Instability and multiple steady states in a meridional-plane model of the thermohaline circulation

By J. MAROTZKE,[1] P. WELANDER[2] and J. WILLEBRAND,[1]
[1] *Institut für Meereskunde an der Universität Kiel, Düsternbrooker Weg 20, D-2300 Kiel, FRG,*
[2] *School of Oceanography, WB-10, University of Washington, Seattle, WA 98195, USA*

ABSTRACT

A meridional-plane model of the thermohaline circulation with a simple friction force and advection and vertical diffusion of the T–S field has been used to demonstrate the instability and existence of multiple steady states associated with "mixed" T–S boundary conditions (specified temperature, flux condition for salinity). With forcing and geometry symmetric to the equator, the symmetric solution was found to be unstable to infinitesimal perturbations, and an asymmetric pole-to-pole circulation was the end-result in all cases. The structure obtained for the meridional-plane stream function and for the poleward heat flux are in qualitative agreement with those obtained by Bryan (1986). Convective overturning caused by static instability was not found to be essential for the transition to the asymmetric steady state. The study suggests that certain aspects of the ocean circulation, in particular those related to the ocean climate, may be profitably explored by use of two-dimensional, zonally averaged models.

1. Introduction

From the observed mass field and the wind-stress forcing, the oceanic currents and associated vertical motions can be calculated according to diagnostic methods, in different approximations. To lowest order, the dynamics can be assumed linear, and in this case the velocity field is uniquely determined (excluding certain "pathological" cases).

The situation is, however, different when the mass field and the velocity field are calculated simultaneously by a predictive model, prescribing only surface boundary conditions (the wind-stress, plus conditions for the temperature and salinity). In this case, an essential non-linear process, the advection of temperature and salinity, comes into play. The strength of this process relative to the turbulent diffusion (the latter usually modeled in a linear approximation) is obvious from the observed alignment of the isotherms and isohalines along the ocean

currents, and from the shallowness of the thermocline in relation to the total depth. For this non-linear system uniqueness of the solution cannot generally be proved, and specific examples of multiple steady solutions, which satisfy identical boundary conditions, can be given. Such a behavior can have important consequences in the context of climate and climate variations on time scales where the oceans play an important rôle.

The interest in the problem of multiple steady states in the ocean circulation, and specifically in its thermohaline component, is relatively recent. It was stimulated by Rooth (1982), who investigated the stability of a simple three-box model of the oceans, comprising one equatorial box which is heated and from which freshwater is removed, and two polar boxes, which are cooled and receive the freshwater. When the heating and cooling is achieved through a Rayleigh-type transfer law (heat flux proportional to the difference between a prescribed temperature and

the box temperature), while the freshwater flux is fixed, Rooth suggested that a symmetric solution would become unstable, and develop into a single pole-to-pole cell under sustained symmetric forcing. This idea was tested by Bryan (1986) in a three-dimensional numerical circulation model, with wind and thermohaline forcing symmetric to the equator. The symmetric solution was indeed found to be unstable to a finite amplitude perturbation: specifically, the perturbed initial state had a 2‰ salinity excess or deficit in the high latitude surface waters of one hemisphere. The instability eventually produced a single pole-to-pole cell, as suggested by Rooth (1982).

A similar box model was examined by one of us (Welander, 1986), and it was found that the instability mechanism did not require a non-linearity in the equation of state, as assumed by Rooth, but that it related to a thermohaline phenomenon studied 25 years ago by Stommel (1961) in a two-box version. Under given symmetric forcing, the three-box model, representing two "back-to-back" Stommel models, could produce no less than 9 steady states, 4 of which are stable. These stable steady states give the symmetric and the two unsymmetric cells, plus a second symmetric cell, with reversed circulation. This last steady state is associated with a very slow circulation, allowing the flux condition for salinity to build up a strong equator-to-pole salinity contrast.

In the real oceans, a strongly asymmetric meridional cell is well documented in the Atlantic. The deep water is mainly formed in the North Atlantic and moves south across the equator into the circumpolar region (only a smaller amount of bottom water moves the other way), while warmer water in the thermocline crosses the equator northward, causing a substantial heat flux of order 10^{15} W into the North Atlantic. The Pacific Ocean represents another type of thermohaline circulation, with only small amounts of deep water produced at the northern end and a more symmetric, shallow circulation. The difference between the Atlantic and Pacific thermohaline circulations could be due to differences and asymmetries in atmospheric forcing and to geometry (e.g., Warren (1983) suggests that the low salinity in the North Pacific, which causes the absence of deep-water formation, is due to less evaporation due to lower surface

temperature, ultimately caused by a different wind field). However, the possibility of an internal mechanism which sustains one or the other of these circulation types has been explored theoretically by Walin (1985). He shows that a salinity perturbation, along the line of the numerical experiment by Bryan (1986), can be amplified by increased inflow of high-salinity equatorial water, and gives a corresponding stability criterion. The equations on which Walin bases the discussion actually turn out to be the same as those of Stommel's (1961) two-box model, although Walin has cast his problem in a somewhat different form.

Finally, the geochemists have found evidence for dramatic shifts in the thermohaline circulation in the past, based on paleo-climatic data. Among others, Broecker et al. (1985) have discussed this evidence, and proposed that variations seen on time-scales of less than a thousand years represent switches between different, distinct steady state regimes.

It seems obvious that the phenomenon of multiple steady states in the large-scale thermohaline circulation deserves further studies along a variety of lines. We have attempted to contribute to the studies by exploring a specific model which is more realistic than the "box-and-pipe" models by Stommel (1961), Rooth (1982), and Welander (1986), but much simpler than the large numerical model used by Bryan (1986). It predicts the temperature, salinity, and circulation in a meridional–vertical plane, due to advection and vertical diffusion, but the dynamics is highly idealized. It can be considered as a crude model for the zonally averaged circulation. In its non-dimensionalized form, the model has two free parameters, one expressing the strength of advection relative to diffusion and one representing the relative contributions from salinity and temperature to the buoyancy. We will investigate certain aspects of the instability mechanism, especially the question of whether or not the initial perturbation has to exceed a finite amplitude to trigger the transition to the asymmetric steady state. A related question is whether or not the overall pole-to-equator salinity contrast must exceed a certain threshold, as argued by Walin (1985). Another focus is the rôle of convective overturning in high latitudes which was found to be crucial in Bryan's (1986) experiments.

2. Model formulation

Let us consider the thermohaline circulation in an ocean of constant depth H and zonal width D. We will assume, at this point, that the motion is everywhere in geostrophic balance, except for some (small) vertical friction. The zonal average of the horizontal momentum balance is given by

$$-f\bar{v} = -\frac{p_E - p_W}{D} + A\bar{u}_{zz}, \tag{1a}$$

$$+f\bar{u} = -\bar{p}_y + A\bar{v}_{zz}. \tag{1b}$$

Here ($^-$) denotes a zonal average, and p_E (p_W) the pressure at the eastern (western) boundary. By eliminating $\bar{u}$ we obtain

$$f^2\bar{v} + A^2\bar{v}_{zzzz} = \frac{f}{D}(p_E - p_W) + A\bar{p}_{yzz}, \tag{2}$$

from which the mean meridional velocity $\bar{v}$ could in principle be determined in terms of the pressure field (and appropriate boundary conditions). However, the zonally averaged equations are not closed as they contain explicitly the east–west pressure difference. The main balance in (2) is between the respective first terms on either side of (2), corresponding to the geostrophic part of (1a), except within the equatorial boundary layer where the situation is more complex (see Philander and Pacanowski, 1986). Hence the east–west pressure difference plays a crucial rôle for the zonally averaged flow. In the absence of wind forcing, it must be related to the meridional density (or pressure) gradient; however, there is no compelling reason to assume that such a relation would be local and linear.

In order to close the problem, we will nevertheless replace (2) by a local and linear relation between mean meridional velocity and pressure gradient:

$$A_*\bar{v}_{zz} = \bar{p}_y. \tag{3}$$

Formally, (3) could be considered as the non-rotating limit of (1b); it does therefore correspond to a possible physical system. It will turn out that, except for the equatorial boundary layer, many qualitative aspects of both equations agree. In particular, both systems lead to surface flow towards regions of higher density, and return flow in deeper levels. However, (3) is not, in a quantitative sense, an approximation to (2).

It is clear that in order to consider an ocean circulation problem, A_* in (3) should not be identified with a physically realistic friction. A comparison with (2) would rather suggest the scaling

$$A_* \sim A\left(1 + \left(\frac{fh^2}{A}\right)^2\right), \tag{4}$$

where h is a vertical length scale. As $A/fh^2 \ll 1$, it follows that we must have $A_* \gg A$ in order to resemble the order-of-magnitude relation between transport and pressure gradient.

Introducing a stream function in the meridional plane according to $\bar{w} = \psi_y$, $\bar{v} = -\psi_z$, invoking the hydrostatic relation and assuming a linear equation of state we obtain

$$\frac{\partial^4\psi}{\partial z^4} = \frac{g}{A_*}\left(-\alpha\frac{\partial T}{\partial y} + \beta\frac{\partial S}{\partial y}\right), \tag{5}$$

with α and β being coefficients of heat and salt expansion, respectively.

The zonally integrated heat and salt budgets are given by (overbars omitted from now on)

$$\frac{\partial}{\partial t}\begin{Bmatrix}T\\S\end{Bmatrix} + v\frac{\partial}{\partial y}\begin{Bmatrix}T\\S\end{Bmatrix} + w\frac{\partial}{\partial z}\begin{Bmatrix}T\\S\end{Bmatrix} = \begin{Bmatrix}q_T\\q_S\end{Bmatrix}, \tag{6}$$

where the source terms q_T, q_S include the Reynolds terms arising from zonal averaging, and also the average of small-scale fluxes. Assuming that the "gyre transport" of heat and freshwater is small compared to the "overturning" component, we neglect the horizontal flux divergence and approximate the source terms by

$$\begin{Bmatrix}q_T\\q_S\end{Bmatrix} = k\frac{\partial^2}{\partial z^2}\begin{Bmatrix}T\\S\end{Bmatrix} + \begin{Bmatrix}q_T^{con}\\q_S^{con}\end{Bmatrix}, \tag{7}$$

where k is a (constant) vertical diffusivity. The 2nd term on the r.h.s. of eq. (7) indicates convective adjustment which results in complete vertical mixing whenever the stratification is locally unstable. Some of the numerical experiments described below are performed without this adjustment mechanism.

At the ocean surface, $z = 0$, we have assumed the boundary conditions:

$$\psi = 0,$$
$$\psi_{zz} = 0 \text{ (no wind stress)}, \tag{8a}$$

whereas at the bottom, $z = -H$,

$$\psi = 0,$$
$$\psi_z = 0 \text{ (no slip)}. \tag{8b}$$

At the northern and southern walls, no influx of mass is permitted so that

$$\psi = 0, \quad y = \pm L. \tag{8c}$$

However, the latter condition conflicts with (5) which generally results in a nonzero stream function at the boundaries. This conflict can be resolved by incorporating a small horizontal friction in (3) and (5), respectively, which would leave (5) approximately valid in the interior, with boundary layers at both walls whose width is controlled by friction. For the numerical experiments described below, we have assumed that the boundary layer width coincides with the grid resolution so that no explicit incorporation of lateral friction is necessary.

The thermohaline boundary conditions contain the essential mechanism that is investigated here. At the surface, the ocean is driven by fluxes of heat and freshwater. While in a laboratory experiment, those fluxes in principle could be prescribed independent of the state of the system, this is not true for the ocean–atmosphere system. Both fluxes depend strongly on the sea surface temperature but very little on surface salinity. The typical time scale for (mid-latitude) surface temperature anomalies is a few months (e.g., Davis, 1976). For much longer time scales, the ocean surface temperature and the evaporation are effectively prescribed by the atmosphere. We

therefore apply the boundary conditions:

$$T = T_0(y), \quad S = S_0(y) \tag{9a}$$

for the spin-up experiment, and

$$T = T_0(y), \quad kS_z = Q_S(y), \tag{9b}$$

for all subsequent experiments. Here $T_0(y)$, $S_0(y)$ are prescribed surface temperature and salinity, respectively. We have chosen the form

$$\{T_0, S_0\} = \{T_*, S_*\}(1 + \cos \pi y/L),$$

where $2T_*(2S_*)$ is the pole-to-equator temperature (salinity) contrast. The choice for the freshwater flux (or equivalent salinity flux) $Q_S(y)$ is discussed in Section 3. No fluxes are allowed through the bottom, i.e., $T_z = S_z = 0$ at $z = -H$.

A picture of the meridional plane model, showing the grid and (differential) boundary conditions, is given in Fig. 1. We adopt the following scaling:

$$y \to y/L, \quad z \to z/H,$$
$$T \to T/T_*, \quad S \to S/S_*,$$
$$\psi \to \psi/\psi_* \quad \text{with } \psi_* = g\alpha T_* H^4/A_* L,$$
$$t \to LH/\psi_* \quad \text{(advective time scale)}.$$

Eqs. (5) and (6) read in dimensionless form

$$\frac{\partial^4 \psi}{\partial z^4} = -\frac{\partial T}{\partial y} + \delta \frac{\partial S}{\partial y} = r(y, z, t), \tag{10}$$

$$\frac{\partial \begin{Bmatrix} T \\ S \end{Bmatrix}}{\partial t} + J\left(\psi, \begin{Bmatrix} T \\ S \end{Bmatrix}\right) = \varepsilon \frac{\partial^2}{\partial z^2} \begin{Bmatrix} T \\ S \end{Bmatrix}, \tag{11}$$

with the dimensionless parameters

$$\delta = \frac{\beta S_*}{\alpha T_*} \quad \text{and} \quad \varepsilon = \frac{kL}{\psi_* H}. \tag{12}$$

Here δ describes the relative importance of the salinity field on density and ε is a dimensionless measure of diffusion relative to advection. Due to the non-exponential character of (10), in conjunction with the homogeneous boundary conditions, typical values of ψ are $O(10^{-3})$ rather than $O(1)$. For that reason, ε is not identical to an inverse Peclet-number.

The prognostic equations (11) for temperature and salinity were solved by finite difference methods, using forward time differences and upstream spatial differences on a rather coarse grid (16 levels vertical, 32 horizontal grid points). Parameter values are $H = 4$ km, $L = 6000$ km,
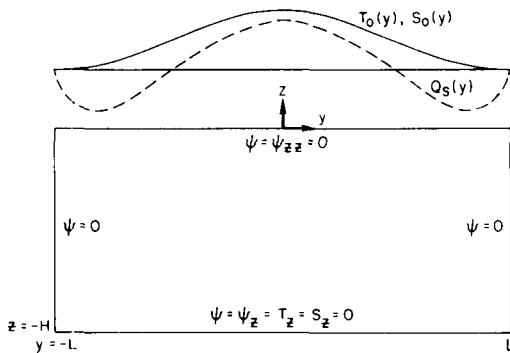

Fig. 1. The model ocean in a meridional plane, indicating boundary conditions for the stream function ψ, and bottom conditions for temperature T and salinity S. At the surface, $T_0(y)$, $S_0(y)$ (full drawn curve) are prescribed; after a symmetric steady state is obtained, the condition for surface salinity is changed, with the resulting freshwater flux $Q_S(y)$ (dashed curve) prescribed.

$2T_* = 25°C$, $2S_* = 2‰$, $k = 10^{-4}$ m^2 s^{-1}, and $A_* = 6 × 10^2$ m^2 s^{-1} (the latter choice gives, in dimensional units, $\psi_{max} = 1.6$ m^2 s^{-1} for the standard experiment which would correspond to 10 Sv of meridional transport in a basin of 6000 km zonal width). The standard values for the non-dimensional parameters are thus $\delta = 0.32$ and $\varepsilon = 10^{-4}$. For the purpose of illustration, we will display some of our results below in physical units, based on the above choice of parameters.

3. Experiments and results

The aim of our numerical experiments is to explore the possible existence of several steady states in our symmetrically forced model and find the minimum amplitude of perturbations giving the transition from an initial symmetric steady solution to a pole-to-pole cell solution. We also want to explore the physical mechanism responsible for the instability, in particular, the rôle of the T–S advection.

The general procedure in our (much simpler) experiments is very similar to the one followed by Bryan (1986). Firstly, we integrate the system with temperature and salinity prescribed in a symmetric way at the top. When the steady state is reached, we diagnostically calculate the surface salinity flux $Q_S(y)$ necessary to maintain that equilibrium. The meridional structure of $Q_S(y)$ is illustrated in Fig. 1.

This salinity flux, together with the surface temperature, serves as boundary condition thereafter. Thus a different mathematical system (i.e., the one with T and Q prescribed at the surface) by construction has a steady state in common with the original system (with T, S prescribed at the surface). We can now investigate how that steady state, which is stable in the original case, reacts to small perturbations.

Fig. 2 shows stream function and salinity for the corresponding symmetric steady state in our standard experiment. The salient feature is the existence of two symmetrical cells, with strong downwelling in the polar regions and weaker upwelling everywhere else. The pronounced boundary layers at the side boundaries are confined to one grid cell, as is to be expected due to the lack of explicit horizontal friction. Temperature and density distributions are identical to that
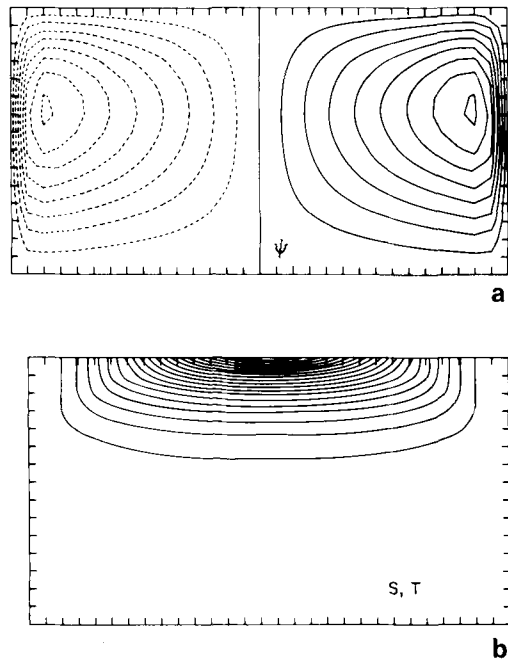
Fig. 2. Fields of stream function (a) and salinity (b) in the symmetric steady state, with standard parameter values $\varepsilon = 10^{-4}$, $\delta = 0.32$. Maximum value of ψ is 1.6 m^2 s^{-1}, which corresponds to a total transport of 10 Sverdrup in a basin of 6000 km width. Contour intervals are 0.2 m^2 s^{-1} (a) and 0.1‰ (b).

of salinity in the symmetric case. The thickness of the thermocline is controlled by vertical diffusion as there is no Ekman pumping in our model. The choice $\varepsilon = 10^{-4}$ provides for a strongly stratified top layer with a depth resembling the mean depth of the oceanic main thermocline. It can be shown that the e-folding depth scale, based on a balance between vertical advection and diffusion, is proportional to $\varepsilon^{1/2}$. Much larger values of ε result in a strongly diffusive solution, whereas for much smaller values the stratification is confined to the uppermost grid level, with completely homogeneous water below. We have chosen the value $\varepsilon = 10^{-4}$ for all subsequent experiments.

At the time when we shifted over to mixed boundary conditions, an initial positive salt anomaly of 0.5‰ was added in the top layer in the northern half of the northern hemisphere. 500 years after the anomaly was added, the northern cell has increased in strength and extent, whereas
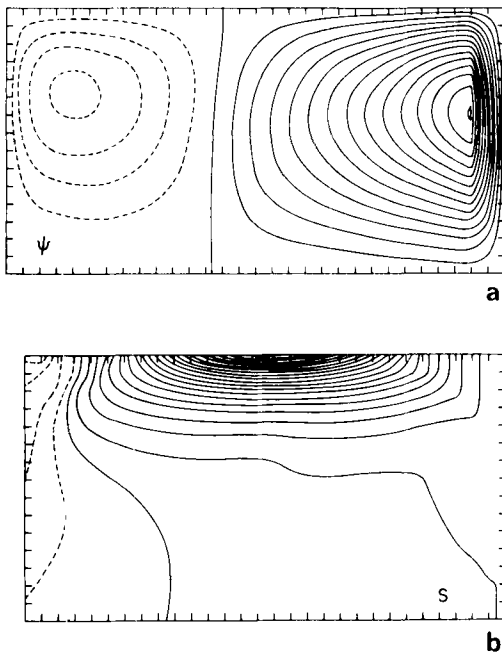
Fig. 3. Stream function (a) and salinity (b) 500 years after a positive salt anomaly of 0.5‰ was added in the top level in the northern (right) half of the northern hemisphere to the symmetric solution shown in Fig. 2. Contour intervals are 0.15 m² s⁻¹ (a) and 0.1‰ (b).

the southern cell is much reduced (Fig. 3a). The stronger downwelling in the north leads to an increase in the deep salinity (Fig. 3b). The southern cell has decreased, leading to fresher values at the surface which subsequently are advected downward. Although the anomaly was added to the northern hemisphere, at this time the decrease of the circulation in the southern hemisphere is already twice as large as the corresponding increase in the northern hemisphere, indicating that the direct effect of the initial anomaly is no longer dominating. It might appear that the stronger decrease in the south is due to a reduction in convective mixing caused by lower surface salinities, while the higher surface salinities in the north do not cause a corresponding increase in convective mixing as the stratification is already nearly homogeneous. However, this is probably not the case, as is discussed below.

In the final steady state, the northern cell is completely dominating, causing a substantial

equatorward heat flux in the southern hemisphere, and the southern cell has almost vanished (Fig. 4a). The salinity is nearly homogenized in the northern hemisphere and has a strong gradient towards much lower surface values in the southern hemisphere, whereas the temperature distribution (Fig. 4c) is quite similar to the symmetric steady state (cf. Fig. 2b).

It is interesting to note that the boundary layer in the circulation field (Fig. 4a) is detached from the physical boundary, leaving behind a (very
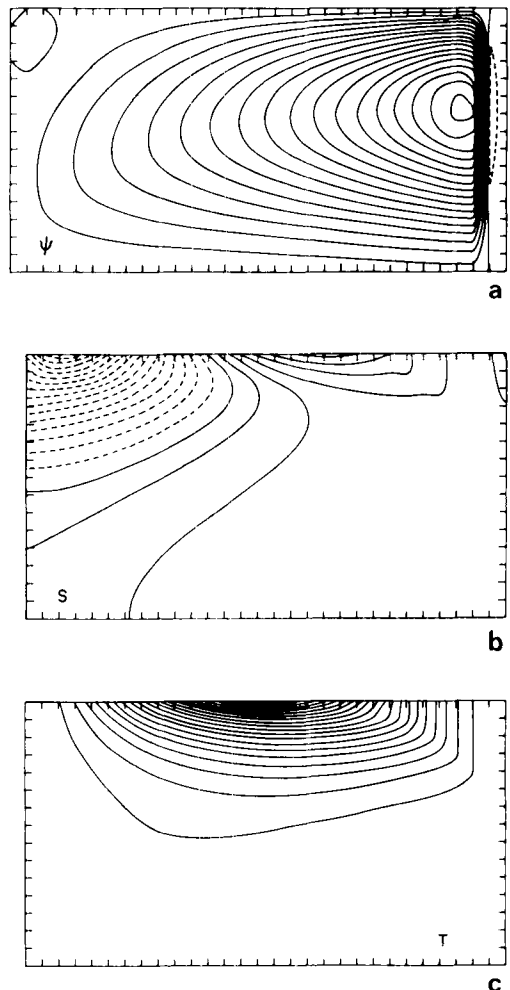

Fig. 4. Stream function (a), salinity (b), and temperature (c) after a new steady state has been achieved in the perturbation experiment. The maximum value of ψ is increased to 2.6 m² s⁻¹. Contour intervals are 0.15 m² s⁻¹ (a), 0.2‰ (b), and 1.25°K (c).

weak) reverse circulation that is confined to the northernmost gridpoints. The reason for this behavior is that the surface density has its maximum two gridpoints away from the northern boundary, due to a salinity maximum. In a steady state, that density maximum must prevail at all depths due to convective adjustment so that $\partial\rho/\partial y = 0$ which implies $\psi = 0$, i.e., such a point effectively acts as a boundary. Close to that point, the isopycnal surfaces outcrop, resulting in rather large values of $\partial\rho/\partial y$. The width of the corresponding boundary layer can be influenced by lateral diffusion of heat and salt as well as by lateral friction. We have neglected both mechanisms, and hence the boundary layer width coincides with the horizontal grid resolution.

In order to investigate the rôle of the convective overturning for the transition, we have repeated the experiment with the same parameter set but without the convective adjustment mechanism in eq. (7). The symmetric equilibrium circulation (Fig. 5a) is not very different
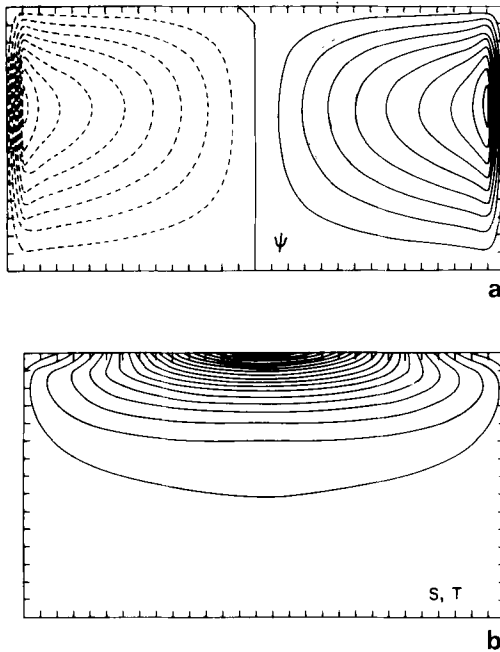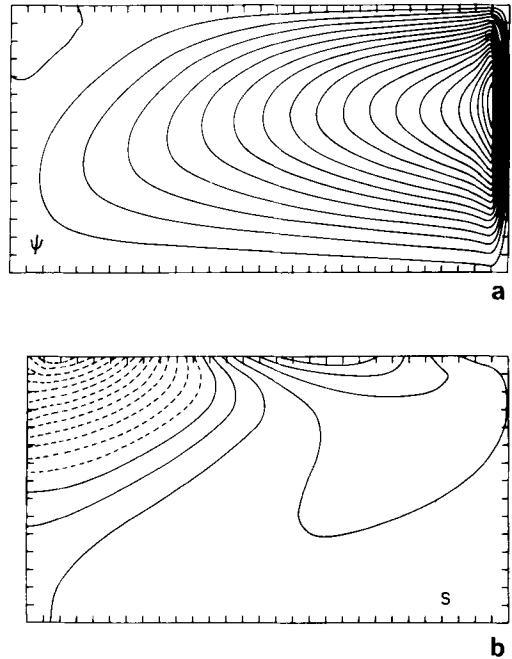

Fig. 6. Stream function (a) and salinity (b) in the asymmetric steady state obtained by perturbing the symmetric state shown in Fig. 5, again without convective adjustment. Contour intervals are 0.15 m^2 s^{-1} (a) and 0.2‰ (b).

Fig. 5. Stream function (a) and salinity (b) in the symmetric steady state with previous boundary conditions, but without convective adjustment. The changes in ψ are small, changes in T and S are concentrated to the high latitudes. Contour intervals are 0.2 m^2 s^{-1} (a) and 0.1‰ (b).
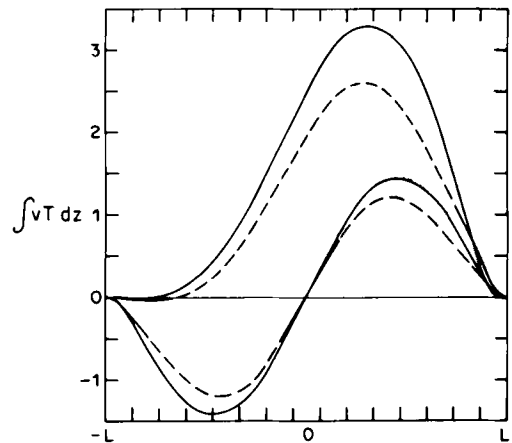

Fig. 7. Meridional heat transport as function of y in the symmetric steady state (lower curves), and asymmetric steady state (upper curves). The full drawn curves show the result with and the dashed curves the result without convective adjustment. Units are dimensionless and would correspond to 0.15×10^{15} W for the standard parameters and a basin of 6000 km zonal width.

from the convective case (Fig. 2a). As is to be expected, the density field exhibits statically unstable regions near the surface in high latitudes, and the deep temperatures and salinities are somewhat higher. The transition to a new steady state after addition of the anomaly occurred in a similar way (though somewhat slower) and the final asymmetric steady state (Fig. 6a, b) is very much like the convective case, with some minor differences (e.g., the boundary layer in Fig. 6a is not detached from the boundary as in the corresponding case with convection, and regions of static instability are now confined to the northern hemisphere).

The total poleward heat flux was calculated for both steady states (note that the poleward salinity flux is fixed directly by the freshwater boundary conditions). It is seen from Fig. 7 that while the symmetric solution has an obvious anti-symmetry in heat flux (from equator to pole in both hemispheres), the asymmetric solution gives a strongly enhanced poleward heat flux in the same direction everywhere, toward the "downwelling pole". The convective adjustment increases the heat transport by about 15% but does not change the curve shapes noticeably. Comparing these heat flux curves with the ones obtained by Bryan (1986; see Fig. 6), one finds a good qualitative resemblance of the curves. The two-dimensional model seems able to represent some integral features of the three-dimensional circulation surprisingly well. Although the absence of convective adjustment leads to some regions with unstable stratification and hence to a physically unrealistic state, the experiment demonstrates that convective overturning is *not* an essential ingredient for the instability mechanism.

Therefore, the only physical interpretation of the phenomenon is based on the advective positive feedback mechanism as outlined by Walin (1985) and Welander (1986). The circulation is driven by the density difference $\Delta\rho$ between equatorial and polar water, the latter having approximately the same density as deep water. As temperature and salinity have opposite influences on the density, the temperature difference between pole and equator drives the circulation, while the salinity difference acts as a brake. Increasing salinity in the polar regions decreases the braking action of the salinity difference so that the circulation is slightly

strengthened in one hemisphere and transports more warm and salty water from the equator poleward. If temperature and salinity interacted equally strongly with the atmosphere, this additionally transported water would arrive at the pole with at most the same density as its surroundings. However, as the temperature signal is removed rapidly due to the strong interactions with the atmosphere, while the salinity signal is only slowly modified by the prescribed freshwater flux on its way to the pole, there is a positive density signal arriving at the pole due to the strengthened circulation. This enhanced density further strengthens the circulation, thus providing for a positive feedback.

A similar advective feedback mechanism is likely to be responsible for the stronger decrease of the circulation in the south, as mentioned above. The same feature appears at approximately the same stage of development in the case without convective adjustment. This is readily explained by observing that a reduction in circulation leads to longer exposure of water parcels to the freshwater flux through the surface at high latitudes. Freshening increases the braking action of the salinity contrast between pole and equator, and the circulation is weakened even more. That this mechanism is faster than the complementary one in the northern hemisphere may be due to the reduced size of the southern circulation cell. The same amount of freshwater flux is able to bring about a greater change in salinity.

Whereas convective adjustment does not play an important part in the mechanisms considered up to now, the picture is somewhat different in the case of a finite negative anomaly ($-0.5\%_0$, figure not shown here). Now the reaction is faster and more drastic than with a finite positive anomaly, which is in qualitative agreement with the results of Bryan (1986).

After we have demonstrated that the symmetric steady state is unstable to salinity perturbations of order $\Delta S = 0.5\%_0$, we have repeated the experiments (both with and without convection) for various smaller values of ΔS. Fig. 8 shows for the non-convective case the time evolution of the stream function at one point in the interior after the anomaly has been added. While for $\Delta S = 0.5\%_0$ the response is almost immediate, and the transition is completed after 2000 years, there seems to be no initial response to an
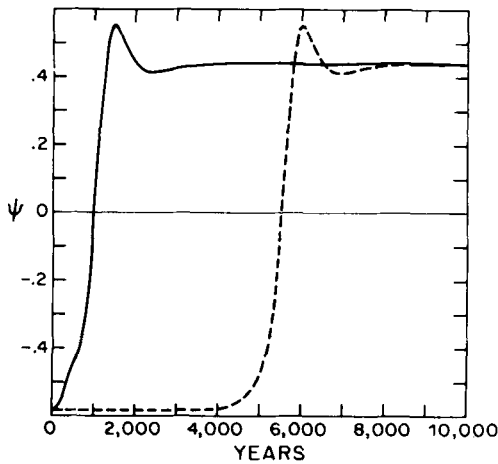
Fig. 8. The stream function at one interior point as function of time from the moment a perturbation is applied. Full line: 0.5‰ anomaly, dashed line: 10^{-6}‰ anomaly. Units correspond to m^2 s^{-1}.

Fig. 9. Basin average surface heat flux as function of time with convective adjustment (full drawn curve) and without convective adjustment (dashed curve), for the perturbation experiment with the 10^{-6} salinity anomaly (see caption to Fig. 8, case (b)). Dimensionless units correspond to 0.22 W m^{-2} for standard parameters.

anomaly $\Delta S = 10^{-6}$‰. However, after approximately 5000 years the system flips over into the asymmetric equilibrium, the details of the transitions being almost identical! For every oceanographic purpose, 10^{-6}‰ is of course an infinitesimal amplitude, and we conclude that the symmetric steady state is unstable to infinitesimal perturbations although its initial growth rate is rather small.

The time-development of the basin-averaged surface heat flux (Fig. 9) illuminates the different reaction of the convective and non-convective system to infinitesimal perturbations. In the convective case the transition occurs somewhat earlier and is more violent. It includes various oscillatory patterns which most likely are related to the heat–salt oscillator mechanism as described by Welander (1982). The sign of the initial anomaly determines which of the two possible asymmetric states is approached. Negative anomalies cause a reversed transition (i.e., mirrored at the equator), resulting in an anticlockwise cell with downwelling at the south pole. For infinitesimal anomalies, this is the only change, and in fact the curves in Fig. 9 are identical for both negative and positive anomalies.

Finally, we have investigated the stability of the symmetric circulation with respect to the parameter δ, the relative contribution of salinity
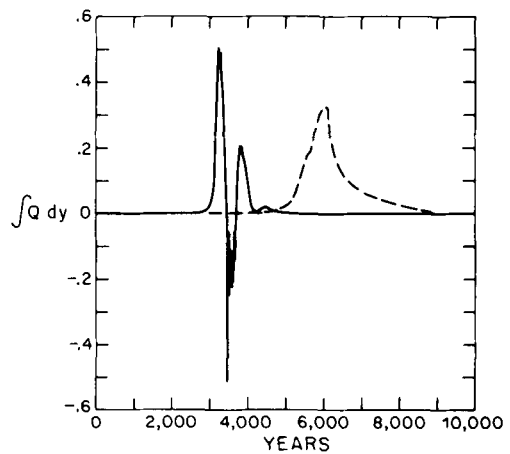
to the density field. Walin (1985) has argued that below a certain threshold which he estimated to be of order 0.5 the instability mechanism would close, and the box-model also requires a finite δ for the transition (Welander, 1986). In our experiments we could not confirm the existence of such a critical value for δ. Fig. 10 shows the final meridional heat flux of various anomaly experiments with different δ-values. The anomaly added was $\Delta S = 0.5$‰. We see that for δ-values
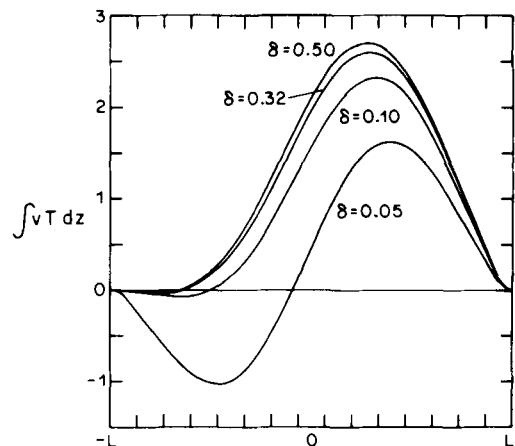

Fig. 10. Meridional heat transport in the asymmetric steady state for various values of δ. Units as in Fig. 7.

above 0.1, there is little difference in the final state. The time necessary to undergo the transition increases with decreasing δ. For even smaller δ, the remaining anticlockwise circulation is stronger and more extended in space, but there still is a pronounced asymmetry in the circulation pattern at $\delta = 0.05$. The transition is much slower and smoother, and for $\delta = 0.01$ the time scales become uninterestingly large. The fact that the asymmetric steady state changes continuously as $\delta \rightarrow 0$ (in the limit $\delta = 0$ it represents a symmetric steady state) was unexpected, and the reason for it is not yet understood.

4. Concluding remarks

The occurrence of multiple equilibrium states apparently is a very robust property of the thermohaline models with "mixed" top boundary conditions (temperature and surface water flux prescribed): it is found in models of greatly varying complexity, from simple box models over the intermediate model considered in this paper up to full-blown general circulation models. All three types of models predict that a circulation pattern which is symmetric with respect to the equator is inherently unstable. The instability mechanism is an advective one, and convective adjustment at high latitudes is an important but not crucial ingredient.

Whether or not a finite amplitude anomaly is required to trigger the instability is less clear. Here our findings are in contrast to the results of box models which do require a finite amplitude. The 3-dimensional model of Bryan (1986) remained stable for 1300 years of integration with a very small anomaly (provided by computational round-off noise). While this result is not in disagreement with our findings (see Fig. 7), it is possible that lateral mixing provided by horizontal gyre motions could stabilize the symmetric circulation. However, as observed surface salinity anomalies have amplitudes up to 1‰, persist over several years, and extend over scales of more than 1000 km in the North Atlantic (J. Meincke, personal communication), the probability that the symmetric steady state would occur seems rather small.

More relevant for the problem of climate change is the question of under what conditions the asymmetric state could flip over into the reversed asymmetric state. It has been speculated that those changes occurred simultaneously but with opposite sign in the Atlantic and Pacific oceans, respectively (Broecker et al., 1985). Hence it is likely that one has to consider the interactions between both ocean basins which are coupled by the Antarctic Circumpolar Current and, perhaps more importantly, via the hydrological cycle in the atmosphere. We believe that the intermediate model which we have discussed here will be a useful tool to address those questions in an exploratory way.

5. Acknowledgements

REFERENCES

Broecker, W. S., Peter, D. M. and Rind, D. 1985. Does the ocean–atmosphere system have more than one stable mode of operation? *Nature 315*, 21–26.

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

Davis, R. E. 1976. Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr. 6*, 249–266.

Philander, S. G. H. and Pacanowski, R. C. 1986. The mass and heat budget in a model of the tropical Atlantic Ocean. *J. Geophys. Res. 91*, 14,212–14,220.

Rooth, C. 1982. Hydrology and ocean circulation. *Progr. Oceanogr. 11*, 131–149.

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

Walin, G. 1985. The thermohaline circulation and the control of ice ages. *Palaeogeogr., Palaeoclimatol., Palaeoecol. 50*, 323–332.

Warren, B. A. 1983. Why is no deep water formed in the North Pacific? *J. Mar. Res. 41*, 327–347.

Welander, P. 1982. A simple heat–salt oscillator. *Dyn. Atmos. Oceans 6*, 233–242.

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