

Instability of the thermohaline ocean circulation on interdecadal time scales

Lianke A. te Raa and Henk A. Dijkstra
Institute for Marine and Atmospheric Research Utrecht
Department of Physics and Astronomy
Utrecht University
Utrecht, the Netherlands

Revised for Journal of Physical Oceanography

Version of April 4, 2001

Corresponding Author:

Henk A. Dijkstra
Institute for Marine and Atmospheric research Utrecht
Department of Physics and Astronomy
Utrecht University
Princetonplein 5, 3584 CC Utrecht
The Netherlands
Phone: -31-30-2533858; Fax: -31-30-2543163
Email: dijkstra@phys.uu.nl

Abstract

The stability of three-dimensional thermally driven ocean flows in a single hemispheric sector basin is investigated using techniques of numerical bifurcation theory. Under restoring conditions for the temperature, the flow is stable. However, when forced with the associated heat flux, an interdecadal oscillatory time scale instability appears. This occurs as a Hopf bifurcation when the horizontal mixing coefficient of heat is decreased. The physical mechanism of the oscillation is described by analyzing the potential energy changes of the perturbation flow near the Hopf bifurcation. In the relatively slow phase of the oscillation, a temperature anomaly propagates westwards near the northern boundary on a background temperature gradient, thereby changing the perturbation zonal temperature gradient, with corresponding changes in meridional overturning. This is followed by a relatively fast phase in which the zonal overturning reacts to a change in sign of the perturbation meridional temperature gradient. The different responses of zonal and meridional overturning cause a phase difference between the effect of temperature and vertical velocity anomalies on the buoyancy work anomaly, the latter dominating the changes in potential energy. This phase difference eventually controls the time scale of the oscillation.

1 Introduction

Climate variability on decadal and interdecadal time scales has recently received a lot of attention. Indications for decadal to interdecadal variability have been found in many climate and climate proxy data, for example global surface air temperatures (Ghil and Vautard, 1991; Schlesinger and Ramankutty, 1994), variations in the formation rate of North Atlantic Deep Water (Roemmich and Wunsch, 1984) and ice core data (Hibler and Johnson, 1979). Knowledge of the patterns and amplitude of natural climate variability on these time scales is of crucial importance in the interpretation of climate change. The short length of the time series of sea surface temperature (SST) and sea level pressure (SLP) (150 years at the most) causes strong uncertainties in reconstructions of patterns of interdecadal variability in these data.

A nice overview of the data that have been analyzed since 1990 and the methods and terminology used can be found in Moron *et al.* (1998). In Deser and Blackmon (1993), the first empirical orthogonal function (EOF) of wintertime mean SST anomalies in the North Atlantic over the period 1900-1989 displays a basin scale SST pattern with strongest positive anomalies in the Gulf Stream region. The time series of this EOF indicates that this region was colder than average over the period 1900-1940 and warmer over the remaining period. The second EOF is a dipole-like pattern with positive (negative) anomalies in the northern (southern) part of the basin with variability in the time series on decadal scales. Using more than 100 years of SST, SLP and wind data from the COADS dataset, Kushnir (1994) showed that SST, SLP and surface winds exhibit interdecadal variability, with a basin scale SST pattern having maxima in the Labrador Sea and northeast of Bermuda. Moron *et al.* (1998) applied multi-channel singular-spectrum analysis (MSSA) to a monthly time series (1901-1994) of SST anomaly fields to find a 13-year oscillation. Two maxima of opposite signs occur near Cape Hatteras and south of the Denmark Straits, with a peak-to-peak amplitude of about 0.5°C . Recently, Delworth and Greatbatch (2000) have identified surface temperature variability with a dominant time scale of 70 years with a pattern mostly concentrated in the North Atlantic.

The uncertainty in the patterns of observed variability and the lack of understanding of the physics of this type of variability have stimulated many model studies. In the coupled GFDL model, interdecadal variability was found and analyzed by Delworth *et al.* (1993); Delworth and Mann (2000). The irregular oscillation has an average period of about 50 years and its pattern roughly corresponds to that found in observations by Kushnir (1994). It was proposed to be associated with density anomalies in the sinking region (with much smaller anomalies of the opposite sign in the broad rising region) and their coupling to anomalies in the circulation.

While it has been suggested that the Delworth *et al.* (1993) variability is a coupled ocean-atmosphere phenomenon (Weaver and Valcke, 1998), the study by Delworth and Mann (2000) suggests that the oscillation is mainly caused by the low frequency component of the atmospheric noise, combined with thermohaline feedbacks. However, decadal to interdecadal variability has also been found in many ocean-only model studies (Weaver *et al.*, 1993; Greatbatch and Zhang, 1995; Weaver and Sarachik, 1991a,b; Chen and Ghil, 1995). Greatbatch and Zhang (1995) find

a regular oscillation with a period of 50 years in a single hemispheric sector ocean model in which the flow is driven only by a time-independent heat flux. The SST anomaly pattern shows very good similarities to that in Delworth *et al.* (1993). Temperature anomalies are advected into the sinking region and change the strength of the overturning circulation with a certain phase lag between temperature and circulation anomalies. Also in other ocean-only model studies driven by a steady buoyancy flux, decadal to interdecadal variability has been found (Winton, 1996; Cai *et al.*, 1998; Huck *et al.*, 1999).

In models in which also the effect of salinity was taken into account under mixed boundary conditions (Weaver *et al.*, 1993; Chen and Ghil, 1995; Yin and Sarachik, 1995), also oscillations with periods between 10 and 50 years have been found. In the proposed mechanism of these oscillations, a phase difference between density anomalies and changes in convection and overturning strength is involved (Yin, 1995). But despite all modelling efforts, the physical processes that cause this phase difference are still unclear. Winton (1996) suggests that viscous boundary-trapped waves are a key to thermohaline oscillations. Huck *et al.* (1999) argue that, instead of boundary-trapped waves, internal potential vorticity waves are important and that the phase difference is caused by both advection and adjustment.

Chen and Ghil (1995) show that the interdecadal oscillation likely arises as a critical parameter value is crossed, i.e. through a Hopf bifurcation. Following this idea, an impressive amount of work was done in a planetary geostrophic model by Huck *et al.* (1999) and Colin de Verdière and Huck (1999), who used the same set-up as Greatbatch and Zhang (1995). Indeed, they find that oscillations occur if the horizontal diffusivity is decreased below a certain value. Guided by the vertical structure of the anomaly fields, baroclinic instability is suggested as a cause of the oscillatory behavior. Different conceptual models have also been put forward to explain the oscillatory behavior and the processes controlling the period (Yin, 1995; Colin de Verdière and Huck, 1999). However, in these simple models the phase difference between the velocity and temperature fields is more or less imposed such that oscillations on interdecadal time scales occur.

Recently, it has become possible to compute directly three-dimensional steady flows of the thermohaline circulation in single hemispheric models and to assess their linear stability. In this paper, we consider flows which are forced by a meridional temperature gradient in a single hemispheric idealized basin (Greatbatch and Zhang, 1995); the model is described in section 2. Within this model, we show that steady states become unstable through a supercritical Hopf bifurcation as a critical value of the horizontal diffusivity is crossed (section 3). The physical mechanism of this oscillation is reconsidered by looking at the relevant terms in the perturbation potential energy balance (section 4). In the discussion (section 5), this physical mechanism is set in context with other mechanisms which have appeared in the literature.

2 Formulation of the model

2.1 Model equations

The model that is used here is a fully implicit model of the three-dimensional ocean circulation as described by Dijkstra *et al.* (2000). The governing equations of this model are the Boussinesq equations in spherical coordinates ϕ , θ and z , with application of the hydrostatic approximation. In the version used here, salinity is not considered (as in Greatbatch and Zhang (1995)) and the flow is only forced by a downward heat flux. The flow domain is a sector $[\phi_W, \phi_E] \times [\theta_S, \theta_N]$ of constant depth D on a sphere with radius r_0 , which rotates with angular velocity Ω .

The ocean velocities in eastward and northward directions are indicated by u_* and v_* , the vertical velocity is indicated by w_* , the pressure by p_* and the temperature by T_* . The density is assumed to depend linearly on the temperature, with expansion coefficient α_T , reference temperature T_0 and reference density ρ_0 . The non-dimensional temperature T and pressure p are defined through $T_* = T_0 + \Delta T T$ and $p_* = -\rho_0 g z + 2\Omega r_0 U \rho_0 p$ and non-dimensional velocities are written as u , v , and w . A characteristic horizontal velocity is indicated by U , and the governing equations are further non-dimensionalized using scales r_0 , D , U , DU/r_0 and r_0/U for horizontal length, vertical length, horizontal velocity, vertical velocity and time, respectively and become

$$\begin{aligned} \varepsilon_R \left(\frac{\partial u}{\partial t} + \mathbf{u} \cdot \nabla u - uv \tan \theta \right) - v \sin \theta &= -\frac{1}{\cos \theta} \frac{\partial p}{\partial \phi} + E_V \frac{\partial^2 u}{\partial z^2} \\ &+ E_H \left(\nabla_H^2 u - \frac{u}{\cos^2 \theta} - \frac{2 \sin \theta}{\cos^2 \theta} \frac{\partial v}{\partial \phi} \right) \end{aligned} \quad (1a)$$

$$\begin{aligned} \varepsilon_R \left(\frac{\partial v}{\partial t} + \mathbf{u} \cdot \nabla v + u^2 \tan \theta \right) + u \sin \theta &= -\frac{\partial p}{\partial \theta} + E_V \frac{\partial^2 v}{\partial z^2} \\ &+ E_H \left(\nabla_H^2 v - \frac{v}{\cos^2 \theta} + \frac{2 \sin \theta}{\cos^2 \theta} \frac{\partial u}{\partial \phi} \right) \end{aligned} \quad (1b)$$

$$\frac{\partial p}{\partial z} = Ra T \quad (1c)$$

$$0 = \frac{\partial w}{\partial z} + \frac{1}{\cos \theta} \left(\frac{\partial u}{\partial \phi} + \frac{\partial(v \cos \theta)}{\partial \theta} \right) \quad (1d)$$

$$\begin{aligned} \frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T &= \nabla_H \cdot (P_H \nabla_H T) + \frac{\partial}{\partial z} \left(P_V \frac{\partial T}{\partial z} \right) \\ &+ Q_T G(z) \end{aligned} \quad (1e)$$

with

$$\begin{aligned} \mathbf{u} \cdot \nabla &= \frac{u}{\cos \theta} \frac{\partial}{\partial \phi} + v \frac{\partial}{\partial \theta} + w \frac{\partial}{\partial z} \\ \nabla_H \cdot (P_H \nabla_H) &= \frac{1}{\cos \theta} \left[\frac{\partial}{\partial \phi} \left(\frac{P_H}{\cos \theta} \frac{\partial}{\partial \phi} \right) + \frac{\partial}{\partial \theta} \left(P_H \cos \theta \frac{\partial}{\partial \theta} \right) \right] \\ \nabla_H^2 &= \nabla_H \cdot \nabla_H \end{aligned}$$

In the right hand side of the temperature equation (1e) a (dimensionless) source term Q_T has been included, which needs further clarification. Under restoring boundary conditions, the dimensional heat flux Q_{T^*} at the surface is proportional to the temperature difference between the ocean surface temperature and a prescribed atmospheric temperature T_{S^*} , i.e.

$$Q_{T^*} = B_T(\eta_T T_{S^*} - T_*) \quad (2)$$

with B_T an ocean-atmosphere exchange coefficient (Haney, 1971) and η_T a dimensionless parameter, introduced to control the amplitude of T_{S^*} . The transfer of heat from the surface downwards occurs in a thin boundary layer, similar to the Ekman layer for momentum transfer and cannot be resolved explicitly. Hence, as in low resolution ocean general circulation models, the surface forcing is distributed as a body forcing over a certain depth of the upper ocean, with thickness H_m , using a vertical profile function $G(z)$. In this way, the coefficient B_T is related to the restoring time scale τ_T through $B_T = C_p \rho_0 H_m / \tau_T$, with C_p the specific heat capacity. Using the scaling as above, the thermal forcing becomes

$$Q_T = B(\eta_T T_S - T) \quad (3)$$

where $B = r_0 / (U \tau_T)$.

Using this source term, the boundary condition for temperature at the ocean-atmosphere boundary is changed into a no-flux condition. This guarantees that the surface integral of the heat flux (3) is zero for each steady solution (Weaver and Hughes, 1996). On the lateral walls, no-slip conditions are prescribed and the normal heat flux is zero. The bottom of the ocean $z = -1$ is assumed to be isolated and satisfies slip conditions. The non-dimensional boundary conditions are hence formulated as

$$z = 0, -1 \quad : \quad w = 0, \quad \frac{\partial u}{\partial z} = \frac{\partial v}{\partial z} = \frac{\partial T}{\partial z} = 0 \quad (4a)$$

$$\phi = \phi_W, \phi_E \quad : \quad u = v = w = \frac{\partial T}{\partial \phi} = 0 \quad (4b)$$

$$\theta = \theta_S, \theta_N \quad : \quad u = v = w = \frac{\partial T}{\partial \theta} = 0 \quad (4c)$$

The parameters in equations (1) and (3) are the Rossby number ε_R , the Rayleigh number Ra , the vertical and horizontal Ekman numbers E_V and E_H , the vertical and horizontal inverse Péclet numbers P_V and P_H and the Biot number B . Expressions for these parameters are

$$Ra = \frac{\alpha_T \Delta T g D}{2\Omega U r_0} ; E_V = \frac{A_V}{2\Omega D^2} ; E_H = \frac{A_H}{2\Omega r_0^2}$$

$$\varepsilon_R = \frac{U}{2\Omega r_0} ; P_H = \frac{K_H}{U r_0} ; P_V = \frac{K_V r_0}{U D^2} ; B = \frac{r_0}{U \tau_T}$$

Apart from the parameter η_T in the forcing function, the system appears to contain 7 parameters. However, only 6 of these are independent; the characteristic horizontal velocity U can be chosen as a function of other parameters.

Since convection, which occurs in case of an unstable stratification, is not resolved by the hydrostatic model, an explicit representation is needed to obtain stably stratified solutions. Here, lacking other differentiable alternatives, we use implicit mixing as the form of convective adjustment (Yin and Sarachik, 1994). This means that when the flow becomes unstably stratified, the vertical mixing coefficient of heat is increased, i.e.

$$P_V = P_V^0 + P_V^c \mathcal{H}\left(-\frac{\partial T}{\partial z}; \epsilon_H\right) \quad (5)$$

where P_V^0 is the background inverse Péclet number, P_V^c is the convective inverse Péclet number, which is much larger than P_V^0 . \mathcal{H} is a continuous approximation to the Heaviside function, for which we use

$$\mathcal{H}(x; \epsilon_H) = \frac{1}{2}\left(1 + \tanh \frac{x}{\epsilon_H}\right) \quad (6)$$

with $\epsilon_H = 0.1$.

2.2 Numerical implementation

The equations and boundary conditions are implemented within a continuation code as is described in more detail by Dijkstra *et al.* (2000). The set of partial differential equations (1) with boundary conditions (4) is discretized on a $N \times M \times L$ grid. To increase the resolution near the surface, the grid was stretched in the z -direction, through the use of the mapping

$$z = \frac{\tanh(q_z \bar{z})}{\tanh(q_z)} \quad (7)$$

Here q_z is a stretching factor, the grid in \bar{z} is equidistant and the grid in z is non-equidistant. The vertical profile function $G(z)$ in (1e) is chosen as

$$G(z) = \mathcal{H}(z - z_{L-1}; \epsilon_H) \quad (8)$$

with z_{L-1} being the depth of the level just below the top and \mathcal{H} as in (6) with $\epsilon_H = 10^{-6}$.

After discretization a system of nonlinear differential equations with algebraic constraints results, which can be written as

$$\mathbf{M} \frac{d\mathbf{u}}{dt} = \mathbf{F}(\mathbf{u}, \boldsymbol{\lambda}) \quad (9)$$

Here \mathbf{u} is the d -dimensional state vector ($d = 5 \times N \times M \times L$), consisting of the unknowns (u, v, w, p, T) at each gridpoint, $\boldsymbol{\lambda}$ is the λ -dimensional vector of parameters, \mathbf{F} is a nonlinear mapping from $R^d \times R^\lambda \rightarrow R^d$ and \mathbf{M} is a linear operator. Stationary solutions satisfy the equation

$$\mathbf{F}(\mathbf{u}, \boldsymbol{\lambda}) = \mathbf{0} \quad (10)$$

which is a system of d nonlinear algebraic equations. To compute a branch of stationary solutions in a control parameter, say μ , a pseudo-arclength method (Keller, 1977) is used. The branches of stationary solutions $(\mathbf{u}(s), \mu(s))$ are parameterized by an 'arclength' parameter s . Since this introduces an extra unknown, an additional equation is needed and the tangent is normalized along the branch, i.e.

$$\dot{\mathbf{u}}_0^T(\mathbf{u} - \mathbf{u}_0) + \dot{\mu}_0^T(\mu - \mu_0) - \Delta s = 0 \quad (11)$$

where Δs is the step length, the superscript T denotes the transpose and a dot indicates differentiation to s . The Newton-Raphson method is used to converge to the branch of stationary solutions. This method finds isolated steady solutions, regardless of their stability. The linear systems are solved with the GMRES method (an iterative linear systems solver) using an MRILU (a multigrid oriented) preconditioning technique (Dijkstra *et al.*, 2000).

When a steady state is determined, the linear stability of the solution is considered and transitions that mark qualitative changes such as transitions to multiple equilibria (pitchfork bifurcations or limit points) or periodic behavior (Hopf bifurcations) can be detected. The linear stability analysis amounts to solving a generalized eigenvalue problem of the form

$$\alpha \mathcal{A} \mathbf{u} = \beta \mathcal{B} \mathbf{u} \quad (12)$$

where \mathcal{A} is the Jacobian matrix (the derivative of \mathbf{F} to \mathbf{u}) and $\mathcal{B} = -\mathbf{M}$. The matrices \mathcal{A} and \mathcal{B} are in general non-symmetric matrices, and α and β are complex numbers. If \mathcal{B} is nonsingular, the problem reduces to an ordinary eigenvalue problem for the matrix $\mathcal{B}^{-1}\mathcal{A}$. Because only real matrices are considered, there are d eigenvalues which are either real or occur as complex conjugate pairs. However, if \mathcal{B} is singular, the eigenvalue structure may be more complicated; the set of eigenvalues may be finite, empty or even the whole complex plane (Golub and Van Loan, 1983). In the particular model here, \mathcal{B} is a singular diagonal matrix because time derivatives are absent in the continuity equation and vertical momentum equation. The problem (12) is solved by the Jacobi-Davidson QZ (JDQZ) method (Sleijpen and Van der Vorst, 1996). With this method one can compute several eigenvalues and optionally eigenvectors near a specified target. Details of the method are described in Sleijpen and Van der Vorst (1996) and the implementation of JDQZ in an earlier version of our continuation code in Van Dorselaer (1997). In case $\alpha \neq 0$ in (12), we will use the notation $\sigma = \sigma_r + i\sigma_i = \beta/\alpha$ to indicate the eigenvalue.

3 Results

The domain is a single hemispheric 64° -wide sector in longitude (with $\phi_W = 286^\circ$ and $\phi_E = 350^\circ$) between latitudes $\theta_S = 10^\circ\text{N}$ and $\theta_N = 74^\circ\text{N}$ and with constant depth $D = 4000$ m. Solutions have been computed on a $16 \times 16 \times 16$ grid with a vertical stretching factor $q_z = 2$, which yields a horizontal resolution of $4^\circ \times 4^\circ$ and a vertical resolution ranging from 41 m near the surface to 516 m near the bottom. The resolution dependence of the solutions in this model was considered in Dijkstra *et al.* (2000).

2Ω	$=$	$1.4 \cdot 10^{-4}$	$[s^{-1}]$	r_0	$=$	$6.4 \cdot 10^6$	$[m]$
D	$=$	$4.0 \cdot 10^3$	$[m]$	U	$=$	$1.0 \cdot 10^{-1}$	$[ms^{-1}]$
ρ_0	$=$	$1.0 \cdot 10^3$	$[kgm^{-3}]$	g	$=$	9.8	$[ms^{-2}]$
α_T	$=$	$1.0 \cdot 10^{-4}$	$[K^{-1}]$	τ_T	$=$	$7.5 \cdot 10^1$	$[days]$
A_H	$=$	$1.6 \cdot 10^7$	$[m^2s^{-1}]$	A_V	$=$	$1.0 \cdot 10^{-3}$	$[m^2s^{-1}]$
K_H	$=$	$1.5 \cdot 10^3$	$[m^2s^{-1}]$	K_V	$=$	$2.3 \cdot 10^{-4}$	$[m^2s^{-1}]$
T_0	$=$	15.0	$[K]$	K_V^c	$=$	$3.3 \cdot 10^{-3}$	$[m^2s^{-1}]$
ΔT	$=$	1.0	$[K]$				
Ra	$=$	$4.2 \cdot 10^{-2}$		P_H^0	$=$	$2.3 \cdot 10^{-3}$	
E_H	$=$	$2.7 \cdot 10^{-3}$		P_V^0	$=$	$9.2 \cdot 10^{-4}$	
E_V	$=$	$4.3 \cdot 10^{-7}$		B	$=$	$1.0 \cdot 10^1$	
P_V^c	$=$	$1.3 \cdot 10^{-2}$		η_T	$=$	$1.0 \cdot 10^1$	
ε_R	$=$	0.0					

Table 1: Standard values of parameters used in the numerical calculations.

3.1 Forcing conditions and parameter choices

The prescribed surface temperature T_S is idealized as

$$T_S(\theta) = \cos\left(\pi \frac{\theta - \theta_S}{\theta_N - \theta_S}\right) \quad (13)$$

Note that, because of the introduction of the parameter η_T in (2), the dimensional meridional temperature difference over the sector is equal to $2\eta_T\Delta T$.

Most of the parameters are fixed at values as used in low resolution ocean general circulation models and these values are listed in Table 1. For these parameter values the Rossby number ε_R is small ($O(10^{-4})$) and hence has been set to zero in all our calculations. Note that the horizontal friction coefficient A_H is rather large. The value of A_H is bounded from below by the thickness of the boundary layers which develop near the continents. Near the western boundary, the Munk frictional boundary layer thickness at a latitude θ_0 scales with $(A_H/\beta_0)^{1/3}$, where $\beta_0 = 2\Omega \cos \theta_0/r_0$ monitors the variation of the Coriolis parameter. With a typical horizontal resolution of 4° , this leads to a typical lower bound of $A_H = 5 \cdot 10^5 m^2s^{-1}$ at $\theta_0 = 45^\circ$. However, the Ekman layers near the continental walls have a typical width of $(A_H/f_0)^{1/2}$, where $f_0 = 2\Omega \sin \theta_0$, which restricts the value of A_H to be larger than $1 \cdot 10^7 m^2s^{-1}$. To be on the safe side, we took a value about twice the latter one. In typical ocean models, much smaller values are taken, but it has been shown that this leads to numerical waves near the boundaries (Winton, 1996; Killworth, 1985) which show up as wiggles in the steady state solutions (Dijkstra *et al.*, 2000).

3.2 The interdecadal oscillation

A branch of steady state solutions under restoring boundary conditions is computed by increasing the parameter η_T from 0 to its standard value $\eta_T = 10.0$. The steady

state at $\eta_T = 10.0$ has a maximum meridional overturning of $20 Sv$. The overturning streamfunction is plotted in Fig. 1a and shows the typical unicellular structure with sinking confined to the northernmost part of the domain. In Figs. 1c and e vector plots of the horizontal circulation for certain sections are shown, superposed on contour plots of the vertical velocity. The surface circulation is anti-cyclonic (Fig. 1c) with upward vertical velocities at the western part of the basin. A reversed flow occurs near the bottom (Fig. 1e), consistent with the overturning flow.

A section of temperature in a north-south vertical plane shows a 'thermocline' in the upper 1000 m , with slight static instabilities in the northern part of the domain (Fig. 1b). Surface temperatures show small advective departures (Fig. 1d) from the zonally uniform state, while at depth there is only very little variation (Fig. 1f). Note that the difference between the maximum and minimum temperature in the latter panel is only 0.7 degree.

The surface heat flux Q_{Td} of the steady state in Fig. 1 is shown in Fig. 2. The maximum amplitude of the surface heat flux is $45 Wm^{-2}$, which is of the right order of magnitude for the North Atlantic domain (Oberhuber, 1988). The heat flux is negative (positive) in the northern (southern) half of the basin with a slight signature of the western intensification of the ocean flow. The particular state in Fig. 1 is also a solution of the steady equations when the flow is forced by the prescribed flux Q_{Td} . In other words, this heat flux is needed to maintain the circulation and a surface temperature which closely matches the imposed temperature T_S .

If one considers the stability of this steady state under restoring boundary conditions (Q_T given by equation (3)), such that temperature perturbations are considerably damped at the surface, it turns out that this state is linearly stable, because all eigenvalues have negative real part. The eigenvalue with the largest real part has $\sigma_r = -0.0092$ and $\sigma_i = 0$, and is associated with a mode with basin scale warming. The next least stable eigenmode has a complex pair of eigenvalues, corresponding to an oscillation period $\mathcal{P} \approx 450$ years. This mode is associated with basin scale temperature and circulation anomalies and is similar to the overturning oscillations in Dijkstra and Molemaker (1997). A real eigenvalue and another complex pair, the latter associated with a mode having a period $\mathcal{P} \approx 320$ years, are next in the spectrum. Although the patterns may be interesting, because these modes may be excited by stochastic noise, we will not consider them further, since their oscillation time scale is much larger than interdecadal.

One can also consider the stability of the steady state under the prescribed heat flux forcing Q_{Td} . In this way, the temperature anomalies are not damped at the surface (Greatbatch and Zhang, 1995). Note that in this case, the temperature is determined up to an additive constant and hence a zero eigenvalue $\sigma = 0$ must appear in the spectrum. Under this forcing condition, the state in Fig. 1 is unstable to an eigenmode with a complex pair of eigenvalues $\sigma = 0.012 \pm 0.1966 i$, which corresponds to an oscillation period $\mathcal{P} = 2\pi r_0 / (U\sigma_i) \approx 65$ years. The imaginary and real part of the eigenvector $\mathbf{x} = \mathbf{x}_R + i\mathbf{x}_I$ corresponding to this oscillatory pair provide the time periodic disturbance structure $\mathbf{P}(t)$ with angular frequency σ_i and

growth rate σ_r to which the steady state is unstable, i.e.

$$\mathbf{P}(t) = e^{\sigma_r t} [\mathbf{x}_R \cos(\sigma_i t) - \mathbf{x}_I \sin(\sigma_i t)] \quad (14)$$

The propagation of the perturbations can be followed by looking for example at $\mathbf{P}(\frac{-\pi}{2\sigma_i}) = \mathbf{x}_I$ and then at $\mathbf{P}(0) = \mathbf{x}_R$.

Pictures of typical patterns corresponding to \mathbf{x}_I and \mathbf{x}_R are shown in Figs. 3 and 4, respectively. Note that since these patterns are derived from eigenvectors, the absolute amplitude is arbitrary; only the relative amplitude of the fields within one eigenvector is fixed. To characterize the eigenmode, meridional overturning streamfunction, surface and deep velocities and three slices of the temperature field are plotted. The mode is clearly not localized but has quite a global structure, although the anomalies reach their maximum amplitude in the north and near the surface. At $t = -\pi/2\sigma_i$, the meridional overturning perturbation is positive, indicating a strengthening of the northward flow near the surface (Fig. 3a). There is anomalous upwelling in the northwest and anomalous downwelling in the northeast of the basin (Fig. 3c), with a similar pattern at the bottom of the domain (Fig. 3e). Near the surface there is a negative temperature anomaly in the northwestern corner, which extends down to the bottom due to 'convective mixing' (Figs. 3b, d and f). Otherwise, the temperature anomaly has quite a baroclinic character. A quarter of a period later, at $t = 0$, the meridional overturning is positive over almost the whole basin and the small oppositely rotating cell in the southern part of the domain that was present at $t = -\pi/2\sigma_i$, has disappeared (Fig. 4a). The vertical velocity perturbations have now switched location, with anomalous downwelling (upwelling) in the western (eastern) part of the domain (Figs. 4c, e). The temperature perturbation is now positive over most of the northern part of the basin, with a maximum in the northeast (Figs. 4b, d and f).

Clearly, there is propagation of both temperature and velocity anomalies. For later reference and for comparison with earlier work, the perturbation vertical velocities (Fig. 5) and temperature anomalies (Fig. 6) near the surface are plotted over half of the oscillation period. These fields are computed from the real and imaginary part of the eigenvector according to (14), where the exponential growth factor is not taken into account. The time scale in the figure caption is now dimensional; remember that the period of the oscillation is 65 years. Clearly, the vertical velocity anomalies have their largest amplitudes near the northern boundary and propagate westwards (Fig. 5). The positive temperature anomaly at the surface, present at $t = 0$, follows the same propagation as the vertical velocities near the northern boundary (Fig. 6). Along the southern boundary, the anomalies are relatively weak and propagate eastwards.

The pattern and propagation behavior of this interdecadal eigenmode resemble those of the 50-year oscillation found by Greatbatch and Zhang (1995). Also Colin de Verdière and Huck (1999) have found an interdecadal oscillation with a period of about 30 years with similar patterns in surface temperature and velocity anomalies. Apart from the western third of the domain, where their temperature anomalies are stationary, they find a westward propagation of temperature anomalies in the northern part of the domain. Huck *et al.* (1999) describe two types of

interdecadal oscillations. The first type of oscillatory behavior is characterized by westward propagation of temperature anomalies in the northern part of the basin, while oscillations of the second type show stationary temperature anomalies in the northwest part of the domain. They argue that, depending on the structure of the prescribed fluxes and the magnitude of the diffusion, westward propagation is either stronger than advection by the mean eastward flow, leading to westward propagating anomalies, or that both effects more or less compensate each other, leading to more stationary behavior. The interdecadal oscillation we find here seems to be of the first type, like the one found by Greatbatch and Zhang (1995) and Colin de Verdière and Huck (1999), as there is clear westward propagation in the northern part of the basin.

3.3 Regime diagram

The results in Chen and Ghil (1995) and Colin de Verdière and Huck (1999) clearly suggest that the unstable interdecadal mode can be stabilized by increasing the horizontal mixing coefficient of heat (K_H). To investigate the behavior of the growth rate and period with K_H , we have computed steady states under restoring conditions by using K_H as control parameter. For five different states, the surface heat flux was diagnosed and the stability of the steady state determined under prescribed flux conditions (as explained above). The growth rate and period corresponding to the interdecadal mode for the five different values of K_H are shown in Fig. 7. The growth rate crosses the zero-axis at $K_H = 1670 \text{ m}^2\text{s}^{-1}$. This indicates that a supercritical Hopf bifurcation occurs with decreasing K_H ; the period at criticality is about 69 *yr*. For $K_H > 1670 \text{ m}^2\text{s}^{-1}$, the steady state is (linearly) stable, but for $K_H < 1670 \text{ m}^2\text{s}^{-1}$ it is unstable. The growth rate increases for smaller K_H and the period shortens slightly, being about 50 years at $K_H = 800 \text{ m}^2\text{s}^{-1}$.

The location of the Hopf bifurcation defines the parameter value of K_H , at fixed K_V , bounding a steady flow regime and an oscillatory regime. By following the path of this Hopf bifurcation in another parameter, a regime diagram in a two-parameter plane is obtained. In the (K_V, K_H) parameter plane such a diagram (based on only a limited number of points) is plotted in Fig. 8. The regime below the curve marked with the open squares is the oscillatory regime. Slightly below this curve, periodic orbits of interdecadal period are expected. Increasing K_H stabilizes the interdecadal mode, while increasing K_V has a destabilizing effect. Changes in K_V have mainly an effect through changes of the steady state overturning, which increases with increasing K_V . The point labelled with a diamond in Fig. 8 indicates standard conditions and is located in the oscillatory regime. The period of the oscillation increases with decreasing K_V , but remains in the interdecadal range.

The use of convective adjustment is not essential for the occurrence of the oscillations. Convective adjustment decreases the critical K_H -value at Hopf bifurcation; without convective adjustment ($P_V^c = 0$) the critical K_H -value is about $1400 \text{ m}^2\text{s}^{-1}$. The destabilizing effect of convective adjustment corresponds to the destabilizing effect of K_V on the oscillation (note that convective adjustment corresponds to a local increase of K_V). Convective adjustment decreases the period, again in corre-

spondence with the effect of K_V . Although we have not explicitly considered the influence of bottom topography, it is expected that bottom topography stabilizes the interdecadal mode (Winton, 1997), since it also decreases the overturning.

The supercritical Hopf bifurcation is in correspondence with the results in Chen and Ghil (1995) and Colin de Verdière and Huck (1999). Huck *et al.* (1999) found that the critical K_H -value varied between 800 and 2500 m^2s^{-1} , depending on the model and horizontal resolution used. They also found that convective adjustment was not essential for the oscillation to occur.

4 The physics of the interdecadal oscillation

There have been several attempts to describe the physical mechanism of this type of low frequency (interdecadal) oscillations. To our opinion, to obtain a satisfactory description one has to separate growth of perturbations under unstable conditions from the physical mechanism which causes the oscillatory behavior.

4.1 Growth of perturbations

To understand why the interdecadal mode is unstable under a forcing with prescribed flux conditions, we must consider the growth of perturbations (below indicated by quantities with a tilde) on a particular steady state (below indicated by quantities with a bar). This is described by the equation for temperature perturbations

$$\frac{\partial \tilde{T}}{\partial t} + \tilde{\mathbf{u}} \cdot \nabla \tilde{T} + \tilde{\mathbf{u}} \cdot \nabla \bar{T} + \tilde{\mathbf{u}} \cdot \nabla \tilde{T} = P_H \nabla_H \cdot (\nabla_H \tilde{T}) + P_V \frac{\partial}{\partial z} \left(\frac{\partial \tilde{T}}{\partial z} \right) + \tilde{Q}_T G(z) \quad (15)$$

where $\mathbf{u} = (u, v, w)^T$ and \tilde{Q}_T is given by

$$\tilde{Q}_T = \begin{cases} -Bi \tilde{T} & \text{restoring} \\ 0 & \text{prescribed flux} \end{cases}$$

Multiplying by \tilde{T} , integration over the flow domain and averaging over one oscillation period gives

$$\frac{1}{2} \frac{\partial \langle \tilde{T}^2 \rangle}{\partial t} = -\langle \tilde{T} \tilde{\mathbf{u}} \cdot \nabla \bar{T} \rangle + \langle \tilde{T} \tilde{Q}_T G(z) \rangle - \langle \mathcal{D}_T \rangle \quad (16)$$

where \mathcal{D}_T is the buoyancy dissipation, brackets $\langle \cdot \rangle$ denote volume integration and a long bar denotes averaging over the period.

Note that $\langle \tilde{T} \mathbf{u} \cdot \nabla \tilde{T} \rangle = \frac{1}{2} \langle \mathbf{u} \cdot \nabla \tilde{T}^2 \rangle = 0$ because of kinematic boundary conditions. This equation is essentially the volume integrated equation for the available potential energy (Huang, 1998). Since $\langle \mathcal{D}_T \rangle$ is always positive definite, growth can only occur if the first term on the right hand side is positive, i.e. if $\langle \tilde{T} \tilde{\mathbf{u}} \cdot \nabla \bar{T} \rangle$ is negative. The latter term is interpreted as the change of available potential energy due to interaction of the buoyancy perturbation and the anomalous buoyancy

advection. The restoring boundary condition introduces an extra damping term $-\langle Bi \tilde{T}^2 G(z) \rangle$, which is absent under flux conditions.

The relative magnitudes of the terms in equation (16) under both restoring and flux conditions are given in Table 2. Upon a switch from restoring to flux conditions, the damping term from the restoring boundary condition disappears, but only this is not enough to make $\partial \langle \tilde{T}^2 \rangle / \partial t$ positive. Also the slight changes in the pattern of the eigenmode that occur upon a switch from restoring to flux conditions are important, as these cause an increase in the advective contribution, relative to the dissipation.

	$-\langle \tilde{T} \tilde{\mathbf{u}} \cdot \nabla \tilde{T} \rangle$	$\langle \tilde{T} \tilde{Q}_T G(z) \rangle$	$-\langle \mathcal{D}_T \rangle$	$\frac{1}{2} \langle \tilde{T}^2 \rangle_t$
Restoring	0.73	-0.32	-1.0	-0.59
Flux	1.12	0.0	-1.0	0.12

Table 2: Dimensionless terms in equation (16) under restoring and flux boundary conditions. Values are scaled with the absolute value of the dissipation.

The changes in growth rate with varying K_H and K_V (section 3.3) can also be understood by looking at the terms in equation (16). For several values of K_H and K_V the relative magnitudes of these terms are given in Table 3. With increasing

$K_H (m^2 s^{-1})$	$K_V (m^2 s^{-1})$	$-\langle \tilde{T} \tilde{\mathbf{u}} \cdot \nabla \tilde{T} \rangle$	$-\langle \mathcal{D}_T \rangle$	$\frac{1}{2} \langle \tilde{T}^2 \rangle_t$
796.0	$2.3 \cdot 10^{-4}$	1.36 (0.82, 0.54)	-1.0 (-0.61, -0.39)	0.36
1460.0	$2.3 \cdot 10^{-4}$	1.12 (0.80, 0.32)	-1.0 (-0.81, -0.19)	0.12
1800.0	$2.3 \cdot 10^{-4}$	0.99 (0.75, 0.24)	-1.0 (-0.89, -0.11)	-0.01
$K_H (m^2 s^{-1})$	$K_V (m^2 s^{-1})$	$-\langle \tilde{T} \tilde{\mathbf{u}} \cdot \nabla \tilde{T} \rangle$	$-\langle \mathcal{D}_T \rangle$	$\frac{1}{2} \langle \tilde{T}^2 \rangle_t$
1460.0	$2.3 \cdot 10^{-4}$	1.12 (0.80, 0.32)	-1.0 (-0.81, -0.19)	0.12
1460.0	$3.7 \cdot 10^{-4}$	1.20 (0.87, 0.33)	-1.0 (-0.81, -0.19)	0.20
1460.0	$7.0 \cdot 10^{-4}$	1.66 (1.35, 0.31)	-1.0 (-0.75, -0.25)	0.66

Table 3: Dimensionless terms in equation (16) for different values of K_H and K_V under prescribed flux conditions. Values are scaled with the absolute value of the total dissipation. The numbers between brackets in the third column are the values of the horizontal and vertical component of $-\langle \tilde{T} \tilde{\mathbf{u}} \cdot \nabla \tilde{T} \rangle$, respectively and the numbers between brackets in the fourth column are the horizontal and vertical dissipation, respectively. The terms for standard values ($K_H = 1460 \text{ m}^2 \text{ s}^{-1}$, $K_V = 2.3 \cdot 10^{-4} \text{ m}^2 \text{ s}^{-1}$) are given twice for convenience.

K_H , advection becomes less important with respect to the total dissipation (Table 3), mainly due to a relative decrease in the vertical advection. Eventually, the growth rate for $K_H = 1800 \text{ m}^2 \text{ s}^{-1}$, $K_V = 2.3 \cdot 10^{-4} \text{ m}^2 \text{ s}^{-1}$ becomes negative, in correspondence with Fig. 7. If K_V is increased, the advection increases relative to the total dissipation. A larger vertical diffusivity causes a stronger circulation, resulting in a relatively larger contribution of the horizontal advective term (Table 3), which has a destabilizing effect.

The term $\overline{\langle \tilde{T} \tilde{\mathbf{u}} \cdot \nabla \bar{T} \rangle}$ was considered in Colin de Verdière and Huck (1999) and used to demonstrate the growth of perturbations leading to the interdecadal oscillation in their model. While this approach and their conclusions are correct, there are two subtleties.

- (i) Colin de Verdière and Huck (1999) consider as perturbation the difference between the equilibrated periodic orbit at supercritical conditions and a reference state. In this case, either this reference state and/or the time mean state may not satisfy exactly the steady equations, so that other production terms appear in (16). These terms will become larger as one is further away from critical conditions. In our approach here, this problem does not occur, because the underlying unstable steady state and the eigenvectors are available.
- (ii) The production term $\overline{\langle \tilde{T} \tilde{\mathbf{u}} \cdot \nabla \bar{T} \rangle}$ may explain growth of perturbations but, apart from its cumbersome interpretation, it is also not mechanistically selective. Every type of instability must have this term negative in the unstable regime.

Hence, it is important to distinguish between the growth of perturbations and the physical mechanism driving the oscillation. In understanding the mechanism, the crucial point is to explain the phase difference between active fields in the oscillation (i.e. temperature and velocity fields). This will also indicate which physical processes determine the time scale of oscillation. In previous papers, this is recognized (Colin de Verdière and Huck, 1999), but a clear description of the physics of this phase difference has not been given.

4.2 Buoyancy work changes

As a means of analyzing what happens during the oscillation, the changes in potential energy along an oscillation cycle are considered. The balance for the volume integrated potential energy $\mathcal{U} = \langle -zT \rangle$ is directly obtained from the temperature equation (1e) and becomes

$$\frac{d\mathcal{U}}{dt} = - \langle wT \rangle + \langle z(P_V T_z)_z \rangle + \langle z Q_T G(z) \rangle \quad (17)$$

where the subscript z denotes differentiation. If one considers infinitesimal perturbations on a steady state, then the potential energy balance of the perturbations becomes

$$\frac{d\tilde{\mathcal{U}}}{dt} = -(\langle \bar{w}\tilde{T} \rangle + \langle \tilde{w}\bar{T} \rangle) + \langle z(P_V \tilde{T}_z)_z \rangle \quad (18)$$

The first term on the right hand side of (18) denotes the production (or destruction) of potential energy due to the effect of temperature perturbations on the mean flow, while the second term denotes the production (or destruction) of potential energy due to the effect of flow perturbations on the background stratification. The last term in (18) provides the change in potential energy due to changes in the stratification. The interpretation of $\langle \bar{w}\tilde{T} \rangle$ is as follows: in a situation of steady state

downwelling ($\bar{w} < 0$) and a negative temperature perturbation ($\tilde{T} < 0$), relatively cold water is transported downwards, thereby decreasing the potential energy of the flow. Similar interpretations hold for the other terms (Huang, 1998). The first two terms on the right hand side of (18) turn out to be dominant in the perturbation potential energy budget. The maximum value of the term $\langle z(P_V \tilde{T}_z)_z \rangle$ is only 18% of the maximum of $\langle \tilde{w} \tilde{T} \rangle$ and 14% of the maximum of $\langle \bar{w} \tilde{T} \rangle$.

The first two terms of the right hand side in (18) are directly related to the mechanical energy balance of the flow. This balance can be obtained by multiplying the momentum equations by the velocity field and integrating over the domain which gives

$$\varepsilon_R \frac{d\mathcal{E}}{dt} = \langle wT \rangle - \langle \mathcal{D}_M \rangle \quad (19)$$

where $\mathcal{E} = \frac{1}{2} \langle u^2 + v^2 \rangle$ is the volume integrated kinetic energy of the flow and \mathcal{D}_M is the dissipation of kinetic energy. Because the momentum equations are approximately diagnostic (and in our computations fully diagnostic, as we chose $\varepsilon_R = 0$), changes in the buoyancy work $\langle wT \rangle$ induce changes in the kinetic energy, but these are instantaneously (with respect to the large time scale of the oscillation) balanced by dissipation. As there are no other production terms in (19), it are the changes in $\langle wT \rangle$ which control the oscillatory behavior of the kinetic energy of the full three-dimensional flow.

In Fig. 9, the terms $\langle \bar{w} \tilde{T} \rangle$ and $\langle \tilde{w} \tilde{T} \rangle$ are plotted during one oscillation cycle for the oscillation described in section 3.2 (with $K_H = 1460 \text{ m}^2 \text{ s}^{-1}$). A phase difference exists between both terms of the perturbation buoyancy work. About 26 years after $\langle \bar{w} \tilde{T} \rangle$ has reached its minimum amplitude, $\langle \tilde{w} \tilde{T} \rangle$ is at a minimum, after which it takes about 6 years before $\langle \bar{w} \tilde{T} \rangle$ reaches a maximum, after which the second half of the oscillation occurs in reversed order. The phase difference between these two fields drives the oscillatory behavior of the basin integrated buoyancy work and hence through the coupling to the mechanical energy balance causes the oscillatory changes in the full three-dimensional flow.

The spatial patterns of $\bar{w} \tilde{T}$ and $\tilde{w} \tilde{T}$ in a horizontal plane near the surface have been plotted for different phases of the oscillation in Figs. 10 and 11, respectively. The pattern of $\bar{w} \tilde{T}$ has a basin-wide structure, with the largest amplitudes near the northern boundary (Fig. 10), as the background vertical velocity is much larger there than in the rest of the basin. Temperature anomalies propagate in northward and westward direction, which can be seen in the pattern of $\bar{w} \tilde{T}$. Near the northern boundary the anomalies extend to the bottom due to convection, but in the rest of the basin the signal has a baroclinic character. In contrast to the basin scale patterns of $\bar{w} \tilde{T}$, the spatial pattern of $\tilde{w} \tilde{T}$ (Fig. 11) seems to be confined to a band in the northern part of the basin, in which the vertical velocity anomalies propagate westwards. To establish the effect of the propagation of the w - and T - anomalies on the fields of $\bar{w} \tilde{T}$ and $\tilde{w} \tilde{T}$ more clearly, Hovmöller diagrams of $\bar{w} \tilde{T}$ and $\tilde{w} \tilde{T}$ are plotted along a line at the surface near the northern boundary (Fig. 12). Westward propagation of w -anomalies can be seen in the pattern of $\tilde{w} \tilde{T}$ during the whole oscillation, while the propagation of T -anomalies occurs mainly between $t = 15 \text{ yr}$ and $t = 30 \text{ yr}$ and between $t = 45 \text{ yr}$ and $t = 60 \text{ yr}$ and is seen in $\bar{w} \tilde{T}$. Hence, to

understand the phase difference between $\langle \bar{w}\tilde{T} \rangle$ and $\langle \tilde{w}\bar{T} \rangle$, we have to go back to the evolution of the temperature and vertical velocity anomalies (Fig. 5 and 6).

4.3 Origin of the phase difference

To explain how the propagation of the velocity and temperature perturbations induces the phase difference in the buoyancy work terms, we start with the situation at $t = 0$ yr (but we may take any point along the oscillation). At this stage, a positive temperature anomaly is present in the northern part of the basin (Fig. 6a) and propagates slowly westwards. In Fig. 13, the different advective terms in the perturbation temperature equation (15) are plotted along a zonal section near the northern boundary at the surface. The dominant advective terms, except near the eastern boundary, are given by

$$\frac{\partial \tilde{T}}{\partial t} \approx - \left(\tilde{v} \frac{\partial \bar{T}}{\partial \theta} + \frac{\bar{u}}{\cos \theta} \frac{\partial \tilde{T}}{\partial \phi} \right) \quad (20)$$

Using this balance, an estimate of the time scale of propagation can be derived (generalizing earlier work in Colin de Verdière and Huck (1999)) by using approximate analytic expressions for the perturbation meridional geostrophic velocity. This velocity can be expressed in terms of the zonal derivative of the temperature perturbation, as is derived in the appendix. The temperature perturbation in the north-central part of the basin satisfies to a reasonable approximation an exponential decrease, i.e.

$$\tilde{T}(\phi, \theta, z, t) = e^{\kappa z} \hat{T}(\phi, \theta, t) \quad (21)$$

with κ a constant. Using this approximation, the integrals over depth in (30a) in the appendix can be evaluated exactly. This gives

$$\frac{\partial \hat{T}}{\partial t} - \frac{c}{\cos \theta} \frac{\partial \hat{T}}{\partial \phi} \approx 0 \quad (22)$$

where c is a phase speed of propagation of temperature anomalies ($c > 0$ indicates westward propagation), which is at the surface given by

$$c = - \left[\bar{u} + \alpha \frac{Ra}{\sin \theta} \frac{\partial \bar{T}}{\partial \theta} \right] \quad (23)$$

where $\alpha = 1/\kappa - 1/\kappa^2 + e^{-\kappa}/\kappa^2$. This phase speed is characteristic of the propagation of temperature anomalies on a background basic state temperature (or potential vorticity) gradient. For example, a warm anomaly in a negative mean temperature gradient will induce northward (southward) perturbation velocities west (east) of the center of the anomaly. The perturbation velocities advect warm (cold) water northwards (southwards) west (east) of the initial anomaly, thereby moving the anomaly to the west (Colin de Verdière and Huck, 1999).

We can use equation (23) to estimate the order of magnitude of the phase speed. As the phase speed is a relatively small difference between two large terms, we can

only try to determine an estimate of the maximum speed in both eastward and westward direction. The anomalies reach their maximum amplitude in the northern part of the basin, so that the magnitude of \bar{u} and $\partial\bar{T}/\partial\theta$ can be estimated by averaging over the horizontal area $[292^\circ, 344^\circ] \times [60^\circ, 72^\circ]$. Furthermore the anomalies are mostly confined to the upper 500 m. This yields estimates of $\bar{u} = 0.13$, $\partial\bar{T}/\partial\theta = -12$ near the surface, at $z = -19$ m and $\bar{u} = 0.06$, $\partial\bar{T}/\partial\theta = -10$ at $z = -557$ m, while averages over the upper 7 model layers (the upper 557 m) give $\bar{u} = 0.09$, $\partial\bar{T}/\partial\theta = -11$. Exponential fits of the temperature perturbation in the north-central part of the basin at $t = 0$ give $\alpha = 0.25$ (at $(\phi, \theta) = (320^\circ, 72^\circ)$) and $\alpha = 0.21$ (at $(\phi, \theta) = (316^\circ, 72^\circ)$). Combination of these estimates yields that the dimensional phase speed can range between $c^* = 7 \cdot 10^{-3} \text{ ms}^{-1}$ in westward direction and $c^* = 4 \cdot 10^{-3} \text{ ms}^{-1}$ in eastward direction. With such phase speeds, a westward propagating temperature anomaly along the northern boundary needs at least 10 years to cross the basin from east to west. This results in oscillation periods on decadal or longer timescales.

In the appendix, expressions are derived for approximations of the meridional and zonal perturbation overturning streamfunctions $\tilde{\Psi}_M$ and $\tilde{\Psi}_Z$, respectively, in terms of the temperature anomalies, for example,

$$\tilde{\Psi}_M(\theta, z, t) = - \int_{-1}^z \left[\int_{\phi_W}^{\phi_E} \frac{Ra}{\sin\theta} f_\phi(\phi, \theta, z', t) d\phi \right] dz' \quad (24a)$$

$$f_\phi(\phi, \theta, z, t) = \int_{-1}^z \frac{\partial\tilde{T}}{\partial\phi} dz' - \int_{-1}^0 \left(\int_{-1}^z \frac{\partial\tilde{T}}{\partial\phi} dz' \right) dz \quad (24b)$$

with a similar expression relating the zonal overturning to the average meridional temperature gradient (see (32) and (33) in the appendix). It is clear from equation (24) that the perturbation meridional overturning is directly related to the east-west perturbation temperature difference and similarly, that the perturbation zonal overturning is related to the north-south perturbation temperature difference. In Fig. 14, the vertically integrated north-south perturbation temperature difference ΔT_{N-S} and the vertically integrated east-west perturbation temperature difference ΔT_{E-W} are plotted during one oscillation cycle. During the slow propagation of the positive temperature anomaly westwards (say from $t = 0$ yr to $t = 15$ yr), ΔT_{N-S} hardly changes (Fig. 14) and hence the zonal overturning perturbation remains anti-clockwise. This can be seen in Fig. 15, where the zonal overturning streamfunction is plotted for several phases during the oscillation.

On the other hand, during this time interval ΔT_{E-W} decreases rapidly (Fig. 14). At this stage, the meridional overturning therefore decreases rapidly and changes from positive to negative around $t = 10$ yr, as can be seen in Fig. 16. This change in meridional overturning induces anomalous upwelling at the northern boundary of the domain compatible with the increase and westward propagation of vertical velocity anomalies (Figs. 5c and d). It also induces downwelling anomalies in the southern part of the basin. There, the evolution of the temperature anomalies is dominated by vertical exchange processes (Fig. 17), since the basic state vertical temperature gradient is relatively large. Hence, through anomalous downwelling

a positive temperature perturbation results in the south (Fig. 6f), which reduces the initially positive north-south perturbation temperature difference. The latter changes sign around $t = 22 \text{ yr}$, causing the zonal overturning perturbation to become positive (Fig. 15g). The positive zonal overturning perturbation after $t = 22 \text{ yr}$ induces downwelling along the eastern boundary and upwelling in the northwest (Figs. 5g and h). This increases the vertically integrated east-west temperature difference (Fig. 14) and the second half of the oscillation starts. Note that during the second half of the oscillation, a cold anomaly will lead to a positive zonal overturning anomaly, which causes upwelling of cold water in the west and downwelling of warm water in the east, again leading to westward propagation.

Now the processes controlling the phase difference between $\langle \bar{w}\tilde{T} \rangle$ and $\langle \tilde{w}\bar{T} \rangle$ can be explained. At $t = 0 \text{ yr}$, $\langle \bar{w}\tilde{T} \rangle$ is at a minimum (Fig. 9). The sign of $\langle \bar{w}\tilde{T} \rangle$ turns out to be mainly determined by the surface temperature perturbation in the northeast, so that looking at the surface fields of both \tilde{T} and the steady state shows that this term is dominated by steady state downwelling of anomalously warm water, thus increasing the potential energy (both $\langle \bar{w}\tilde{T} \rangle$ and $\langle \tilde{w}\bar{T} \rangle$ appear with a minus sign in (18)). As the temperature anomaly propagates westwards, where the basic state downwelling is weaker, $\langle \bar{w}\tilde{T} \rangle$ increases. At the same time, \tilde{T} in the northeast decreases through upwelling of cold water (which is a reaction to the negative zonal overturning from the previous phase of the oscillation), so that $\langle \bar{w}\tilde{T} \rangle$ increases even more and becomes positive after $t = 16 \text{ yr}$. When the temperature perturbation in the whole northern part of the basin has become negative, $\langle \bar{w}\tilde{T} \rangle$ reaches a maximum around $t = 33 \text{ yr}$, so that downwelling of anomalously cold water causes a decrease of potential energy. The sign of the term $\langle \tilde{w}\bar{T} \rangle$ is mainly determined along the northern boundary, so we see from Fig. 1 and Fig. 5 that, at $t = 0 \text{ yr}$, the anomalous downwelling in the northwest in the region of cold steady state temperatures causes $\langle \tilde{w}\bar{T} \rangle$ to be positive. This term is therefore decreasing the potential energy. The upwelling which occurs along the northern boundary due to the rapid change in the meridional overturning causes $\langle \tilde{w}\bar{T} \rangle$ to decrease and become negative around $t = 10 \text{ yr}$. This upwelling has spread out along most of the northern boundary at about $t = 24 \text{ yr}$, so that $\langle \tilde{w}\bar{T} \rangle$ reaches a minimum around $t = 27 \text{ yr}$.

In summary, the phase difference between $\langle \bar{w}\tilde{T} \rangle$ and $\langle \tilde{w}\bar{T} \rangle$ thus comes from the phase difference between \tilde{w} and \tilde{T} , which in turn originates from the westward propagation of the temperature anomalies and the interplay of changing zonal and meridional temperature gradients with subsequent responses of the zonal and meridional overturning (Fig. 18).

5 Discussion

It has been shown that a steady state obtained under a prescribed heat flux will become unstable once the horizontal thermal diffusion is small enough. Note that the shape of the heat flux is not important for the destabilization process, because it does not affect the linear stability problem. Only the fact that damping of temperature

anomalies is turned off is important. Obviously, this only holds at criticality and the shape of the heat flux will influence the periodic orbit arising from the instability in a way described by Huck *et al.* (1999). The new element in our study has been that the linear stability of these complex three-dimensional steady states has been computed and that it is demonstrated that the interdecadal mode arises through a supercritical Hopf bifurcation. A regime diagram in the (K_H, K_V) space shows that the stability characteristics change with these parameters, mainly through changes in the mean state, but that the time scale of the oscillatory mode remains interdecadal.

In correspondence with other studies, this shows that the interdecadal mode introduces a very robust preferred pattern in the buoyancy driven ocean circulation. Winton (1997) shows that interdecadal thermohaline oscillations that occur in a model with a flat bottom under flux boundary conditions disappear when a bowl-shaped bottom topography is included. As bottom topography is expected to change the overturning, the results here suggest that the Hopf bifurcation is shifted in parameter space due to the effect of bottom topography. This is in agreement with the reappearance of the oscillation, under bottom topography, if the forcing strength is increased (Winton, 1997).

The oscillation also exists under restoring boundary conditions, but in this case it is stable. Probably, the large spatial scale of the SST-anomalies as seen in interdecadal variability will have significant influence on the atmospheric temperature, so that a restoring boundary condition with standard restoring coefficients imposes an unrealistically fast relaxation. Chen and Ghil (1996) coupled a simple ocean model to an energy balance model of the atmosphere and showed that, at low frequencies, the ocean sees almost a constant heat flux. However, even if damping by the atmosphere is present, the existence of the interdecadal mode is not affected, but only its stability properties are changed.

Different ideas to explain the phase difference responsible for the oscillation have been proposed in the literature. In Greatbatch and Zhang (1995) it was suggested, but not analyzed further, that convective mixing was involved and in combination with advective processes, leads to a phase difference between the temperature and velocity anomalies. Both the results in Huck *et al.* (1999) and our results clearly indicate that the oscillation still exists under absence of convective mixing. This is also compatible with the mechanism described in the previous section. Although convective mixing may affect the phase difference between the two buoyancy production fields, it is not important for the existence of the oscillation. As convective mixing changes the vertical temperature gradient, it may reduce κ in (21). For $\kappa \gg 1$, the constant α is approximately inversely proportional to κ . A reduction of κ in this case will lead to an increase in α and hence a larger value of c . However, for values of $0 < \kappa < 1$, the effect of convective adjustment on the phase speed can be either way and it is not clear how convective adjustment would affect the phase difference and the period of the oscillation.

In Winton (1996) and Greatbatch and Peterson (1996), viscous boundary-trapped waves which propagate as Kelvin waves are suggested to be responsible for the phase difference between velocity and temperature field. In the model here, the vertical velocity anomalies are indeed boundary trapped. From the analysis of Winton (1996)

it follows that when the boundary-wave mechanism is dominant, the oscillation period should change significantly with the value of A_H (and scale with $A_H^{-1/2}$) when the steady state is fixed. In this case, the period should also be fairly insensitive to the buoyancy forcing Ra . We determined the linear stability of the steady state in Fig. 1 for different values of the parameters E_H and Ra , with the effect of changes in parameters only through the perturbation balances (the steady state remains the same). For $Ra = 0.5Ra_s$, $Ra = Ra_s$ and $Ra = 2Ra_s$, the dimensionless frequency σ_i is 0.14, 0.19 and 0.22, respectively, where Ra_s is the standard value. This increase in frequency is expected from (23), but does not support the viscous boundary-wave mechanism. For $E_H = E_{Hs}$, $2E_{Hs}$ and $5E_{Hs}$, we find frequencies of 0.19, 0.21 and 0.20, respectively which indicates a weak dependence on E_H . While effects of boundary-wave propagation may still be present, these results suggest a dominant role for the propagation mechanism as suggested in section 4.

The mechanism proposed here is also fairly well in agreement with that suggested in Colin de Verdière and Huck (1999). Indeed, the propagation of temperature anomalies and the subsequent response of the flow is crucial. Here, we have given a detailed mechanistic view of how the response of the velocity field through both zonal and meridional overturning streamfunctions is related to the propagation of the temperature anomalies. Since only the buoyancy work terms monitor the actual changes in the mechanical energy balance and hence in the kinetic energy of the flow, it is not sufficient to show only the phase difference in the responses of the two streamfunctions (zonal and meridional), but one has to explain how the phase difference between the buoyancy work terms arises.

There is another argument to use an interpretation in terms of buoyancy work. In Colin de Verdière and Huck (1999), the instability has been referred to as baroclinic instability. Growth of perturbations in baroclinic instability indeed must be controlled by the same production term in the available potential energy balance as growth of perturbations in the instability leading to the interdecadal oscillation. However, one may ask how the phase difference between the density field and the velocity field arises in a classical case of baroclinic instability, such as the Eady problem (Pedlosky, 1987). The change in potential energy of the perturbations is then given by

$$\frac{d\tilde{\mathcal{U}}}{dt} = \langle \bar{w}\tilde{\rho} \rangle + \langle \tilde{w}\bar{\rho} \rangle \quad (25)$$

where ρ is the density. For the dimensionless basic state $\bar{u} = z + 1$, $\bar{v} = \bar{w} = 0$ and $\bar{\rho} = y$, the first term on the right hand side is identically zero. In quasi-geostrophic theory, the vertical velocity perturbations at first order in the Rossby number are given by

$$\tilde{w} = \varepsilon_R \mathcal{S}^{-1} \left(\frac{\partial \tilde{\rho}}{\partial t} + \tilde{\mathbf{u}} \cdot \nabla \bar{\rho} + \bar{\mathbf{u}} \cdot \nabla \tilde{\rho} \right) \quad (26)$$

where \mathcal{S} is the Burger number (the square of the ratio of the length scale of the flow and the internal Rossby deformation radius), which is taken constant. Substituting

the basic state, we find for the potential energy change of the perturbations

$$\frac{d\tilde{\mathcal{U}}}{dt} = \varepsilon_R \mathcal{S}^{-1} \left(\langle y \frac{\partial \tilde{\rho}}{\partial t} \rangle + \langle y \tilde{v} \rangle + \langle y(z+1) \frac{\partial \tilde{\rho}}{\partial x} \rangle \right) \quad (27)$$

There exists a phase difference between these three terms, which is responsible for the oscillation. However, the basic state upwelling is not involved in this phase difference since it is identically zero.

For the interdecadal instabilities, however, the existence of a basic state with a non-zero vertical velocity is crucial to the existence of the instability. Such a state also appears necessary in the three-layer model suggested in Colin de Verdière and Huck (1999) where long time scale instabilities occur. To our opinion, one could view this type of interdecadal instability as a 'generalized' baroclinic instability, but then many instabilities could be labelled with the same name. For example, also the overturning oscillations found in two-dimensional thermohaline models (Dijkstra and Molemaker, 1997), where the phase difference is determined by the terms $\langle wT \rangle$ and $\langle wS \rangle$, where S is the salinity, could then be called baroclinic instability. If one wishes to distinguish oscillations according to the processes leading to the phase difference, then we think the interdecadal instability falls in a separate class of thermohaline instabilities. In this way, the terms in the buoyancy work are mechanistically selective.

In this context, it would be interesting to investigate the origin of the phase difference driving the interdecadal oscillation found by Chen and Ghil (1995). If it is mainly determined by the temperature, while salinity plays only a minor role, the mechanism behind this oscillation is basically the same as the one we propose here. On the other hand, salinity can turn out to be the important factor in determining the phase difference of the oscillation, for example it could be between $\langle \bar{w}\tilde{S} \rangle$ and $\langle \tilde{w}\bar{S} \rangle$, or between one of these terms and either $\langle \bar{w}\tilde{T} \rangle$ or $\langle \tilde{w}\bar{T} \rangle$. In these latter cases, this oscillation would be different from the one we have found.

The results here motivate a different approach for looking at observations. A nice element in this study is that the changes in the zonal and meridional overturning streamfunctions can be related to changes in vertically integrated meridional and zonal temperature differences, respectively (according to (24)). This motivates to look at the phase differences between east-west and north-south temperature differences over a long time in observations. Using these phase differences, a clear relationship between the interdecadal oscillations in this model and the variability found in observations can be established. In this way, the importance of instabilities of the thermohaline circulation can be assessed, which may lead to modifications of current theories which attribute a substantial role to the low-frequency atmospheric variability (Delworth and Mann, 2000; Delworth and Greatbatch, 2000).

Acknowledgements

This work was supported by the Netherlands Organization for Scientific Research (NWO) under a PIONIER grant to HD. The authors thank Wilbert Weijer and

Maurice Schmeits for stimulating discussions and Hakan Öksüzöglu for his help with the code and the post-processing. The useful comments of two anonymous reviewers are gratefully acknowledged.

Appendix: Diagnostic momentum equations

Since the momentum equations are linear when inertia is neglected, approximate expressions of the perturbation geostrophic velocities can be obtained in terms of temperature (or more general buoyancy) perturbations. Consider the set of diagnostic equations for the perturbation velocities, i.e.

$$-\tilde{v} \sin \theta = -\frac{1}{\cos \theta} \frac{\partial \tilde{p}}{\partial \phi} + E_H L_u(\tilde{u}, \tilde{v}) + E_V \frac{\partial^2 \tilde{u}}{\partial z^2} \quad (28a)$$

$$\tilde{u} \sin \theta = -\frac{\partial \tilde{p}}{\partial \theta} + E_H L_v(\tilde{u}, \tilde{v}) + E_V \frac{\partial^2 \tilde{v}}{\partial z^2} \quad (28b)$$

$$\frac{\partial \tilde{p}}{\partial z} = Ra \tilde{T} \quad (28c)$$

$$0 = \frac{\partial \tilde{w}}{\partial z} + \frac{1}{\cos \theta} \left(\frac{\partial \tilde{u}}{\partial \phi} + \frac{\partial(\tilde{v} \cos \theta)}{\partial \theta} \right) \quad (28d)$$

where $L_u(\tilde{u}, \tilde{v})$ and $L_v(\tilde{u}, \tilde{v})$ are linear operators denoting horizontal friction. Integration over the total depth shows that, since wind forcing is absent, the vertically averaged pressure is constant. Hence the pressure can be explicitly determined from the temperature fields, with the result

$$\tilde{p} = Ra \left[\int_{-1}^z \tilde{T} dz' - \int_{-1}^0 \left(\int_{-1}^z \tilde{T} dz' \right) dz \right] \quad (29)$$

Using this expression and neglecting friction, one obtains from the geostrophic relations that

$$\tilde{v} = \frac{Ra}{\sin \theta \cos \theta} \left[\int_{-1}^z \frac{\partial \tilde{T}}{\partial \phi} dz' - \int_{-1}^0 \left(\int_{-1}^z \frac{\partial \tilde{T}}{\partial \phi} dz' \right) dz \right] \quad (30a)$$

$$\tilde{u} = \frac{-Ra}{\sin \theta} \left[\int_{-1}^z \frac{\partial \tilde{T}}{\partial \theta} dz' - \int_{-1}^0 \left(\int_{-1}^z \frac{\partial \tilde{T}}{\partial \theta} dz' \right) dz \right] \quad (30b)$$

With these expressions and the definition of the perturbation meridional and zonal overturning streamfunction

$$\frac{\partial \tilde{\Psi}_M}{\partial z} = - \int_{\phi_W}^{\phi_E} \tilde{v} \cos \theta d\phi \quad ; \quad \frac{\partial \tilde{\Psi}_Z}{\partial z} = - \int_{\theta_S}^{\theta_N} \tilde{u} d\theta \quad (31)$$

expressions can be derived for $\tilde{\Psi}_M$ and $\tilde{\Psi}_Z$ in terms of the temperature anomalies, namely

$$\tilde{\Psi}_M(\theta, z, t) = - \int_{-1}^z \left[\int_{\phi_W}^{\phi_E} \frac{Ra}{\sin \theta} f_\phi(\phi, \theta, z', t) d\phi \right] dz' \quad (32a)$$

$$f_\phi(\phi, \theta, z, t) = \int_{-1}^z \frac{\partial \tilde{T}}{\partial \phi} dz' - \int_{-1}^0 \left(\int_{-1}^z \frac{\partial \tilde{T}}{\partial \phi} dz' \right) dz \quad (32b)$$

and

$$\tilde{\Psi}_z(\phi, z, t) = \int_{-1}^z \left[\int_{\theta_s}^{\theta_N} \frac{Ra}{\sin \theta} f_\theta(\phi, \theta, z', t) d\theta \right] dz' \quad (33a)$$

$$f_\theta(\phi, \theta, z, t) = \int_{-1}^z \frac{\partial \tilde{T}}{\partial \theta} dz' - \int_{-1}^0 \left(\int_{-1}^z \frac{\partial \tilde{T}}{\partial \theta} dz' \right) dz \quad (33b)$$

References

- Cai, W., Greatbatch, R. J., and Zhang, S. (1998). Interdecadal variability in an ocean model driven by a small, zonal redistribution of the surface buoyancy flux. *J. Phys. Oceanography*, **25**, 1998–2010.
- Chen, F. and Ghil, M. (1995). Interdecadal variability of the thermohaline circulation and high-latitude surface fluxes. *J. Phys. Oceanography*, **22**, 161–167.
- Chen, F. and Ghil, M. (1996). Interdecadal variability in a hybrid coupled ocean-atmosphere model. *J. Phys. Oceanography*, **26**, 1561–1578.
- Colin de Verdière, A. and Huck, T. (1999). Baroclinic instability: an oceanic wave-maker for interdecadal variability. *J. Phys. Oceanography*, **29**, 893–910.
- Delworth, T. L. and Greatbatch, R. J. (2000). Multidecadal thermohaline circulation variability driven by atmospheric surface flux forcing. *J. Climate*, **13**, 1481–1495.
- Delworth, T. L. and Mann, M. E. (2000). Observed and simulated multidecadal variability in the Northern Hemisphere. *Climate Dynamics*, **16**, 661–676.
- Delworth, T. L., Manabe, S., and Stouffer, R. J. (1993). Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model. *J. Climate*, **6**, 1993–2011.
- Deser, C. and Blackmon, M. L. (1993). Surface climate variations over the north atlantic ocean during winter: 1900–1989. *J. Climate*, **6**, 1743–1753.
- Dijkstra, H. A. and Molemaker, M. J. (1997). Symmetry breaking and overturning oscillations in thermohaline-driven flows. *J. Fluid Mech.*, **331**, 195–232.
- Dijkstra, H. A., Öksüzöglu, H., Wubs, F. W., and Botta, E. F. F. (2000). A fully implicit model of the three-dimensional thermohaline ocean circulation. *Journal of Computational Physics*, **xxx**, submitted.
- Ghil, M. and Vautard, R. (1991). Interdecadal oscillations and the warming trend in global temperature time series. *Nature*, **350**, 324–327.
- Golub, G. H. and Van Loan, C. F. (1983). *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, U.S.A.
- Greatbatch, R. J. and Peterson, K. A. (1996). Interdecadal variability and oceanic thermohaline adjustment. *J. Geophysical Res.*, **101**, 20,467–20,482.
- Greatbatch, R. J. and Zhang, S. (1995). An interdecadal oscillation in an idealized ocean basin forced by constant heat flux. *J. Climate*, **8**, 82–91.
- Haney, R. L. (1971). Surface thermal boundary conditions for ocean circulation models. *J. Phys. Oceanography*, **4**, 241–248.

- Hibler, W. D. and Johnson, S. J. (1979). The 20-yr cycle in greenland ice core records. *Nature*, **280**, 481–483.
- Huang, R. X. (1998). Mixing and available potential energy in a boussinesq ocean. *J. Phys. Oceanography*, **28**, 669–678.
- Huck, T., Colin de Verdière, A., and Weaver, A. J. (1999). Interdecadal variability of the thermohaline circulation in box-ocean models forced by fixed surface fluxes. *J. Phys. Oceanography*, **29**, 865–892.
- Keller, H. B. (1977). Numerical solution of bifurcation and nonlinear eigenvalue problems. In P. H. Rabinowitz, editor, *Applications of Bifurcation Theory*. Academic Press, New York, U.S.A.
- Killworth, P. D. (1985). A two-level wind and buoyancy driven thermocline model. *J. Phys. Oceanography*, **15**, 1414–1432.
- Kushnir, Y. (1994). Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *J. Phys. Oceanography*, **7**, 141–157.
- Moron, V., Vautard, R., and Ghil, M. (1998). Trends, interdecadal and interannual oscillations in global sea-surface temperature. *Climate Dynamics*, **14**, 545–569.
- Oberhuber, J. M. (1988). *The Budget of Heat, Buoyancy and Turbulent Kinetic Energy at the Surface of the Global Ocean*. Max Planck Institute für Meteorologie Hamburg report nr. 15, Hamburg, Germany.
- Pedlosky, J. (1987). *Geophysical Fluid Dynamics*. Springer-Verlag, New York, U.S.A.
- Roemmich, D. H. and Wunsch, C. (1984). Apparent changes in the climatic state of the deep north atlantic ocean. *Nature*, **307**, 447–450.
- Schlesinger, M. E. and Ramankutty, N. (1994). An oscillation in the global climate system of period 65-70 years. *Nature*, **367**, 723–726.
- Sleijpen, G. L. G. and Van der Vorst, H. A. (1996). A jacobi-davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, **17**, 410–425.
- Van Dorsselaer, J. J. (1997). Computing eigenvalues occurring in continuation methods with the Jacobi-Davidson QZ method. *J. Comp. Physics*, **138**, 714–733.
- Weaver, A. J. and Hughes, T. M. (1996). On the incompatibility of ocean and atmosphere and the need for flux adjustments. *Climate Dynamics*, **12**, 141–170.
- Weaver, A. J. and Sarachik, E. S. (1991a). Evidence for decadal variability in an ocean general circulation model: an advective mechanism. *Atmosphere-Ocean*, **29**, 197–231.

- Weaver, A. J. and Sarachik, E. S. (1991b). The role of mixed boundary conditions in numerical models of the ocean's climate. *J. Phys. Oceanography*, **21**, 1470–1493.
- Weaver, A. J. and Valcke, S. (1998). On the variability of the thermohaline circulation in the gfdl coupled model. *J. Climate*, **11**, 759–767.
- Weaver, A. J., Marotzke, J., Cummings, P. F., and Sarachik, E. S. (1993). Stability and variability of the thermohaline circulation. *J. Phys. Oceanography*, **23**, 39–60.
- Winton, M. (1996). The role of horizontal boundaries in parameter sensitivity and decadal-scale variability of coarse-resolution ocean general circulation models. *J. Phys. Oceanography*, **26**, 289–304.
- Winton, M. (1997). The damping effect of bottom topography on internal decadal-scale oscillations of the thermohaline circulation. *J. Phys. Oceanography*, **27**, 203–208.
- Yin, F. L. (1995). A mechanistic model of ocean interdecadal thermohaline oscillations. *J. Phys. Oceanography*, **25**, 3239–3246.
- Yin, F. L. and Sarachik, E. S. (1994). An efficient convective adjustment scheme for ocean general circulation models. *J. Phys. Oceanography*, **24**, 1425–1430.
- Yin, F. L. and Sarachik, E. S. (1995). Interdecadal thermohaline oscillations in a sector ocean general circulation model: advective and convective processes. *J. Phys. Oceanography*, **25**, 2465–2484.

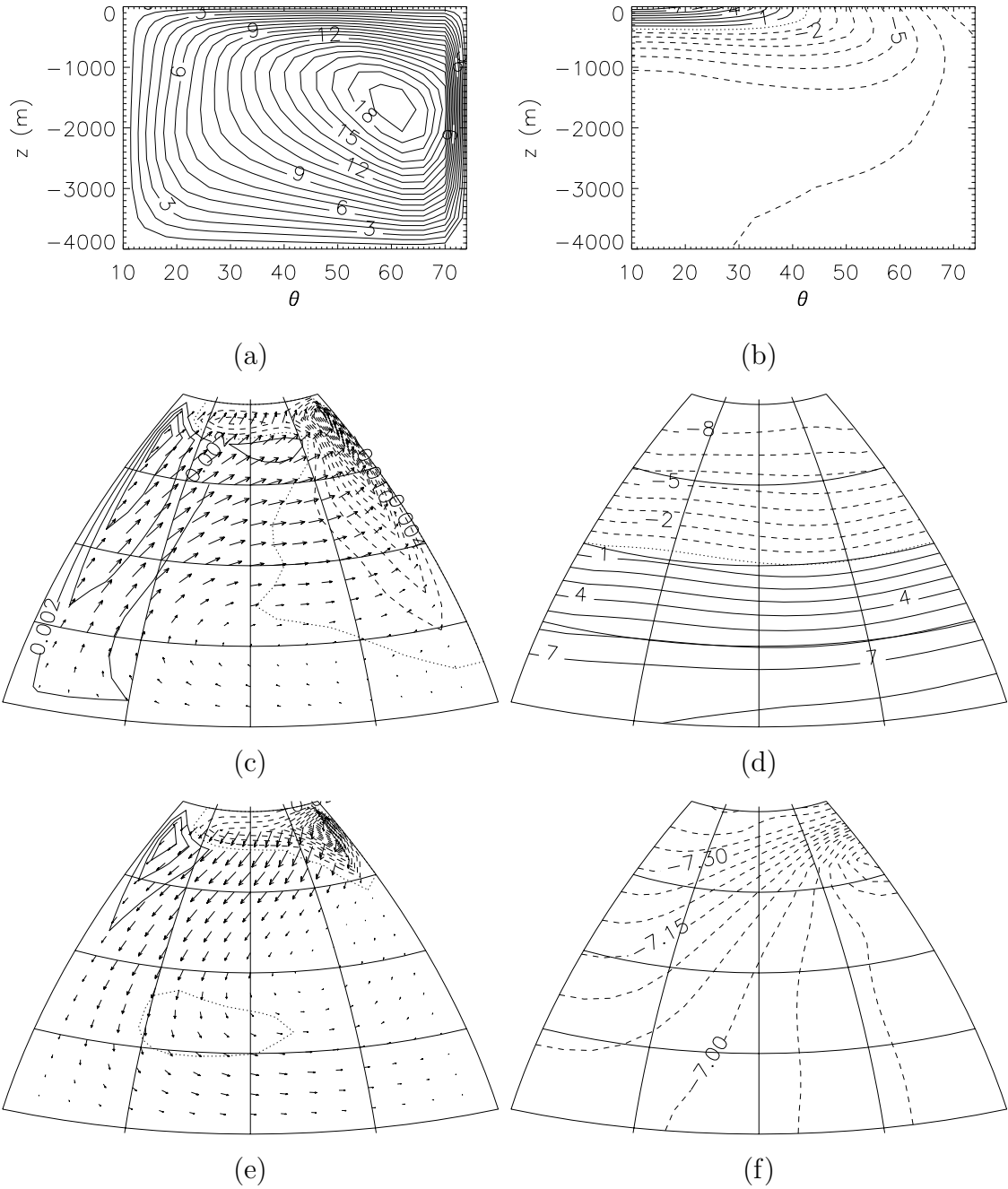


Figure 1: *Steady state solution. (a) Meridional overturning streamfunction (in Sverdrups). (b) Temperature (dimensionless) for a north-south vertical plane through the middle of the basin ($\phi = 318^\circ$). The dimensional temperature can be obtained from $T_* = 15.0 + T$. (c) Velocity (dimensionless) near the surface (at 41 m depth). In this plot, vectors indicate the horizontal velocity, (u, v) and contours represent the dimensionless vertical velocity, w . Solid lines represent upwelling (flow out of the plane), dashed lines downwelling (flow into the plane). The maximum dimensional horizontal velocity is $1.7 \cdot 10^{-2} \text{ ms}^{-1}$, the maximum amplitude of the vertical velocity is $1.8 \cdot 10^{-6} \text{ ms}^{-1}$ (downwelling). (d) Temperature near the surface. (e) Velocity at $z = -3200$ m. Maxima are $5.5 \cdot 10^{-3} \text{ ms}^{-1}$ for the horizontal and $2.1 \cdot 10^{-5} \text{ ms}^{-1}$ (downwelling) for the vertical velocity. (f) Temperature at $z = -3200$ m.*

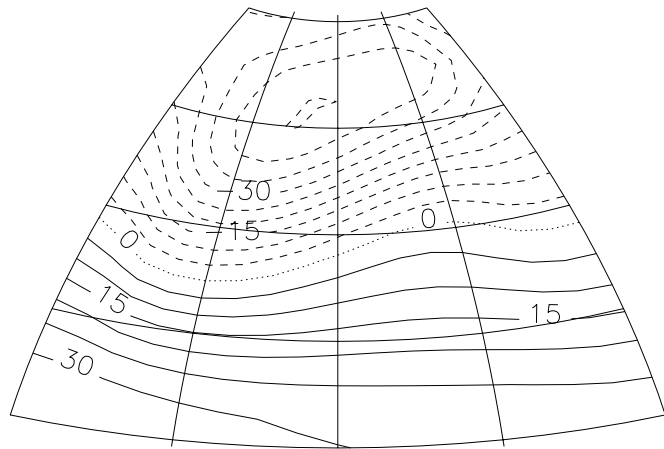


Figure 2: *Surface heat flux (in $W m^{-2}$), diagnosed from the solution obtained under restoring boundary conditions from Fig. 1. Solid lines represent heat gain from the atmosphere, dashed lines heat loss to the atmosphere.*

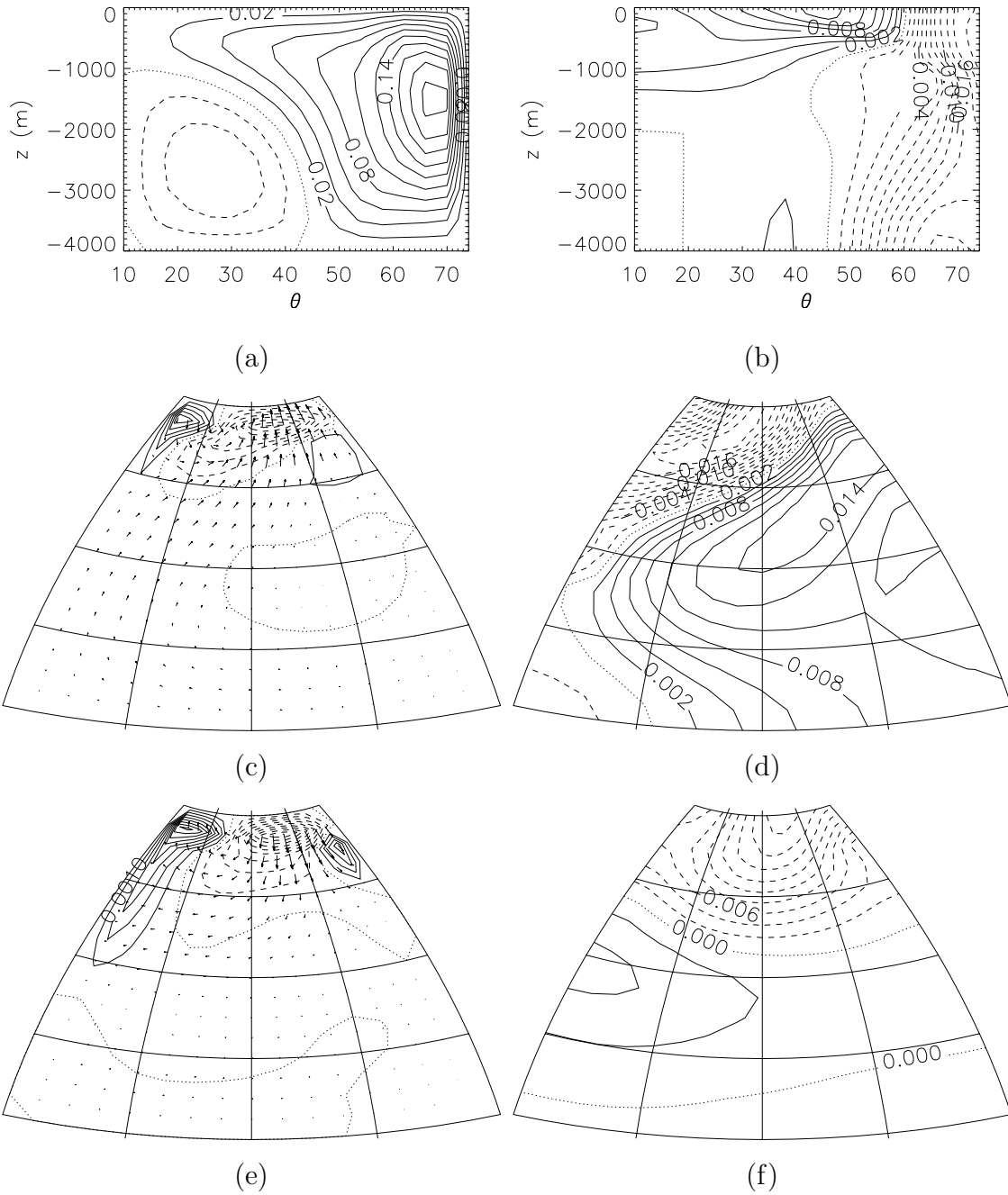


Figure 3: *Imaginary part of the eigenvector corresponding to the interdecadal mode ($\sigma_r = 0.0116$, $\sigma_i = 0.1966$). (a) Meridional overturning streamfunction. (b) Temperature for a north-south vertical plane through the middle of the basin. Solid lines represent positive values, dashed lines negative values. (c) Velocity at 41 m depth. (d) Temperature at 41 m depth. (e) Velocity at $z = -3200$ m. (f) Temperature at $z = -3200$ m. Note that the amplitude is arbitrary.*

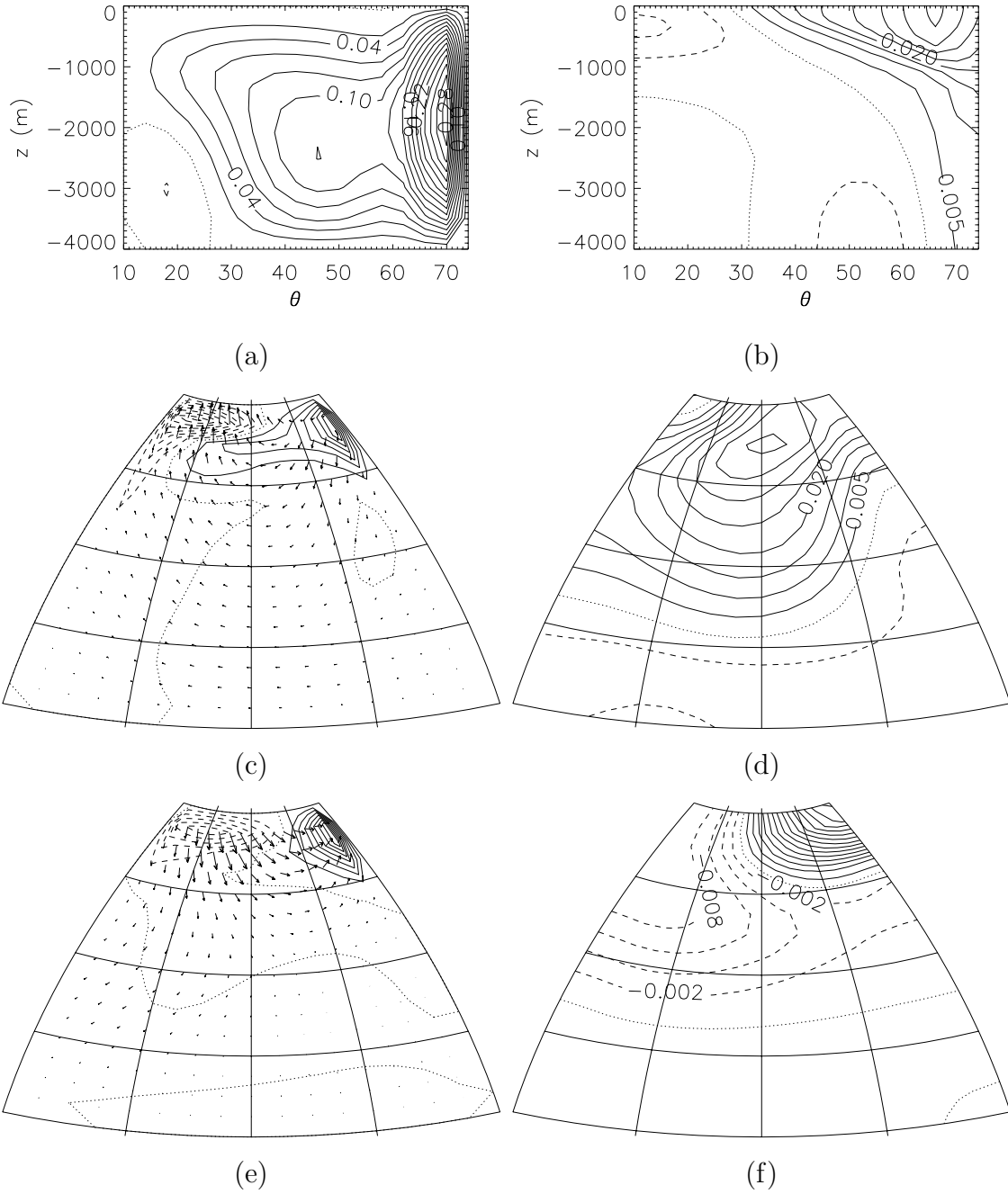


Figure 4: *Real part of the eigenvector corresponding to the interdecadal mode. (a) Meridional overturning streamfunction. (b) Temperature for a north-south vertical plane through the middle of the basin. (c) Velocity at 41 m depth. (d) Temperature at 41 m depth. (e) Velocity at $z = -3200$ m. (f) Temperature at $z = -3200$ m. Format is similar as in Fig. 3.*

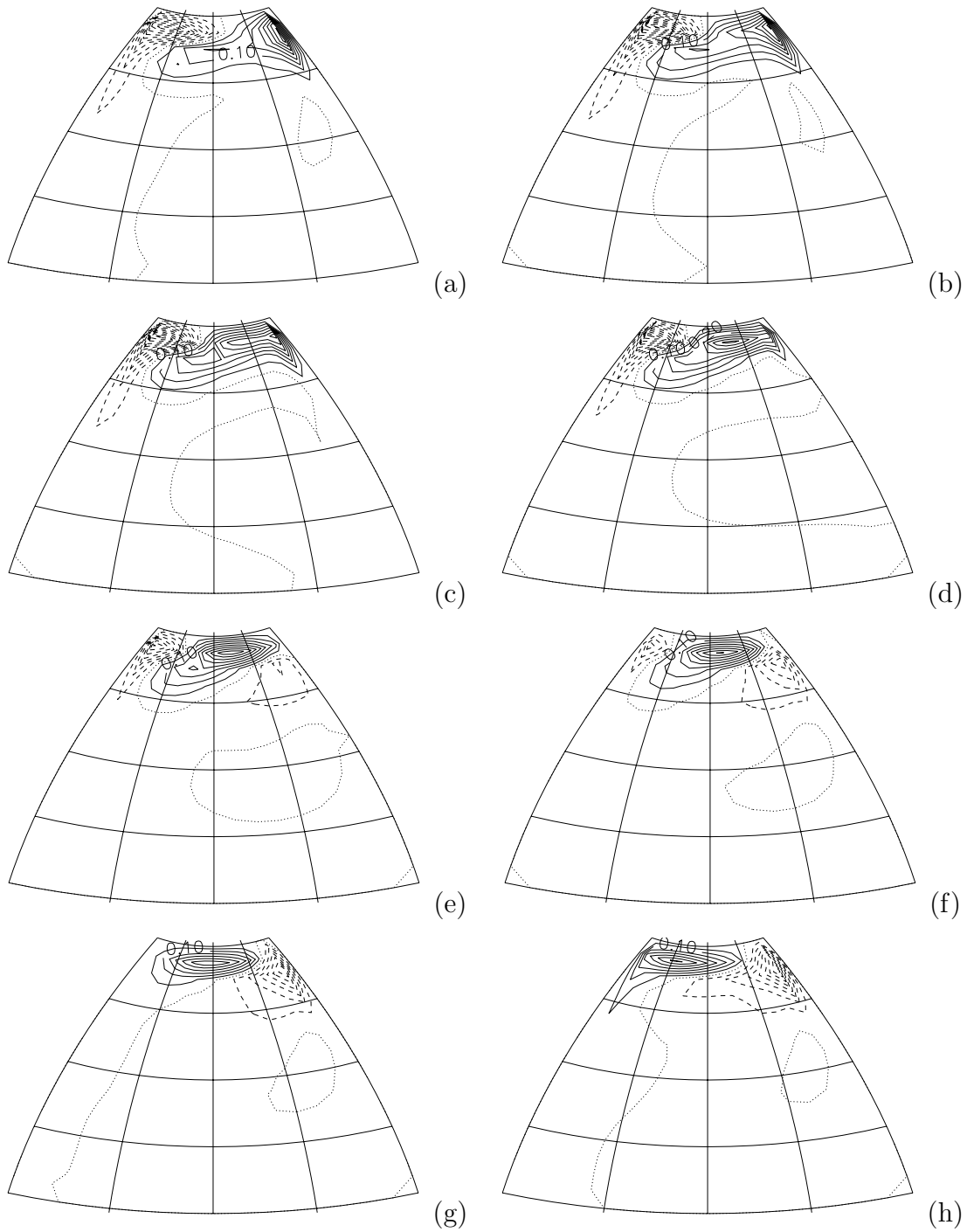


Figure 5: Vertical velocity perturbations at $z = -19$ m at (a) $t = 0$ yr, (b) $t = 4.1$ yr, (c) $t = 8.1$ yr, (d) $t = 12.2$ yr, (e) $t = 16.3$ yr, (f) $t = 20.3$ yr, (g) $t = 24.4$ yr and (h) $t = 28.4$ yr. The plots are $1/16$ th period apart. Time $t = 0$ corresponds with the real part of the eigenvector (Fig. 4).

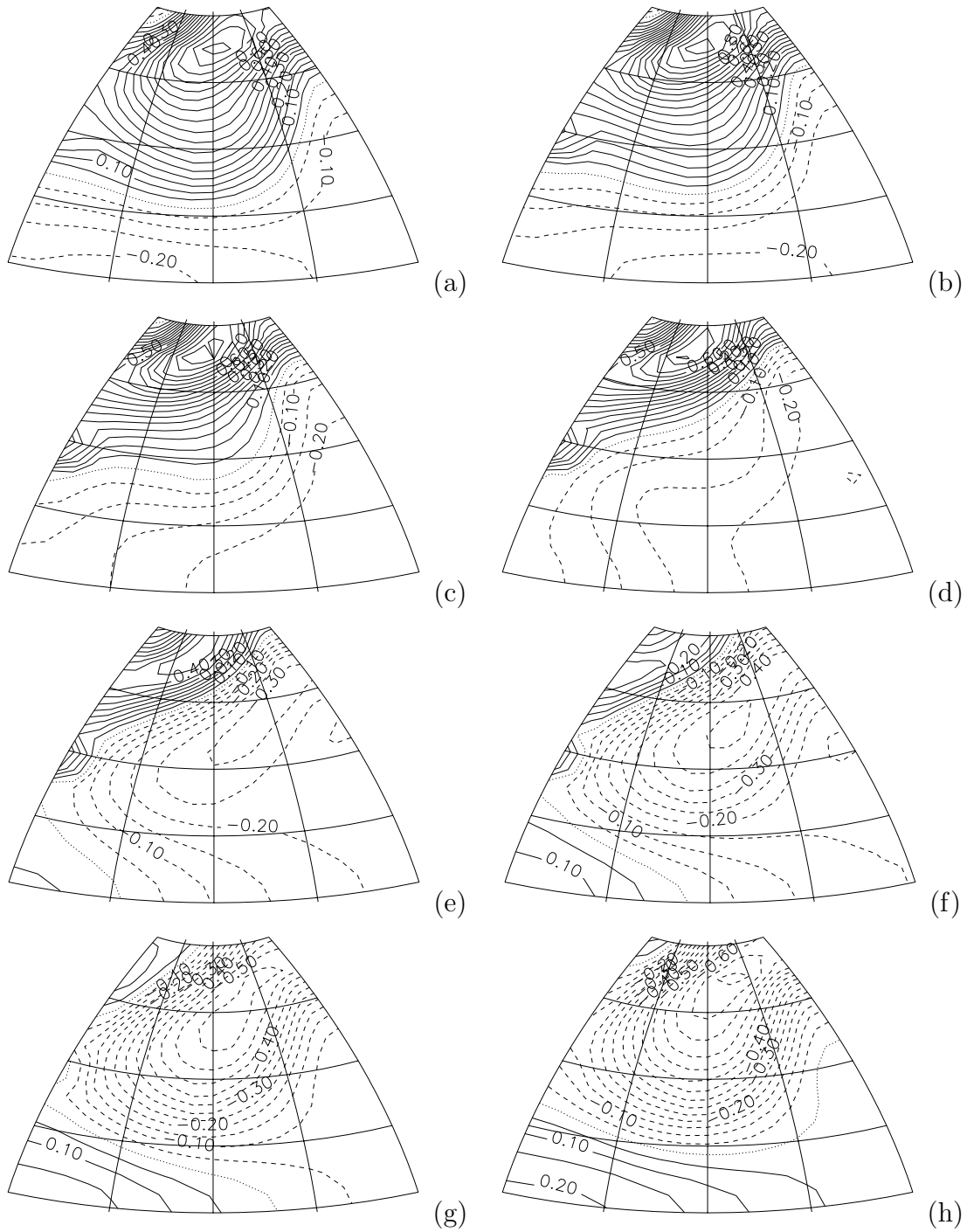


Figure 6: *Temperature perturbations at $z = -19$ m at (a) $t = 0$ yr, (b) $t = 4.1$ yr, (c) $t = 8.1$ yr, (d) $t = 12.2$ yr, (e) $t = 16.3$ yr, (f) $t = 20.3$ yr, (g) $t = 24.4$ yr and (h) $t = 28.4$ yr. The plots are $1/16$ th period apart.*

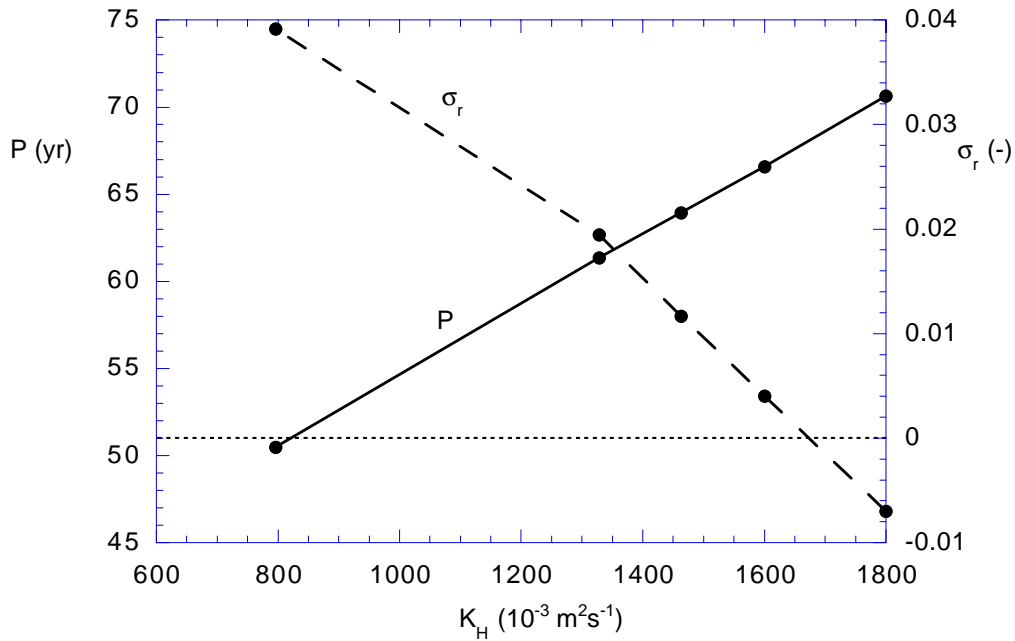


Figure 7: Period (in years, solid line) and growth rate (dimensionless, dashed line) as a function of the horizontal diffusivity K_H . The left vertical axis is for P , the right vertical axis for σ_r .

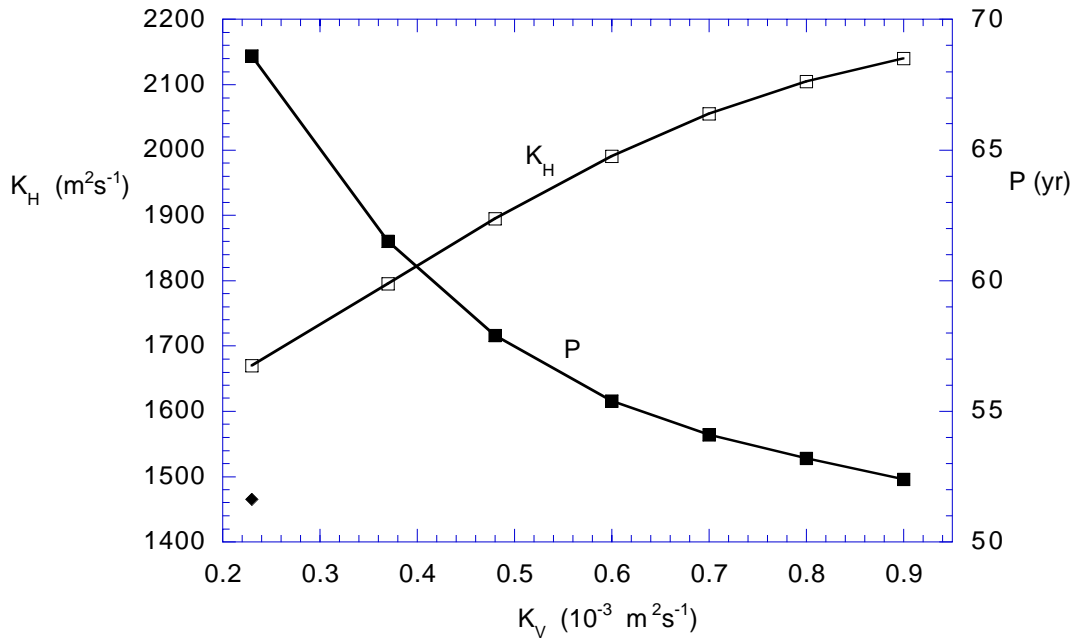


Figure 8: Regime diagram in the K_V - K_H plane for the interdecadal oscillation under prescribed flux conditions. Open squares denote the values of K_H for which the Hopf bifurcation occurs. Below this curve, the steady state is unstable and oscillatory behavior is found; above this curve it is stable. The filled squares give the oscillation period at Hopf bifurcation. The point labelled with a diamond indicates the standard values of K_V and K_H .

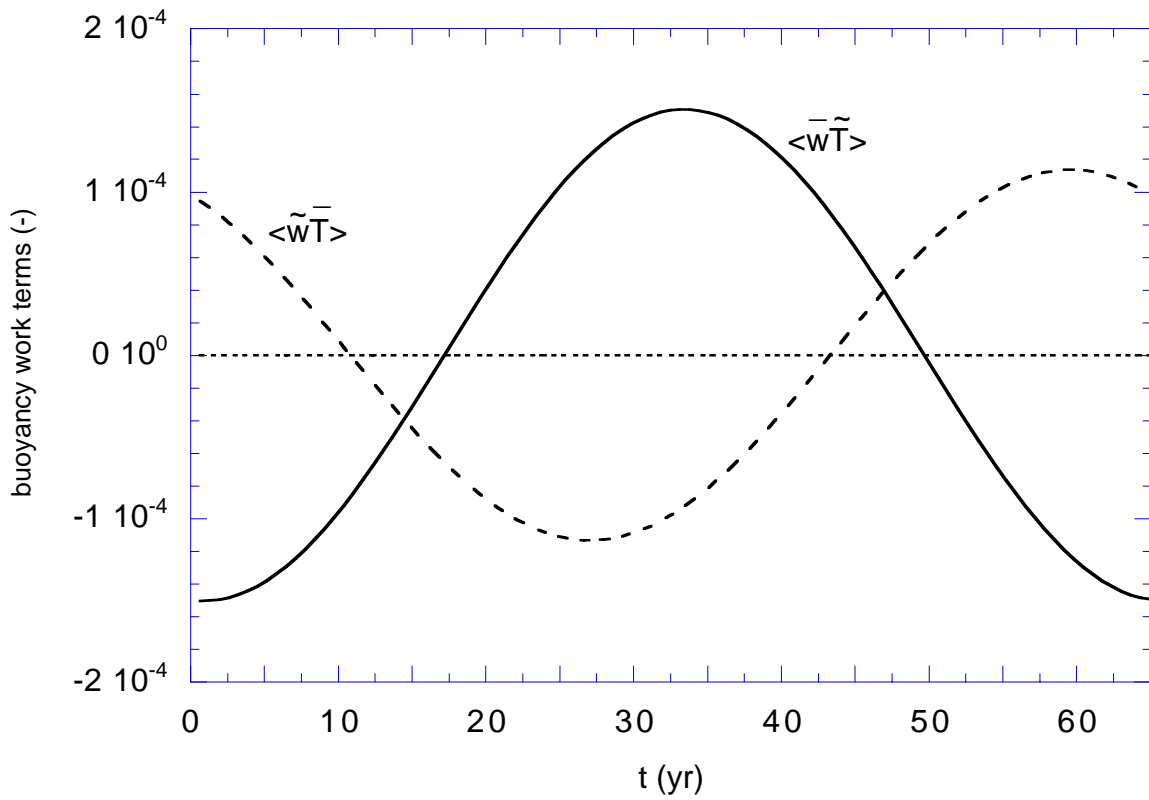


Figure 9: Dimensionless terms $\langle \bar{w}\tilde{T} \rangle$ (solid line) and $\langle \tilde{w}\bar{T} \rangle$ (dashed line) as a function of time (in years) for one oscillation cycle.

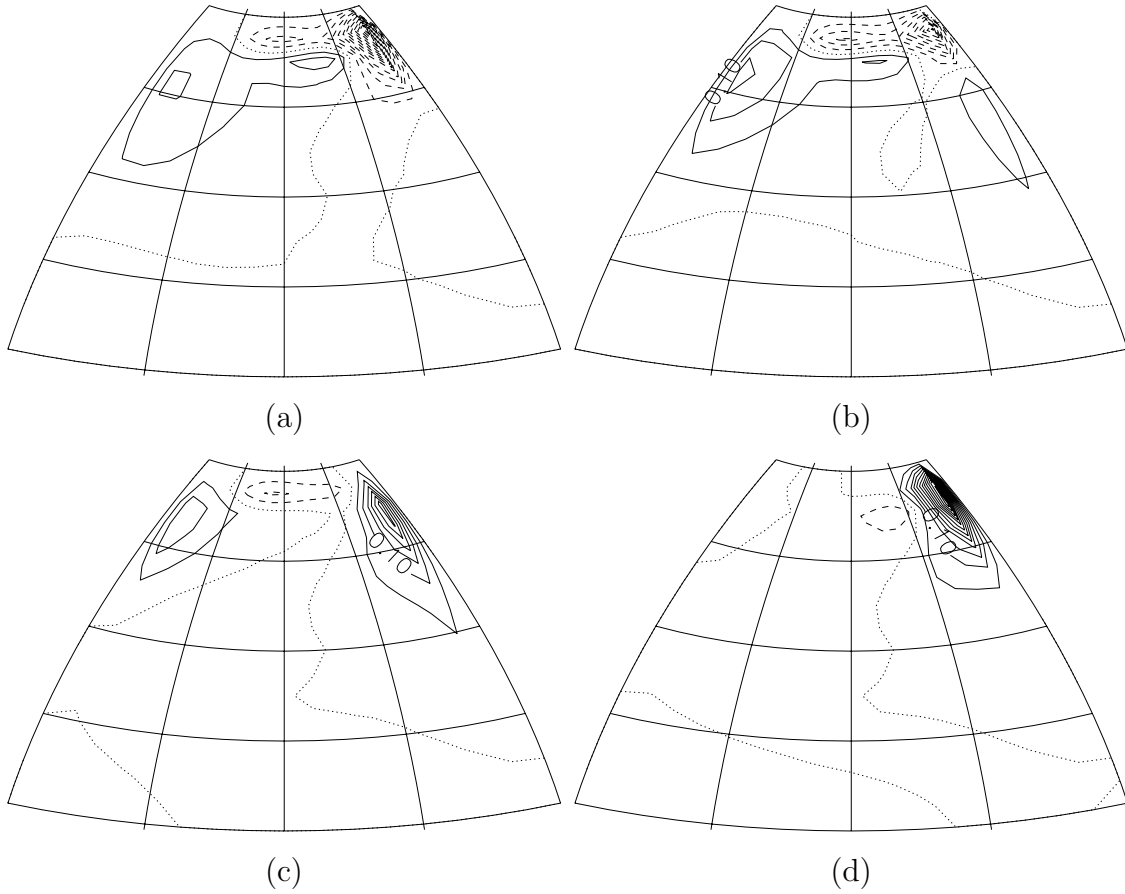


Figure 10: Pattern of $\bar{w}\tilde{T}$ at a horizontal slice just below the surface (at $z = -19$ m) at (a) $t = 0$ yr, (b) $t = 8.1$ yr, (c) $t = 16.3$ yr and (d) $t = 24.4$ yr. The time interval between the plots is $1/8$ th period and the variables are normalized with the maximum value.

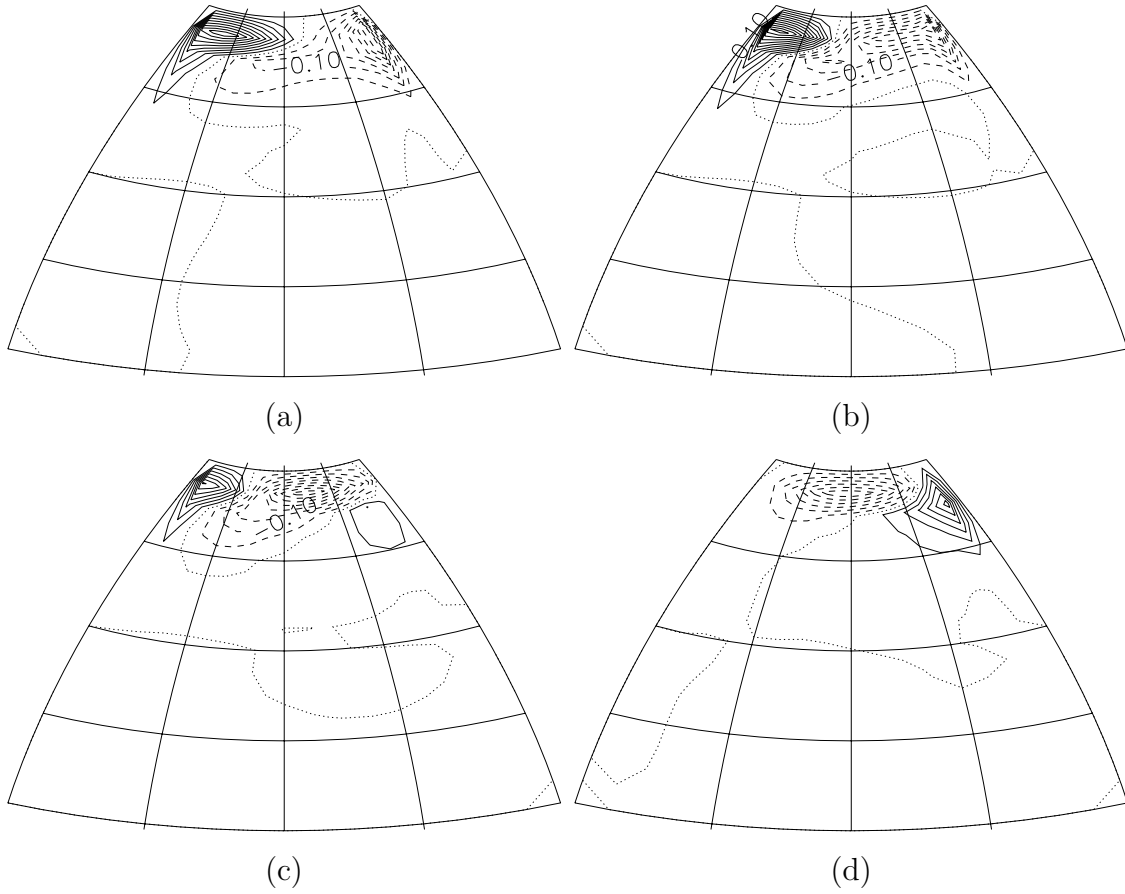


Figure 11: Pattern of $\tilde{w}\bar{T}$ at a horizontal slice just below the surface (at $z = -19$ m) at (a) $t = 0$ yr, (b) $t = 8.1$ yr, (c) $t = 16.3$ yr and (d) $t = 24.4$ yr. Format as in Fig. 10.

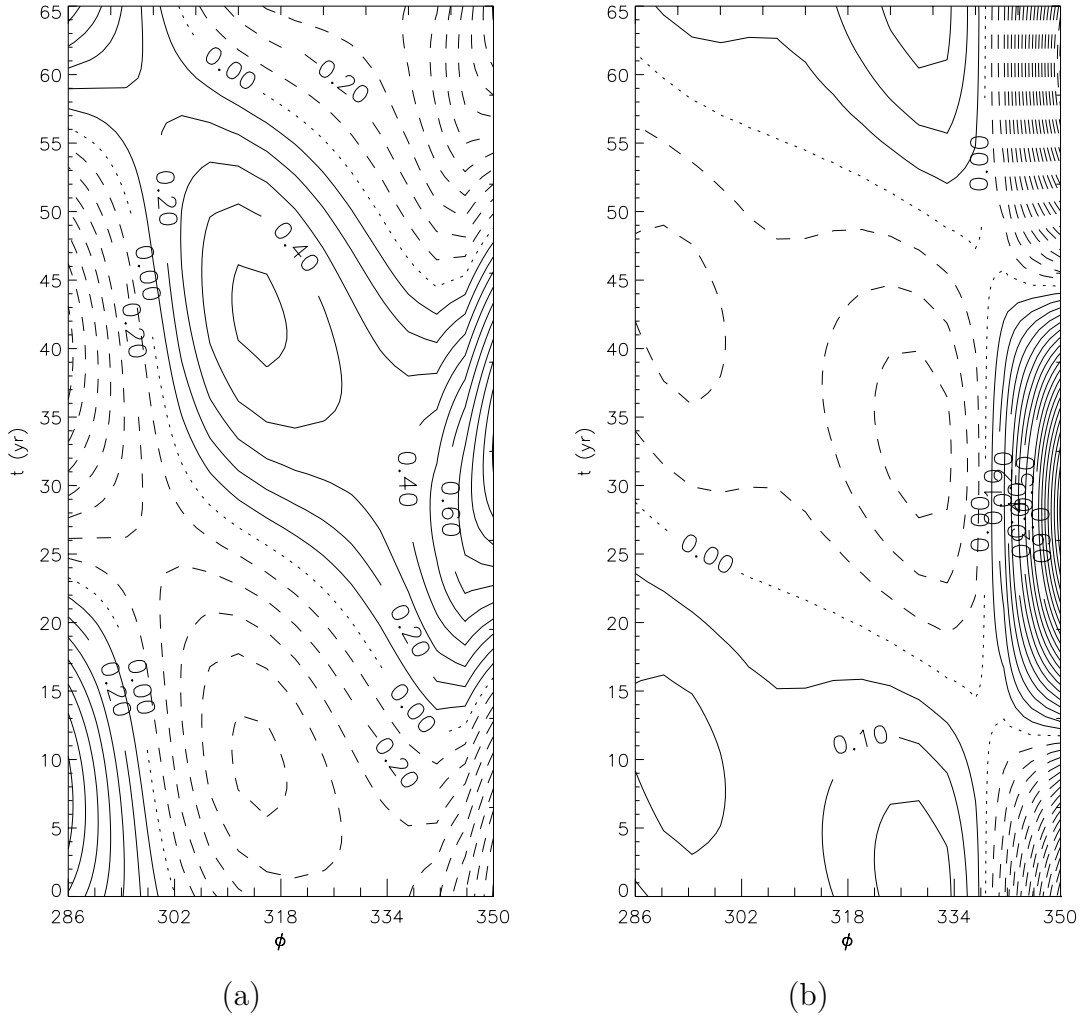


Figure 12: Hovmöller diagram of (a) $\tilde{w}\bar{T}$ and (b) $\tilde{w}\tilde{T}$ just below the surface (at $z = 19$ m depth) at $\theta = 68^\circ$. Both plots are normalized with their maximum values.

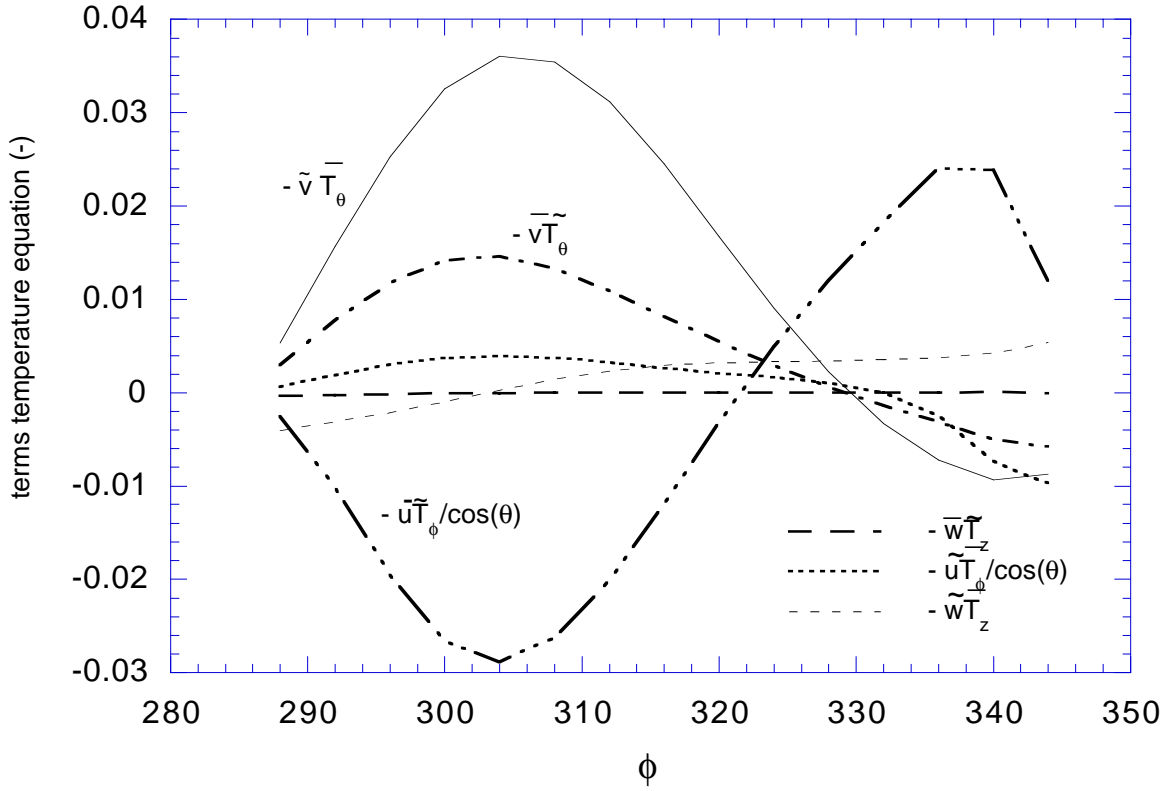


Figure 13: *Advective terms in the perturbation temperature equation as a function of longitude along the northern boundary ($\theta = 68^\circ$) at $t = 0$ yr.*

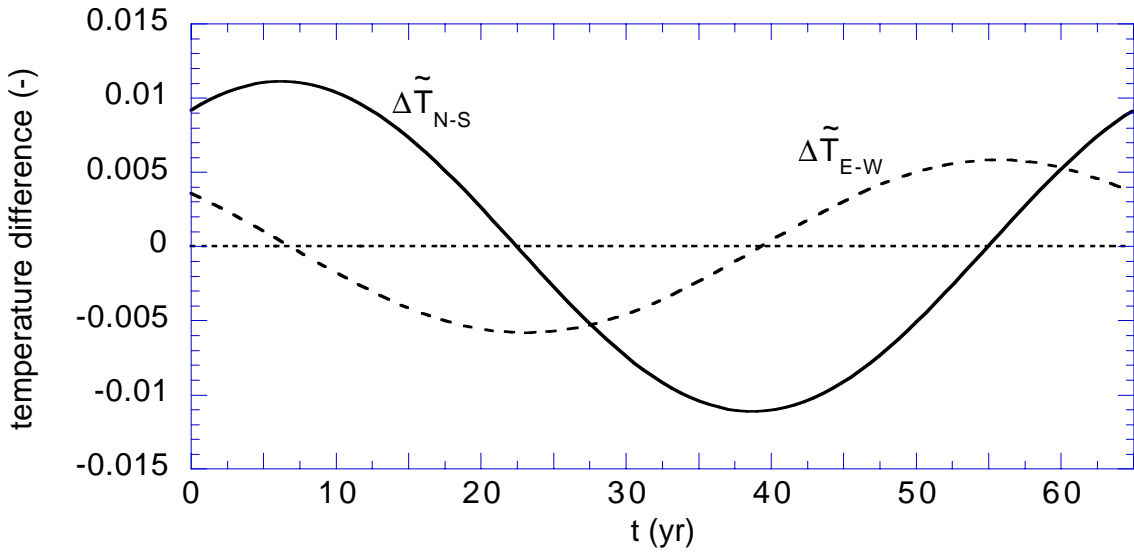


Figure 14: *Zonally averaged north-south temperature difference ΔT_{N-S} (solid line) and meridionally averaged east-west temperature difference ΔT_{E-W} (dotted line), both integrated over the upper 1525 m, as a function of time during one oscillation cycle.*

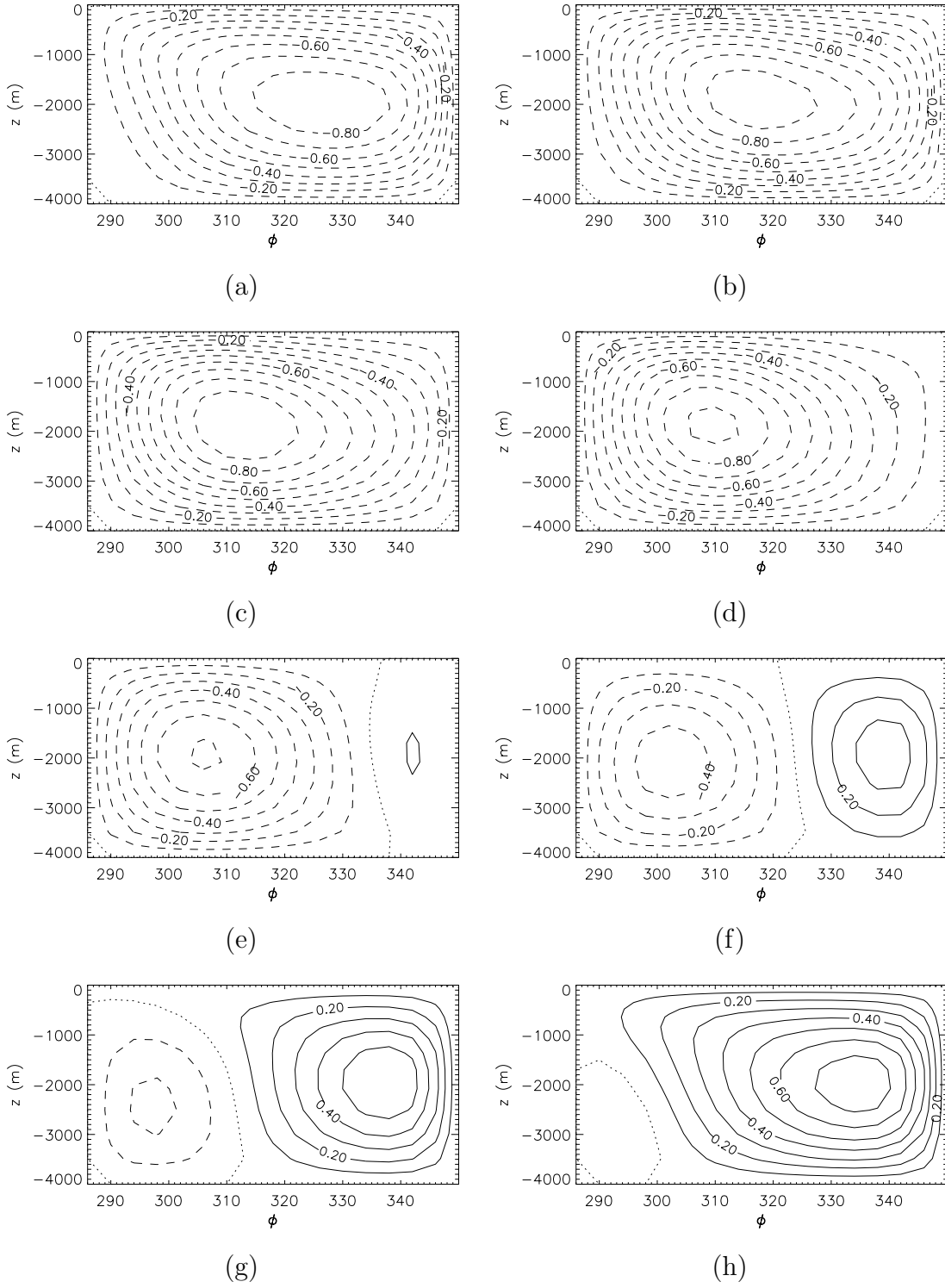


Figure 15: Zonal overturning perturbations at (a) $t = 0$ yr, (b) $t = 4.1$ yr, (c) $t = 8.1$ yr, (d) $t = 12.2$ yr, (e) $t = 16.3$ yr, (f) $t = 20.3$ yr, (g) $t = 24.4$ yr and (h) $t = 28.4$ yr. The plots are $1/16$ th period apart.

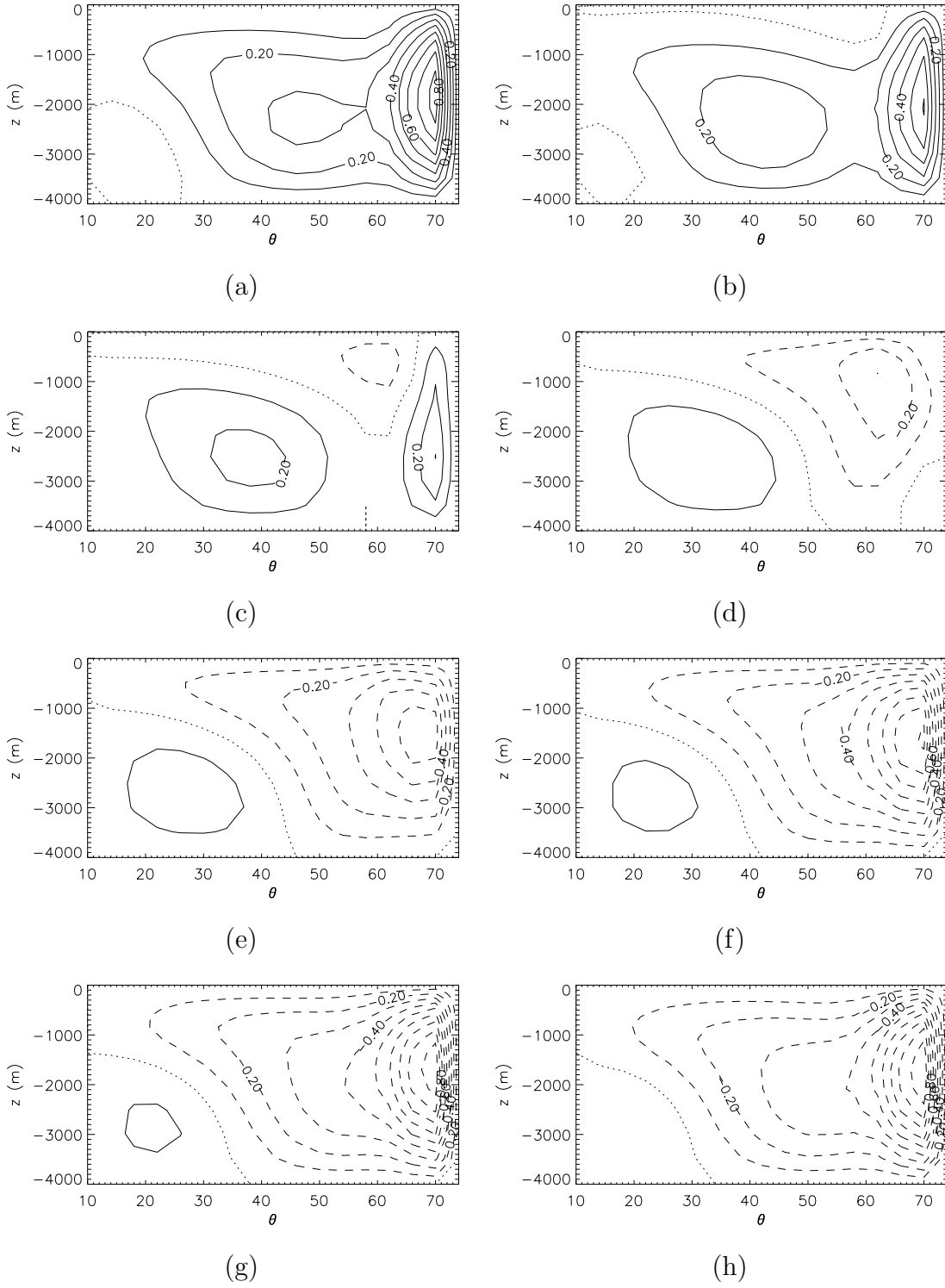


Figure 16: *Meridional overturning perturbations at (a) $t = 0$ yr, (b) $t = 4.1$ yr, (c) $t = 8.1$ yr, (d) $t = 12.2$ yr, (e) $t = 16.3$ yr, (f) $t = 20.3$ yr, (g) $t = 24.4$ yr and (h) $t = 28.4$ yr. The plots are $1/16$ th period apart.*

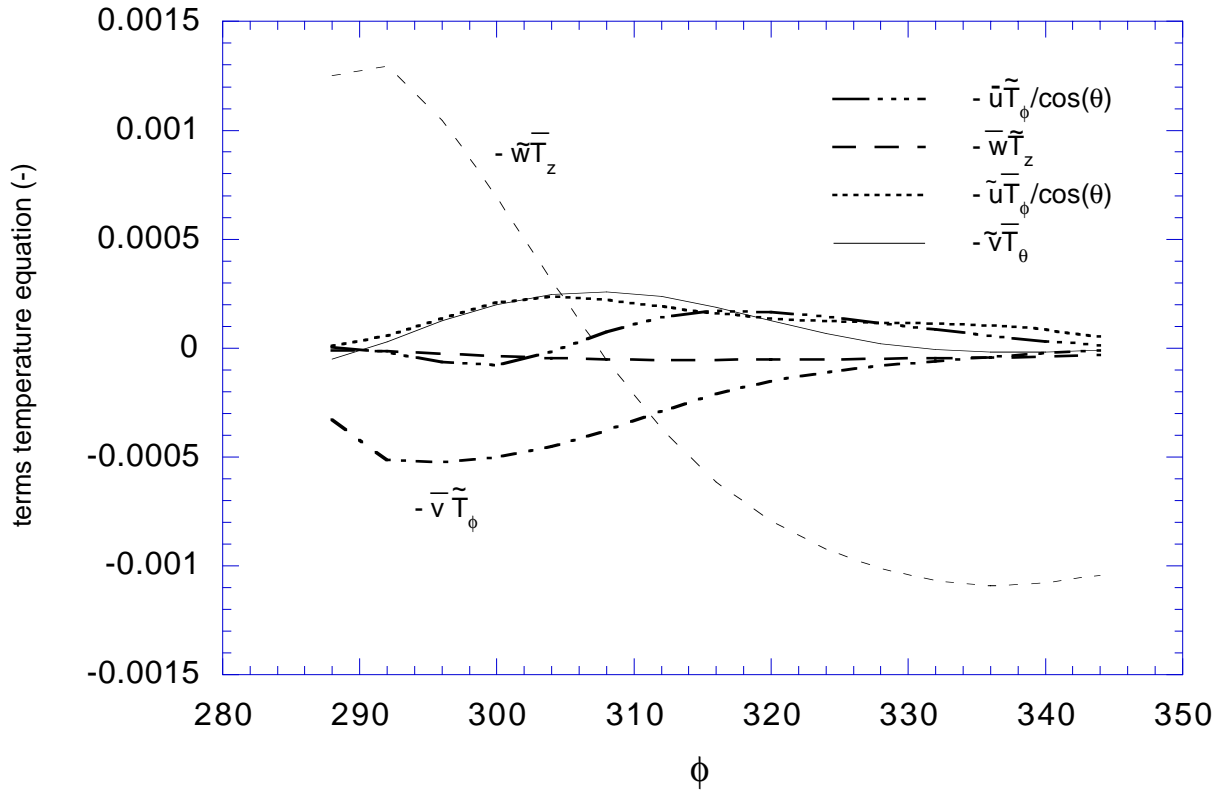


Figure 17: *Advective terms in the perturbation temperature equation as a function of longitude along the southern boundary ($\theta = 16^\circ$) at $t = 0$ yr.*

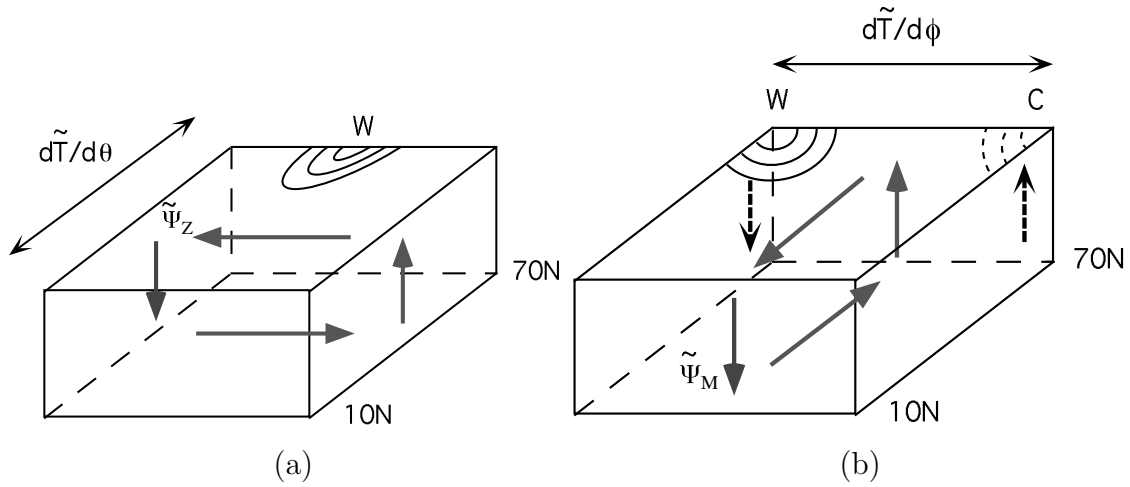


Figure 18: *Schematic diagram of the oscillation mechanism: a warm anomaly in the north-central part of the basin causes a positive meridional perturbation temperature gradient (a), which induces a negative zonal overturning perturbation. The anomalous upwelling and downwelling associated with this zonal overturning are consistent with westward propagation of the warm anomaly, while a cold anomaly appears in the east (b). Due to the westward propagation of the warm anomaly, the east-west temperature difference decreases and becomes negative, inducing a negative meridional overturning perturbation. The resulting upwelling and downwelling perturbations along the northern and southern boundary reduce the north-south perturbation temperature difference, causing the zonal overturning perturbation to change sign and the second half of the oscillation starts.*