Tellus (2005), 57A, 100–115
Printed in UK.

TELLUS

Statistical hydrodynamics of the thermohaline circulation in a two-dimensional model

By GREGORY L. EYINK*, *Applied Mathematics and Statistics, The Johns Hopkins University, Baltimore, MD 21218, USA*

(Manuscript received 25 July 2003; in final form 19 May 2004)

ABSTRACT

We study a simple conceptual model of the climate-mean ocean thermohaline circulation, the two-dimensional Boussinesq model of Thual and McWilliams, generalized to include a random surface salinity flux, representing synoptic variability in freshwater sources. In a small aspect ratio limit previously considered by Cessi and Young for the deterministic model, we show that the dynamics reduces to a stochastically-driven one-dimensional Cahn–Hilliard equation (with non-conserving boundary conditions). If the random force is space–time white noise, a non-linear fluctuation–dissipation relation permits an exact determination of the statistical equilibrium distribution. The effect of the spatial distribution of the systematic (non-random) salinity flux on the stability of multiple equilibria is studied in detail. We determine a flux distribution for which the model caricatures the circulation states in the Atlantic during the last glacial period. For this case, we determine the multimodal equilibrium distributions associated to the multiple states of circulation. We also study the dynamics of random state-transitions by a simulation of the stochastic Cahn–Hilliard equation. Finally, with a weak, periodic modulation in the freshwater forcing of the correct frequency, the model is shown to exhibit stochastic resonance, i.e. the transitions synchronize with the periodic force. The stochastic resonant frequency is calculated analytically in the model by the Langer formula for the mean residence time and compared with numerical solutions.

1. Introduction

It has been realized for some time that the oceanic circulation may have multiple states of stable equilibrium. This was first pointed out by Stommel (1961) in a study of a simple two-box model, with one equatorial box which was heated and lost freshwater and one polar box which was cooled and received freshwater. It was in this same work that Stommel coined the term 'thermohaline circulation' for the meridional ocean circulations apparently driven by temperature and salinity differences between the equatorial and polar regions. In the two-box model, one circulation was thermally dominated, with cold, fresh water sinking at the pole and another was salinity dominated, with warm, salty water sinking at the equator. Since this pioneering paper, similar multiple equilibrium states have been observed in a variety of more complicated box models (Welander, 1988), in GCMs (Bryan, 1986; Manabe and Stouffer, 1988), and in climate models of intermediate complexity (Ganopolski et al., 1998).

These multiple equilibria seem to have considerable bearing on the subject of ice-age cycles and abrupt climate change. There is accumulated evidence that the ocean circulation has existed in alternate choices of these stable equilibria during the past, and transitions between them have been correlated with rapid warming and cooling events in the climate record (Stocker, 2000; Clark et al., 2002; Rahmstorf, 2002). Although the circulation states are locally stable, transient jumps between them may be triggered by perturbations in the salinity or temperature forcing. This set of climate problems has motivated the study of statistical mechanical models of the thermohaline circulation, in which rapid, synoptic-scale variations in the forcings are represented by a random noise. Such studies have been carried out in box models (Cessi, 1994; Bryan and Hansen, 1995; Timmermann and Lohmann, 2000; Vélez-Belchí et al., 2001) and also in more realistic climate models (Ganopolski and Rahmstorf, 2002).

A principal focus of Vélez-Belchí et al. (2001) and Ganopolski and Rahmstorf (2002) was to investigate the role of 'stochastic resonance' in initiating transitions between ocean circulation states. The phenomenon of stochastic resonance was first proposed as an explanation of Quaternary 100K glacial cycles by Benzi et al. (1981, 1982). According to their theory, weak periodic perturbations in solar energy flux associated with changes in the Earth's orbital eccentricity were enhanced by random noise. While this explanation of ice-age cycles remains open to question, the phenomenon of stochastic resonance has been conclusively established in a wide variety of biological,

*Corresponding author.
e-mail: eyink@ams.jhu.edu

neurophysiological, electronic, and other systems (Gammaitoni et al., 1998). The role of stochastic resonance in climate change was revived in the context of rapid warming events during the most recent glacial period by Alley et al. (2001a,b). These so-called Dansgaard–Oeschger (DO) oscillations occurred with a rough periodicity of 1450 yr during the last ice age but not during the current interglacial (Hinnov et al., 2002). It was shown by an analysis of temperature proxies from ice core data in Alley et al. (2001a,b) that the statistical characteristics of these DO events are consistent with an explanation by stochastic resonance. The simulations of the stochastic models of the thermohaline circulation in Vélez-Belchí et al. (2001) and Ganopolski and Rahmstorf (2002) have lent some further support to this interpretation by demonstrating stochastic resonance in the numerical solutions, with results similar to those observed in the ice core records.

In the present paper we introduce and study a simple conceptual model for statistical mechanics of the thermohaline circulation. It is a straightforward generalization of a deterministic model of the thermohaline circulation studied previously by several authors (Marotzke et al., 1988; Thual and McWilliams, 1992; Quon and Ghil, 1992; Cessi and Young, 1992; Dijkstra and Molemaker, 1997), a two-dimensional (2-D) Boussinesq model of thermal and saline convection in the vertical-meridional plane. We have extended this model to include a stochastic component to the surface salinity flux, similar to what has been done in the box models (Cessi, 1994; Bryan and Hansen, 1995; Timmermann and Lohmann, 2000; Vélez-Belchí et al., 2001) and climate models of intermediate complexity (Ganopolski and Rahmstorf, 2002). Like the box models, our model makes no pretense to detailed verisimilitude but is intended to be useful for conceptual purposes. However, unlike multibox models, the present partial differential equation (PDE) model is analytically tractable and permits the derivation of many exact results. This was already demonstrated for the deterministic case by Cessi and Young (1992), where many analytical results on dynamics, multiple equilibria and their stability were obtained in a small aspect-ratio limit. We shall follow that work closely for the stochastic version of the model and, in the same limit, obtain a variety of analytical results. These include equations for the dynamical evolutions, for the statistical steady-state distribution, and for mean residence times in the domains of stable equilibria. In particular, the phenomenon of stochastic resonance may be demonstrated to exist in the model with an analytical calculation of the resonant frequency. Such exact results should prove useful in further exploration of the statistical physics of the ocean circulation. One of our own immediate motivations for developing the model is to use it as a test case for statistical data assimilation and inverse modeling schemes. The analytical results demonstrated in this work will be extremely useful as benchmarks for testing approximate estimation procedures.

Another advantage of the present conceptual model over simpler box models, already pointed out in Cessi and Young (1992), is that it is possible to extend it easily to include additional physical effects. There is some evidence from bifurcation studies (Weijer and Dijkstra, 2001) that refinements such as Earth rotation, three-dimensionality (zonal variations), and wind stress forcing will not change the qualitative character of our results. Of various effects, we regard turbulence from unresolved scales as the most crucial to include and to understand in the near future. A number of recent works have pointed out the necessity of external energy sources to generate vertical (or diapycnal) turbulent mixing of density in a stably-stratified ocean (Munk and Wunsch, 1998; Huang, 1999; Paparella and Young, 2002). These papers extend and clarify the old thermodynamical arguments of Sandström (1908) against horizontal convection, i.e. convection driven by heat sources at the same (or greater) height than the sources of cooling. It has been shown that horizontal convection cannot self-generate the turbulence necessary for significant mixing of density across pycnal surfaces (Paparella and Young, 2002). We shall consider in some detail below the limitations imposed on our model by our neglect of explicit turbulent processes and the energy sources required to maintain them. We shall argue that the results obtained here can still be regarded as qualitatively correct, with an appropriate interpretation of vertical mixing rates in our model. However, this is certainly the key area for further extensions and physical explorations of the model assumptions.

The detailed contents of this paper are as follows. In Section 2 we introduce the model equations and discuss their physical interpretation and limitations. In Section 3 we introduce the small aspect-ratio limit, which is considered in the remainder of the paper. In Section 4 we discuss the relation of stability of deterministic fixed points with the spatial distribution of surface salinity flux. In Section 5 we derive the statistical equilibrium distribution and explain its features in terms of the dynamical time series. In Section 6 we discuss stochastic resonance. Section 7 contains our concluding remarks.

2. The model and its interpretation

We consider the 2-D Boussinesq model of Thual and McWilliams (1992) in a vertical plane, $-\ell < y < \ell$, $0 \leq z \leq d$, where 2ℓ is the pole-to-pole distance and d is the depth. The analysis uses a stream function formulation with $v = -\partial_z \psi$, $w = \partial_y \psi$. The dynamical equations for stream function ψ, temperature T, and salinity S are

$$\partial_t \nabla^2 \psi + J(\psi, \nabla^2 \psi) = g(\alpha_T \partial_y T - \alpha_S \partial_y S) + \nu \nabla^4 \psi \tag{2.1}$$

$$\partial_t T + J(\psi, T) = \kappa_T \nabla^2 T \tag{2.2}$$

$$\partial_t S + J(\psi, S) = \kappa_S \nabla^2 S \tag{2.3}$$

where the Jacobian $J(A, B) = (\partial_y A)(\partial_z B) - (\partial_z A)(\partial_y B)$. As discussed further in Thual and McWilliams (1992), these equations are supposed to be interpreted as time- and zonally-averaged, relevant for the climate time-scale and the global space-scale meridional circulation. The viscosity ν and the conductivities κ_T and κ_S are supposed to represent turbulent or 'eddy' diffusivities. The buoyancy force $g(\alpha_{\mathrm{T}}T - \alpha_{\mathrm{S}}S)$ assumes a linear equation of state, where α_{T} and α_{S} are thermal and saline expansion coefficients, respectively.

Following Thual and McWilliams (1992), we adopt no-stress boundary conditions on the velocity, which correspond to free-slip conditions for the stream function:

$$\psi = 0, \qquad \partial_n^2 \psi = 0. \tag{2.4}$$

Other boundary conditions for the velocity, such as no-slip, might be used as well, with similar results. The boundary conditions on temperature and salinity are

$$T(y, d) = \Delta T \theta(y), \qquad \partial_z S(y, d) = \Delta S F(y)/d \tag{2.5}$$

$$\partial_z T(y, 0) = 0, \qquad \partial_z S(y, 0) = 0 \tag{2.6}$$

$$\partial_y T(\pm\ell, z) = 0, \qquad \partial_y S(\pm\ell, z) = 0. \tag{2.7}$$

The last four equations indicate that there is zero net flux of heat and salt at the ocean bottom and at the poles. The surface boundary conditions for temperature and salinity are distinct, because the ocean temperature at the surface tends to adjust to an equilibrium value but the salt flux at the surface is imposed independently of the surface salinity (Bryan, 1986). A consequence of the salinity flux condition is that the dynamics is invariant under a shift $S \to S + S_0$, for any constant S_0. Thus, the equations are insensitive to the net salt content of the oceans but depend only upon salinity differences.

The shape functions $\theta(y)$ and $F(y)$ represent the meridional distribution of the surface forcings of temperature and salinity. Thual and McWilliams took both $\theta(y) = F(y) = \cos(\pi y/\ell)$, because the equator is warm and the poles cool, and evaporation gives a net positive salt flux near the equator while precipitation, continental runoff, and glacial melting in the high latitudes give a net positive freshwater flux. We shall keep the temperature forcing the same as in Thual and McWilliams (1992)

$$\theta(y) = \cos(\pi y/\ell) \tag{2.8}$$

but generalize the salinity forcing to be time-dependent

$$\Delta S F(y)/d \to (\Delta S/d)\bar{F}(y, t) + \Sigma_0 \widetilde{F}(y, t) \tag{2.9}$$

where $\bar{F}(y, t)$ is the systematic salinity flux, specified later, and $\widetilde{F}(y, t)$ is a random salinity flux. The latter is the fundamentally new feature of our model compared with that of Thual and McWilliams (1992). It represents variability in the surface salinity (freshwater) flux on short, synoptic time-scales, similar to what was considered in Cessi (1994), Bryan and Hansen (1995), Timmermann and Lohmann (2000), Vélez-Belchí et al. (2001) and Ganopolski and Rahmstorf (2002). This random component has strength determined by the constant Σ_0, multiplying a zero-mean, Gaussian, space–time white-noise field with covariance

$$\langle \widetilde{F}(y, t)\widetilde{F}(y', t')\rangle = \delta(y - y')\delta(t - t'). \tag{2.10}$$

As discussed further below, we expect that our main results are qualitatively independent of this choice, but it is analytically convenient.

We now discuss some limitations of our model. It seems to realize the paradigm of ocean circulation as a heat engine, driven by meridional inhomogeneities in temperature and salinity forcing. We shall see below that the model indeed develops equilibrium solutions with global circulations. As pointed out long ago by Jeffreys (1925) in his response to Sandström (1908), this is permitted by the existence of vertical diffusion of heat and salt. However, the magnitudes of κ_T and κ_S required to maintain the abyssal density stratification against the global upwelling implied by observed rates of deep-water formation are of the order of 1 $\mathrm{cm}^2\ \mathrm{s}^{-1}$ (Munk and Wunsch, 1998). This is about three orders of magnitude larger than the molecular diffusivities of water (Batchelor, 1967, Appendix 1), consistent with the interpretation of κ_T and κ_S as eddy diffusivities produced by turbulent motions at unresolved scales. Therefore, one may accept the model as a crude, conceptual description of the ocean circulation, with the appropriate interpretation of the transport coefficients as turbulent diffusivities.

However, this interpretation requires some caution. To quote from a well-known turbulence textbook (Tennekes and Lumley, 1972, Section 1.4): 'If we use an effective diffusivity, we tend to treat turbulence as a property of a fluid rather than as a property of a flow. Conceptually, this is a very dangerous approach.' In particular, molecular diffusivities will not change if the ocean circulation makes a transition from one stable equilibrium to another, but turbulent diffusivities very well may because the turbulence-generating mechanisms (such as stirring by flow over ocean ridges and bottom topography, wind mixing, etc.) will depend upon the details of the circulation itself. Another caveat is that horizontal convection – that is, convection due to surface sources of heat or salt alone with only horizontal gradients – is incapable of generating by itself the required level of turbulence. This follows from the 'antiturbulence' theorem of Paparella and Young (2002), which can be regarded as an improved version of Sandström's theorem. As discussed in several recent works (Munk and Wunsch, 1998; Huang, 1999), the generation of turbulent diffusion against the abyssal stable stratification of density requires external sources of energy. Thus, the paradigm of the ocean meridional circulation as a 'mechanical conveyor belt' of heat and salt seems to have merit at least equal to that of circulation as a 'heat engine' (Wunsch, 2002).

The model considered may be easily extended to take account of these remarks. Additional equations may be added to represent the small-scale turbulent kinetic energy K and energy dissipation ε (as in Huang, 1999, Appendix A). These equations are not closed, but even relatively crude closures of $K - \varepsilon$ type would be useful to explore the essential physics. Necessary stirring sources can be added to maintain the mean flow against turbulent dissipation, which feeds energy into the small-scale turbulence. For example, a time-dependent gravitational potential can be added to represent tidal forcing, which is one mechanical energy source of probable importance (Munk and Wunsch, 1998). It would also be easy to replace the zero-stress boundary conditions (2.4) at the surface with conditions that contain a mean wind stress from the atmospheric dynamics. Our zero stress conditions assume a lack of persistent wind-stress forcing of the meridional flow at the ocean upper surface on climate time-scales. Indeed, the meridional component of the mean wind contributes negligibly to the work by the wind stress on the oceanic general circulation, which is dominated instead by the work of the mean zonal wind in the Southern Ocean (Wunsch, 1998). However, this work is also an important possible source of mechanical energy to maintain vertical mixing (Munk and Wunsch, 1998). Geothermal heating through the sea floor is another pertinent energy source, which is of smaller magnitude but which might be particularly important as a generator of vertical convection (Wunsch, 1998). This effect can be represented in the model by modifying the zero heat-flux boundary condition (2.6) at the ocean bottom to include a heat source there.

As we shall argue in detail in the following, we expect that most of our results below will not change qualitatively by refinements of the type considered above. However, it would be interesting to verify this by an explicit inclusion into the model of such effects. We hope to return to this in future work.

3. Small aspect-ratio expansion

We shall now turn to an analytical treatment of the model, asymptotically in the limit of small aspect ratio

$$\epsilon \equiv \frac{\pi d}{\ell} \ll 1. \tag{3.1}$$

Such a limit is easy to motivate for the ocean general circulation, where ℓ is of the order of 100–1000 km and d is 1–5 km. Our treatment follows closely that of Cessi and Young (1992) for the deterministic case. We introduce non-dimensionalized variables (indicated temporarily by $\hat{}$) as

$$\begin{aligned} (y, z) &= d\left(\frac{\hat{y}}{\epsilon}, \hat{z}\right), \qquad t = \frac{d^2}{\kappa_T}\hat{t}, \qquad \psi = \frac{\kappa_T}{\epsilon}\hat{\psi}, \\ T &= \frac{\nu\kappa_T}{g\alpha_T d^3\epsilon^2}\hat{T}, \qquad S = \frac{\nu\kappa_T}{g\alpha_S d^3\epsilon^2}\hat{S}. \end{aligned} \tag{3.2}$$

In these new variables (with $\hat{}$ now removed) the flow domain becomes $0 \le z \le 1$, $-\pi < y < \pi$, and the equations are

$$P^{-1}[\partial_t\zeta + J(\psi, \zeta)] = \partial_y T - \partial_y S + \left(\partial_z^2 + \epsilon^2\partial_y^2\right)\zeta \tag{3.3}$$

$$\partial_t T + J(\psi, T) = \left(\partial_z^2 + \epsilon^2\partial_y^2\right) T \tag{3.4}$$

$$L^{-1}[\partial_t S + J(\psi, S)] = \left(\partial_z^2 + \epsilon^2\partial_y^2\right) S \tag{3.5}$$

with zonal vorticity $\zeta = (\partial_z^2 + \epsilon^2\partial_y^2)\,\psi$ and Prandtl and Lewis numbers $P = \nu/\kappa_T$, $L = \kappa_S/\kappa_T$. The surface boundary conditions for T and S are now

$$T(y, 1) = a \cdot \theta(y), \quad \partial_z S(y, t, 1) = b\bar{F}(y, \epsilon^2 t) + c\widetilde{F}(y, t) \tag{3.6}$$

with thermal and saline Rayleigh numbers

$$a = \frac{g\alpha_T \Delta T d^3\epsilon^2}{\nu\kappa_T}, \qquad b = \frac{g\alpha_S \Delta S d^3\epsilon^2}{\nu\kappa_T}. \tag{3.7}$$

We have made a further hypothesis on the systematic salinity flux $\bar{F}$, which was originally assumed by Cessi and Young (1992) to be independent of time. We now take it to be time-dependent but only on a slow, diffusive scale. This is the correct rate to match with the evolution in the slow amplitude equation below. Another dimensionless group also appears

$$c = \frac{g\alpha_S \Sigma_0(\epsilon d)^{5/2}}{\nu\kappa_T^{1/2}}, \tag{3.8}$$

that determines the magnitude of the stochastic flux term.

To obtain a non-trivial limit, Cessi and Young (1992) assumed a scaling

$$a = \epsilon a_1, \qquad b = \epsilon^3 b_3, \tag{3.9}$$

that is, small Rayleigh numbers. Of course, in Earth's oceans, Rayleigh numbers based upon molecular diffusion coefficients are large. Because the parameters a and b in eq. (3.7) represent instead turbulent Rayleigh numbers, they can be legitimately taken to be much smaller. Based on these scaling assumptions, Cessi and Young (1992) introduced a multiscale expansion for the dynamical fields

$$(\psi, T, S) = \epsilon(\psi_1, T_1, S_1) + \epsilon^2(\psi_2, T_2, S_2) + \cdots. \tag{3.10}$$

Implicit in the asymptotics of Cessi and Young (1992) are several distinct time-scales, with different dependences on the aspect ratio ϵ. The shortest is the vertical mixing time which is $t = O(1)$. The advection times, both vertical and meridional, are longer, going as $t = O(\epsilon^{-1})$. Longest is the meridional mixing time which is $t = O(\epsilon^{-2})$. In the distinguished limit considered by Cessi and Young (1992) the dominant balance in the momentum equation corresponding to eq. (3.3) is between the buoyancy force and the vertical mixing. In eqs. (3.4) and (3.5) for temperature and salinity the fast vertical mixing completely homogenizes these fields in depth, which only vary in latitude to leading order. The first-order temperature field T_1 is then just proportional to its surface boundary condition. However, the first-order salinity field S_1 with the flux boundary condition is still undetermined at this stage, and a solvability condition emerges only at third order in the expansion. The time-scale on which the salinity field evolves turns out to be that of the slow meridional mixing. This led Cessi and Young (1992) to introduce a new time variable $t_2 = \epsilon^2 t$, for which the rate of evolution is order unity.

We can follow exactly the same procedure as Cessi and Young (1992) for our stochastic model. So that the random flux term will appear at third order, we must take

$$c = \epsilon^2 c_2. \tag{3.11}$$

The difference from the deterministic scaling arises because of the relation for white noise that $\tilde{F}(y, t) = \epsilon \tilde{F}(y, t_2)$. Defining $\tau = Lt_2$ and $\sigma(y, \tau) = a_1^{-1} S_1$, the solvability equation for our problem is

$$\partial_\tau \sigma - \mu^2 \partial_y \left[\partial_y \sigma (\partial_y \sigma - \partial_y \theta)^2\right] = r\bar{F}(y, \tau) + \sigma_0 \widetilde{F}(y, \tau) + \partial_y^2 \sigma - \gamma^2 \partial_y^4 \sigma. \tag{3.12}$$

As in Cessi and Young (1992), $\mu^2 = L^{-2} a_1^2 \int_0^1 W^2(z) \mathrm{d}z$, where $W(z)$ is a shape function that arises for no-stress boundary conditions, and $r = b_3/a_1$ is the ratio of the strengths of the salinity forcing and the thermal forcing. Likewise, $\sigma_0 = L^{1/2} c_2/a_1$. The hyperdiffusion term with $\gamma \sim \epsilon$ has been promoted from fifth order in the expansion to third order. Without this term, sharp fronts would develop in the solutions and eq. (3.12) could only be interpreted weakly, i.e. in the distributional sense. The hyperdiffusion term is a convenient device to avoid dealing with such singular solutions.

Defining meridional thermal and salinity gradients $\eta \equiv \partial_y \theta$, $\chi \equiv \partial_y \sigma$, the solvability equation can be finally rewritten as the following stochastic PDE

$$\partial_\tau \chi = \partial_y^2 \left[\mu^2 \chi (\chi - \eta)^2 + \chi - \gamma^2 \partial_y^2 \chi - r\bar{f}(y, \tau)\right] + \partial_y \widetilde{F}(y, \tau). \tag{3.13}$$

An antiderivative $\bar{f}(y, \tau) = -\int_{-\pi}^{y} \bar{F}(\bar{y}, \tau)\, \mathrm{d}\bar{y}$ has been introduced for the systematic flux, so that it may be brought inside the Laplacian ∂_y^2. On the other hand, the random flux is kept outside. It can be taken to be a Gaussian, zero-mean white noise with covariance

$$\langle \widetilde{F}(y, \tau) \widetilde{F}(y', \tau') \rangle = \sigma_0^2 \cdot \delta(y - y') \delta(\tau - \tau') \tag{3.14}$$

after absorbing a factor of σ_0 into the definition of $\widetilde{F}$.

An important observation in Cessi and Young (1992) was that the deterministic part of the dynamics has a variational structure. In fact, eq. (3.13) can be rewritten as

$$\partial_\tau \chi = \partial_y^2 \frac{\delta \Phi[\chi]}{\delta \chi(y)} + \partial_y \widetilde{F}(y, \tau) \tag{3.15}$$

where

$$\Phi[\chi] = \int_{-\pi}^{\pi} \left\{ V[\chi(y), y] + \frac{1}{2} \gamma^2 |\partial_y \chi(y)|^2 \right\} \mathrm{d}y \tag{3.16}$$

with quartic potential

$$V(\chi, y) = \mu^2 \left[\frac{1}{4}\chi^4 - \frac{2}{3}\chi^3 \eta(y)\right] + \frac{1}{2}\chi^2 [1 + \mu^2 \eta^2(y)] - r\bar{f}(y)\chi. \tag{3.17}$$

Note that the functional derivative in eq. (3.15) is defined by

$$\frac{\delta \Phi}{\delta \chi(y)}[\chi] = \lim_{h \to 0} \frac{\Phi[\chi + h\delta(\cdot - y)] - \Phi[\chi]}{h}$$

(see also Gardiner, 1985, p. 305). The deterministic equation is of a form known in statistical mechanics and materials science as the Cahn–Hilliard equation. It is used to describe the phase-separation process of spinodal decomposition, occurring when a chemical mixture or metallic alloy is quenched into an unstable state (see Cahn and Hilliard, 1958, or Bray, 1994 for a review). The version we have derived with an extra term that is a divergence of a white noise is called the stochastic Cahn–Hilliard equation or, in the theory of dynamic critical phenomena, ‘Model B’ (Hohenberg and Halperin 1977). The divergence of the noise is necessary in order to keep the equation formally conserving the ‘order parameter’ $\int_{-\pi}^{\pi} \mathrm{d}y\, \chi(y)$. See Cardon-Weber (2001) for a recent mathematical study of this stochastic PDE. However, our equation differs in two respects from the standard version of the stochastic Cahn–Hilliard equation. First, it is inhomogeneous, because the potential $V(\chi, y)$ in eq. (3.17) depends explicitly upon space point y. Secondly and more importantly, the boundary conditions for our equation are

$$\chi(\pm\pi) = \partial_y^2 \chi(\pm\pi) = 0, \tag{3.18}$$

which follow directly from the second equation in eq. (2.7). See Cessi and Young (1992). These are not the standard boundary conditions for the Cahn–Hilliard equation (where first and third derivatives are set to zero instead) and do not conserve the ‘order parameter’.

4. Existence and stability of multiple equilibria

As was discussed in detail by Cessi and Young (1992), the deterministic version of eq. (3.15) – without the noise term – may have multiple stable equilibria. Throughout this section we shall consider only the autonomous, deterministic system with $\bar{F}(y, \tau) \equiv \bar{F}(y)$, $\widetilde{F}(y, \tau) \equiv 0$. The steady states satisfy the equation

$$0 = \frac{\delta \Phi[\chi]}{\delta \chi(y)} = V_\chi(\chi, y) - \gamma^2 \partial_y^2 \chi = \mu^2 \chi [\chi - \eta(y)]^2 - r\bar{f}(y) + \chi - \gamma^2 \partial_y^2 \chi. \tag{4.1}$$

For certain values of μ^2, r, $\eta(y)$, and $\bar{f}(y)$, the cubic $V_\chi(\chi, y)$ may have three real roots

$$\chi_{\mathrm{A}}(y) < \chi_{\mathrm{B}}(y) < \chi_{\mathrm{C}}(y). \tag{4.2}$$

For this to be true, it is necessary that $\mu^2 \eta^2(y) > 3$. If $\eta(y) = -\sin y$, then this can occur only for mid-latitude ranges of y, i.e. near 45N or 45S, and outside those intervals only one branch of roots survives. To have three real roots for a point in either of these two mid-latitude intervals, it is necessary and sufficient that r take on an intermediate set of values, neither too large nor too small. Let us define

$$\bar{r}(y) = r\bar{f}(y)/\eta(y), \tag{4.3}$$

which is the ratio of the local strengths of salinity forcing and thermal forcing. Also, we define

$$\bar{r}_*(y) = \frac{2}{3}\left[1 + \frac{1}{9}\mu^2\eta^2(y)\right] \tag{4.4}$$

and

$$\Delta(y) = \frac{2}{27\mu^2|\eta(y)|}[\mu^2\eta^2(y) - 3]^{3/2}, \tag{4.5}$$

assuming that the latter is real. Then, the interval with three real roots corresponds exactly with the interval of r and y for which

$$|\bar{r}(y) - \bar{r}_*(y)| < \Delta(y). \tag{4.6}$$

See Appendix A for proofs of all these results. It is in this intermediate range that competition between salinity forcing and thermal forcing allows there to be multiple equilibrium states.

The Laplacian term $\gamma^2\partial_y^2\chi$ in eq. (4.1) only becomes important when there are rapid spatial changes in a solution. Because the three branches of roots in eq. (4.2) are smooth functions of y, the Laplacian term only becomes important near a change of branch. The analysis in Cessi and Young (1992, Appendix B) shows that solution branches, e.g. $\chi_{\rm A}$, $\chi_{\rm C}$, can be patched together only at critical points y_*

$$V(\chi_{\rm A}(y_*), y_*) = V(\chi_{\rm C}(y_*), y_*), \tag{4.7}$$

where the solutions have equal values of the potential V. This possibility to switch from one branch to another at critical points leads to the existence of multiple equilibria. The stability of these various equilibrium solutions is determined by properties of the functional $\Phi[\chi]$. According to eq. (4.1), the equilibria are stationary points of this functional. In fact, $\Phi[\chi]$ is a Lyapunov functional for the deterministic dynamics, because it is bounded below and

$$\partial_\tau \Phi = -\int_{-\pi}^{\pi}\left|\partial_y \frac{\delta\Phi[\chi]}{\delta\chi(y)}\right|^2 {\rm d}y \le 0. \tag{4.8}$$

As discussed in detail in Cessi and Young (1992, Appendix A), local maxima of Φ are unstable, while local minima are (locally) stable. Because the Laplacian term $\gamma^2\partial_y^2\chi$ only becomes important at critical points with switching from one branch to another, stability is largely determined by the potential term V. Because the middle root $\chi_{\rm B}(y)$ is always a local maximum of the quartic potential, it leads generally to a solution which is unstable. The extreme roots $\chi_{\rm A}(y)$, $\chi_{\rm C}(y)$ are always at least local minima and thus have non-empty domains of stability. Which of these is the global minimum is controlled by the ratio of salinity to thermal forcing $\bar{r}(y)$. It follows from Cessi and Young (1992, Appendix B) that the critical value of $\bar{r}(y)$ for which eq. (4.7) holds coincides with $\bar{r}_*(y)$ defined in eq. (4.4) above. (See also our own Appendix A.) When $\bar{r}(y) < \bar{r}_*(y)$ and thermal forcing dominates, then stability of $\chi_{\rm A}(y)$ is favored. The latter corresponds to a strong circulation state with weak salinity gradients. On the other hand, when $\bar{r}(y) > \bar{r}_*(y)$ and salinity forcing dominates, then stability of $\chi_{\rm C}(y)$ is favored. This corresponds conversely to a weak circulation state with strong salinity gradients.

In the paper of Cessi and Young (1992) the forcing functions $\theta(y) = \bar{F}(y) = \cos y$ are employed. With those forcings, there are four stable states, which correspond to choosing between solution branches $\chi_{\rm A}$, $\chi_{\rm C}$ at mid-latitudes in both the Southern and Northern Hemispheres. We may label these four states as AA, AC, CA, and CC, by listing the branches selected in the order south–north. See fig. 7 in Cessi and Young (1992) for plots of the stream functions of these states. In the notations of Thual and McWilliams (1992), AA corresponds to the thermally-driven circulation TH, CC to the salinity-driven circulation SA, CA to the north pole-to-pole circulation NPP, and AC to the south pole-to-pole circulation SPP. Thus, in the thin ocean limit considered by Cessi and Young (1992) the 'superposition principle' of Welander (1988) is exactly realized, according to which asymmetric pole-to-pole circulations can be viewed as 'superpositions' of half of an SA cell in one hemisphere and half of a TH cell in the other. Of course, as remarked already in Cessi and Young (1992), it is an unrealistic feature of the thin-ocean limit that its AC and CA equilibria no longer consist of a single pole-to-pole circulation but, instead, have the equator as a streamline separating the two cells.

Our model is intended for conceptual purposes and not for realistic modeling of the ocean circulation in any climate period. However, it is fruitful to compare the present simplified dynamics with more complex and realistic ones in some basic features. Bifurcation studies of three-dimensional primitive equation models with rotation (Weijer and Dijkstra, 2001) have shown that, as freshwater flux strength increases, symmetric TH solutions lose their stability in a supercritical pitchfork bifurcation to a pair of asymmetric NPP/SPP solutions. This is not true in our model, because the TH = AA equilibrium and the NPP/SPP = CA/AC equilibria are all simultaneously coexisting and stable. With addition of full wind stress, the model of Weijer and Dijkstra (2001) does have a range of freshwater forcing where four stable states coexist, but these consist rather of 'strong' NPP/SPP and 'weak' NPP/SPP states. In more realistic models of the present (Holocene) climate, both coupled GCMs (Bryan, 1986; Manabe and Stouffer, 1988) and Earth models of intermediate complexity (EMICs; Ganopolski et al., 1998; Ganopolski and Rahmstorf, 2001), there seem also to be only two stable states, not four. These correspond to a strong circulation state with North Atlantic Deep Water (NADW) formation and a state in which the deep ocean circulation is reversed in direction, shallower and weaker. These are roughly analogous in the simpler models to the NPP/SPP states, respectively, except that they are not just mirror images of each other. Such states may remain under ice-age conditions, but the climate models then exhibit also various types of NPP states. For example, see figs. 2b and d in Ganopolski and Rahmstorf (2001) for plots of the zonally-averaged stream functions for two such NPP states in the CLIMBER-2 model, which differ primarily in the latitude of Atlantic deep-water formation. Transitions between such NPP states – from a stable state with lower-latitude NADW formation to a 'metastable' state with higher-latitude formation – have been implicated in the DO rapid-warming events during the last ice age (Stocker, 2000; Ganopolski and Rahmstorf, 2001). Unlike transitions between NPP/SPP states, such mode switches have little impact on the export of NADW into the Southern Ocean and there is no bipolar 'seesaw' effect on the climate of the two hemispheres.

We would like to modify the surface forcings used in the original papers of Thual and McWilliams (1992) and Cessi and Young (1992) in order to caricature the latter type of transition. To do so, we must break the symmetry between the two hemispheres. We shall keep $\theta(y) = \cos(y)$, $\eta(y) = -\sin\ y$ to represent heating at the equator and cooling at the poles. However, we shall take

$$\bar{r}(y) = \begin{cases} \frac{2}{3}\left[1 + \frac{\mu^2}{9}\eta^2(y)\right][1 + \eta^2(y)] & -\pi < y < 0 \\ \frac{2}{3}\left[1 + \frac{\mu^2}{9}\eta^2(y)\right] & 0 < y < \pi \end{cases} \tag{4.9}$$

unlike $\bar{r}(y) = r$, a constant, as in Cessi and Young (1992). For $-\pi < y < 0$, $\bar{r}(y) = \bar{r}_*(y)[1 + \eta^2(y)]$ and it is easy to check that $\bar{r}_*(y)\eta^2(y) > \Delta(y)$ in the range where $\mu^2\eta^2(y) > 3$. Thus, there is now only one stable state in the Southern Hemisphere, C-type, with a weak circulation and upwelling at the pole. This corresponds roughly to the state of both the present-day and the ice-age South Atlantic, whose dominant water mass is the Antarctic Circumpolar Deep Water (ACDW), an area of mass upwelling (Sugden, 1982). On the other hand, for $0 < y < \pi$, $\bar{r}(y) = \bar{r}_*(y)$, which is the critical value. Thus, with this choice of forcing, there are two circulation states in the mid-latitudes of the Northern Hemisphere, exactly bistable. The one state, of A-type, has strong downwelling at high northern latitudes, roughly corresponding to NADW, and the other state of C-type has instead weak upwelling at high latitudes.

Figure 1 plots the surface salinity flux which follows from eq. (4.9) using

$$\bar{F}(y) = -\frac{\mathrm{d}\bar{f}}{\mathrm{d}y}(y)$$

for the same value $\mu^2 = 10$ that was used in Cessi and Young (1992). Although we have chosen the mathematical form of the salinity-to-thermal forcing ratio in eq. (4.9) largely for analytical convenience, in many respects the resulting salinity flux distribution shown in Fig. 1 is more realistic than the original one employed by Thual and McWilliams (1992) and Cessi and Young (1992). It may be compared with fig. 1.7 in Tomczak and Godfrey (2003), which plots annual mean precipitation minus evaporation (*P–E*) values across the Earth's oceans under current climate conditions. As shown there, the salinity flux attains greater values in the southern Atlantic than in the northern. Furthermore, the salt flux has a local minimum – in fact, a negative value, or net freshwater input – in the equatorial Atlantic. Unlike the original salinity flux model of Thual and McWilliams (1992)

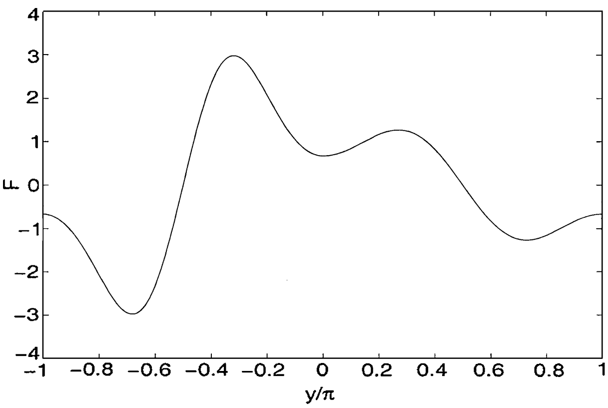
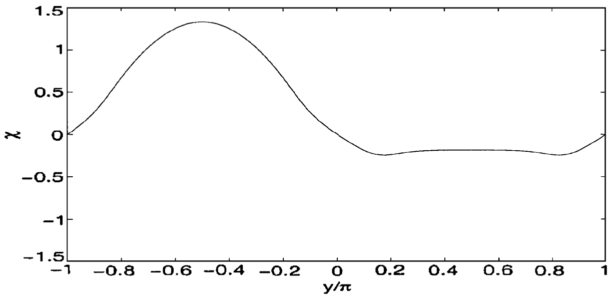


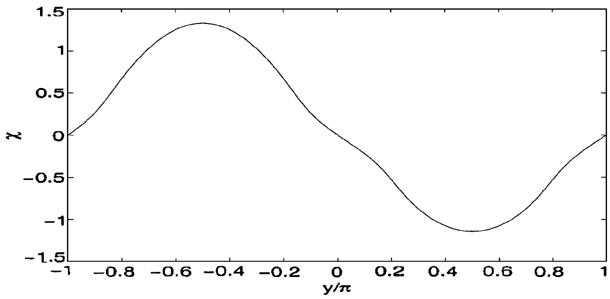
Fig 1. The surface salinity flux $\bar{F}(y)$ adopted in the model, plotted as a function of y.

and Cessi and Young (1992), our model in Fig. 1 has both larger amplitude in the Southern Hemisphere and an equatorial local minimum. (In fact, the value at the latter could easily be made negative with no change in our main results.) Just as for the original forcing, there is no net flux of salt into either hemisphere in our model, as required for the existence of equilibrium solutions. Hence, a spatial redistribution of the salinity flux suffices to change the stability properties of the flow, and no change in the net salt flux is required. This is a feature of our idealized model that we expect to be true more generally.

Of the four states in the model with the original forcing, only the states CA and CC remain stable with the new boundary conditions. We may refer to these equilibria more simply as just A and C, distinguishing only between alternatives selected in the Northern Hemisphere. Figures 2a and 2b plot the salinity gradient fields χ^*_{A}, χ^*_{C} for the two equilibria. These fields were obtained by discretizing eq. (4.1) with $\mu^2 = 10$, $\gamma = 0.1$ using 149 values of y in the interval $[-\pi, \pi]$. Details of the numerical discretization are given in Appendix B. The resulting set of coupled non-linear equations were solved by the Newton–Raphson method, using as initial guesses the corresponding branches of roots χ_{A}, χ_{C} of the cubic. Figures 3a and 3b plot the first-order stream functions in the vertical plane for the two solutions, using $\psi_1(y, z) = a_1(\chi - \eta)W(z)$ [eq. (5.6) of Cessi and Young, 1992]. Clearly, equilibrium A has a strong clockwise circulation in the Northern Hemisphere which efficiently transports salt and smooths out the equator-to-pole salinity gradient. In this state, the north polar waters are fresh and cold, the equatorial waters warm and salty, and the south polar waters very fresh and very cold. Equilibrium C has instead a weak counterclockwise circulation in the mid-latitudes of the Northern Hemisphere, as well as secondary, clockwise circulation cells near the equator and the pole. As a consequence of the weaker circulation, the high northern latitudes are now fresher and colder than in equilibrium A. In this respect, the two equilibria are similar to the two distinct NPP states observed in ice-age climate models and believed to



(a)



(b)

Fig 2. The stable equilibrium solutions for the meridional salinity gradient $\chi(y)$ of the deterministic model: (a) thermally-driven state χ^*_{A} and (b) salinity-driven state χ^*_{C}.

be associated with DO transitions (Stocker, 2000; Ganopolski and Rahmstorf, 2001). Admittedly, the analogy is not perfect. However, the states A and C in the present model share also with the two ice-age NPP states both the insensitivity of the Southern Hemispheric circulation to mode changes in the north and the primary role of freshwater forcing at the northern latitudes in altering their stability. The analogy could be improved by choosing a salt flux in the Northern Hemisphere to prevent instability except in a very narrow, high-latitude band. However, the present choice yields a caricature that is sufficient for our purposes and theoretically tractable.

5. Statistical equilibrium and state transitions

We now return to the study of the stochastic Cahn–Hilliard equation, eqs. (3.14) and (3.15). This stochastic PDE has an equivalent formulation in terms of a functional Fokker–Planck equation for the probability distribution $P[\chi, \tau]$ of the field $\chi(y, \tau)$

$$\frac{\partial P}{\partial \tau}[\chi, \tau] = -\int \mathrm{d}y \frac{\delta}{\delta\chi(y)} \{A[y, \chi]P[\chi]\} + \frac{1}{2}\int \mathrm{d}y \int \mathrm{d}y' B(y, y') \frac{\delta^2}{\delta\chi(y)\delta\chi(y')} P[\chi, \tau] \tag{5.1}$$

with drift

$$A[y, \chi] = \partial_y^2 \frac{\delta\Phi}{\delta\chi(y)}[\chi] \tag{5.2}$$

and diffusion

$$B(y, y') = -\sigma_0^2 \partial_y^2 \delta(y - y'). \tag{5.3}$$

For a general discussion of such equations, see Gardiner (1985, Section 8.1), and for eqs. (5.1)–(5.3) in the context of 'Model B' of critical dynamics, see Hohenberg and Halperin (1977). A Fokker–Planck equation will hold whenever the random forcing is white noise in time, so that the time evolution of $\chi(y, \tau)$ is a Markov diffusion process. The delta function appears in eq. (5.3) because the random component of the salinity flux was assumed also to be spatial white noise. In general, the diffusion kernel will be the spatial covariance of the random forcing term $\partial_y \tilde{F}(y, \tau)$, which may be correlated in space; see Gardiner (1985, Section 8.1).

However, the spatial white-noise forcing has special properties. In this case, the functional Fokker–Planck equation may be rewritten as a continuity equation for the probability density in function space

$$\frac{\partial P}{\partial \tau}[\chi, \tau] + \int \mathrm{d}y \frac{\delta}{\delta\chi(y)} J[y, \chi] = 0 \tag{5.4}$$

with the probability current given by

$$J[y, \chi] = \partial_y^2 \left[\frac{\delta\Phi}{\delta\chi(y)}[\chi]P + \frac{1}{2}\sigma_0^2 \frac{\delta P}{\delta\chi(y)} \right]. \tag{5.5}$$

A statistical equilibrium distribution $P_*[\chi]$ is reached if $J[y, \chi] \equiv 0$, and this occurs precisely when

$$\mathcal{P}_*[\chi] \propto \exp\left(-\frac{2\Phi[\chi]}{\sigma_0^2}\right). \tag{5.6}$$

In general, for other choices of the random noise it is not possible to write down a simple analytical formula such as eq. (5.6) for the stationary probability distribution of the process. That such a formula exists when the forcing is space–time white-noise is a fluctuation-dissipation relation for the generalized gradient dynamics (3.15). For example, see Hohenberg and Halperin (1977) and Gardiner (1985, Section 8.2.2). The vanishing of the current in eq. (5.5), $J[y, \chi] \equiv 0$, is also a special case of the so-called potential conditions for the stationary distribution (5.6); see Gardiner (1985, Section 5.3.3).

The existence of the exact distribution formula (5.6) permits the statistics of the stationary distribution to be efficiently sampled by Monte Carlo methods. For example, we consider the equilibrium distributions of the following two variables

$$\bar{\chi}_{\mathrm{N}} \equiv \frac{1}{\pi}\int_{\pi}^{0} \mathrm{d}y\chi(y), \qquad \bar{\chi}_{\mathrm{S}} \equiv \frac{1}{\pi}\int_{-\pi}^{0} \mathrm{d}y\chi(y), \tag{5.7}$$

which are the pole-to-equator mean salinity gradients in the Northern and Southern Hemispheres, respectively. In Fig. 4 we plot the distributions of these two variables, obtained by 8×10^7 steps of the standard Metropolis algorithm (Gilks et al., 1996) applied to the formula (5.6) for $\mu^2 = 10$, $\gamma = 0.1$, $\sigma_0 = 0.114$ and forcing (4.9). As can be seen, the distribution of $\bar{\chi}_{\mathrm{S}}$ is a unimodal distribution, centered around high values. This reflects

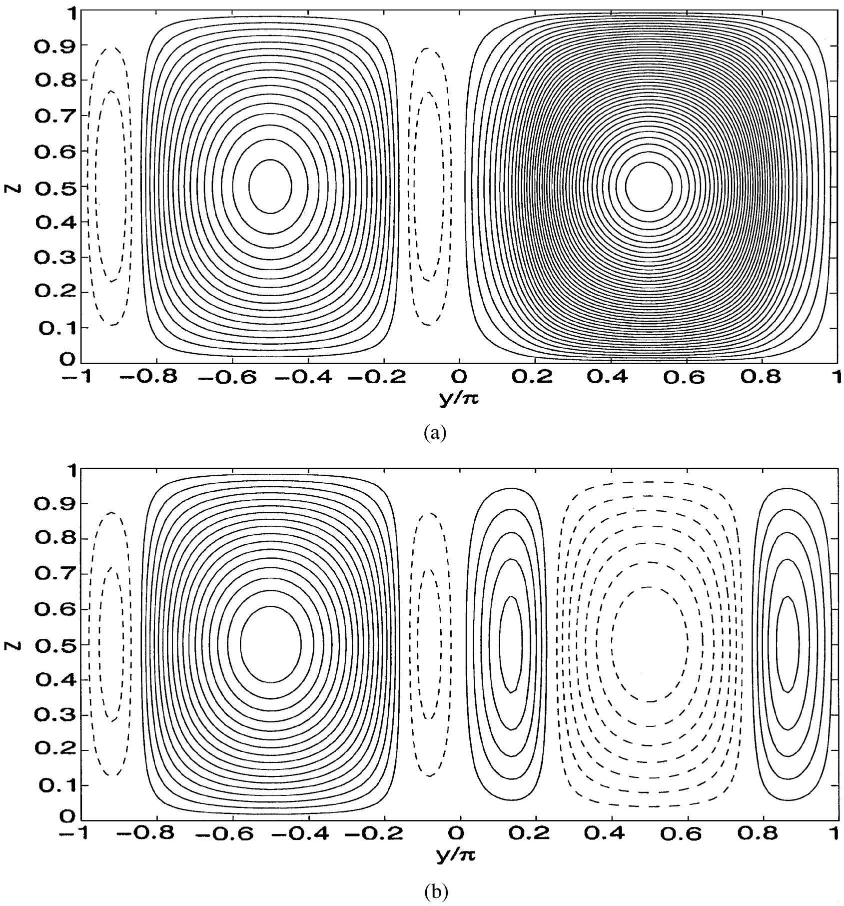


Fig 3. Contour plot of the first-order stream function $\psi_1(y, z)$, corresponding (a) to the thermally-driven state χ_A^* and (b) to the salinity-driven state χ_C^*. Solid lines denote positive values and dashed lines negative values.

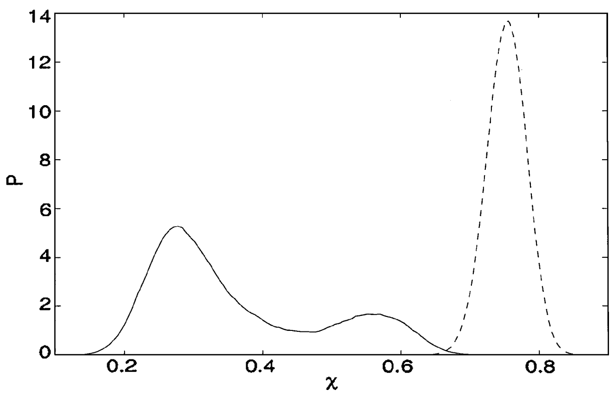


Fig 4. The probability density functions of pole-to-equator salinity gradients $\bar{\chi}_N$ and $\bar{\chi}_S$. The solid line is for the Northern Hemisphere and the dashed line is for the Southern Hemisphere.

the large difference in salinity between the equator and the south pole. On the other hand, the distribution of $\bar{\chi}_N$ is supported at smaller values and is bimodal. The values tend to cluster around those typical of the fixed points χ_A^* and χ_C^*, which represent the strong and weak circulation states at northern mid-latitudes. Although the salinity forcing was selected so that these two states are nearly bistable, i.e. $\Phi[\chi_A^*] \approx \Phi[\chi_C^*]$, we see that the peak at low values of $\bar{\chi}_N$ is higher, indicating that the system spends more time in the vicinity of χ_A^* than in the vicinity of χ_C^*. This is verified by analytical calculations below.

To obtain further insight about the dynamics we plot in Fig. 5 the time series $\bar{\chi}_S(\tau)$ and $\bar{\chi}_N(\tau)$. These were obtained by a numerical integration of the stochastic Cahn–Hilliard equation, using a semi-implicit Euler scheme discussed in Appendix C. It can be seen that the dynamical variable $\bar{\chi}_S(\tau)$ exhibits small, nearly Gaussian fluctuations around its high mean value. By contrast, the variable $\bar{\chi}_N(\tau)$ makes rapid switching transitions between levels with values near χ_A^* and χ_C^*. It is also apparent that the upper plateaux are somewhat shorter than the lower ones, indicating that $\bar{\chi}_N(\tau)$ spends less time in the vicinity of χ_C^* than in the vicinity of χ_A^*. It was proposed by Cessi and Young (1992) that these rapid transitions would occur by a nucleation process, in which a critical droplet of one 'phase' arises as a fluctuation inside the other 'phase' and then expands outward to fill the whole system. This is known to be the optimal path for phase ordering of the system in infinite volume, or, more generally, for the situation where the domain wall thickness $\sim \gamma$ is small compared to the system size 2π. However, a number of recent works (Maier and Stein, 2001; E et al., 2004) have shown that, when γ

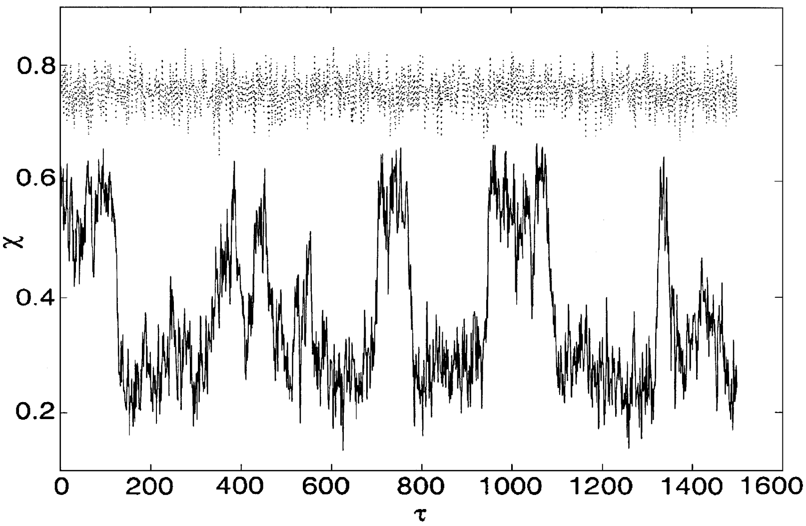


Fig 5. Time series of the pole-to-equator salinity gradients $\bar{\chi}_N(\tau)$ and $\bar{\chi}_S(\tau)$, in the stochastic model. The solid line is for the Northern Hemisphere and the dotted line is for the Southern Hemisphere.

exceeds a certain critical value, then the transition occurs instead by the nucleation of a single front or domain wall at one side of the system, which sweeps across to fill the whole region. For the relatively large value $\gamma = 0.1$ employed in our simulation, we observed the latter type of transition.

This knowledge of the transition process can be used to calculate analytically the mean residence time $\bar{\tau}_S$, i.e. the average time the system remains in the domain of attraction of each stable state χ_S^*, S = A, C. This time can also be written as $\bar{\tau}_S = 1/k_S$, where k_S is the escape rate from the domain of χ_S^*, S = A, C. Note from eqs. (3.8) and (3.11) that $\sigma_0 \propto \epsilon^{1/2}\Sigma_0$. Thus, the noise will be weak if Σ_0 is held fixed in the small aspect ratio limit. In that case, the escape rate k_S is given asymptotically as $\epsilon \to 0$ by the formula of Langer (1969):

$$k_S = \sum_{i=1}^{n} \frac{\lambda_{U_i}^{+}}{2\pi} \left| \frac{\mathrm{Det}\,\mathbf{L}_S}{\mathrm{Det}\,\mathbf{L}_{U_i}} \right|^{1/2} \exp\left(-\frac{2\Delta\Phi_{U_i|S}}{\sigma_0^2} \right). \tag{5.8}$$

Here $\chi_{U_i}^*$, i=1, . . ., ,n are the minimum-potential, unstable saddle points through which the transition occurs, $\Delta\Phi_{U_i|S} = \Phi[\chi_{U_i}^*]$ -$\Phi[\chi_S^*]$ is the barrier height, $\mathbf{L}_X$ are the linearized evolution operators at fixed point χ_X^* with kernels

$$\begin{aligned} L_X(y, y') &= \partial_y^2 \frac{\delta^2\Phi}{\delta\chi(y)\delta\chi(y')}[\chi_X^*] \\ &= \partial_y^2\left[\left\{\partial_y^2 + V_{\chi\chi}[\chi_X^*(y), y]\right\}\delta(y - y')\right] \end{aligned} \tag{5.9}$$

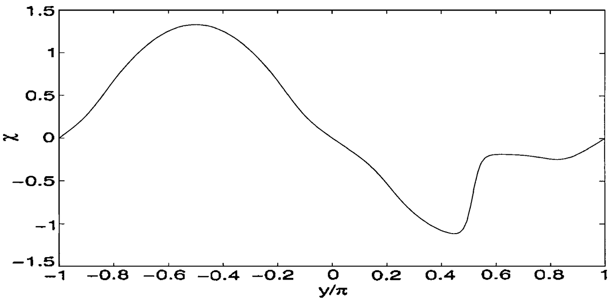
and $\lambda_{U_i}{}^{+}$ is the unique, positive (unstable) eigenvalue of $\mathbf{L}_{Ui}$. For a derivation of eq. (5.8), see Hänggi et al. (1990) and also Gardiner (1985, Section 9.3).

As discussed earlier, the transitions in the model with our choice of parameters appear to occur by moving fronts which sweep across the system from either the right or the left, entering one 'phase' and leaving the other 'phase' behind. These fronts correspond to equilibrium solutions of the deterministic model with one unstable direction, which are called 'domain walls' in statistical physics and 'instantons/anti-instantons' in the field theory literature. We have found such stationary solutions of our model numerically, by spatially discretizing on the interval $-\pi < y < \pi$ with 149 internal grid points, as discussed in Appendix B. We began with an analytical approximation to the fronts or boundary layer solutions, derived by Cessi and Young (1992, Appendix B), in the limit of narrow fronts. We then employed the Newton–Raphson algorithm to determine the exact fixed points, initialized with the approximate fronts from the Cessi–Young formula located at the northern mid-latitude ($y = \pi/2$). The resulting fixed points are plotted below, the 'instanton' $\chi_{U_+}^*$ in Fig. 6a and the 'anti-instanton' $\chi_{U_-}^*$ in Fig. 6b. To six decimal places, we find that $\Phi[\chi_{U_+}^*] = \Phi[\chi_{U_-}^*] = -2.19142$.

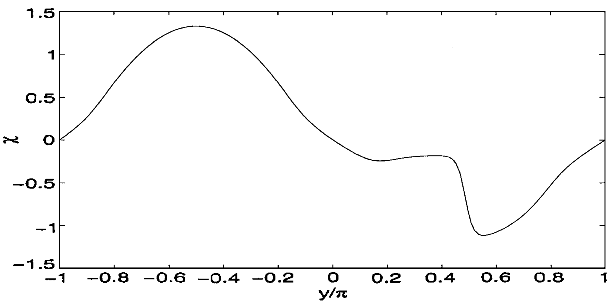
Both of these states appear in the sum formula (5.8) for the escape rates, with $n = 2$, for each of the stable fixed points, χ_A^* and χ_C^*. Note that $\chi_{U_-}^*$, $\chi_{U_+}^*$ are saddle points, on the boundary between the domains of attraction of χ_A^* and χ_C^*, and that, passing through them, one can return to either stable equilibrium. The determinants and eigenvalues for the discretizations of the operators appearing in eq. (5.8) have been calculated by MATLAB™. The resulting formulae for the mean residence time near each equilibrium solution are

$$\begin{aligned} \bar{\tau}_A &\approx (0.6099)\exp\left(0.07458/\sigma_0^2\right), \\ \bar{\tau}_C &\approx (0.9135)\exp\left(0.05742/\sigma_0^2\right). \end{aligned} \tag{5.10}$$

These are rather rapidly changing functions of the noise strength σ_0, although their ratio is more slowly varying. In Fig. 7 we plot $\bar{\tau}_A/\bar{\tau}_C$ over the interval $0.11 < \sigma_0 < 0.12$, where it can be seen that $\bar{\tau}_A/\bar{\tau}_C \approx 2.5$. Thus, for this set of parameters, the dynamics spends about twice as much time in the vicinity of χ_A^* as in the vicinity of χ_C^*. The primary reason for this is the slightly different values of the 'energy' functional for the two equilibria: $\Phi[\chi_A^*] = -2.22871$ and $\Phi[\chi_C^*] = -2.22013$. Although the forcing was constructed so that the 'potential energy' terms in eq. (3.16) are the same for the two states, the gradient or 'kinetic energy' term lifts the degeneracy.



(a)



(b)

Fig 6. Unstable saddle-point solutions of the deterministic model: (a) the 'instanton' $\chi^*_{\mathrm{U}+}(y)$ and (b) the 'anti-instanton' $\chi^*_{\mathrm{U}-}(y)$.

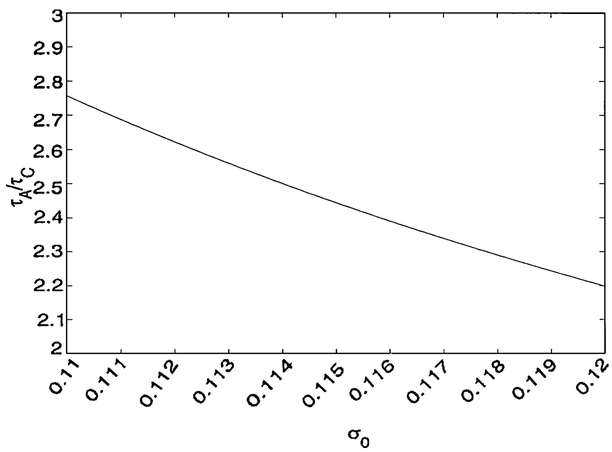


Fig 7. Ratio of mean residence times τ_{A}, τ_{C} in the domains of the two stable equilibria χ^*_{A} and χ^*_{C}, as a function of noise strength σ_0.

To recapitulate, we have derived in this section two principal results of the stochastic model: first, the exact formula (5.6) for the steady-state distribution, with its bimodal properties evidenced in Fig. 4, and, secondly, the formula (5.8) for the residence times near the equilibrium points, concretely calculated in eq. (5.10) under an assumption about the optimal transition paths. It is worth considering how sensitively these results might depend upon the various model assumptions, in so far as this may be plausibly argued. Broadly speaking, we expect that the results for the steady state are more robust and generic, while those for the stochastic transitions are likely to be more dependent quantitatively or even qualitatively upon details of the model. For example, the choice of parametrization of unresolved chaotic or turbulent processes by simple 'eddy diffusivities' in eqs. (2.1)–(2.3) or by more sophisticated modeling will likely have a negligible effect upon the steady-state distribution. This statistical distribution and its multimodal behavior derives from the existence of multiple stable equilibria of the deterministic model and on the local properties of the dynamics in the vicinity of these fixed points. In particular, the flows that are characteristically represented near each distribution peak differ not greatly from the equilibria themselves and the same turbulence parametrization probably suffices for all of them. By contrast, the transition from one equilibrium to another depends upon the properties of an unstable saddle point on the boundary of the domains of each fixed point, as seen in eq. (5.8), and its flow properties could be very different from those of either equilibrium. It is quite probable that a different turbulent parametrization must be used for this transient state. This will not affect the existence of the transitions themselves, as long as there is random noise to initiate them, but it can quantitatively influence such features as residence times that depend upon the transition pathway. Our assumption of a space–time white-noise field to represent rapid, synoptic-scale variability in freshwater fluxes is also not likely to be a critical factor in determining the properties of the steady-state distribution, as long as whatever noise is active in the spatial regions with multiple states is relatively weak. On the other hand, work on simple model systems shows that space–time correlations in the random forcing, i.e. 'colored noise', could possibly affect the stationary distribution but even more strongly the properties of transitions (see Hänggi et al., 1990, Section VIII-C). State-dependent or 'multiplicative' noise, if it were present, could change the stationary distribution as well (Horsthemke and Lefever, 1984). For a study of these issues in a two-box model of the ocean circulation, see Timmermann and Lohmann (2000).

6. Stochastic resonance

As mentioned in the introduction, one of the motivations of recent studies (Vélez-Belchí et al., 2001; Ganopolski and Rahmstorf, 2002) of statistical models of ocean circulation was to demonstrate the existence of stochastic resonance. As we show now, the present model exhibits this phenomenon in an analytically tractable form. Indeed, the stochastic version of the Cessi–Young dynamics is very similar to a prototype model that been used to illustrate stochastic resonance in spatially-extended systems (Benzi et al., 1985).

A simple explanation of the phenomenon of stochastic resonance has been offered in Benzi et al. (1981, 1982); see also Gammaitoni et al. (1998). Suppose that a weak periodic force is imposed upon a system exhibiting stochastic state transitions, of the sort we have considered in the previous section, and that

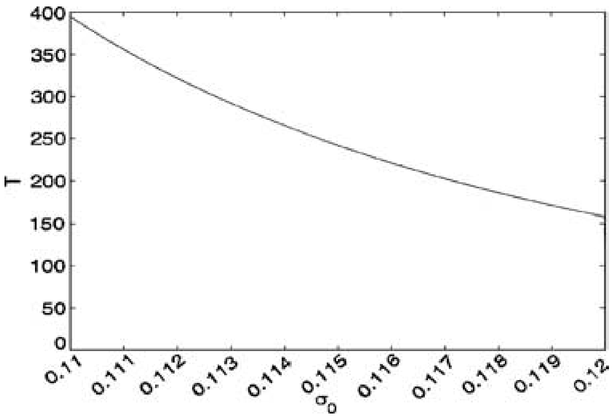


Fig 8. (Mean) stochastic resonant period T as a function of noise strength σ_0.

the force biases the stability of one or the other state as it cycles through its period. Maximum bias will occur twice each period, alternately for each state. In that case, even if the periodic forcing is subthreshold and cannot trigger a transition on its own, it will tend to facilitate a stochastic transition if the period of the forcing T is related to the mean residence time $\bar{\tau}$ of the stable states by the relation $T = 2\bar{\tau}$. With such a resonance condition, the force will lower the barrier to a transition precisely at the time that the system is inclined, on average, to execute a random jump. Therefore, there will tend to be a synchronization of the stochastic transitions with the periodic driving. Somewhat off-resonance, the synchronization will be less perfect, as some optimal jump times will be missed and the system must then wait a whole period or more for the optimal conditions to re-arrive. In this case, there will be no single, dominant residence time $T/2$ but there will be a tendency instead for residence times to be odd multiples of $T/2$. [In fact, this is closer to what is observed in the ice-core data of the last 100 ka, as reported in Alley et al. (2001a,b).]

In our model, there are different mean residence times to be near the two stable states χ_A^* and χ_C^*. However, these times do not differ on order of magnitude. It is therefore plausible to take as the resonant period an average of the values $T_A = 2\bar{\tau}_A$ and $T_C = 2\bar{\tau}_C$ for the two equilibria:

$$T = \frac{1}{2}(T_A + T_C) = \bar{\tau}_A + \bar{\tau}_C, \tag{6.1}$$

using eq. (5.10). In Fig. 8 we plot this stochastic resonant period in our model, as a function of the noise strength σ_0. It varies from 150–400 over the interval $0.11 < \sigma_0 < 0.12$. At the value used in our simulations, $\sigma_0 = 0.114$, we obtain $T \approx 265$.

To illustrate the phenomenon of stochastic resonance in our model, we add a time-periodic modulation of the surface salinity flux. Our choice is to alter the ratio of salinity to thermal forcing as follows:

$$\bar{r}(y,\tau) = \begin{cases} \frac{2}{3}\left[1 + \frac{\mu^2}{9}\eta^2(y)\right](1+\eta^2(y)) \\ \qquad\qquad -\pi < y < 0 \\ \frac{2}{3}\left\{1 + \frac{\mu^2}{9}\eta^2(y)[1 + A\cos(\omega\tau + \varphi)]\right\}. \\ \qquad\qquad 0 < y < \pi \end{cases} \tag{6.2}$$

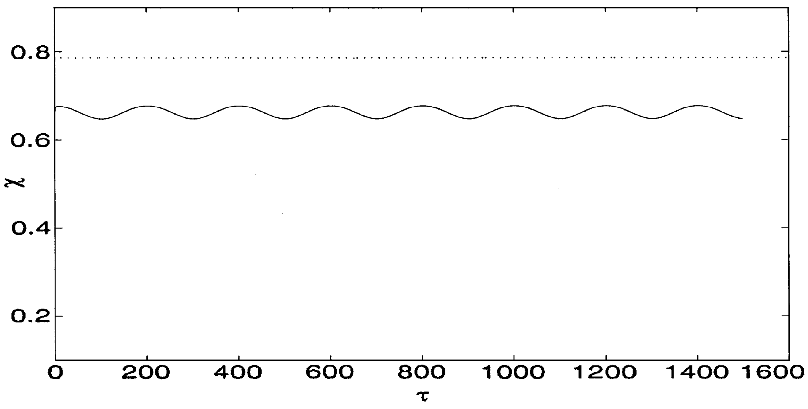
We adopt the angular frequency $\omega = 2\pi/T$ with $T = 200$, from our calculation above. We take the amplitude and phase of the modulation to be $A = 0.1$ and $\varphi = 0$. With these choices the forcing is subthreshold. For example, see in Fig. 9a the time series of the model with this periodic driving alone and no random force. Started in the state χ_C^* the system always stays in the domain of attraction of that equilibrium solution, executing small periodic oscillations about it. However, in Fig. 9b see the time series with both random and periodic driving. Despite the weakness of the time-dependent modulation, the stochastic transitions are now strongly correlated with its period. This is the hallmark of stochastic resonance.

It is interesting to observe that, as σ_0 decreases, the mean residence times for the two equilibria χ_A^*, χ_C^* rapidly 'detune', as a consequence of the slight asymmetry in their stability. For example, for $\sigma_0 = 0.08$, $\bar{\tau}_A/\bar{\tau}_C \approx 9.75$ and the two times already differ by about an order of magnitude. This means that, as σ_0 decreases further, the transitions out of the two states will be in resonance with periodic driving at quite different frequencies. Thus, stochastic resonance should be possible in our model only when noise levels are large enough and resonant frequencies high enough that the two times are comparable. This same restriction should apply also in natural cases with asymmetric stability, such as the 'cold' and 'warm' states in ice-age climate.

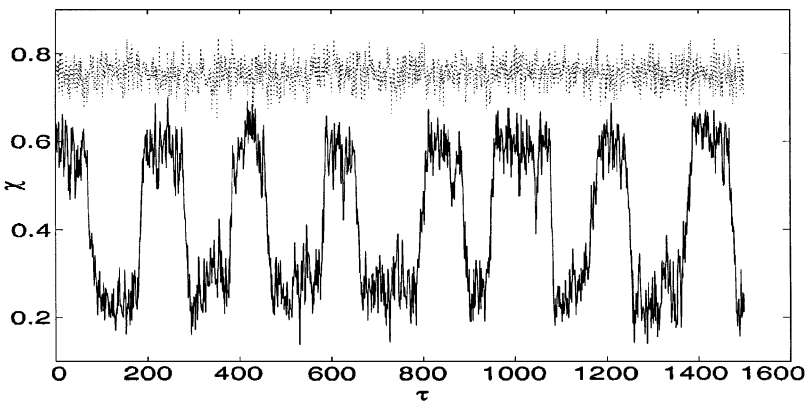
7. Conclusion

In the present work we have studied a simple stochastic model of the meridional overturning current of the ocean general circulation on climate time-scales. The model exhibits multiple stable equilibria, which are reflected in multimodal steady-state statistics. The model spends a relatively long period in the vicinity of one of these stable fixed points and then makes rapid transitions from that state to the other. The switching transitions are triggered by stochastic perturbations in the surface freshwater flux and are randomly distributed in time. However, these abrupt, random changes may be synchronized to a weak periodic perturbation of the correct frequency, the so-called stochastic resonance phenomenon.

Models with the features discussed here – multimodal statistics and rapid, stochastic transitions between regimes – are known to be difficult cases for most current methods of data assimilation or statistical inverse modeling (Miller et al., 1994, 1999). Such methods are necessary to extract the best estimates of past or future climate from the combination of data records



(a)



(b)

Fig 9. Time series of the pole-to-equator salinity gradients $\bar{\chi}_{\mathrm{N}}(\tau)$ and $\bar{\chi}_{\mathrm{S}}(\tau)$ with subthreshold periodic driving: (a) the deterministic model; (b) the stochastic model. The solid line is for the Northern Hemisphere and the dotted line is for the Southern Hemisphere.

and dynamical models. A very fruitful application of such methods in paleoclimatology would be to develop optimal hindcasts of past climate using oxygen isotope records from polar ice and deep-sea cores (Hinnov et al., 2002) together with coupled climate models (Ganopolski et al., 1998; Ganopolski and Rahmstorf, 2001). An application of even more immediate practical interest would be to develop optimal forecasts of future climate based upon available observations. An ongoing collaboration, the Arctic–Subarctic Ocean Flux (ASOF) study, aims to measure and model the variability of fluxes between the Arctic and Atlantic Oceans (Dickson and Boscolo, 2002). This study has detected a widespread and sustained freshening of the deep and abyssal North Atlantic over the last four decades (Dickson et al., 2002), although it is unclear whether these changes reflect the onset of global change. Efforts are already underway to forecast future changes in the Atlantic meridional circulation combining *in situ* and remote observations with GCMs using variational assimilation techniques (T. Haine and D. Lea, private communication).

However, some of the currently most popular methods such as the least-squares estimator or 4DVar method and the ensemble Kalman filter and smoother, which are derived assuming Gaussian statistics, fail badly in multimodal systems. In particular, they either miss the transitions entirely (Miller et al., 1994) or else grossly underestimate the rapidity of climate change (Miller et al., 1999). Methods which employ more realistic non-Gaussian, multimodal representations of the system statistics, either by moment closure methods (Eyink and Restrepo, 2000) or by empirical ensembles (Kim et al., 2003), are far more successful. One of the intended uses of the present model is to provide a benchmark system for the development of advanced methods of data assimilation in multimodal non-linear systems.

8. Acknowledgements

We wish to thank R. B. Alley, T. W. N. Haine, L. A. Hinnov, D. J. Lea, J. M. Restrepo, and C. Wunsch for helpful discussions. This work was supported by NSF grant no. DMS-0113649.

9. Appendix A: Roots of the cubic

The cubic polynomial

$$V_\chi(\chi, y) = \mu^2 \chi[\chi - \eta(y)]^2 + \chi - r\bar{f}(y) \quad \text{(A1)}$$

may be converted to a standard form without quadratic term (Uspensky, 1948)

$$\xi^3 + p\xi = q, \tag{A2}$$

by means of the substitution

$$\chi = \frac{2}{3}\eta + \xi. \tag{A3}$$

Simple calculation gives

$$p = \left(1 - \frac{\mu^2}{3}\eta^2\right) \Big/ \mu^2 \tag{A4}$$

$$q = (27r\bar{f} - 18\eta - 2\mu^2\eta^3)/27\mu^2. \tag{A5}$$

The cubic has three distinct real roots precisely when its discriminant is negative (Uspensky, 1948):

$$\mathrm{Dis} \equiv \frac{1}{27}p^3 + \frac{1}{4}q^2 < 0. \tag{A6}$$

Using the above values for p, q, this becomes the condition

$$4(\mu^2\eta^2 - 3)^3 \geq \mu^2(27r\bar{f} - 18\eta - 2\mu^2\eta^3)^2. \tag{A7}$$

To satisfy this condition it is clearly necessary that $\mu^2\eta^2 > 3$. When this is true, it is easy to use the definitions (4.3)–(4.5) in the text to show that this condition coincides with

$$[\bar{r}(y) - \bar{r}_*(y)]^2 < \Delta^2(y). \tag{A8}$$

As noted in Cessi and Young (1992) the critical value $\bar{r}_*(y)$ can be characterized by the equality $\chi_B = (2/3)\,\eta$, where the latter is the inflection point of the cubic polynomial, at which $V_{\chi\chi\chi} = 0$. From the above change of variables, it follows that the inflection point is a root precisely when $\xi = 0$ is a root. This is true if and only if $q = 0$, which leads to eq. (4.4) in the text.

10. Appendix B: Numerical methods

We discuss here the discretization of the stochastic eq. (3.15) in space and time. Let us first consider the discretization in space, keeping time continuous. The interval $-\pi \leq y \leq \pi$ is represented by $N + 2$ grid points $y_j = -\pi + jh$, $j = 0, 1, 2, \ldots, N + 1$ for $h = 2\pi/(N + 1)$. The dynamical eq. (3.15) now becomes a set of N coupled stochastic ordinary differential equations

$$\dot{\chi}_j = \frac{1}{h^2}\left[\left(\frac{\delta\Phi}{\delta\chi}\right)_{j+1} + \left(\frac{\delta\Phi}{\delta\chi}\right)_{j-1} - 2\left(\frac{\delta\Phi}{\delta\chi}\right)_j\right] + \frac{1}{h}(\widetilde{F}_{j+1} - \widetilde{F}_j), \tag{B1}$$

for variables χ_j, $j = 1, 2, \ldots, N$. The boundary conditions (3.18) become

$$\chi_{-1} = \chi_{N+1} = 0, \quad \chi_{-2} + \chi_0 = \chi_N + \chi_{N+2} = 0. \tag{B2}$$

The random force $\widetilde{F}_j$, $j = 1, \ldots, N$ (with $\widetilde{F}_{N+1} \equiv 0$) is a zero-mean Gaussian with covariance

$$\langle \widetilde{F}_j(t)\widetilde{F}_m(t')\rangle = \frac{\sigma_0^2}{h}\delta_{jm}\delta(t - t') \tag{B3}$$

for $j, m = 1, \ldots, N$. The functional Φ in eq. (3.16) has been discretized as

$$\Phi[\chi] = \sum_{j=1}^{N} h\left[\frac{1}{2}\gamma^2\left(\frac{\chi_{j+1} - \chi_j}{h}\right)^2 + V_j(\chi_j)\right] \tag{B4}$$

with

$$\begin{aligned}\left(\frac{\delta\Phi}{\delta\chi}\right)_j &= \frac{1}{h}\frac{\partial\Phi}{\partial\chi_j}\\ &= -\gamma^2\left(\frac{\chi_{j-1} + \chi_{j+1} - 2\chi_j}{h^2}\right) + V_j'(\chi_j)\\ &= -\gamma^2 \triangle \chi_j + V_j'(\chi_j)\end{aligned} \tag{B5}$$

where $\triangle$ is the lattice Laplacian with Dirichlet boundary condition. The space-discretization considered has attractive properties: in particular, the obvious lattice analogs hold of the formula (5.6) for the stationary distribution and of eq. (5.8) for the escape rate. The spatial step size h must be chosen $\sim \epsilon$ to resolve the boundary layers with thickness $\gamma \sim \epsilon$. Equation (B1) can also be written as

$$\dot{\chi}_j = -\gamma^2 \triangle^2 \chi_j + \triangle V_j'(\chi_j) + \frac{1}{h}(\widetilde{F}_{j+1} - \widetilde{F}_j), \tag{B6}$$

where $\triangle^2$ is the lattice biLaplacian with the boundary condition (B2). This corresponds to the pentadiagonal matrix

$$\triangle^2 = \frac{1}{h^4}\begin{bmatrix} 5 & -4 & 1 & & & & & \\ -4 & 6 & -4 & 1 & & & & \\ 1 & -4 & 6 & -4 & 1 & & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & 1 & -4 & 6 & -4 & 1 \\ & & & & 1 & -4 & 6 & -4 \\ & & & & & 1 & -4 & 5 \end{bmatrix}. \tag{B7}$$

The time-discretization of the dynamics uses $t_i = ik$, $i = 0, 1, 2, \ldots$ for spacing k, with the history denoted by χ_j^i, $j = 1, \ldots, N$, $i = 0, 1, 2, \ldots$. Like the deterministic Cahn–Hilliard equation, our stochastic version is very stiff, with a wide range of time-scales to be resolved (see, for example, Stuart and Humphries, 1994). To achieve stability, a natural time-discretization of eq. (B6) is a semi-implicit Euler scheme

$$\frac{\chi_j^{i+1} - \chi_j^i}{k} = -\gamma^2 \triangle^2 \chi_j^{i+1} + \triangle V_j'(\chi_j^i) + \frac{1}{h}\left(\widetilde{F}_{j+1}^i - \widetilde{F}_j^i\right), \tag{B8}$$

where now

$$\langle \widetilde{F}_j^i \widetilde{F}_m^n \rangle = \frac{\sigma_0^2}{kh} \delta_{jm} \delta^{in}. \tag{B9}$$

This scheme has been proved to converge in probability to the solution of the stochastic PDE, uniformly in space and time. At each time-step of the numerical integration the linear system

$$[I + \gamma^2 k \triangle^2] \chi_j^{i+1} = \chi_j^i + k \triangle V_j'(\chi_j^i) + \frac{k}{h} \left(\widetilde{F}_{j+1}^i - \widetilde{F}_j^i \right) \tag{B10}$$

was solved using a stored factorization of the pentadiagonal matrix $[I + \gamma^2 k \triangle^2]$, following the method in Cheney and Kincaid (2003). Invertibility of this matrix requires a time-step $k \sim \epsilon^2$. In our integrations we used $N = 149$ or $h \approx 4.19 \times 10^{-2}$ and $k = 9.375 \times 10^{-5}$.

References

Alley, R. B., Anandakrishnan, A. and Jung, P. 2001a. Stochastic resonance in the North Atlantic. *Paleoceanography* **16**, 190–198.

Alley, R. B., Anandakrishnan, A., Jung, P. and Clough, A. 2001b. Stochastic resonance in the North Atlantic, further insights. In: *The Oceans and Rapid Climate Change, Past, Present, and Future* (eds D. Seidov, B. J. Haupt, and M. Maslin). American Geophysical Union.

Batchelor, G. K. 1967. *An Introduction to Fluid Mechanics*. Cambridge University Press, Cambridge.

Benzi, R., Parisi, G., Sutera, A. and Vulpiani, A. 1982. Stochastic resonance in climate change. *Tellus* **34**, 10–16.

Benzi, R., Sutera, A. and Vulpiani, A. 1981. The mechanism of stochastic resonance. *J. Phys. A: Math. Gen.* **14**, L453–L457.

Benzi, R., Sutera, A. and Vulpiani, A. 1985. Stochastic resonance in the Landau–Ginzburg equation. *J. Phys. A: Math. Gen.* **18**, 2239–2245.

Bray, A. J. 1994. Theory of phase-ordering dynamics. *Adv. Phys.* **43**, 357–459.

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

Bryan, K. and Hansen, F. C. 1995. A stochastic model of North Atlantic climate variability on a decade to century time-scale. In: *The Natural Variability of the Climate System on 10–100 Year Time Scales*, National Academy Press, pp. 355–364.

Cahn, J. W. and Hilliard, J. E. 1958. Free energy for a non-uniform system, I. interfacial free energy. *J. Chem. Phys.* **2**, 258–267.

Cardon-Weber, C. 2001. Cahn–Hilliard stochastic equation, existence of the solution and its density. *Bernoulli* **7**, 777–816.

Cessi, P. 1994. A simple box model of stochastically forced thermohaline flow. *J. Phys. Oceanogr.* **24**, 1911–1920.

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

Cheney, W. and Kincaid, D. 2003. *Numerical Mathematics and Computing*, 5th Edition. Brooks/Cole Publishing Co., Monterey, CA.

Clark, P. U., Pisias, N., Stocker, T. F. and Weaver, A. J. 2002. The role of the thermohaline circulation in abrupt climate change. *Nature* **415**, 863–869.

Dickson, R. and Boscolo, R. 2002. The Arctic–Subarctic ocean flux study (ASOF), rationale, scope and methods. *CLIVAR Exchanges* **25**, 64–65.

Dickson, R., Yashayaev, I., Meincke, J., Turrell, B., Dye, S. et al. 2002. Rapid freshening of the deep North Atlantic ocean over the past four decades. *Nature* **416**, 832–836.

Dijkstra, H. A. and Molemaker, M. J. 1997. Symmetry breaking and overturning oscillations in thermohaline-driven flows. *J. Fluid Mech.* **331**, 195–232.

E, W., Ren, W. and Vanden-Eijnden, E. 2004. Minimum action method for the study of rare events. *Commun. Pure Appl. Math.* **LVII**, 637–656.

Eyink, G. L. and Restrepo, J. M. 2000. Most probable histories for non-linear dynamics, tracking climate transitions. *J. Stat. Phys.* **101**, 459–472.

Gammaitoni, L., Hänggi, P., Jung, P. and Marchesoni, F. 1998. Stochastic resonance. *Rev. Mod. Phys.* **70**, 223–287.

Ganopolski, A. and Rahmstorf, S. 2001. Rapid changes of glacial climate simulated in a coupled climate model. *Nature* **409**, 153–158.

Ganopolski, A. and Rahmstorf, S. 2002. Abrupt glacial climate changes due to stochastic resonance. *Phys. Rev. Lett.* **88**, 038501/1–4.

Ganopolski, A., Rahmstorf, S., Petoukhov, V. and Claussen, M. 1998. Simulation of modern and glacial climates with a coupled global model of intermediate complexity. *Nature* **391**, 351–356.

Gardiner, C. W. 1985. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, 2nd Edition. Springer-Verlag, Berlin.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. 1996. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London.

Hänggi, P., Talkner, P. and Borkovec, M. 1990. Reaction-rate theory, fifty years after Kramers. *Rev. Mod. Phys.* **62**, 252–341.

Hinnov, L., Schulz, M. and Yiou, P. 2002. Interhemispheric space–time attributes of the Dansgaard-Oeschger oscillations between 100 and 0k̃a. *Q. Sci. Rev.* **21**, 1213–1228.

Hohenberg, P. C. and Halperin, B. I. 1977. Theory of dynamical critical phenomena. *Rev. Mod. Phys.* **49**, 435–479.

Horsthemke, W. and Lefever, R. 1984. *Noise-Induced Transitions. Theory and Applications in Physics, Chemistry, and Biology*. Springer-Verlag, Berlin.

Huang, R. X. 1999. Mixing and energetics of the oceanic thermohaline circulation. *J. Phys. Oceanogr.* **29**, 727–746.

Jeffreys, H. 1925. On fluid motions produced by differences of temperature and humidity. *Q. J. R. Meteorol. Soc.* **51**, 347–356.

Kim, S., Eyink, G. L., Restrepo, J. M., Alexander, F. J. and Johnson, G. 2003. Ensemble filtering for non-linear dynamics. *Mon. Wea. Rev.* **131**, 2586–2594.

Langer, J. S. 1969. Statistical theory of the decay of metastable states. *Ann. Phys.* **54**, 258–275.

Maier, R. S. and Stein, D. L. 2001. Droplet nucleation and domain wall motion in a bounded interval. *Phys. Rev. Lett.* **87**, 270601/1–4.

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

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

Miller, R. N., E. F. Carter, J. and Blue, S. T. 1999. Data assimilation into non-linear stochastic models. *Tellus* **51A**, 167–194.

Miller, R. N., Ghil, M. and Gauthiez, P. 1994. Advanced data assimilation in strongly non-linear dynamical systems. *J. Atmos. Sci.* **51**, 1037–1056.

Munk, W. and Wunsch, C. 1998. Abyssal recipes II, energetics of tidal and wind mixing. *Deep-Sea Res.* **45**, 1977–2010.

Paparella, F. and Young, W. R. 2002. Horizontal convection is non-turbulent. *J. Fluid Mech.* **466**, 205–214.

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

Rahmstorf, S. 2002. Ocean circulation and climate during the past 120 000 years. *Nature* **419**, 207–214.

Sandström, J. W. 1908. Dynamische versuche mit meerwasser. *Annals in Hydrodynamic Marine Meteorology*, 6 (in German).

Stocker, T. F. 2000. Past and future reorganizations in the climate system. *Q. Sci. Rev.* **19**, 301–319.

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

Stuart, A. M. and Humphries, A. R. 1994. Model problems in numerical stability theory for initial value problems. *SIAM Rev.* **36**, 226–257.

Sugden, D. 1982. *Arctic and Antarctic, A Modern Geographical Synthesis*. Blackwell, Oxford.

Tennekes, H. and Lumley, J. L. 1972. *A First Course in Turbulence*. MIT Press, Cambridge, MA.

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

Timmermann, A. and Lohmann, G. 2000. Noise-induced transitions in a simplified model of the thermohaline convection. *J. Phys. Oceanogr.* **30**, 1891–1900.

Tomczak, M. and Godfrey, J. S. 2003. *Regional Oceanography, An Introduction*, 2nd Edition. Daya Publishing House, Delhi.

Uspensky, J. 1948. *Theory of Equations*. McGraw-Hill, New York.

Vélez-Belchí, P., Alvarez, A., Colet, P., Tintoré, J. and Haney, R. L. 2001. Stochastic resonance in the thermohaline circulation. *Geophys. Res. Lett.* **28**, 2053–2056.

Weijer, W. and Dijkstra, H. A. 2001. A bifurcation study of the three-dimensional thermohaline ocean circulation, the double hemisphere case. *J. Marine Res.* **59**, 599–631.

Welander, P. 1988. Thermohaline effects in ocean circulation and related simple models. In: *Large-Scale Transport Processes in Oceans and Atmosphere* (eds J. Wildebrand, and D. T. L. Anderson). D. Reidel Publishing, Dordrecht, pp. 163–200.

Wunsch, C. 1998. The work done by the wind on the oceanic general circulation. *J. Phys. Oceanogr.* **28**, 2332–2340.

Wunsch, C. 2002. What is the thermohaline circulation? *Science* **298**, 1179–1180.