records, may be attributed to large-scale changes of the thermohaline flow pattern in the Atlantic basin.

These observations strongly motivate studies of the thermohaline ocean circulation using a hierarchy of ocean models. Using a very simple box model, (Stommel 1961) showed that the presence of heat and salt, with their different influence on the density field, may lead to different stable steady flow patterns. Multiple steady ocean circulation patterns were indeed found in extended box models and thermohaline loop models (Welander 1986), two-dimensional Boussinesq models (Thual & McWilliams 1992), other zonally averaged models (Wright & Stocker 1991) and general circulation models (GCMs) (Bryan 1986). The stability of the present large-scale ocean circulation has therefore become an important issue in climate research.

Central to the issue of multiple steady states is the use of mixed boundary conditions (Haney 1971), i.e. a prescribed temperature and fresh-water flux at the ocean surface. In a 2-hemispheric ocean GCM, Bryan (1986) obtained a symmetric 2-cell flow pattern by prescribing both an equatorially symmetric salinity and temperature at the ocean surface. By diagnosing the fresh-water flux of this state, he showed that it became unstable under mixed boundary conditions. By perturbing the symmetric state with finite-amplitude fresh-water anomalies, Bryan (1986) further showed that aside from the symmetric 2-cell state, also two 1-cell pole-to-pole asymmetric stable steady states existed.

The feedback processes responsible for the transition from the symmetric towards the asymmetric states were studied in more detail in a zonally averaged 2-hemispheric model (Marotzke & Willebrand 1988). When a high-latitude positive salinity anomaly is imposed on the symmetric 2-cell state with equatorial upflow, an advective feedback process between the strength of the circulation and the associated salt transport induces the transition to an asymmetric state. If a high-latitude negative salinity anomaly is added, a fresh-water cap is generated, shutting down the high-latitude convection before moving equatorward to give an asymmetric state; this feedback process is called convective. In single-hemispheric models, a 1-cell state with deep water formation (downflow) at high latitudes may also become unstable to high-latitude fresh-water anomalies. Under mixed boundary conditions it may develop into a state with deep water formation at the equator giving rise to warm deep water (Marotzke, Welander & Willebrand 1988). This transition is referred to as a polar halocline catastrophe, below abbreviated as PHC. Eventually, the latter state changes rapidly into the original state because the warm deep water causes high-latitude convection again; this rapid change is referred to as a flush.

In a 1-hemispheric ocean model comparison study Weaver *et al.* (1993) found multiple equilibria and two types of sustained oscillations. One of those oscillations is related to the horizontal gyre scale advection of salinity anomalies and is essentially a three-dimensional phenomenon and the other oscillation is related to the above-mentioned flushes. The mechanism of the first oscillation, which has a decadal time scale, was described in Weaver & Sarachik (1991). The second type of oscillation was investigated by Winton & Sarachick (1993). In the latter work, another type of oscillation was also found at strong fresh-water flux forcing having a time scale comparable to that of the meridional overturning. The latter type of oscillation has also been found in the response of other models, for instance the Hamburg Large Scale Geostrophic model (Mikolajewicz & Maier-Reimer 1990).

In view of simulations with coupled climate models, it is important to know whether an ocean equilibrium state spun up with prescribed surface salinity and temperature is stable under mixed boundary conditions and/or when it is coupled to an atmosphere model. The sensitivity of the PHC to the formulation of the air–sea interaction was demonstrated by Zhang, Greatbach & Lin (1993). No PHC was found when both heat

flux and fresh-water flux were prescribed, whereas this transition did occur under mixed boundary conditions. Using a global ocean model, Tziperman *et al.* (1994) found that even the restoring time towards the specified salt distribution used to obtain the ocean mean state may affect its stability under mixed boundary conditions. A consistency criterion was proposed to choose the restoring times which give a stable climate state. The latter result indicates that the stability of steady states is strongly sensitive to the spatial pattern of the surface fresh-water forcing.

Instead of following a few trajectories of a particular system of equations for different parameter sets and/or boundary conditions, much more insight into the physics of a particular system can be obtained by a systematic search of equilibria of the system and the critical points separating regions of different qualitative behaviour (bifurcation points). This methodology of dynamical systems theory (Guckenheimer & Holmes 1983) has already been applied successfully to many physical systems. It has been used in the study of atmospheric variability (Legras & Ghil 1988) and been applied to problems of coupled ocean–atmosphere interaction in the Tropics (Dijkstra & Neelin 1995) and the stability of the wind-driven ocean circulation (Speich, Dijkstra & Ghil 1995). When looking at the results of ocean models, for thermohaline circulation, from a dynamical systems point of view three questions arise. Do the multiple steady states result from only one or several bifurcation points as the strength of the fresh-water flux is changed? Second, do oscillations arise as Hopf bifurcations of these steady states or are they for example associated with isolated limit cycles? Third, is the origin of the sensitivity to surface boundary conditions caused by a fast movement of the bifurcation points in parameter space (i.e. a quantitative change), or by a complete change in the structure of these points (i.e. a qualitative change)? By systematically computing bifurcation diagrams in parameter space and their change with different boundary conditions these questions are addressed in this paper.

The steady-state structure of simple box models of the thermohaline circulation has been systematically analysed in several cases (Thual & McWilliams 1992). Box models may serve as 'toy' models to investigate one particular issue in isolation, for example the effect of the salinity restoring time (Tziperman *et al.* 1994). The main disadvantage of box models is that the coupling between boxes may not converge to that of the spatial coupling between relevant degrees of freedom in the governing system of partial differential equations. This may lead to many artificial steady states and artificial oscillations. Coupling between spatial degrees of freedom is present in two-dimensional models, for example the zonally averaged models Wright & Stocker (1991) or non-rotating Boussinesq models (Cessi & Young 1992, Quon & Ghil 1992, Thual & McWilliams 1992). It is the latter type of models on which we focus attention in this paper. Although these models are highly idealized as ocean models, the results obtained so far indicate that much of the underlying dynamical structure – of phenomena which have a two-dimensional character – in the large-scale ocean models is also present in these simple models. For example, the stable steady states of this model have shown a number of similarities with those of complex three-dimensional ocean models (Weaver & Hughes 1992). Although the relation between the solutions of both types of models is still unclear, the model under study is one in which the physics of the thermohaline transport is captured. It has clearly the advantage that the results can be analysed in detail. Therefore, the results are valuable in the interpretation of more detailed models. As such, this type of model has recently been used as the ocean component of a coupled ocean–atmosphere model (Saravanan & McWilliams 1995).

Thual & McWilliams (1992) showed that the transition from restoring to mixed boundary conditions is not essential to obtain multiple equilibria. By increasing the strength of a prescribed fresh-water flux there is a regime where multiple stable states appear. A bifurcation diagram was proposed but could not completely be verified, because only stable states could be reached with transient integration. Asymptotic analysis in the large-aspect-ratio (ratio of width to height) limit by Cessi & Young (1992) showed that signatures of multiple steady states already occur in the small forcing limit. Their asymmetric solutions are not globally defined and the amplitude equation they derive does not seem able to capture the solution structure found in Thual & McWilliams (1992). Using the same model, Quon & Ghil (1992) found multiple steady states and obtained an approximate bifurcation diagram. They proposed a supercritical pitchfork bifurcation as the origin of symmetry breaking of the thermally dominated solution and the occurrence of asymmetric states. Furthermore, they presented the path of this bifurcation point in the plane spanned by the Rayleigh number and the fresh-water flux strength. In a follow-up paper (Quon & Ghil 1995), oscillatory solutions were found in the large-aspect-ratio regime, in a model with anisotropic diffusivities. Quon & Ghil (1992, 1995) applied the switch from restoring to mixed boundary conditions at each value of the Rayleigh number. The spatial pattern of the fresh-water flux therefore depends on the Rayleigh number, which hinders a straightforward comparison with the results in Thual & McWilliams (1992) and Cessi & Young (1992).

In this paper, we present a bifurcation study of solutions in the two-dimensional Boussinesq model by solving for branches of steady states in parameter space using path-following techniques. These techniques are superior to transient integration, because both stable and unstable branches of solutions can be computed. Moreover, the bifurcation points at these branches can be computed accurately and consequently the regions of different qualitative behaviour can be identified as paths of these points. The first focus of the paper is the origin of symmetry breaking and its sensitivity to parameter changes and different surface boundary conditions. The second focus is the occurrence of oscillatory instabilities as Hopf bifurcations of particular steady states in this system and their sensitivity to the spatial pattern of the prescribed fresh-water flux. The third focus is the description of physical mechanisms of symmetry breaking and oscillatory instabilities through a study of the interaction of the perturbations (near criticality) and the equilibrium state.

In §2, the model and the numerical techniques used are briefly described. In §3, the bifurcation structure is determined for the relatively small-aspect-ratio case as considered in Quon & Ghil (1992). Integral properties of the solutions are investigated as well as the physics of the instabilities that are found. The effects of the type of boundary conditions on the bifurcation diagram are studied in §4. In the next section, the bifurcation structure is followed into the large-aspect-ratio limit and the results are connected to those in Cessi & Young (1992) and Thual & McWilliams (1992). This is followed by a discussion in §6 on what types of trajectories one can expect in these models and the relevance of these results to the behaviour of ocean models.

2. Formulation

2.1. *Model equations and non-dimensionalization*

The Boussinesq model of the zonally averaged thermohaline circulation is similar to that used previously in Cessi & Young (1992), Quon & Ghil (1992) and Thual

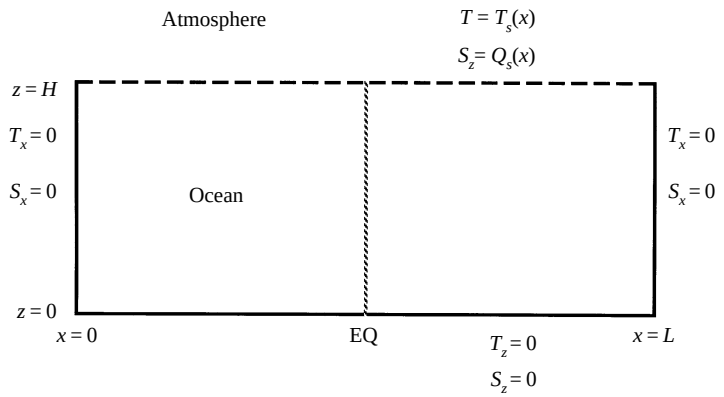

FIGURE 1. Model set-up and boundary conditions for salinity and temperature; EQ indicates the equator.

& McWilliams (1992). It is considered for a two-dimensional pole-to-pole ocean basin of length L and depth H (figure 1). The diffusivities of heat κ_T, salt κ_S, and momentum ν, are assumed constant and must be interpreted as eddy diffusivities. A linear equation of state is assumed, with thermal and solutal coefficients indicated by α_T and α_S.

The governing equations are non-dimensionalized using scales H, κ_T/H, ΔT, and $\Delta S/\lambda$ for length, velocity, temperature and salinity, respectively. Here ΔT and ΔS are characteristic meridional temperature and salinity differences and λ is the buoyancy ratio ($\lambda = \alpha_S \Delta S/(\alpha_T \Delta T)$. With horizontal and vertical velocities u and w, respectively, the governing equations in streamfunction ψ and vorticity ω (with $u = \partial\psi/\partial z$, $w = -\partial\psi/\partial x$ and $\omega = \partial w/\partial x - \partial u/\partial z$) formulation are

$$Pr^{-1}\left(\frac{\partial\omega}{\partial t} + u\frac{\partial\omega}{\partial x} + w\frac{\partial\omega}{\partial z}\right) = \nabla^2\omega + Ra\left(\frac{\partial T}{\partial x} - \frac{\partial S}{\partial x}\right), \tag{2.1a}$$

$$\omega = -\nabla^2\psi, \tag{2.1b}$$

$$\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} + w\frac{\partial T}{\partial z} = \nabla^2 T, \tag{2.1c}$$

$$\frac{\partial S}{\partial t} + u\frac{\partial S}{\partial x} + w\frac{\partial S}{\partial z} = Le^{-1}\nabla^2 S. \tag{2.1d}$$

All boundaries are assumed stress free and the lateral and bottom boundary are isolated and impervious to salt, i.e.

$$x = 0, A: \qquad \psi = \omega = \frac{\partial S}{\partial x} = \frac{\partial T}{\partial x} = 0, \tag{2.2a}$$

$$z = 0: \qquad \psi = \omega = \frac{\partial S}{\partial z} = \frac{\partial T}{\partial z} = 0. \tag{2.2b}$$

At the ocean surface, the usual mixed boundary conditions are given by

$$z = 1: \quad \psi = \omega = 0, \quad T = T_S(x), \quad \frac{\partial S}{\partial z} = \sigma Q_S(x), \tag{2.2c}$$

where the function $T_S(x)$ is a prescribed temperature distribution along the ocean surface. The parameter σ measures the strength of the surface fresh-water flux (the latter is also called a virtual salt flux) and $Q_S(x)$ represents its spatial structure. When

the surface integral of this function is zero, the total salt content is conserved for each steady state.

Besides σ in (2.2c), the equations above contain four other dimensionless parameters, the Prandtl number Pr, the Lewis number Le, the thermal Rayleigh number Ra, and the aspect ratio A defined by

$$Pr = \frac{\nu}{\kappa_T}; \quad Le = \frac{\kappa_T}{\kappa_S}; \quad Ra = \frac{g\ \alpha_T\ \Delta T H^3}{\nu \kappa_T}; \quad A = \frac{L}{H}. \tag{2.3}$$

Apart from parameters appearing in the functions $T_S(x)$ and $Q_S(x)$, the equations form a dynamical system having five parameters (σ, A, Ra, Le, Pr). In addition to Le and Pr, three other parameters, i.e. Ra_T, σ and d are used in Quon & Ghil (1992), whereas in Thual & McWilliams (1992), the parameters a, b and k are used. For convenience and later reference, we give the expressions for these parameters in terms of the ones used here:

$$Ra_T = A^3\ Ra; \quad d = \frac{1}{A}; \quad a = \frac{Ra}{A^2}; \quad b = \sigma a; \quad k = \frac{2\pi}{A}. \tag{2.4}$$

2.2. *Numerical methods and validation of results*

The stationary version of the equations (2.1) and boundary conditions (2.2) are discretized using a control volume method as in Dijkstra (1992) on a (non-) equidistant grid for $i = 0, \ldots, N$; $j = 0, \ldots, M$. This gives a nonlinear system of algebraic equations of the form

$$\boldsymbol{F}_{N,M}(\boldsymbol{u}, \boldsymbol{p}) = 0 \tag{2.5}$$

where $\boldsymbol{u}$ is the $4(N+1)(M+1)$ dimensional solution vector consisting of the unknowns at the gridpoints. The vector $\boldsymbol{p}$ consists of the values of the parameters (σ, A, Ra, Le, Pr). Steady solutions are calculated using pseudo-arclength continuation and the linear stability of a particular solution is determined by solving the corresponding generalized eigenvalue problem. The numerics of the code and its performance for several test problems is described in Dijkstra *et al.* (1995).

There is a special requirement in the mixed boundary condition formulation. Since the salinity field is only determined up to an additive constant, the Jacobian matrix of the nonlinear algebraic system of equations after discretization is singular. The additive constant is determined by an integral constraint expressing the conservation of salt over the total volume. In this case, because the function $Q_S(x)$ is chosen such that the integral over the surface is zero, this constraint is automatically satisfied as the parameters are varied. The system of equations is regularized by substituting an equation, fixing the salt field at a particular point, for one of the algebraic equations in (2.5).

The accuracy of the results was considered for a particular case in Quon & Ghil (1992). In their notation we fix $d = 0.2$ and $Ra_T = 5\ 10^6$. The function $T_S(x)$ given by

$$T_S(x) = \frac{1}{2}\ \left(\cos\left(2\pi\left(\frac{x}{A} - \frac{1}{2}\right)\right) + 1\right) \tag{2.6}$$

and the flux $Q_S(x)$ was computed in the same way as in Quon & Ghil (1992) and is shown as the solid curve in figure 3. The bifurcation diagram is similar to that of figure 16 in Quon & Ghil (1992) and shown for $N = 60$, $M = 30$ in figure 2. The vertical axis in figure 2 is the value of the streamfunction (ψ_c) at the centre of the flow domain. Note that in Quon & Ghil (1992) results were presented using $\sigma/0.32$, instead of σ. Results on the accuracy of the symmetry-breaking pitchfork bifurcation

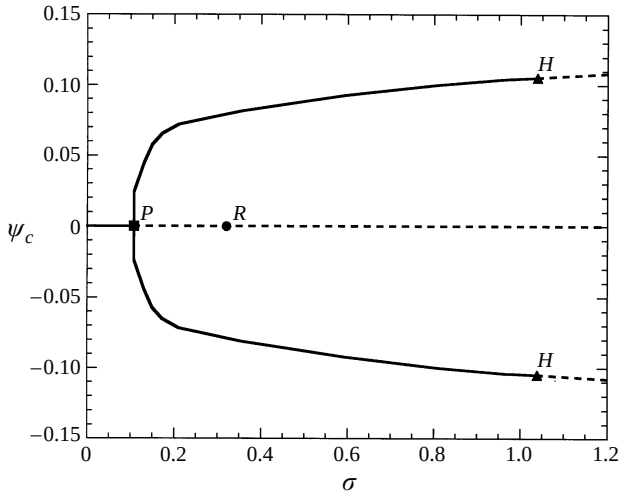

FIGURE 2. Bifurcation diagram for the case considered in figure 16 of Quon & Ghil (1992), with $A = 5.0$ and $Ra = 4.0 \times 10^4$. On the vertical axis, the streamfunction at the centre is plotted. In this plot and similar ones below, solid (dashed) lines indicate stable (unstable) solution branches.

N	M	σ_P
30	15	0.1075
60	30	0.1143
120	60	0.1160
240	120	0.1164

TABLE 1. Values of σ at the symmetry-breaking bifurcation point P in figure 2 for several grids

point P in figure 2 are shown in table 1. A monotonically increasing value of σ_P is found with nearly quadratic convergence of this value with decreasing grid size. A 60×30 grid gives sufficiently accurate results since σ_P differs less than 2% from the high-resolution value; this resolution is used in the computations below. The value of σ_P at the pitchfork bifurcation P is very close to that in Quon & Ghil (1992) although we prescribed slightly different boundary conditions, because we allow slip at both horizontal walls. At large fresh-water forcing, where the circulation is nearly pole to pole, the stable steady states become unstable to time-periodic perturbations at the Hopf bifurcation points H. The latter were not shown in Quon & Ghil (1992), but were found in Quon & Ghil (1995) in a different region of parameter space.

3. Results

The parameters $Pr = 2.25$ and $Le = 1$ are fixed as in Quon & Ghil (1992). The value of Le is reasonable because there are no good arguments why the large-scale heat and salt transport due to (turbulent) subgrid-scale processes should be different. The value of Pr is chosen for comparison with results in Quon & Ghil (1992); the bifurcation diagrams presented below hardly change quantitatively if Pr is increased. The aim of this section is (i) to investigate the structure of stable and unstable steady solutions for the particular case $Ra = 4 \times 10^4$ and $A = 5$, (ii) to study integral properties of the solutions and (iii) to describe the physical mechanisms of symmetry

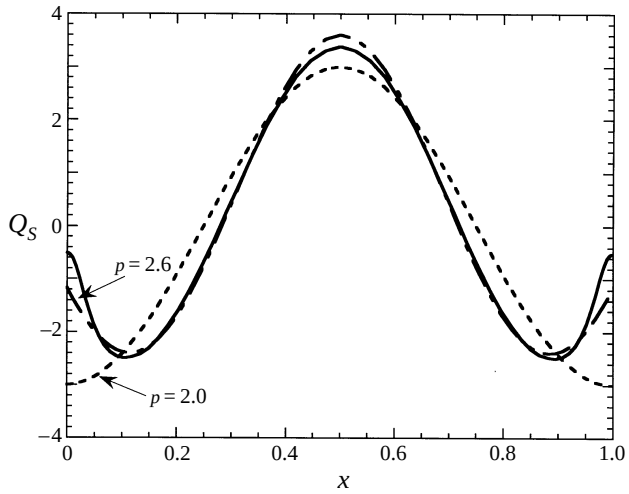

FIGURE 3. The fresh-water flux forcing $Q_S(x)$ as a function of x. The solid line is the fresh-water flux used to compute figure 2. The other curves are the flux according to equation (3.1) in the text for two values of p.

breaking and overturning oscillations associated with the occurrence of pitchfork and Hopf bifurcations.

3.1. *Bifurcation diagram*

For $Ra = 4 \times 10^4$ and $A = 5$, the fresh-water flux used to compute figure 2 is shown as the solid curve in figure 3. This flux can be well-approximated by a function of the form

$$Q_S(x) = 3 \cos\left(p\pi\left(\frac{x}{A} - \frac{1}{2}\right)\right) - \frac{6}{p\pi} \sin\left(p\frac{\pi}{2}\right) \tag{3.1}$$

By varying the parameter p in (3.1), the spatial pattern is changed and two examples ($p = 2.0$ and $p = 2.6$) are shown in figure 3. When p is changed, the fresh-water flux mainly varies in the polar regions.

The bifurcation diagram computed by prescribing (2.2c) with (2.6) and (3.1), the latter with $p = 2.6$, is presented in figure 4. The value at the vertical axis (ψ_{RM}) is that of the streamfunction at the particular gridpoint ($i = 45$, $j = 15$), such that the different branches can be well distinguished in a plot. Bifurcation points along the branches are indicated by symbols: squares indicate pitchfork bifurcations and triangles Hopf bifurcations. Saddle node bifurcations (limit points) are not marked because they are obvious from the shape of the branches. Solid branches indicate linearly stable solutions whereas the solutions along dashed branches are unstable. Corresponding flow patterns (through contour plots of the streamfunction ψ) and salinity fields at several marked points along the branches in figure 4 are shown in figure 5.

For small σ, the bifurcation diagram in figure 4 is qualitatively similar to that in figure 2. The symmetric (figure 5a) thermally dominated 2-cell state – along the TH-branch (Thual & McWilliams 1992) – becomes unstable at the supercritical pitchfork P_1 near $\sigma = 0.13$ at which two symmetry-related asymmetric states stabilize. Both asymmetric states (the southward sinking solution is shown in figure 5b) – PP-branches in Thual & McWilliams (1992) – remain stable up to the Hopf bifurcations

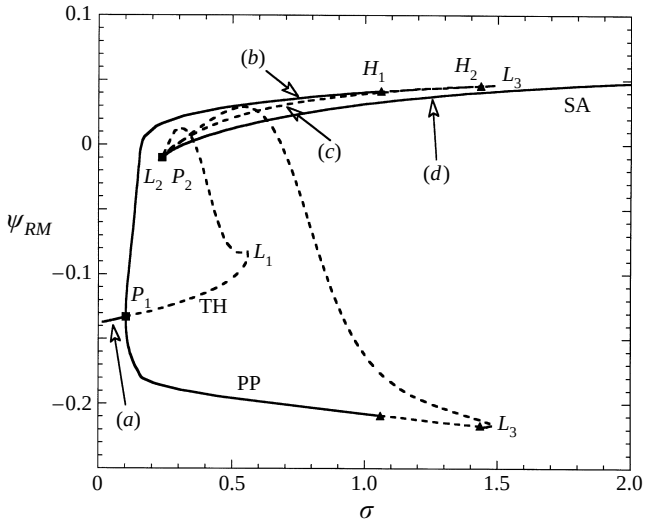

FIGURE 4. Bifurcation diagram for the case $A = 5$, $Ra = 4.0 \times 10^4$ and prescribed $Q_S(x)$ with $p = 2.6$. On the vertical axis, now the streamfunction value at a particular gridpoint $\psi_{RM} = \psi(45, 15)$ is shown.

H_1 at $\sigma = 1.06$ where they becomes unstable. At slightly larger σ, these states stabilize again at the Hopf bifurcations H_2 ($\sigma = 1.44$), but they eventually destabilize at the limit point L_3 ($\sigma = 1.47$). Along an unstable branch, the 1-cell patterns then deform towards 2-cell solutions with equatorial downwelling (figure 5*c*) and connect at a second pitchfork bifurcation P_2 ($\sigma = 0.24$) to the stable branch of a salinity-dominated 2-cell (figure 5*d*) solution, the SA-branch in Thual & McWilliams (1992). The stability properties of this branch are difficult to detect very near P_2. The SA-solution is certainly stable for $\sigma > 0.37$ and remains stable up to large values of σ. For values of $\sigma < 0.37$, Hopf bifurcations may exist along the SA-branch but are hard to detect. The unstable 2-cell TH-solution also connects up to the SA-branch at P_2, after it has gone through two limit points L_1 and L_2, the latter being very close to P_2. Hence, there are three σ-intervals where a unique stable steady state appears, and three intervals where there are multiple stable steady states. The latters intervals are given by the σ-values between both pitchfork bifurcations P_1 and P_2, between the pitchfork P_2 and the Hopf bifurcation H_1 and between the Hopf bifurcation H_2 and the limit point L_3.

3.2. *Integral properties of the solutions*

In the previous section it was found that only four types of solutions are linearly stable. In this section, we investigate integral properties of these solutions. One can anticipate that on the TH-branch, the solution is thermally driven and that on the SA-branch it is saline driven. However, it is not immediately clear what the dominant driving force is on both asymmetric branches, how the overturning transport varies with the driving force and how salt is transported by the overturning flow.

Whether the flow is saline or thermally driven can be decided from the volume-integrated mechanical energy balance. This balance is obtained by taking the inner product of the velocity and the momemtum balance and integrate the resulting

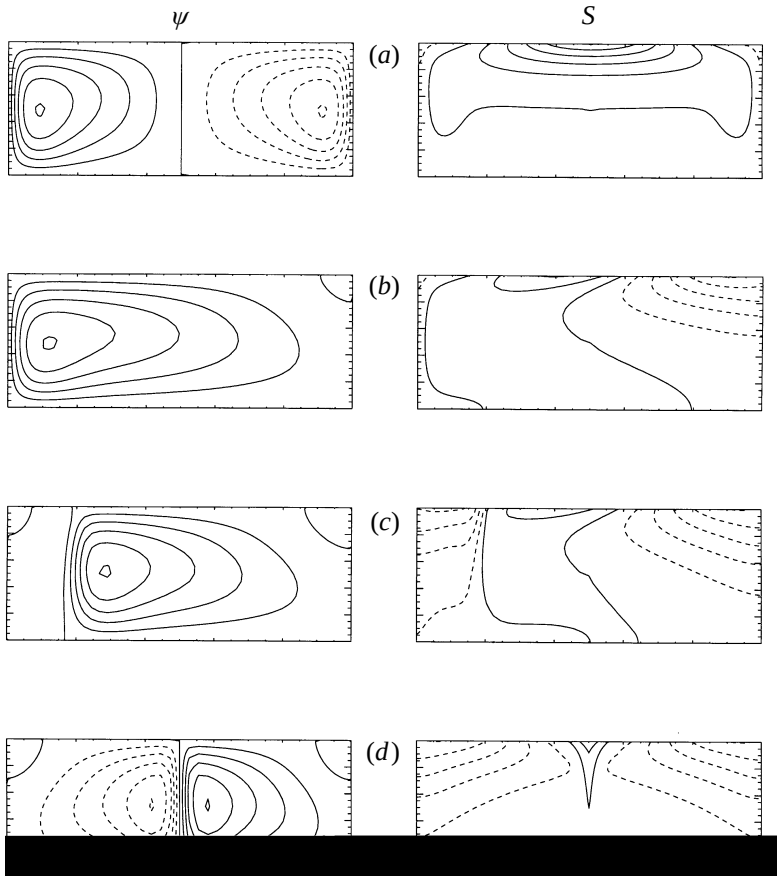

FIGURE 5. Contour plots of streamfunction, and salinity for marked points in figure 4. In these plots and all others below, all fields are scaled with their absolute maximum value and contours are relative to this maximum. Contour interval is 0.2 and dashed contours represent negative values.

equations over the flow domain. This gives with $\boldsymbol{u} = (u, w)$,

$$Pr^{-1}\frac{\mathrm{d}\mathscr{E}}{\mathrm{d}t} = -\langle\mathscr{D}\rangle + Ra\,\langle wT - wS\rangle \tag{3.2a}$$

where the volume-integrated kinetic energy $\mathscr{E}$ is defined as

$$\mathscr{E}(t) = \tfrac{1}{2}\langle\boldsymbol{u}\cdot\boldsymbol{u}\rangle \equiv \frac{1}{2A}\int_0^A\int_0^1 \boldsymbol{u}\cdot\boldsymbol{u}\,\mathrm{d}x\mathrm{d}z \tag{3.2b}$$

and $\mathscr{D}$ is the dissipation function,

$$\mathscr{D} = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial z}\right)^2. \tag{3.2c}$$

Because the dissipation $\mathscr{D}$ is always positive, it follows from (3.2a) that in steady state $\langle wT - wS\rangle \geqslant 0$. Hence, we can distinguish the solutions according to the sign of $\langle wT\rangle$ and $\langle wS\rangle$. The solution is thermally driven and inhibited by fresh-water forcing if $\langle wT\rangle > 0$ and $\langle wS\rangle > 0$. It is saline driven and inhibited by thermal forcing if $\langle wT\rangle < 0$ and $\langle wS\rangle < 0$ and it is driven by both mechanisms if $\langle wT\rangle > 0$ and $\langle wS\rangle < 0$. If $\langle wT\rangle < 0$ and $\langle wS\rangle > 0$ then there can be no steady flow.

In figure 6(*a*), the values of $\langle wT\rangle$, $\langle wS\rangle$ and $\langle wB\rangle = \langle wT\rangle - \langle wS\rangle$ are plotted along the TH-branch from $\sigma = 0$ until the point P_2 in figure 4. Along this branch, both $\langle wT\rangle$ and $\langle wS\rangle$ are positive, indicating that the solutions are thermally driven and inhibited by fresh-water forcing. At the limit point L_1, both quantities go through maximal values, whereas $\langle wB\rangle$ assumes a minimal value. A similar plot (figure 6(*b*)) along the SA branch in figure 4 shows that both $\langle wT\rangle$ and $\langle wS\rangle$ are negative along this branch. Hence, the solutions are saline driven and inhibited by thermal forcing and $\langle wB\rangle$ increases along the branch. On the PP-branches originating from the symmetry-breaking bifurcation, $\langle wS\rangle$ is negative and $\langle wT\rangle$ decreases monotically and even becomes negative at P_2, where the branch connects to the SA-branch (figure 6*c*). At the bifurcation point P_1, the value of $\langle wS\rangle$ is still small whereas the value of $\langle wT\rangle$ is positive. The stable pole-to-pole solutions are therefore driven by both thermal and saline forcing. It is remarkable that at the limit point L_2, the value of $\langle wB\rangle$ goes through a maximum. One can therefore characterize the limit points in the bifurcation diagram by the extrema of these integral properties.

A second issue of interest is the relation between the overturning strength and the driving force of the circulation. In recent work (Rahmstorf, Marotzke & Willebrand 1995), it was found that within an ocean general circulation model, there is a near linear relation between these quantities. In the model used here, the overturning strength is conveniently measured by the maximum of the streamfunction ψ_M. For the density difference driving the flow, one could take a difference in the density between the north and south ends of the basin, for instance at the surface or vertically integrated. However, in that case it turns out that the result is quite sensitive to the choice of the horizontal positions from which the density gradient is computed. Alternatively, the volume-integrated energy balance (3.2*a*) provides a definition for the driving force. Since $\langle wB\rangle$ drives the flow, a characteristic vertical velocity w_o can be defined by the overturning strength together with an overall density difference $\Delta\rho$ responsible for driving the flow as

$$w_o\Delta\rho = -\langle wB\rangle \tag{3.3a}$$

If $\mathscr{L}$ is the (dimensionless) horizontal scale over which downwelling occurs for any solution, then $w_o \approx \psi_M/\mathscr{L}$. Once the relation between $\langle wB\rangle$ and the overturning strength is determined, the relation between $\Delta\rho$ and ψ_M follows from (3.3*a*). In figure 7, we have plotted $\langle wB\rangle^{1/2}$ against ψ_M along the branches in figure 4. It turns out that on each branch the relations are not exactly linear but do not deviate too much. These deviations might result from the influence of the sidewalls, which can be substantial at the relatively small aspect ratio considered in figure 4. The relation indeed becomes better in the large-aspect-ratio limit, where one term in the dissipation (i.e. $(\partial u/\partial z)^2$) becomes dominant. In this case, we find from (3.2*a*) and (3.3*a*) that

$$w_o\Delta\rho \approx -\left\langle \left(\frac{\partial^2\psi}{\partial z^2}\right)^2 \right\rangle. \tag{3.3b}$$

As an estimate of the right-hand side of (3.3*b*), we can approximate the second derivative at the horizontal location of maximum overturning by central differences and use the fact that ψ is zero at the top and bottom of the domain. This rough estimate then leads, with the previous scaling for w_o, to a linear relation $\Delta\rho \approx \psi_M$. It would be interesting to investigate this relation in more realistic ocean models using a definition of $\Delta\rho$ based on the overall energy balance.

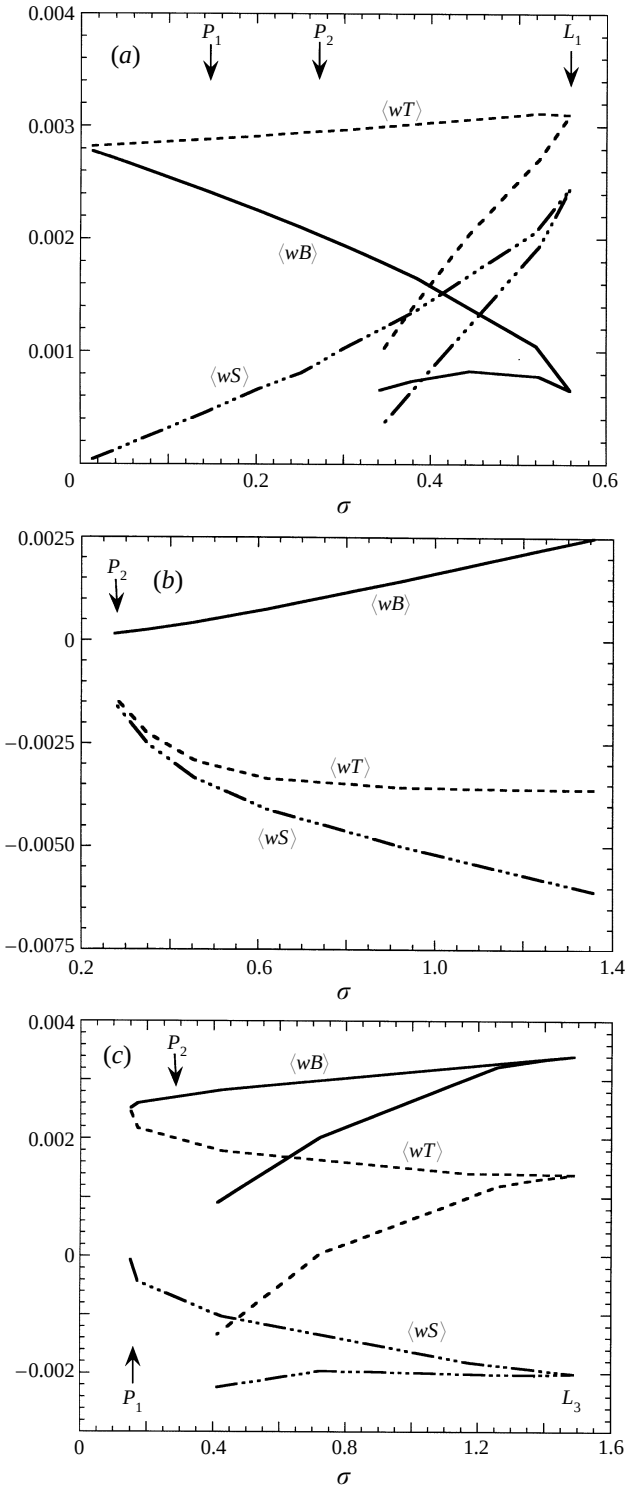

FIGURE 6. Plots of the integral quantities $\langle wT\rangle, \langle wS\rangle$ and $\langle wB\rangle$ along the branches in figure 4. (*a*) TH-branch; (*b*) SA-branch; (*c*) PP-branches

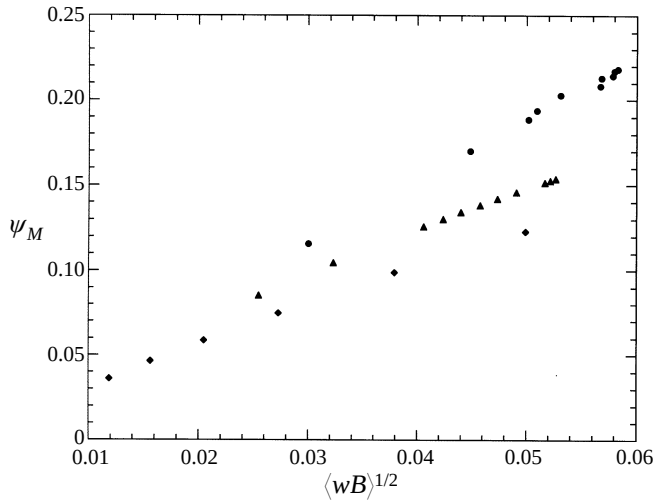

FIGURE 7. Plot of the maximum streamfunction ψ_M along the branches in figure 4 against a measure of the buoyancy forcing $\langle wB \rangle^{1/2}$.

A third issue is the transport of salt by the overturning circulation for the different solutions. For each type of solution, the vertically averaged horizontal salt transport is solely determined by the surface boundary conditions. If for a particular steady state, (2.1*d*) is integrated over the domain $[x_0, A] \times [0, 1]$ and the boundary conditions (2.2) are applied, we find (with $Le = 1$)

$$\int_{x_0}^{A} Q_S(x)\mathrm{d}x = -\int_0^1 \left(uS - \frac{\partial S}{\partial x}\right)(x_0, z)\mathrm{d}z \equiv \Phi(x_0). \tag{3.4}$$

The salt flux Φ consists of a diffusive part (in ocean models due to subgrid-scale processes) and an advective part. If Φ is positive, salt is exported into the region $[0, x_0]$; if Φ is negative, salt is exported into the region $[x_0, A]$. Hence, if $Q_S(x)$ is chosen to be positive at the equator and negative at higher latitudes and symmetric, e.g. as in (3.1), then $x_0 < \frac{1}{2}A \Leftrightarrow \Phi > 0$ and hence for *every* solution, independent of whether it is driven thermally, by salinity or driven by both agents, there will be export of salt into the region $[0, x_0]$. Although in realistic ocean models, the flux $Q_S(x)$ is not perfectly symmetric, it certainly has the same features as in (3.1), with net equatorial evaporation and poleward freshening. Also in these models there must be a net export of salt at the south end of the Atlantic basin, since there is a similar constraint as in (3.4). However, the advective part then consists of both the density-driven overturning component and a wind-driven flux. The latter may substantially modify the salt transport, but the net export is always constrained by the surface forcing as in the current model.

3.3. *Physical mechanisms of transition: symmetry breaking*

Both the TH-branch and the SA-branch undergo symmetry-breaking bifurcations. Both 2-cell symmetric solutions become unstable to a perturbation having a particular spatial pattern. For instance, for values of σ slightly larger than the value at P_1 in figure 4, the symmetric TH-solution will be unstable to a particular pattern even if this disturbance has a very small ampitude initially. With the continuation method, we can locate the bifurcation point up to a prescribed accuracy and solve for the linear

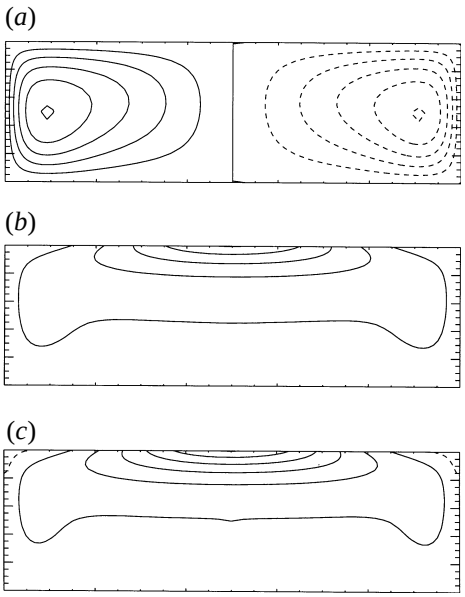

FIGURE 8. Contour plots of the streamfunction (*a*), temperature (*b*) and salinity (*c*) of the steady TH-solution near the bifurcation point P_1 in figure 4.

stability problem. The spatial pattern of the most unstable perturbation is therefore contained in the eigenvector corresponding to a zero eigenvalue, near the bifurcation. To show that a steady state becomes unstable to that particular disturbance, we aim to describe the causal chain of how this disturbance is amplified through its interaction with the steady state. A mechanistic understanding of the physics of this process can be obtained by dividing the instability process into two virtual stages, an initiation stage and a growth stage. Both stages obviously occur simultaneously in reality, but this two-stage view has proved to be useful in other stability problems (Dijkstra & Steen 1991). During the initiation stage, a perturbation is assumed to be present in the system and we describe the causal chain for how the perturbation changes the steady state. During the growth stage, a description is provided of how the changes in the steady state feed back to the perturbations leading to amplification of the original perturbation.

We first consider the instability of the 2-cell symmetric TH-solution at the bifurcation point P_1 (in figure 4). The steady-state streamfunction, temperature and salinity are plotted in the figures 8(*a*), 8(*b*) and 8(*c*), respectively. The patterns of the most unstable perturbation are plotted in figure 9. Consider the salinity perturbation in figure 9(*c*) as the initial disturbance. This salt perturbation is positive over most of the northern part of the basin and negative over the southern part, with substantial gradients near the equator. From the density perturbation (figure 9*d*), we see that salinity determines the sign of the density perturbation. The perturbation salt gradient therefore drives the flow which is seen in figure 9(*a*). The temperature perturbation is compatible with this flow: in the northern hemisphere, warm water is transported northwards and in the southern hemisphere, cold water is transported upwards.

It is observed that the northern cell of the steady state (figure 8*a*) is strengthened by the flow disturbance (figure 9*a*) and the southern cell is weakened. The horizontal perturbation velocities at the surface induce a horizontal salt transport (figure 8*c*).

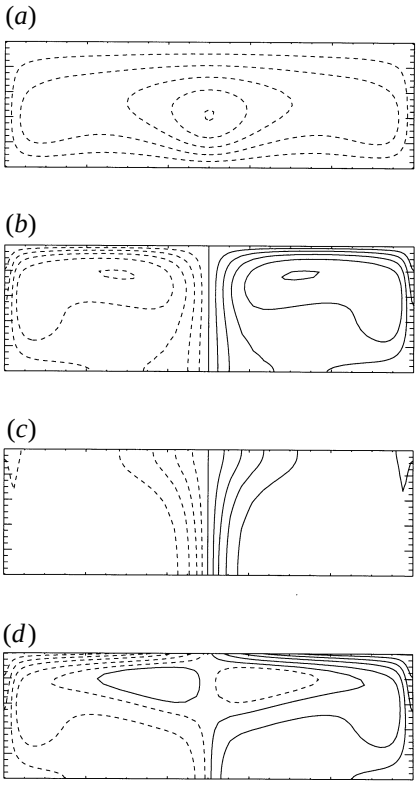

FIGURE 9. Contour plots of the streamfunction (*a*), temperature (*b*), salinity (*c*) and density (*d*) of the perturbation destabilizing the steady solution in figure 8 at P_1.

Note that there is no surface perturbation heat transport, because the temperature perturbation is zero at the surface. In the northern part of the basin, the term $u'\bar{S}_x$, where the bar denotes the steady state and the prime the perturbation, dominates the evolution of the salinity perturbation. Since it is negative, the tendency of the salinity perturbation is positive, leading to an amplification of the original disturbance. Similar reasoning holds for the southern part, where the sign of $u'\bar{S}_x$ is positive. Note that the temperature perturbation plays a rather passive role in this mechanism, except that it weakens the perturbation flow, because of its influence on the perturbation density field. In other words, σ has to be large enough such that the salinity anomaly induces the perturbation flow. However, the thermal field itself is crucial since the temperature field maintains the circulation of the equilibrium state.

Next, the instability of the 2-cell SA-solution at the pitchfork P_2 is considered. Although the instability of the TH-solution was addressed schematically by Marotzke *et al.* (1988), the destabilization of the SA-branch for decreasing values of σ through a pitchfork bifurcation has got little attention, if any. We consider the steady state near the bifurcation point P_2 in figure 4. The streamfunction, temperature and salinity are plotted in figure 10 and the perturbation structures of the mode inducing the symmetry breaking are shown in figure 11. There is a remarkable difference between these perturbation structures and those destabilizing the TH-solution (compare figure 9*a* with figure 11*a*). In this case, the flow perturbation is localized near the equator and does not extend over the whole flow domain. Since the density is again controlled by salinity (figure 11*c*, *d*), the perturbation flow is driven

(*a*)

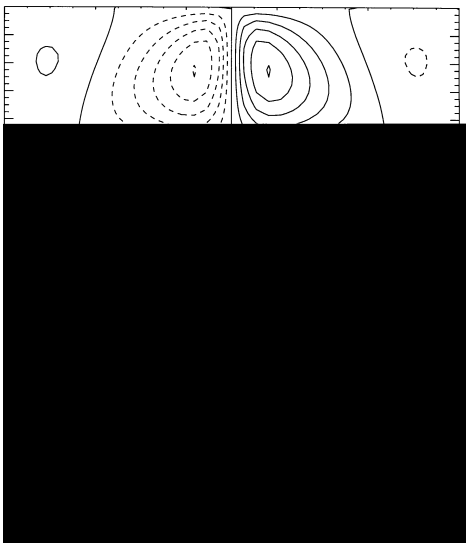

FIGURE 10. Contour plots of the streamfunction (*a*), temperature (*b*) and salinity (*c*) of the steady SA-solution near the bifurcation point P_2 in figure 4.

by the salinity anomaly and the temperature perturbation is again compatible with the perturbation flow. Owing to the perturbation flow, the left-hand circulation cell of the basic state (figure 10*a*) is now strengthened, whereas the right-hand circulation cell is weakened. Note that a broad scale flow anomaly such as in figure 9*a* would cause just the opposite effect, hence the localization of the perturbation is important to induce the symmetry breaking. The perturbation flow leads to an asymmetric salt transport through the transport term $u'\bar{S}_x$ just as it did in the TH-case, leading to the amplification of the original perturbation. For instance, to the right of the equator the horizontal velocity perturbation and the basic-state salt gradient are both negative. Hence the tendency in the salt field $\partial S/\partial t \approx -u'\bar{S}_x$ is negative amplifying the original negative salt perturbation (figure 11*c*).

3.4. *Physical mechanisms of transition: transition to oscillatory flows*

A Hopf bifurcation, for example as found along the asymmetric branch for $p = 2.6$ (H_1 in figure 4), marks the location in parameter space where time-periodic disturbances are about to be amplified through their interaction with the steady state. In this section, we concentrate on the Hopf bifurcation H_1 along the southward sinking branch in figure 4; the equilibrium state at this point is plotted in figure 12. The frequency of the oscillations corresponds to a time scale comparable to the overturning time scale of the steady state. This motivates us to describe the instability mechanism since it might be relevant for overturning oscillations, which have been found in a hierarchy of ocean models and were shown to have a two-dimensional character (Winton & Sarachick 1993).

The Hopf bifurcation is supercritical, since a limit cycle is found for parameter values σ slightly larger than the value at bifurcation. This limit cycle disappears at H_2 (figure 4) in a reverse Hopf–bifurcation. Near the Hopf bifurcation H_1 the corresponding eigenfunctions, say (w_1, w_2), show exactly the time-periodic disturbance structures to which the steady state becomes unstable. The most unstable time-periodic

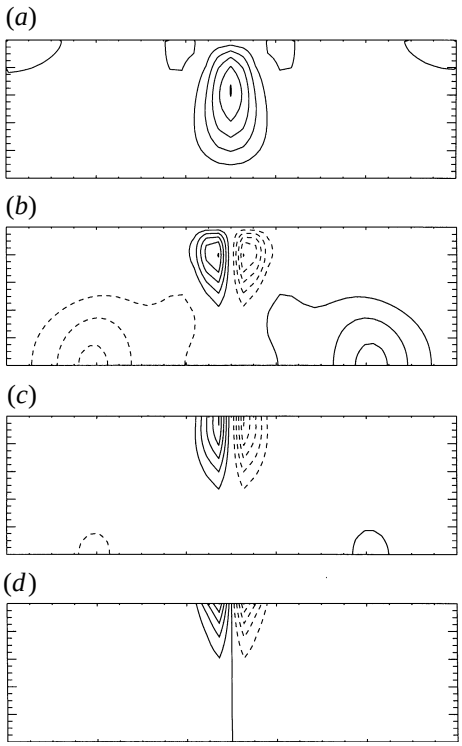

FIGURE 11. Contour plots of the streamfunction (*a*), temperature (*b*), salinity (*c*) and density (*d*) for the perturbation destabilizing the steady solution in figure 10 at P_2.

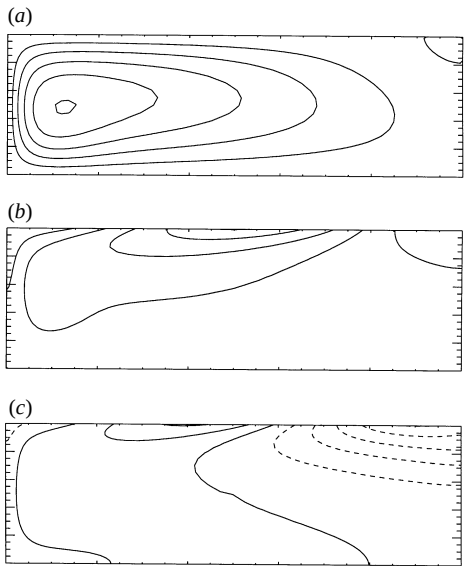

FIGURE 12. Contour plots of the streamfunction (*a*), temperature (*b*) and salinity (*c*) of the steady (southern sinking) PP-solution at the Hopf bifurcation point H_1 in figure 4.

perturbation $\phi(x, y, t)$ is given by

$$\phi(x, y, t) = w_1(x, y)\ \sin(\omega t) + w_2(x, y)\ \cos(\omega t) \tag{3.5}$$

where ω is the angular frequency, corresponding to the imaginary part of the eigenvalue at the Hopf bifurcation. For $\omega t = 0,\ \pi/4,\ \pi/2$ and $3\pi/4$ the streamfunction, density, temperature and salinity corresponding to ϕ are plotted in figure 13(*a–d*), respectively.

Before discussing the oscillation in terms of the changes of the spatial patterns of the perturbation, we first monitor the changes of relevant integral quantities along the oscillation. In figure 14, the terms $\langle wS\rangle' = \langle \bar{w}S' + \bar{S}w'\rangle$, $\langle wT\rangle' = \langle \bar{w}T' + \bar{T}w'\rangle$ and $\langle wB\rangle' = \langle \bar{w}B' + \bar{B}w'\rangle$ are plotted along one period of the oscillation. Here the prime denotes the perturbation quantities (as in figure 13) and the bar to the steady state (as in figure 12). Within the linearized volume-averaged energy balance for the perturbations, the latter term $\langle \bar{w}B' + \bar{B}w'\rangle$ appears as the buoyancy forcing. The buoyancy forcing is mainly determined by the salinity forcing (figure 14), indicating that the salinity perturbation field mainly drives the oscillatory flow. The thermal contribution to the forcing, becoming positive over half a cycle of the oscillation, introduces the phase difference between salinity and buoyancy forcing. In this way, the oscillation resembles a 'thermohaline loop' oscillation as presented in Welander (1986) in his 'oscillator gallery'.

We now describe a more detailed mechanism compatible with the results in figure 14, accepting that we know what the steady state (figure 12) and perturbation structures (figure 13) look like. Hence, we do not try to answer why these perturbation structures have this particular shape. Within the initiation stage one needs to describe the causal chain of why the perturbations return after half a cycle to the initial pattern but with negative sign, i.e. this stage is concerned with the essence of the oscillation. The growth phase is concerned with the amplification of the perturbation after one cycle of the oscillation. We restrict the description to the first stage; we did not succeed in describing the growth stage in a satisfactory way. Suppose that the salt perturbation in figure 13(*d*), applied at $t = 0$, is the initial disturbance. Since it controls the density anomaly (figure 13*b*, $t = 0$) it causes liquid to sink in the south and to rise over the rest of the basin, thereby giving the perturbation flow structure in figure 13*a* ($t = 0$). The temperature perturbation is consistent with this flow giving relatively warmer water in the south through advection and colder water over the central part in the basin (figure 13*c*, $t = 0$). Because the flow perturbation strengthens the equilibrium state circulation (figure 12*a*), it transports salt to the central part of the basin (figure 13*d*, $t = 0.125$). The presence of this heavier water substantially weakens the flow in the northern part (figure 13*a*, $t = 0.125$). In this region of weak flow, the heat transport is dominated by diffusion and colder water appears in the northern region (figure 13*c*, $t = 0.25$). This induces a reverse flow in the northern part of the basin (figure 13*a*, $t = 0.25$). The reverse flow affects the salt perturbation (figure 13*d*, $t = 0.25$), and the upwelling perturbation flow (figure 13*a*, $t = 0.25$) transports salt water northwards in the central part of the basin. Hereby, the positive salt perturbation is extended over the whole basin from south to north (figure 13*d*, $t = 0.375$). As a consequency, the salinity is reduced near the southern boundary and this strengthens the reverse flow perturbation in the basin at half a period of the oscillation (figure 13*a*, $t = 0.375$) and the reverse cycle starts. In summary, the oscillation is a based on a combined advective salt/diffusive heat transport, where the salt perturbation drives the oscillation (compatible with figure 14). The temperature perturbation becomes

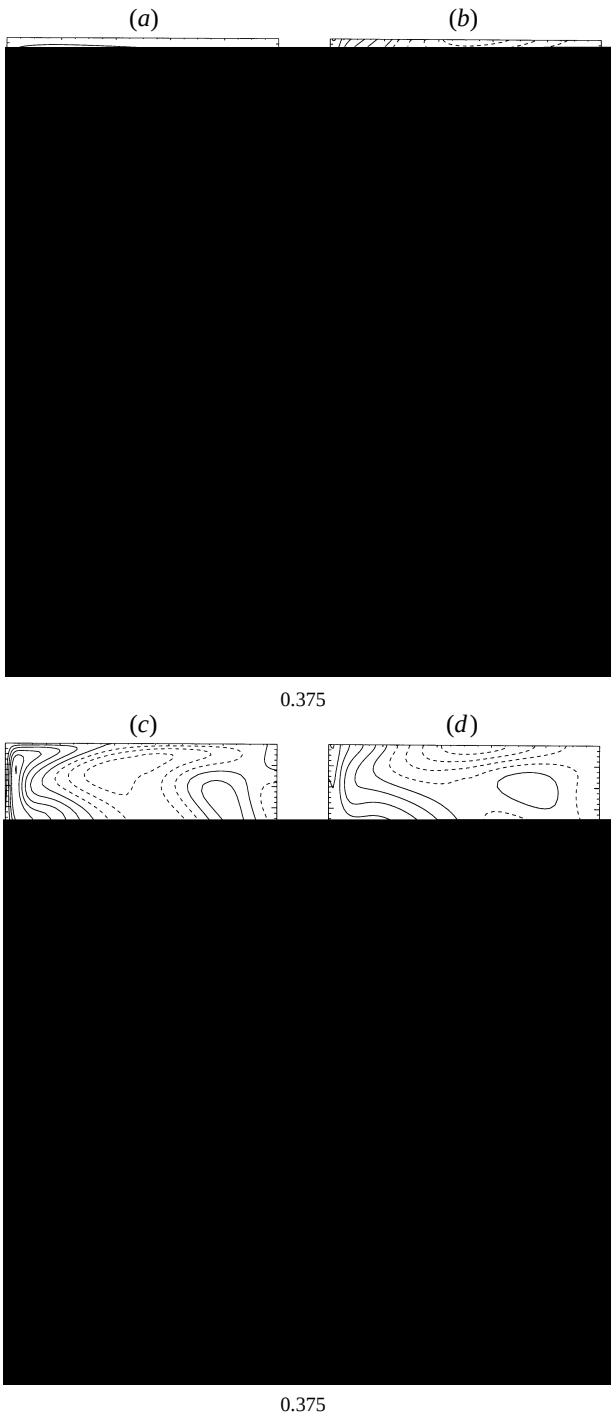

FIGURE 13. Contour plots of the streamfunction (*a*), density (*b*), temperature (*c*) and salinity (*d*) for the periodic disturbance ϕ (cf. equation (3.5)) destabilizing the steady solution in figure 12 at H_1. The patterns are shown for four different times $\omega t/2\pi$ along the orbit. The dimensionless angular frequency $\omega = 0.326$, which implies a period in the order of the overturning time scale of the steady PP-solution.

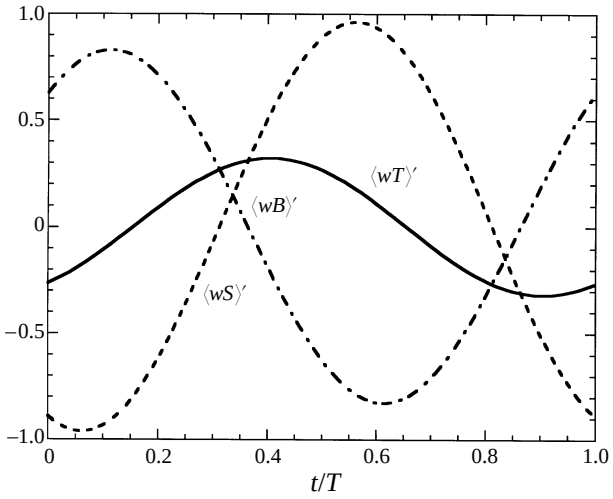

FIGURE 14. Plot of the integral perturbation quantities determining the buoyancy forcing of the perturbation flow at the Hopf bifurcation H_1.

important only in regions where the perturbation flow is weak, inducing the phase difference (figure 14) between salinity and buoyancy forcing.

4. Sensitivity to different surface boundary conditions

4.1. *Fresh-water forcing*

In this section, the shape of the fresh-water flux forcing is varied by decreasing the parameter p from 2.6 to 2.0. For $p = 2.0$, the spatial pattern of the fresh-water flux forcing and the thermal forcing have a similar shape. The bifurcation diagram for the same parameter values as in figure 4 ($A = 5,\ Ra = 4 \times 10^4$) is presented in figure 15 and differs from figure 4 in that the Hopf bifurcations have disappeared and two additional limit points L_4 and L_5 occur on the asymmetric branches. Along each asymmetric branch a complex-conjugate pair of eigenvalues appears as least stable modes in the linear stability analysis, just as in the $p = 2.6$ case, but this pair does not cross the imaginary axis before the limit point L_3 is reached. The oscillatory perturbations are fairly similar to those for the case described in the previous section, and the occurrence of the amplification of the oscillatory perturbations for the case $p = 2.6$ is likely to be associated with the decrease in fresh-water flux polewards (cf. figure 3).

One might conjecture that these overturning oscillations only exist as bifurcations from the PP-solutions if the fresh-water flux has a mimimum within each hemisphere (cf. figure 3). However, careful analysis shows that this is not the case. What happens in the bifurcation diagram is that the second Hopf bifurcation H_2 in figure 4 moves towards H_1 with decreasing p and eventually both coalesce between $p = 2.25$ and $p = 2.3$ and disappear. This demonstrates the sensitivity of the oscillatory instabilities to the spatial pattern of the surface fresh-water flux.

4.2. *Thermal forcing*

As has been discussed in many papers on the subject, mixed boundary conditions are not a good representation of the air–sea interaction processes. In reality, the ocean

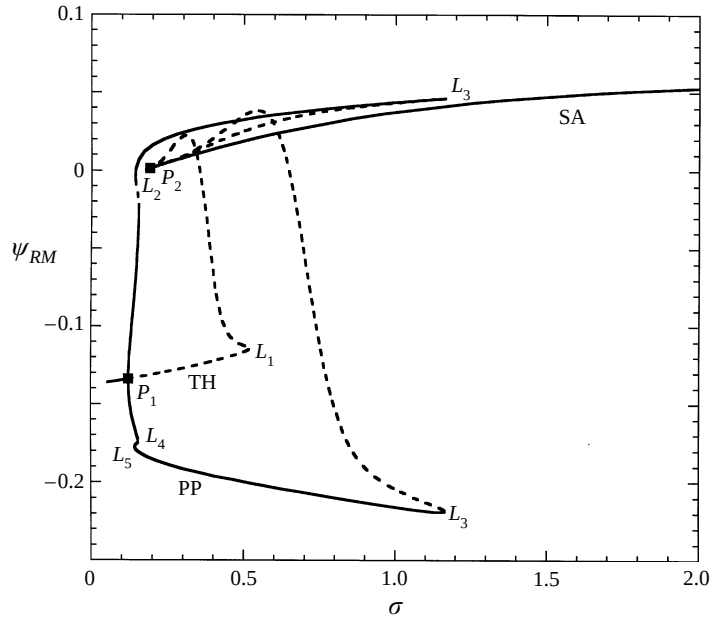

FIGURE 15. Bifurcation diagram as in figure 4 but for $p = 2.0$.

is coupled to an atmosphere and the temperature of the upper ocean layer is not fixed but determined by the local heat balance between ocean and atmosphere. The atmosphere is scale selective with respect to sea surface temperature anomalies in that small-scale anomalies are more damped than large-scale anomalies (Marotzke 1993). Similarly, the fresh-water flux is determined by the vapour transport through the atmosphere and the local salt flux through the ocean–atmosphere surface is zero (Huang 1993).

In this section, the bifurcation structures of the thermohaline flows are determined for a more general form of the thermal boundary condition. Our main aim is to investigate systematically, whether the bifurcation structure as found in the previous section for mixed boundary conditions is robust when more realistic boundary conditions are prescribed. In other words, are the changes due to a more general thermal boundary condition qualitative and/or quantitative ?

The temperature condition at the surface is changed to a Newtonian cooling condition. In large-scale ocean models the ocean–atmosphere heat flux H_f is parameterized as $H_f = \rho_0\, \Gamma\, C_p\, (T^* - T)$, where Γ is interpreted as a ratio of the mixed layer depth and a relaxation time scale and T^* as a radiation equilibrium temperature. Here, it is assumed that this temperature is prescribed and equal to T_S. The non-dimensional heat transfer boundary condition then becomes

$$\frac{\partial T}{\partial z} = -B(T - T_S(x)) \tag{4.1}$$

where $B = H\Gamma/\kappa_T$. The parameter B is a ratio of the diffusive time scale of (vertical) heat transfer and the relaxation time scale H/Γ. Typical values used in large-scale ocean models give values for B of order 100. In the limit $B \to \infty$ the restoring boundary condition for temperature is obtained.

The bifurcation diagram for the case $p = 2$, $A = 10$, $Ra = 10^4$ and $B = 100$ is presented in figure 16. Although the Newtonian cooling boundary condition gives

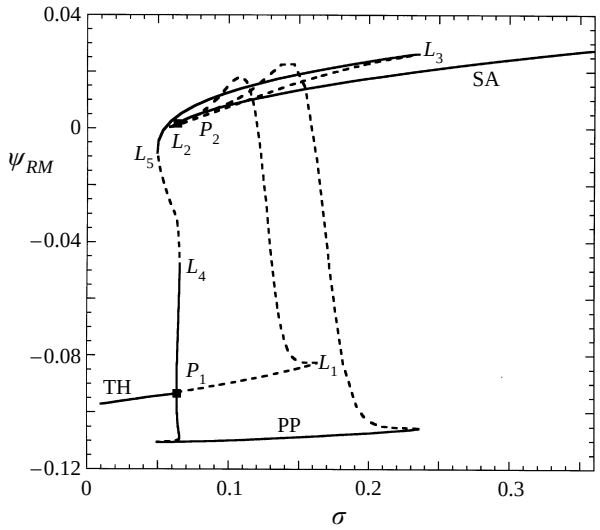

FIGURE 16. Bifurcation diagram as in figure 4 but for $p = 2.0$, $A = 10$, $Ra = 10^4$ and the Newtonian cooling condition, equation (4.1), with $B = 100$.

B	P_1	$P_2 = L_2$	L_1	L_3
∞	0.0639	0.0584	0.1616	0.2329
100	0.0638	0.0585	0.1686	0.2358
10	0.0609	0.0570	0.1555	0.2249
2	0.0505	0.0495	0.0976	0.1302

TABLE 2. Values of σ at the relevant bifurcation points for $A = 10$, $Ra = 10^4$ and different B

a slight shift in the position of the bifurcation points, the qualitative dynamical behaviour remains the same as for mixed boundary conditions for this value of B. In Table 2, the σ-values of the relevant bifurcation points P_1, P_2, L_1, L_2 and L_3 are shown for several values of B. The value of σ at P_1 decreases with decreasing B, since the perturbation temperature gradient at the surface becomes smaller. According to the symmetry-breaking instability mechanism as described above, the salt perturbation is not counteracted that much and can induce an asymmetric flow disturbance at smaller σ. In the limit $B \to 0$, the boundary condition reduces to $\partial T/\partial z = 0$ and therefore the flow is purely saline driven; only the SA-branch remains.

In Zhang *et al.* (1993), flows driven by both prescribed fresh-water and heat flux were considered. For this case, they found no PHC, whereas this transition was found under mixed boundary conditions for the same parameters. In the current model, with a linear equation of state, there is effectively one tracer when both fluxes are prescribed since one can add the temperature and salinity equations and obtain a boundary condition just for the sum. With one tracer, only one branch remains and no symmetry breaking occurs. In this case, there are no bifurcation points between the TH and SA parts of the branch. After a restoring run, one always arrives (after switching to prescibed fluxes) at a unique solution, which is either thermally or saline

driven. Hence, even very large salinity perturbations (as used in experiment 4 in Zhang *et al.* 1993) are not able to induce a transition to another flow pattern. Although in Zhang *et al.* (1993) a nonlinear equation of state is used, the nonlinearity is weak and the large difference of the behaviour of the flow for both boundary conditions, as is seen in their results, is not surprising.

5. The bifurcation structure at large A

In this section, we determine the origin of the multiple equilibria by following the path of the relevant bifurcation points in the three-dimensional parameter space (Ra, A, σ) for $B \to \infty$ and compare these results with those in Cessi & Young (1992) and Thual & McWilliams (1992). From the results above, it follows that there are σ-intervals where multiple stable symmetric equilibria occur. Also there are regions where both stable asymmetric and symmetric equilibria occur simultaneously. Below we demonstrate that the occurrence of symmetric multiple solutions is caused by a cusp structure associated with the coalescence of the limit points L_1 and L_2. The origin of symmetry breaking is another singularity, associated with the coalescence of the two pitchfork bifurcations P_1 and P_2. Since two parameters are involved to obtain this singularity, it has codimension-two (Guckenheimer & Holmes 1983). For large aspect ratio A, each pitchfork bifurcation P_i moves very close to the limit points L_i and eventually the origin of the complete bifurcation structure in parameter space is a codimension-three bifurcation. The unfolding of this singularity in parameter space contains all possible bifurcation diagrams.

For $A = 10$ and two values of Ra (10^2 and 10^3), the bifurcation diagrams in figure 17a are obtained. The value of $A = 10$ corresponds to intermediate k in Thual & McWilliams (1992), i.e. $k = 0.63$. In the small buoyancy forcing limit, $Ra = 10^2$, the bifurcation diagram is quite simple (figure 17*a*) since a unique solution exists for all σ. The thermally dominated solution at small σ deforms smoothly into the salinity dominated solution at large σ; the transition point is a near-motionless solution. Basically, this is the same picture as in Thual & McWilliams (1992) for small a. At $Ra = 10^3$, two bifurcation points appear on the symmetric branch; a close-up of this bifurcation diagram is shown in figure 17(*b*). At both bifurcation points, linearly stable asymmetric PP-solutions arise, which are related by symmetry. The path in the (σ, Ra)-plane of both pitchforks is plotted in figure 17(*c*) by computing bifurcation diagrams for several values of Ra. When Ra is decreased from 10^2 both pitchforks coalesce and disappear (shown as a limit point in figure 17*c*). Since two parameters, σ and Ra, are needed to obtain this type of singularity, say O_1, it is a codimension-two bifurcation. The bifurcation diagram, for example in figure 16 is easily visualized to arise from figure 17(*b*) due both to the movement of symmetry breaking bifurcation points with increasing Ra and the appearance of additional limit points. Indeed, figure 17(*c*) indicates that these limit points are already present at $Ra = 2 \times 10^3$ since the value of σ at P_2 is already smaller than that at P_2.

Thual & McWilliams (1992) were able to determine the path of the limit points L_1 and L_2 within the same model, using transient techniques, but they did not calculate the paths of the pitchfork bifurcations. They proposed that a cusp (associated with the coalescence of two limit points) is central to the occurrence of symmetry breaking and therefore to the appearance of the asymmetric PP-solutions. However, in our computations for $A = 10$, the limit points disappear at values of Ra larger than the value at O_1. In our case, the path of the pitchfork bifurcations (g_1 and g_2 in their figure 3) therefore encloses the cusp instead of appearing within the cusp structure.

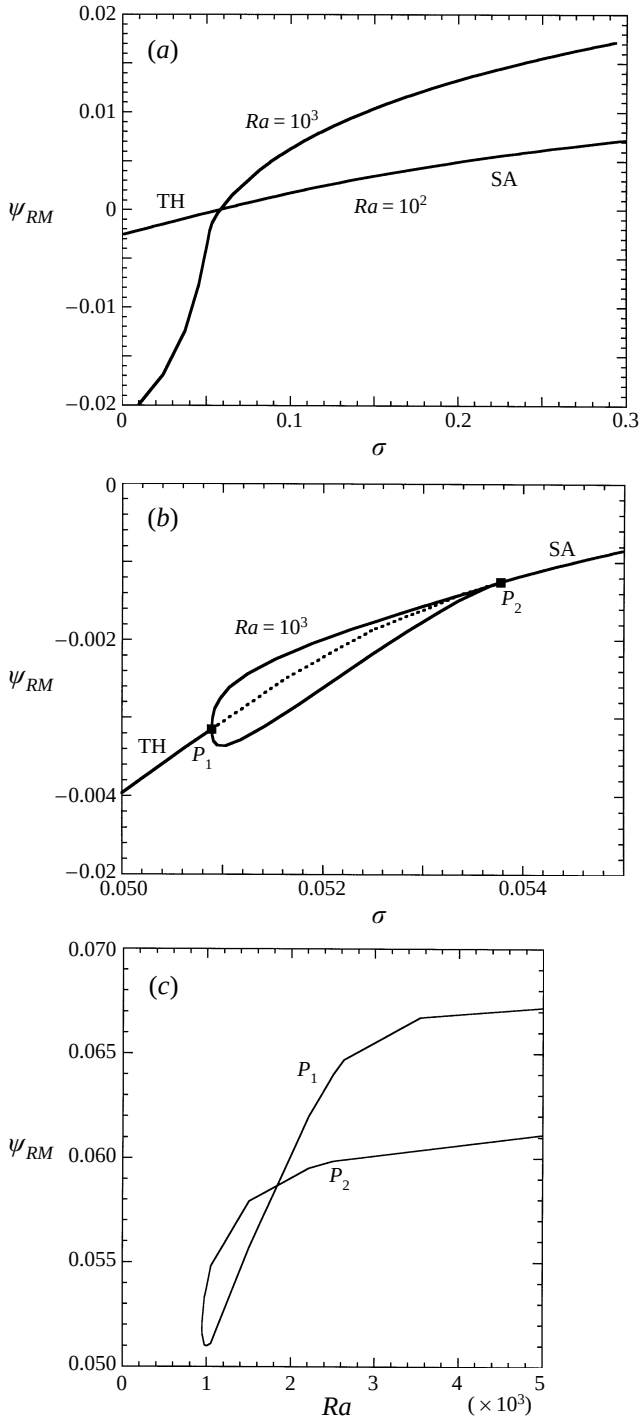

FIGURE 17. (*a*) Bifurcation diagram as in figure 4, but for $A = 10$ and two values of Ra, 10^2 and 10^3. (*b*) Close-up of (*a*) for $Ra = 10^3$. (*c*) Path of both pitchfork bifurcations P_1 and P_2, as shown in (*b*), in the (σ, Ra)-plane.

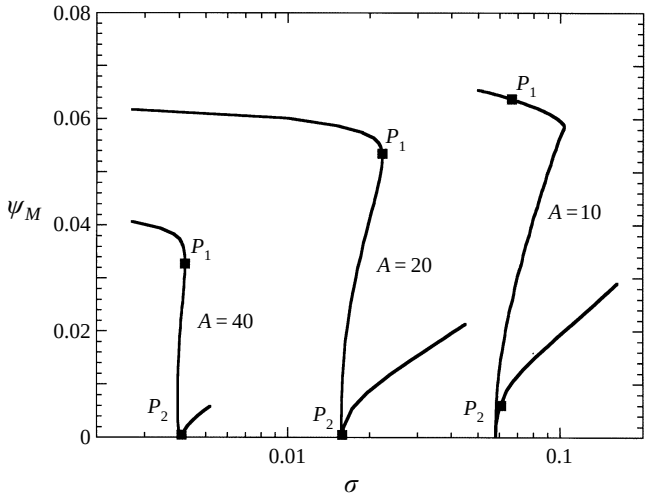

FIGURE 18. Bifurcation diagrams as in figure 17(*a*) (only the symmetric branch) for different aspect ratio A at fixed $Ra = 5 \times 10^3$.

This is not due to the different values of Pr used, since we find the same structure for $Pr = 10^3$. Hence, for relatively small A the sketch in their figure 3 is slightly incorrect since symmetry breaking is not associated with the cusp, but with the coalesence of two pitchfork bifurcations. The latter singularity is the right 'organizing centre' for the symmetry breaking; it is also the generic bifurcation in this $\mathscr{Z}_2$-symmetric system (Golubitsky & Schaeffer 1985).

At larger A, however, the pitchfork bifurcations P_i on the symmetric branch move towards the limit points L_i on this branch. This is seen in figure 18, where the symmetric branch is shown together with the locations of the pitchfork bifurcations for $A = 10,\ 20$ and 40 for $Ra = 5 \times 10^3$. On the vertical axis in figure 18, the maximum value of the streamfunction ψ_M is now plotted. For $A = 40$, the value of σ at the pitchfork and the limit point coincide already up to numerical accuracy. When Ra is decreased for this case, the limit points disappear at the same time as the pitchforks. Hence, this is a codimension-three singularity, say O_2, which is the 'organizing center' of the occurrence of both symmetric and asymmetric multiple equilibria in these types of flows.

In Cessi & Young (1992), an amplitude equation was derived in the large-aspect-ratio limit ($A \to \infty$) and relatively small Ra. In this section we investigate whether this amplitude equation captures the singularity O_2 of the system. With $\epsilon = \pi/A$ we rescale:

$$\tilde{y} = -\pi + 2x\ \epsilon; \quad \tilde{u} = u\ \epsilon; \quad \tilde{w} = w; \quad \tilde{T} = 4\ Ra\ \epsilon^2\ T; \quad \tilde{S} = 4\ Ra\ \epsilon^2\ S \tag{5.1a}$$

to give the dimensionless equations as in Cessi & Young (1992) with

$$a = 4\ Ra\ \epsilon^2;\ b = 4\ \sigma\ Ra\ \epsilon^2. \tag{5.1b}$$

The steady amplitude equation for the first-order (in ϵ) depth-averaged salinity profile $\hat{S}$ becomes

$$\hat{S}'' + \mu^2[\hat{S}'(\hat{S}' - T_S')^2]' + r\ Q_S = \delta^2 \hat{S}'''' \tag{5.2a}$$

with μ proportional to a/ϵ, $r = \sigma/\epsilon^2$ and the primes indicate differentiation to y. The

parameter δ serves to allow for boundary layers in regions of the flow with steep gradients (to regularize the equation (5.2*a*)).

We consider Q_S and T_S given by (2.6) and (3.1), which become rescaled as $Q_S(y) = 3\cos y$, $T_S(y) = 0.5(\cos y + 1)$. These forcing functions give similar results as those originally considered in Cessi & Young (1992), i.e. $Q_S(y) = T_S(y) = \cos y$, for slightly different values of r and μ; this can be shown by rescaling $\hat{S}$ in (5.2*a*). With the latter choice of Q_S and T_S, equation (5.2*a*) can be integrated once in y to give the boundary value problem

$$\delta^2 \chi'' = r \sin y + \mu^2 \chi(\chi + \sin y)^2 + \chi, \tag{5.2b}$$

$$\chi(-\pi) = \chi(\pi) = 0 \tag{5.2c}$$

with $\chi = \hat{S}'$. Solutions of this boundary value problem were computed numerically, by using standard ODE software (Doedel 1981). For $\delta = 0.1$ and $\mu^2 = 7$ the bifurcation diagram with respect to r is shown in figure 19(*a*). This diagram is characterized by two pitchfork bifurcations P_1 and P_2, which connect the TH-, PP- and SA-branches. Similar to the large-A results with the full numerical model, the symmetric branch has two limit points and the location of the pitchforks coincides with those of the limit points up to numerical accuracy. The path of these limit points that in the (r, μ)-plane for several values of δ is shown in figure 19(*b*). It would seem that only two parameters were necessary to obtain this cusp (a codimension-two singularity). However, note that these bifurcation diagrams are valid for very small ϵ (large A) which causes the pitchforks and the limit points to coincide. Hence, ϵ serves as the third parameter and therefore the cusp shown in figure 19(*b*) is in fact a codimension-three singularity, which corresponds to the singularity O_2 obtained earlier. For smaller values of δ this cusp shifts to smaller values of r and μ and converges in the limit $\delta \to 0$ to the diagram presented in Cessi & Young (1992), with the cusp located at $r = \frac{8}{9}$ and $\mu^2 = 3$.

Hence, although the solutions of the amplitude equation in the limit $\delta \to 0$ are not globally defined (as shown in Cessi & Young 1992), they are globally defined at slightly positive δ. The solutions then indeed show boundary layer jumps connecting those solution parts, which already existed locally at $\delta = 0$. The qualitative correspondence between the results of (5.2*b*, *c*) and our numerical results demonstrates that the amplitude equation indeed captures the highest-order singularity leading to multiple equilibria in parameter space.

6. Summary and discussion

We have tried to give a quite complete picture of the qualitative dynamical behaviour of a two-dimensional Boussinesq model of surface-forced thermohaline flows. The use of path-following techniques was demonstrated to be an effective tool to solve for all equilibrium points in parameter space. A quite detailed picture of the connection between solutions and the physics of the transitions between flow regimes is obtained which would be hard to determine by computing trajectories. As is well-known, under restoring boundary conditions for both temperature and salinity, there is only a single branch of solutions connecting salinity-dominated or temperature-dominated solutions when varying the relative forcing strengths. When the parameters are chosen such that the 2-cell thermally dominated pattern is reached for $t \to \infty$, this solution corresponds to a point on the TH-branch (cf. point R in figure 2). When one switches to mixed boundary conditions, the bifurcation diagrams show several possible evolution scenarios if this state is perturbed; we consider those implied by figure 4.

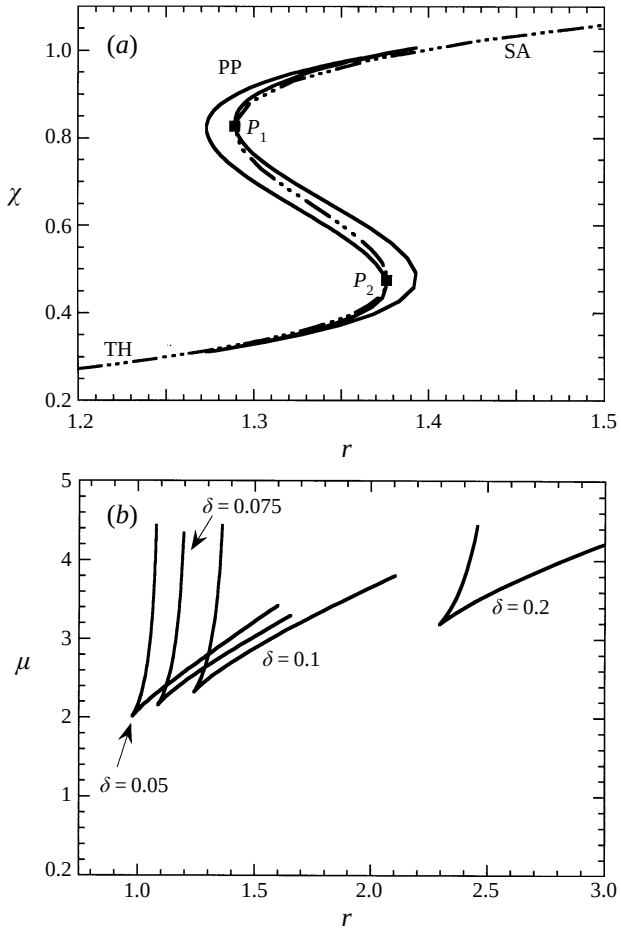

FIGURE 19. Bifurcation diagram obtained with the amplitude equation (5.2) derived in Cessi & Young (1992) for parameters $\delta = 0.1$ and $\mu^2 = 7$. The symmetric branch is dash-dotted, the solid branches are the asymmetric solutions. Path of the pitchfork bifurcations P_1 and P_2 (in *a*) in the (μ, r)-plane for different values of δ.

If the TH-state is stable (values of σ to the left of P_1), then it is unique and any perturbation will eventually decay. If it is unstable, several evolution scenarios are possible. For values of σ between P_1 and P_2, one of the asymmetric states will be reached since these are the only stable states. For values of σ between P_2 and the Hopf bifurcation H_1, the SA-solution can be reached. This would correspond to a complete collapse of the overturning circulation similar to the occurrence of a PHC in a 1-hemispheric ocean model. The occurrence of a PHC is thus connected to an overlap of an (unstable) TH-branch and a (stable) SA-branch. However, one of the asymmetric states can also be reached. In this case, there would be a transition from a symmetric 2-cell flow to a pole-to-pole flow and either direction of these flows would be possible to obtain by providing the right perturbation. This behaviour is similar to that observed in a 2-hemispheric ocean model (Bryan 1986).

The switch from restoring to mixed boundary conditions is not essential; restoring boundary conditions only help the arrival somewhere in state space. However, for each set of parameters one would have a different fresh-water flux, which may modify

the bifurcation structure quantitatively. Hence, in this way no clear sensitivity to parameters in the model is studied, but rather the combined effect of the change of the parameter and the fresh-water flux needed to maintain the restored surface salinity field.

The above-mentioned evolution scenarios are very robust under mixed boundary conditions and not sensitive to the shape of the fresh-water flux. They still occur when a more general interfacial heat transfer condition is used to model the air–sea interaction, for example a Newtonian cooling condition. Although there is much quantitative sensitivity associated with the shift of bifurcation points with parameter changes, the same transitions can still occur but for different parameter values.

The reason for the qualitative correspondence between the results from this two-dimensional model and the results from ocean models is that the details of the momentum balance may not be that crucial compared to the balances of heat and salt. One argument in favour of this is that the physics of the symmetry-breaking phenomenon, responsible for the asymmetric states, does not crucially depend on the details of the velocity perturbation. As long as its north-south component leads to asymmetric flow changes, the original salt perturbation will be amplified. Two-dimensional models, which incorporate the effects of rotation and/or the effect of the wind-driven gyres, also find the same instability mechanism of the TH-solution (Vellinga 1996).

The physical mechanism of symmetry breaking should therefore have some generality and was described in detail as an initiation stage and a growth stage. A salt perturbation drives a perturbation flow which asymmetrically changes the circulation cells of the equilibrium state. The interaction of this perturbation flow with the density distribution causes a lateral salt flux at the surface, whereas the lateral surface heat flux is zero. This leads to asymmetric surface buoyancy transport and amplifies the original perturbation. The same mechanism, albeit with different steady-state and perturbation patterns, also causes the symmetry breaking occurring at the SA-branch. A critical Rayleigh number exists for both bifurcations (figure 17*c*) because the perturbation is otherwise damped by viscosity and/or diffusive effects. A sufficiently large value of σ is required such that the thermal field is not able to counteract the asymmetric flow perturbation caused by the salinity perturbation.

Referring back to figure 4, for values of σ larger than H_1, both the SA-branch and a limit cycle may be reached. The associated periodic orbit corresponds to an overturning oscillation which results from advection of salt. Thermal diffusion acts to give a phase difference between the salinity and buoyancy perturbation fields and is essential to the oscillation. However, the oscillation was found to be sensitive to the fresh-water flux profile. By changing the shape of the fresh-water flux slightly in the polar regions, the Hopf bifurcations easily disappear.

It was demonstrated that the 'organising centre' of the occurrence of the asymmetric solutions is a codimension-two bifurcation, the singularity O_1. This is most clearly seen by considering the flow for relatively small aspect ratio A. The bifurcation O_1 only appears when a critical Rayleigh number is exceeded. There is only one branch of solutions at sufficiently small Ra and no asymmetric solutions exist. At larger Ra, symmetry-breaking bifurcation points appear as an unfolding of O_1. This singularity lies behind the 'superposition principle' proposed in Thual & McWilliams (1992): asymmetric states always arise as a 'superposition' of half a TH-symmetric circulation and half an SA-circulation. Although this qualitative behaviour is in agreement with that proposed by Thual & McWilliams (1992) the highest singularity of importance to symmetry breaking is not the cusp due to the coalescence of limit points, but O_1.

The cusp becomes important for the appearance of multiple symmetric equilibria as seen at large aspect ratio A and the complete bifurcation structure in parameter space arises through an unfolding of the codimension-three singularity O_2. It was shown that the amplitude equation derived in Cessi & Young (1992) indeed captures this 'organising centre'.

Finally, the solution diagrams presented here can be regarded as the underlying perfect bifurcation structures. Any asymmetry posed on the system, for example asymmetric boundary conditions, will destroy the pitchfork bifurcations and force reconnections of the branches (i.e. through limit points). By starting from the symmetric problem, one is able to follow these branches systematically in parameter space. The imperfection structure of the solution diagrams due to flux-correction and asymmetric forcing is currently under study.

All computations were performed on the CRAY-C98 at the Academic Computing Centre (SARA), Amsterdam, the Netherlands within the project SC212. Use of these computing facilities was sponsored by the Stichting Nationale Supercomputer faciliteiten (National Computing Facilities Foundation, NCF) with financial support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization for Scientific Research, NWO). HD thanks J. David Neelin (UCLA, USA) for discussions on the results in this paper and continuing work on the imperfection theory of thermohaline flows, Laurence Fleury and Olivier Thual (CERFACS, France) for ongoing discussions and work on this subject and Will de Ruijter (RUU, NL) for discussions on the description of the mechanism of the overturning oscillation. Finally, we thank A. van der Ploeg and E.F.F. Botta (RUG, NL) for providing their iterative linear systems solvers. We thank one of the (anonymous) referees for pointing out some major deficiencies in the first version of this paper and in particular for the suggestions on how to correct these.

REFERENCES

Broecker, W. S. 1993 The Great Ocean Conveyor. *Oceanography* **4**, 79–89.

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

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

Dijkstra, H. A. 1992 On the structure of cellular solutions in Rayleigh-Bénard-Marangoni flows in small-aspect-ratio containers. *J. Fluid Mech.* **243**, 73–102.

Dijkstra, H. A., Molemaker, M. J., Van der Ploeg, A., Botta, E. F. F. 1995 An efficient code to compute nonparallel flows and their linear stability. *Computers Fluids* **24**, 415–434.

Dijkstra, H. A. & Neelin, J. D. 1995 On the attractors of an intermediate ocean-atmosphere model. *Dyn. Atmos. Oceans* **22**, 19–48.

Dijkstra, H. A. & Steen, P. H. 1991 Thermocapillary stabilization of the capillary break-up of an annular film of liquid. *J. Fluid Mech.* **229**, 205–228.

Doedel, E. J. 1981 AUTO: A program for the automatic bifurcation analysis of autonomous systems. In *Proc. 10th Manitoba Conf. on Numerical Maths and Comput., Univ. Manitoba, Winnipeg, Canada, Cong. Num.* vol. 30, pp. 265–284.

Golubitsky, M. & Schaeffer, D. G. 1985 *Singularities and Groups in Bifurcation Theory*, Part I. Springer.

Guckenheimer, J. & Holmes, P. 1983 *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.* Springer.

Haney, R. L. 1971 Surface thermal boundary conditions for ocean circulation models. *J. Phys. Oceanogr.* **4**, 241–248.

Huang, R. X. 1993 Real freshwater fluxes as a natural boundary condition for the salinity balance and thermohaline circulation forced by evaporation and precipitation. *J. Phys. Oceanogr.* **23**, 2428–2446.

Legras, B. & Ghil, M. 1988 Persistent anomalies, blocking, and variations in atmospheric predictability. *J. Atmos. Sci.* **42**, 433–471.

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

Marotzke, J. 1993 Ocean models in climate problems. In *Ocean Processes in Climate Dynamics: Global and Mediterranean Examples* (ed. P. Malanotte-Rizzoli & A.R. Robinson), pp. 79–109. Kluwer.

Marotzke, J. Welander, P. & Willebrand, P. 1988 Instability and multiple steady states in a meridional-plane model of the thermohaline circulation. *Tellus* **40**, 162–149.

Marotzke, J. & Willebrand P. 1988 Two stable equilibria of a coupled ocean-atmosphere model. *J. Climate* **1**, 841–866.

Mikolajewicz, U. & Maier-Reimer, E. 1990 Internal secular variability in an ocean general circulation model. *Climate Dyn.* **4**, 145–156.

Quon, C. & Ghil, M. 1992 Multiple equilibria in thermosolutal convection due to salt-flux boundary conditions. *J. Fluid Mech.* **245**, 449–484.

Quon, C. & Ghil, M. 1995 Multiple equilibria and stable oscillations in thermosolutal convection at small aspect ratio. *J. Fluid Mech.* **291**, 33–56.

Rahmstorf, S. Marotzke, J. & Willebrand J. 1995 Stability of the thermohaline circulation. In *The Warm Water Sphere of the North Atlantic Ocean* (ed. W. Krauss), Chap. 5. Borntraeger, Stuttgart.

Saravanan, R. & McWilliams, J. C. 1995 Multiple equilibria, natural variability, and climate transitions in an idealized ocean-atmosphere model. *J. Climate* **8**, 2296–2323.

Speich, S. Dijkstra, H. A. & Ghil, M. 1995 Successive bifurcations of a shallow model with applications to the wind driven circulation. *Nonlinear Proc. Geophys.* **2**, 241–268.

Stommel, H. 1961 Thermohaline convection with two stable regimes of flow. *Tellus*, **2**, 244–230.

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

Tziperman, E., Toggweiler, J. R. Feliks, Y. & Bryan, K. 1994 Instability of thermohaline circulation with respect to mixed boundary conditions: Is it really a problem for realistic models? *J. Phys. Oceanogr.* **24**, 217–232.

Vellinga, M. 1996 Instability of two-dimensional thermohaline circulation. *J. Phys. Oceanogr.* **26**, 309–319.

Weaver, A. J. & Hughes, T. M. 1992 Stability and variability of the thermohaline circulation and its link to climate. *Trends in Physical Oceanography*. Research Trends Series, Council of Scientific Research Integration, Trivandrium, India, 56 pp.

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

Weaver, A. & Sarachik, E. S. 1991 Evidence for decadal variability in an ocean general circulation model: an advective mechanism. *Atmos.-Ocean* **29**, 197–231.

Welander, P. 1986, Thermohaline effects in the ocean circulation and related simple models. In *Large Scale Transport Processes in Oceans and Atmosphere* (ed. J. Willebrand & D. L. T. Anderson), pp. 163–200. D. Reidel.

Winton, M. & Sarachik, E. S. 1993 Thermohaline oscillations induced by strong steady salinity forcing of ocean general circulation models. *J. Phys. Oceanogr.* **23**, 1389–1410.

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

Zhang, S., Greatbatch, R. J. & Lin, C. A. 1993 A re-examination of the polar halocline catastrophe and implications for coupled ocean – atmosphere modelling, *J. Phys. Oceanogr.* **23**, 287–299.