Abstract.

Christopher W. Hughes, Christ Church

D.Phil. Michaelmas 1992.

The Effect of Topography on Ocean Flow.

The rôle which topography plays in constraining ocean flow is investigated in several ways, mostly aimed at application to the Southern Ocean where topography is known to be important. The physics of topographic Rossby waves is discussed in some depth and a description of ocean flow in terms of a sum of topographic normal modes is developed. It is shown that the apparent incompleteness of topographic modes can often be circumvented by including a function which absorbs the nett input of potential vorticity. Some subtle problems with this description are dealt with, and a calculation of topographic modes for the Southern Ocean is presented, which shows that the modes are very localised, making the use of them to describe basin-wide flows difficult.

The effect of interactions between stratification and topography is investigated in terms of a quasi-two-dimensional model which deals only with the depth-integrated flow, and the assumptions which go into the model are examined in detail both analytically and by calculating terms of interest from a data set produced by the Fine Resolution Antarctic Model. It is shown that advection of density in the Southern Ocean can be described to a first approximation as being due to a barotropic current with no vertical velocity, the horizontal component of the baroclinic flow producing very little effect. The balance of terms reveals interesting features in the modelled flow in the Southern Ocean, showing the value of this type of analysis.

Finally, insight developed in the course of the investigation allows a simple model to be constructed representing the feedback between density advection and forcing due to density gradients. This model is used to provide an explanation for the fact that the FRAM model spins up linearly, where most simple models would predict a component of quadratic behaviour in the spin-up.

Acknowledgements.

The work in this thesis owes much to my supervisor, Peter Killworth, for his continuous enthuiasm and support, and for his choice of an interesting problem to start me off with. I never found a satisfactory answer to that initial problem, but the contents of this thesis are the story of the search. The journey turned out to be an education even if I never quite arrived at the originally intended destination.

Thanks also to everyone in the Hut for providing a lively atmosphere and interesting discussion, especially on Friday night, and particularly to Jeff Blundell for his patience with my FORTRAN foibles. I never did quite obey the ten commandments but I'm getting better, honest.

The final thank-you must go to my physics tutor, Jack Paton, for believing in me and teaching me the difference between physics and maths.

After the thanks, at last, come the apologies. To all those non-oceanographers and non-scientists who listened patiently to my incomprehensible rantings about potential vorticity and the JEBAR effect, who nodded politely and pretended to be interested, I'm sorry. Maybe one day I'll discover something that I really am capable of explaining without vector calculus.

Contents

Introduction

Bottom topography is essential to the dynamics of the Southern Ocean. This ocean contains a large fraction of the world's water and is unique in that it includes a band of latitudes defined by the Drake Passage (between the tip of South America and the Antarctic peninsula) which is not blocked by land at any longitude. This has two effects. First of all, it means that it forms the main connection between the three major oceans, Atlantic, Indian and Pacific. This means the Southern Ocean is important for carrying heat and salt between these oceans and thus has a large impact on climate the world over. Secondly, it makes the dynamics of the Southern Ocean potentially very different from the Sverdrup balance and western boundary currents found in other extra-tropical regions, since Sverdrup balance involves a pressure difference between the east and west coasts of the ocean, and there is no east or west coast to parts of the Southern Ocean.

Heat transport across the Southern Ocean is also very important in determining climate. If heat is to be transported from equator to pole in the southern hemisphere it must cross the Antarctic Circumpolar Current (ACC), so the amount of heat transport must depend on the dynamics of that current and the mechanism of heat transport across it. The extent of sea-ice formation on the fringes of the Antarctic continent also depends on (and affects) the path of and heat flow in the ACC. Sea-ice has a much higher albedo than the ocean surface, and forms an insulating layer which reduces ocean-atmosphere heat exchange, again producing a potentially enormous effect on climate. It is therefore vital for climate studies that we gain an understanding of how the ACC works. This strong

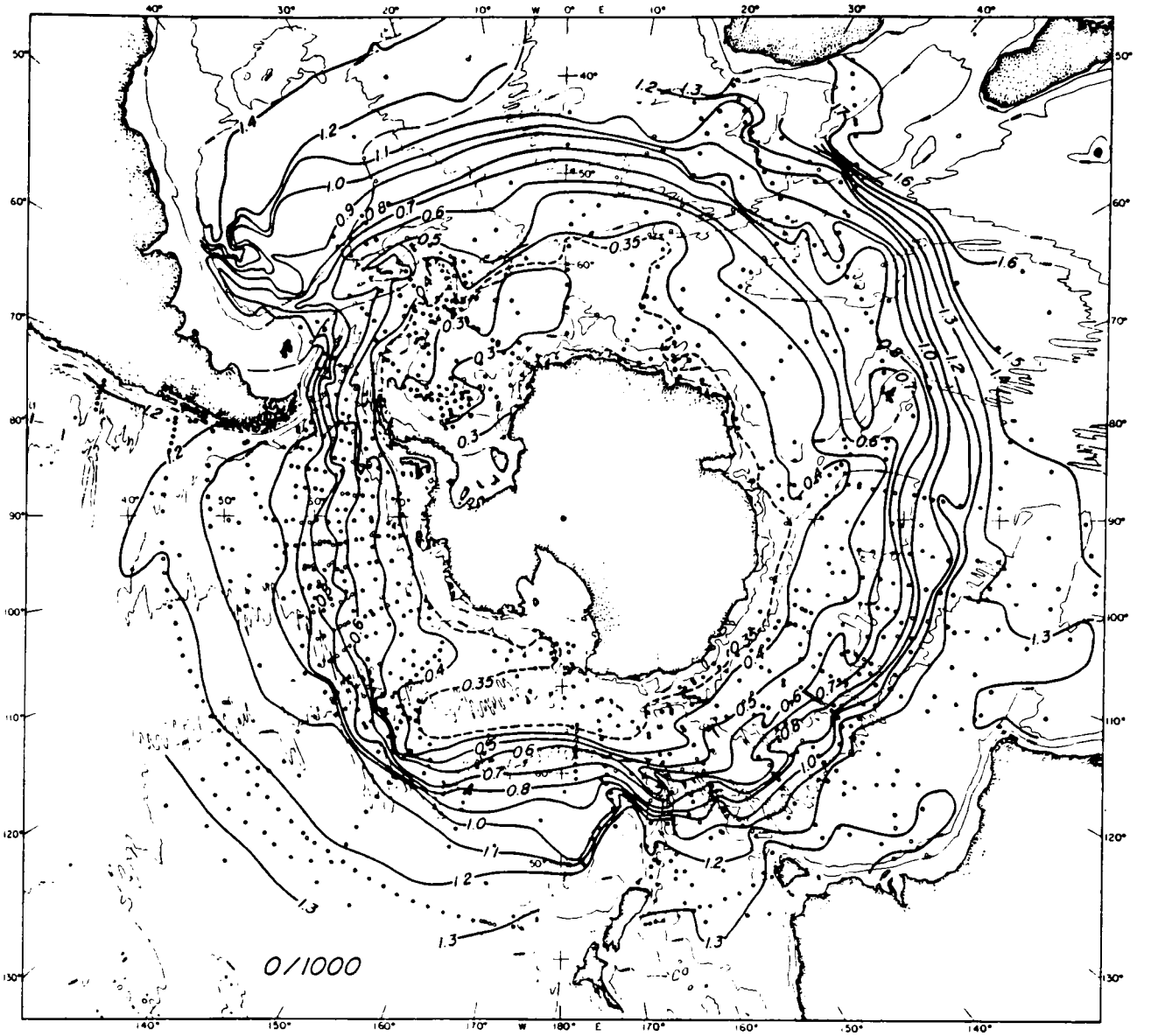


Fig.(0.1). *Sea surface height relative to an equipotential surface in the Southern Ocean, calculated from hydrographic data on the assumption that the flow is geostrophic and there is no flow below about 1000 m (1000 db). The units are geopotential metres, which are approximately equal to metres, and the circumpolar current is shown by the gradient of sea surface height. Taken from Gordon et al., 1978.*

current is clearly seen in hydrographic measurements (fig.(0.1)) and is evidently the most important oceanic feature in the region, if not in the world ocean.

The fact that topography is a vital component of the dynamic balance of the ACC was first pointed out by Munk and Palmén (1951), who looked at the terms which could balance wind stress in a zonally reconnected ocean. They found that the usually accepted values for parameterisations of lateral and vertical eddy viscosities were far too low to provide sinks for the momentum input by the strong eastward winds that blow over almost all the Southern Ocean, without producing a current of order 1000 Sv (Sv=sverdrups, 1 sverdrup=$10^6\,\mathrm{m}^3\,\mathrm{s}^{-1}$). The accepted value for transport in the ACC is closer to 100 Sv. Munk and Palmén (1951) therefore proposed that the momentum input by wind must be taken out by the solid earth below, by a topographic drag produced by higher pressures on the upstream side of ridges than on the downstream side. Their calculations showed that surface height differences of only about 4 cm across the four major topographic features in the Southern Ocean would be enough to account for all the input momentum. This problem of keeping the circumpolar transport down to reasonable values without invoking unacceptably high friction was later referred to by several authors as "Hidaka's dilemma", after Hidaka and Tsuchiya (1953) published the results of numerical calculations confirming Munk and Palmén's analysis for the flat-bottomed ocean.

Numerous other factors were put forward over the following years which could affect the size of the circumpolar current, an excellent review of which can be found in Nowlin and Klinck (1986). Calculation of the size of these effects, however, continued to show that topography was a necessary component of the balance of terms. Gill (1968) looked at the non-zonal nature of the current by introducing partial meridional barriers and solving via

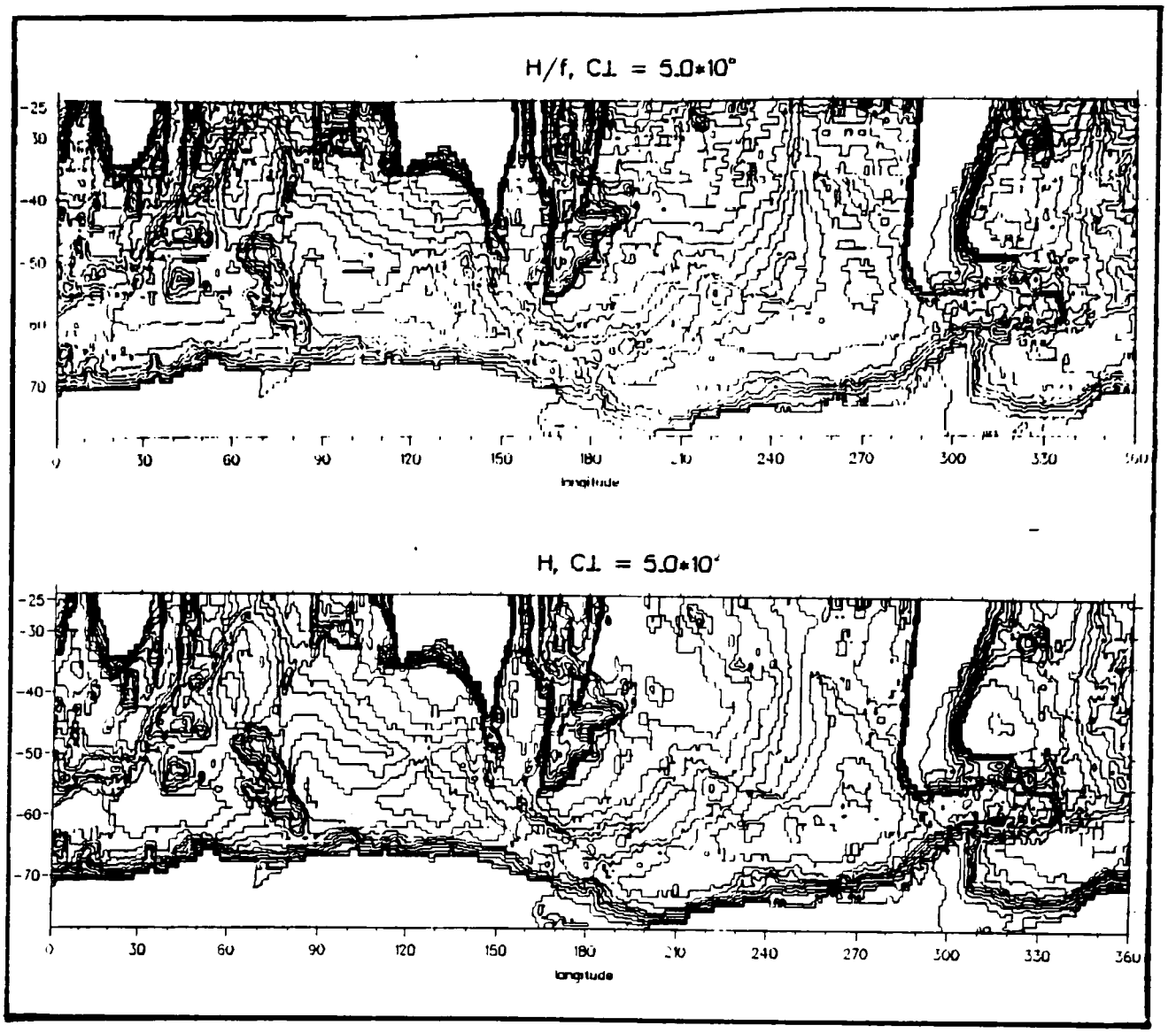


Fig.(0.2). *Plots of H/f (ms) and H (m) taken from FRAM. C.I. stands for Contour Interval. The strong similarity between the two plots shows that gradients of f/H are dominated by the topographic component except in some regions north of about 40 degrees south.*

a boundary layer analysis. His solution reflected some of the characteristics of the ACC (particularly the sharp turn to the north just east of South America), but again required high friction coefficients to produce a reasonable transport.

If topography was accepted as being the missing factor, constant density models which included topography had problems of their own. Schulman (1970) and Bryan and Cox (1972) both produced transports which were far too low (20 Sv and 32 Sv respectively) when topography was included. The reason for the dramatic drop is that, when topography is included in a homogeneous ocean, it is no longer the lines of latitude that are dynamically important but lines of constant f/H where f is the Coriolis parameter $f = 2\Omega \sin\phi$ (Ω is the angular speed of rotation of the earth and ϕ is latitude), and H is the ocean depth. The topography in the Southern Ocean is such that very few of these contours actually pass right round Antarctica (see fig.(0.2)), and flow across these contours is strongly limited by conservation of potential vorticity (see chapter 1). Topography, then, strongly constrains the flow in a homogeneous ocean, and the nett transport is correspondingly small.

To resolve this second dilemma, stratification must be considered. Gill and Bryan (1971) included stratification in an eight-level model in which the thermohaline circulation was forced by imposing the surface temperature. Without topography the circulation was close to that predicted by the linear theory of Gill (1968), but including topography actually increased the transport. The reason for this increase was traced to a temperature difference across the ridge in the Drake Passage which was such as to produce a torque accelerating the flow eastwards. Although criticised for using a Coriolis parameter ten times too small (for the sake of numerical stability), this experiment showed two things: the effect of

effect of stratification is capable of overcoming the constraints which topography imposes on a homogeneous flow, and the effect of thermohaline circulation is to drive a flow in the same direction as that driven by the wind (as a simple picture, consider a deep, equatorward flow of cold water supported against topography. Any return flow above the topography must be frictionally balanced and will acquire a large eastward velocity before friction becomes strong enough to stop the acceleration).

The potential influence of stratification can be gauged from the steady state version of eq.(1.9) derived in chapter 1:

$$J(\Psi, f/H) = \frac{1}{\rho_0} J(E, 1/H)$$

where

$$E = g \int_{-H}^{0} z\rho \, dz$$

and Ψ is a volume flux stream function (see chapter 1 for a detailed definition of terms in these equations). If gradients of f/H are dominated by gradients of $1/H$, as is shown to be the case in most of the Southern Ocean in fig.(0.2), then density gradients produce a cross-isobath flow of order

$$\Psi_x = \frac{1}{\rho_0 f} E_x \sim \frac{gH^2}{2f} \frac{\rho_x}{\rho_0}.$$

Choosing a typical value of $\rho_x/\rho_0 = 5$ parts per thousand /1000 km, and a depth of 4000 m gives a value of $\Psi_x = 4\,\mathrm{Sv/km}$, an order of magnitude larger than what is observed in the Southern Ocean (and three orders of magnitude larger than a similar scaling for the wind-driven flow). Clearly, then, the density field is strongly constrained by interactions with topography and is the primary driving term for the circumpolar current. This second fact was nicely demonstrated recently by Olbers and Wübber (1991). They integrated two

versions of a Southern Ocean model in parallel, one forced by density, wind, non-linear terms and friction, and the second forced only by density taken from the first model. They found that the differences between the two models were small, and much smaller than differences caused by changing the forcing of the thermohaline circulation. Given a density field, then, the flow is well defined. Whether the density field is set up by wind or by thermohaline forcing is another question. Several authors have stated that the flow in the Southern Ocean is "clearly wind driven", and it is certainly true that a large flow develops in model stratified Southern Oceans driven only by wind, but the fact that the wind forcing operates through its effect on the density field surely means that a significant contribution to the flow due to thermohaline forcing should not be ruled out.

The other important factor in the dynamics of the Southern Ocean is the influence of eddies. Because of the expense of running eddy-resolving models, and for the sake of ease of analysis, most investigations of eddies have been in zonal channels with various idealised topographies. McWilliams, Holland and Chow (1978) ran a two-layer, quasigeostrophic model with a flat-bottomed, zonal channel, the same channel with a partial meridional barrier, and with a partial barrier and a topographic ridge blocking the lower part of the gap in the barrier. The presence of the barrier reduced the (wind-driven) transport from 400-600 Sv to 20-80 Sv, and dramatically altered the momentum balance. In the flat-bottomed case, the input of momentum by wind stress was transmitted to the lower layer by transient eddies and removed from the bottom layer by friction. With topography, although transient eddies were still observed, the flux of momentum into the lower layer was dominated by "standing eddies" *i.e.,* permanent deviations from a purely zonal flow. Removal of momentum from the bottom layer then took place by form drag on the ridge. Later work using a three layer version of the model (McWilliams and Chow, 1981) did not

consider topography, concentrating instead on giving a thorough description of turbulence in a flat-bottomed, zonal channel, but Treguier and McWilliams (1990) looked in detail at the interactions between eddies and topography. A number of three and four layer models were compared, with various topographies including random topographies with different scales, isolated topographic features and combinations of the two. The effect of topography was found to depend on scale, with larger features being much more efficient at producing form drag. The effect of topographic scales smaller than the baroclinic rossby radius was confined to the lower layers, in which lateral Reynolds' stress divergences were found to be anticorrelated with bottom stress, effectively smoothing the topography over these scales (although it is noted that the resulting large scale flow is quite different from what is found using a smoothed version of the topography). Again it is the standing eddies that produce most of the downward momentum transfer, although the component of the flow that actually produces a form drag represents only about 5% of the kinetic energy of the large scale flow.

Wolff, Maier-Reimer and Olbers (1991) ran a two-layer model with a variety of simple topographies, and found again that standing eddies contribute most to the downward momentum flux except in regions where contours of f/H run freely around the channel, where transient eddies become important. Their results emphasise that the form stress is not a "drag" in the normal sense, since westward flows sometimes occur at the bottom which would imply a drag force to the east, acting in concert with wind stress. In fact the form stress is found always to oppose the wind stress, although bottom friction (a more conventional type of drag) sometimes acts in the same direction as wind stress. Wolff et al. also ran an experiment with a version of the actual topography of the Macquarie Ridge, scaled down so as to meet the requirements of quasi-geostrophy. They found a balance

in which friction was unimportant, topographic stress balanced wind stress on the large scale, with interfacial form stress due to standing eddies acting to transport momentum between the layers. On the small scale, lateral Reynolds' stress divergences and interfacial form stress were anticorrelated as in (Treguier and McWilliams, 1990).

Although all these studies point to standing eddies being much more important than transient eddies in the vertical transport of momentum where there are large topographic features, one analytical model of the effect of transient eddies is worth a mention. Johnson and Bryden (1989) propose that essentially all the momentum input by wind stress is carried to the bottom by transient eddies. Assuming these eddies are due to baroclinic instability, they apply a parameterisation of the heat flux due to eddies from Green (1970) and Stone (1974), and use the geostrophic relation between heat flux and downward momentum transport to infer the heat flux from the wind stress. Setting these two heat fluxes to be equal at all depths implies, from the parameterisation, a baroclinic structure for the current. Comparison with measurements in the Drake Passage gives fairly good agreement between the two predicted heat fluxes and that which was measured, and the baroclinic structure of the current is thus given fairly well. A weakness of the model is that, contrary to observations, it predicts that there should be no surface Ekman layer.

All these models are quasi-geostrophic and wind-driven, and for the most part involve idealised topographies. Useful as they are in learning about the way in which eddies work in a recirculating flow, they do not tell us what rôle eddies play in the real Southern Ocean. The use of the quasi-geostrophic approximation means that it must be questionable whether their results can be applied to an ocean in which topography extends over a substantial fraction of the ocean depth—a situation in which the use of the quasi-geostrophic

equations cannot be formally justified. To produce reliable information about the balance of terms in the ACC, a model must be eddy-resolving and must use real topography (and cannot therefore be quasi-geostrophic). One such model is the model of Olbers and Wübber which produces the interesting result that the density field forms the primary forcing, although this model contains an artificially short Southern Ocean in which the region from 90° W to 90° E is cut out and reconnected end to end. The other major eddy-resolving, primitive equation model of the Southern Ocean, and the one with which much of this work is concerned, is FRAM—the Fine Resolution Antarctic Model. This is a model of the entire world ocean south of 24° S, and was conceived as part of the U.K. contribution to the World Ocean Circulation Experiment. It was run for the first six years in robust diagnostic mode, *i.e.,* with an additional term in the advection-diffusion equations for temperature and salinity, forcing these quantities towards the values given in the Levitus (1982) data set over a time scale of hundreds of days (for the first 2 years 160 days the time scale was 180 days at depths shallower than 140 m and 540 days below that. For subsequent times the time scale was 360 days at all depths). The aim of this procedure was to bring the model to a state that was dynamically self-consistent, while producing temperatures and salinities as close to observations as possible, thus producing a better picture of the Southern Ocean than can be gained by interpolation and extrapolation of the relatively sparse measurements that are available. Data from the model can then be used in the planning and interpretation of ocean cruise data, and, to this end, the FRAM atlas (Webb et al., 1991) was produced, showing (among other things) temperature and salinity sections from the model along the tracks of proposed WOCE cruises.

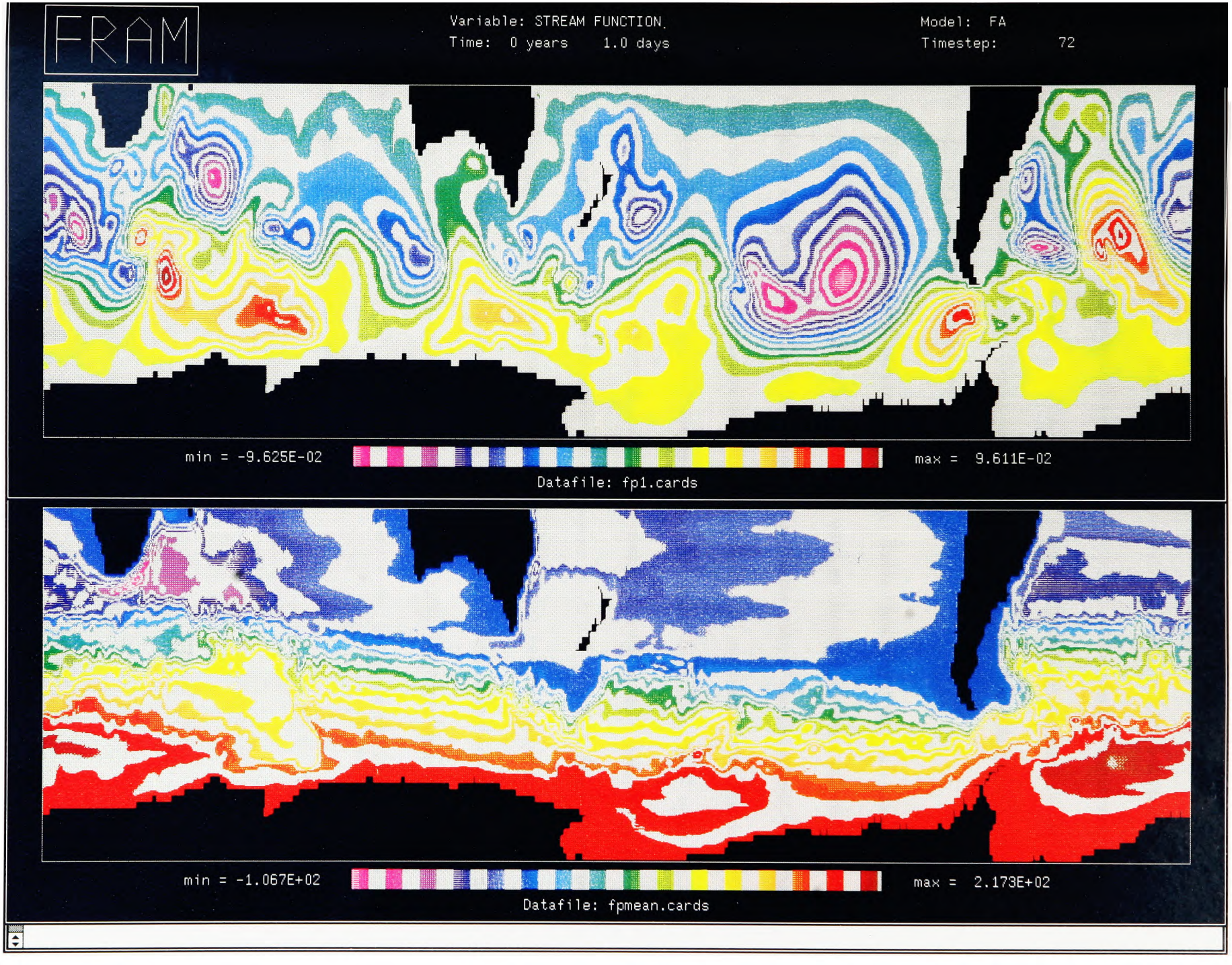
After six years, the relaxation towards Levitus was stopped in all but the top layer, and the model was integrated for a further ten years in this "free" state. A six year mean

data set from years 10 to 16 was compiled by Simon Thompson from IOSDL in Surrey, and David Stevens of the University of East Anglia. This is used as a representation of the steady state in later chapters. It is worth noting that the mean data set used is actually slightly wrong... one of the 72 data sets used in compiling the mean was taken from the wrong time. A correct version has since been calculated and differs only minutely from the incorrect version, but constraints of time mean that it is impossible to repeat the calculations presented in chapter 4 using this corrected dataset.

The model itself is based on that of Semtner (1974) and Cox (1984), and has 32 vertical levels with resolution concentrated at the top, and a horizontal grid spacing of $\frac{1}{2}^\circ$ in the east-west direction and $\frac{1}{4}^\circ$ north-south. The topography is based on the DBDB5 data set and is in most places at 1° resolution, except where steep gradients are liable to lead to instability (Killworth, 1987), where a smoothing has been applied which spreads the gradient producing gradual slopes rather than 1° blocks of constant depth. Viscosity coefficients are $2 \times 10^2\,\mathrm{m^2 s^{-1}}$ horizontally and $10^{-4}\,\mathrm{m^2 s^{-1}}$ vertically. The diffusion coefficients for temperature and salinity are $10^2\,\mathrm{m^2 s^{-1}}$ horizontally and $10^{-4}\,\mathrm{m^2 s^{-1}}$ vertically. A linear bottom friction corresponding to a spin-down time of about 50 days for a 4000 m ocean was used until eight years, eleven months into the run when a quadratic bottom friction and a biharmonic lateral viscosity were introduced. Winds from (Hellerman and Rosenstein, 1983) were used. These were introduced starting after 2 years 6 months and increasing linearly to their full annual mean value at 3 years. The winds remained at this value until the end of six years when seasonality was introduced, the winds changing linearly between their monthly mean values.

Another version of the model was later run from the sixteen year mark, incorporating a sea-ice model. Results from this are not considered in the following work.

The work in this thesis is concentrated on two aspects of FRAM. Firstly, the effect of topography on the spin-up (about the first 200 days), during which the only forcing was the interaction of density gradients and topography, and the strong constraints these jointly impose on the flow caused the stream function to change from its initial response (fig.(0.3)) to a stable, linearly growing state (fig.(2.4)) strongly resembling the final steady state (fig.(0.3)). Secondly, the steady state balance of FRAM is considered in terms of barotropic potential vorticity and density advection, with the powerful constraints topography imposes on these balances being looked at in particular. The first three chapters consider the use of topographic Rossby waves and basin modes to describe the spin-up of a model. The physics of topographic waves is discussed in chapter 1, some problems of using these waves to form a complete description of a flow are addressed in chapter 2 and chapter 3 culminates in the calculation of basin modes for a coarse resolution model of the Southern Ocean. Chapter 4 addresses the steady-state balance in terms of an investigation of the interactions between a flow forced by horizontal density gradients and the advection of density by that flow. Insights gained from this analysis are then used in chapter 5 to return to the spin-up of FRAM and provide a possible explanation for one particularly strange aspect of that spin-up.


FRAM
Variable: STREAM FUNCTION
Time: 0 years 1.0 days
Model: FA
Timestep: 72
min = -9.625E-02
max = 9.611E-02
Datafile: fp1.cards
min = -1.067E+02
max = 2.173E+02
Datafile: fpmean.cards

The influence of transient eddies and questions about interactions between eddies and topography are not addressed because, as I hope the following chapter illustrate, I believe there is enough of interest to be worked out about the steady component of the effect of topography in the Southern Ocean before adding in the complicating factor of eddies, vital though they may be to some aspects of the flow.

Finally, some general points of notation. Units throughout are S.I. except for the sverdrup (Sv), defined as $10^6\,\mathrm{m^3 s^{-1}}$, which is used as the unit of ocean volume flux. The equations are presented in Cartesian coordinates for the sake of clarity, except when finite difference formulations are taken directly from FRAM. All equations work just as well for a spherical surface, they are just rather messier. Vectors, including the ∇ operator, are two dimensional unless otherwise specified by means of a subscript (*e.g.,* ∇_3), and the two dimensional cross product is treated as a scalar: $\mathbf{a}\wedge\mathbf{b} \equiv \hat{k}\cdot(\mathbf{a}_3\wedge\mathbf{b}_3)$. Since all the work invokes the rigid lid approximation, the barotropic velocity field is defined as the vertically averaged horizontal velocity field with an associated vertical velocity to make it non-divergent, the baroclinic velocity is the remainder. Lastly, the term 'geostrophic balance' is taken to mean $f\hat{k}\wedge\mathbf{u} = -(\nabla P)/\rho_0$, and geostrophic flows are therefore horizontally divergent due to the beta effect.

Chapter 1.

The physics of topographic Rossby waves.

One of the earliest descriptions of rotating flow over topography was the Taylor-Proudman theory which states that, for steady, inviscid, slow motion in a rapidly-rotating system, the flow velocity must be independent of the along-axis direction. The flow must therefore be along contours of thickness of the fluid measured in the direction of the axis of rotation. This was predicted by Proudman (1916) and confirmed experimentally by Taylor (1923) in a famous set of experiments which demonstrated the existence of what are now called 'Taylor columns': columns of fluid which are trapped above a piece of topography while the flow around them moves almost independently. For the oceans this suggests the importance of contours of $H/sin\,\phi$ (ϕ is latitude, H is depth of the ocean) for guiding the mean flow. This parameter is also seen to be important from the unforced, inviscid version of the equations generally used to describe ocean flow, the primitive equations. These are

$$\mathbf{u}_t + \mathbf{u}\cdot\nabla\mathbf{u} + w\mathbf{u}_z + f\hat{\mathbf{k}}\wedge\mathbf{u} = \frac{-\nabla P}{\rho_0}, \tag{1.1}$$

$$P_z = -\rho g, \tag{1.2}$$

$$\nabla\cdot\mathbf{u} + w_z = 0, \tag{1.3}$$

where $f = 2\Omega\sin\phi$, Ω is the angular speed of the Earth's rotation, P is pressure, g is acceleration due to gravity, ρ is density and ρ_0 is a mean density. Here only the locally vertical component of the earth's rotation is taken to be important since the joint constraints of hydrostatic and near geostrophic balance mean that vertical velocities are much smaller than horizontal velocities.

Considering steady, geostrophic flow as in Proudman's theory, eq(1.1) becomes

$$f\hat{\mathbf{k}} \wedge \mathbf{u} = \frac{-\nabla P}{\rho_0}. \tag{1.4}$$

Taking the curl of this:

$$\nabla \wedge (f\hat{\mathbf{k}} \wedge \mathbf{u}) = \beta v - f w_z = 0$$

$$\therefore \quad w_z = \frac{\beta v}{f}.$$

Integrating over depth and setting w to zero at the top gives

$$w_B = \frac{-\beta V}{f} \qquad \text{where} \qquad V = \int_{-H}^{0} v\, dz.$$

Integrating over depth also allows us to introduce a stream function Ψ such that

$$U = -\Psi_y, \qquad V = \Psi_x$$

$$\therefore \quad w_B = \frac{-\beta \Psi_x}{f}. \tag{1.5}$$

In the case of a homogeneous fluid, or a fluid with density a function of depth only, the hydrostatic relation (1.2) gives

$$\frac{\partial}{\partial z}(\nabla P) = -g \nabla \rho = 0$$

$$\therefore \quad \frac{\partial \mathbf{u}}{\partial z} = 0,$$

$$\therefore \quad \mathbf{u}_B = \frac{\mathbf{U}}{H} = \frac{\hat{\mathbf{k}} \wedge \nabla \Psi}{H},$$

but we know that the flow at the bottom must be parallel to the bottom, so

$$w_B = -\mathbf{u}_b \cdot \nabla H = \frac{-(\hat{\mathbf{k}} \wedge \nabla \Psi) \cdot \nabla H}{H} = \frac{-1}{H} J(\Psi, H)$$

$$\therefore \quad \frac{1}{H} J(\Psi, H) = \frac{1}{f} J(\Psi, f)$$

$$\dot{.}\dot{.} \quad J(\Psi, f/H) = 0.$$

So contours of $H/\sin\phi$ or of f/H guide the flow as in the Taylor-Proudman theory. We see that, in the geostrophic approximation, the horizontal divergence necessary for bottom flow to cross depth contours must be balanced by the beta effect acting on the equatorward component of the vertically integrated flow. In the homogeneous case, the bottom flow is the same as the barotropic flow, thus providing a strong constraint on the flow direction.

If we assume that both $\mathbf{u}$ and ∇P are independent of depth, we can integrate eq.(1.1) vertically to get

$$\hat{\mathbf{k}} \wedge \nabla\Psi_t - \nabla\Psi\nabla\cdot\left(\frac{\nabla\Psi}{H}\right) + \frac{H}{2}\nabla\left(\left(\frac{\nabla\Psi}{H}\right)^2\right) - f\nabla\Psi = \frac{-H\nabla P}{\rho_0}.$$

Dividing through by H and taking the curl gives

$$\nabla\cdot\left(\frac{\nabla\Psi_t}{H}\right) + J\left(\Psi, \frac{f + \nabla\cdot\left(\frac{\nabla\Psi}{H}\right)}{H}\right) = 0. \tag{1.6}$$

So the quantity $\nabla\cdot\left(\frac{\nabla\Psi}{H}\right)$ can be changed by advection of a function of itself, f and H.

If we take the curl before dividing through by H we get

$$\nabla^2\Psi_t + J\left(\Psi, f + \nabla\cdot\left(\frac{\nabla\Psi}{H}\right)\right) = \frac{1}{\rho_0}J(P,H) + \frac{1}{\rho_0}J\left(\frac{1}{2}\left[\frac{\nabla\Psi}{H}\right]^2, H\right), \tag{1.7}$$

which is much easier to give a physical interpretation to. Consider a cylinder with a flat top and sloping bottom as in fig.(1.1). If we can take the density to be constant for momentum calculation (the Boussinesq approximation, already used in eq.(1.1) in setting the density to ρ_0), then the total angular momentum of this cylinder about its vertical axis is

$$\begin{aligned} M = \rho\int_{\text{vol}} \mathbf{r}\wedge\mathbf{u}\,dV &= \rho\int_0^{r_0} dr\, r\int_0^{2\pi} d\phi\int_{-H}^{0} \mathbf{r}\wedge\mathbf{u}\,dz \\ &= \rho\int_0^{r_0}\int_0^{2\pi} r\mathbf{r}\wedge(\hat{\mathbf{k}}\wedge\nabla\Psi)\,d\phi\,dr \end{aligned}$$

but $(\hat{\mathbf{k}} \wedge \nabla\Psi) \wedge \mathbf{r} d\phi = (\hat{\mathbf{k}} \wedge \nabla\Psi) \cdot d\mathbf{s}$

$$\therefore \quad M = \rho \int_0^{r_0} r \int_0^{2\pi r} (\hat{\mathbf{k}} \wedge \nabla\Psi) \cdot d\mathbf{s}\, dr$$
$$= \rho \int_0^{r_0} r \int_{\text{inside } r} \nabla \wedge (\hat{\mathbf{k}} \wedge \nabla\Psi)\, dA\, dr$$
$$= \rho \int_0^{r_0} r \int_{\text{inside } r} \nabla^2\Psi\, dA\, dr.$$

If $\nabla^2\Psi$ is constant over the area then this gives

$$M = \rho \int_0^{r_0} \pi r^3 \nabla^2\Psi\, dr = \frac{\pi r_0^4}{4} \rho \nabla^2\Psi.$$

So we see that $\nabla^2\Psi$ is a measure of the local angular momentum due to fluid motion. Note that we have made no assumptions other than the Boussinesq and rigid lid approximations.

Consider now the torque acting on the cylinder. For an inviscid flow, there can be no torque due to forces on the curved surface or the top, so it is all due to forces on the bottom surface and is given by

$$T = \hat{\mathbf{k}} \cdot \int \mathbf{r}_3 \wedge \mathbf{F}_3\, dA = \int \mathbf{r}_3 \cdot (\mathbf{F}_3 \wedge \hat{\mathbf{k}})\, dA.$$

Writing the bottom pressure as $P_B = P_0 + \gamma x + \epsilon y$, gives

$$\mathbf{F}_3 = P_B(\hat{\mathbf{k}} \cos\theta - \hat{\imath} \sin\theta)$$

$$\mathbf{F}_3 \wedge \hat{\mathbf{k}} = \hat{\jmath} P_B \sin\theta$$

$$\mathbf{r}_3 \cdot (\mathbf{F}_3 \wedge \hat{\mathbf{k}}) = y P_B \sin\theta$$

and

$$dA = \frac{r\, dr\, d\phi}{\cos\theta}$$

$$\therefore \quad T = \int\int y P_B \tan\theta\, r\, dr\, d\phi = -H_x \int\int y P_B r\, dr\, d\phi.$$

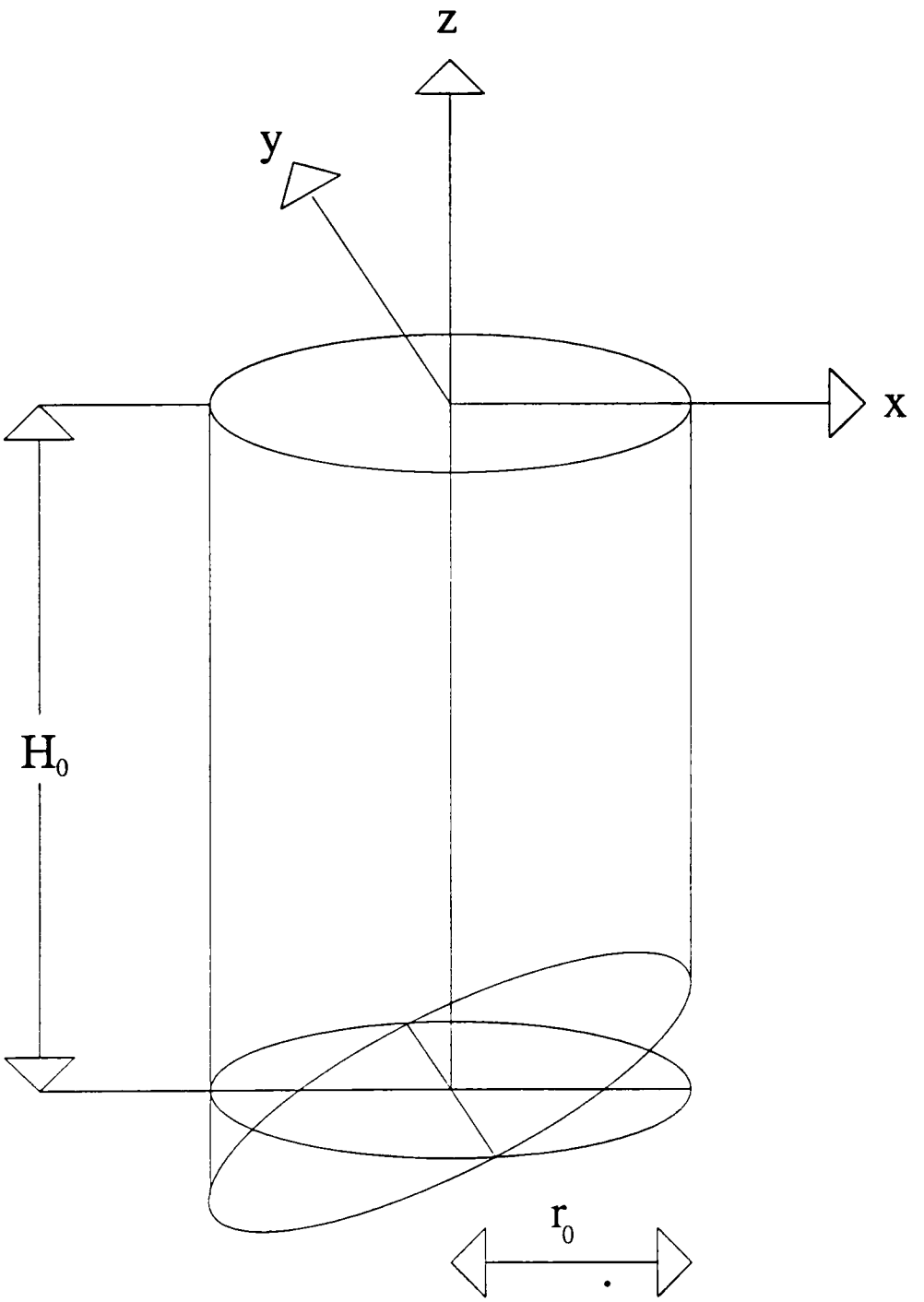


Fig.(1.1). *Cylinder of fluid used to show how the barotropic vorticity equation is related to conservation of angular momentum (see text).*

Terms in P_0 and γx integrate to zero giving

$$\begin{aligned} T &= -H_x \epsilon \int\int y^2 r \, dr \, d\phi \\ &= -H_x P_{By} \int\int r^3 \sin^2 \phi \, dr \, d\phi \\ &= -H_x P_{By} \frac{\pi r^4}{4}. \end{aligned}$$

Since the direction of the slope is arbitrary, this may be generalised to

$$T = \frac{\pi r^4}{4} \nabla P_B \wedge \nabla H = \frac{\pi r^4}{4} J(P_B, H).$$

So, $J(P_B, H)$ is to torque as $\rho_0 \nabla^2 \Psi$ is to angular momentum.

The remaining terms in eq.(1.7) may similarly be calculated in terms of advection of angular momentum (except that the terms non-linear in Ψ require the additional assumption that velocities are independent of depth). Note that depth independence is why the right hand side of eq.(1.7) uses P instead of P_B.), so we can see that eq.(1.7) represents the balance of a local form of vertical angular momentum.

To get from eq.(1.7) to eq.(1.6) we note that, in the depth independent case, ∇P may be rewritten in terms of the stream function, so substituting for ∇P, dividing by H and regrouping terms gives eq.(1.6), showing that eq.(1.6) represents a rather subtle interplay between the dynamics of depth independent flow with conservation of angular momentum.

If we look at the linearised version of either eq.(1.6) or eq.(1.7), with a flat bottom, we get

$$\nabla^2 \Psi_t + \beta \Psi_x = 0, \tag{1.8}$$

an equation which supports Rossby waves. It can be seen that any north-south flow produces a change in $\nabla^2 \Psi$ which induces a circulation and thus alters the flow. A sinusoidal wave of north-south motion is thus propagated with phase speed always to the west. The

physical mechanism for this depends on the curvature of the earth. As a cylinder of fluid moves northwards, there is never any torque acting about its vertical axis. The horizontal torque necessary to twist the cylinder as it moves over the earth's curved surface is, in the rigid lid approximation, provided by the boundary condition in the form of pressure gradients implicitly applied to keep the surface from moving vertically. (If a free surface is allowed then vertical motion occurs which allows the hydrostatic relation to provide this torque. This vertical motion alters the torque necessary since a stretching of the water column changes its moment of inertia thus providing another way of absorbing angular momentum changes. The fact that only small vertical motions can produce large pressure changes means that this is only important over length scales greater than the Rossby radius $a = \sqrt{gH}/f$, which is of order 2000 km in the deep ocean.) The vertical component of angular momentum is made up of two components: f and $\nabla^2\Psi$, so as the cylinder is twisted to follow the curvature of the earth, f changes and in the absence of vertical torques, $\nabla^2\Psi$ must change to compensate.

When the depth varies, there are vertical torques acting on the fluid due to pressure variations at the bottom. Angular momentum is thus not conserved within the fluid since some is transferred to the solid earth beneath, but the relationship between pressure gradients and fluid flow allows eq.(1.6) to be derived. In linearised form:

$$\nabla \cdot \left(\frac{\nabla \Psi_t}{H} \right) + J(\Psi, f/H) = 0 \,,$$

which has a very similar form to eq.(1.8). This tells us that any flow across contours of f/H produces a change in $\nabla \cdot \left(\frac{\nabla\Psi}{H}\right)$ which will induce a circulation allowing propagation in a sort of equivalent-westward direction: the direction with higher values of f/H to the right. The physical mechanism supporting these waves is rather more complicated since there is

a balance between time variation, advection of planetary angular momentum and pressure torques. It is tempting to think of the squashing or stretching of water columns as they move up or down slopes producing a change in rotation rate due to simple conservation of angular momentum, but this is not the whole story, since there are also pressure torques acting on the columns, whose form is related to the form of the flow.

Rossby waves have been known about for a long time (see for example Longuet-Higgins, 1964). If the rigid lid approximation is not made, the full dispersion relation may be derived:

$$\omega = \frac{-\beta k}{k^2 + l^2 + 1/a^2}$$

where a is the Rossby radius. The rigid lid version may be derived by setting g, and therefore a, to infinity. The equation above gives a maximum for ω at $k = -1/a$ corresponding to a wavelength for deep ocean of about 13000 km. For wavelengths of this order or larger the rigid lid approximation produces large errors, but when the wavelength gets down to the Rossby radius, *i.e.,* about 2000 km, the rigid lid approximation is right to less than five percent. For wavelengths less than $2\pi a$ (about 13000 km), waves have westward phase velocity and eastward group velocity in both cases, but for longer wavelengths the free surface waves have westward group velocity and tend towards being non-dispersive, whereas the rigid lid waves continue to have eastward group velocity.

Rossby waves are important in the spin-up of a flat-bottomed ocean and the formation of a western boundary current. As long as viscous and non-linear terms are small, this boundary current gets narrower, so eventually friction and/or non-linear terms must become important. A good description of the spin-up of a flat-bottomed ocean by Rossby waves is given in Anderson and Gill (1975).

Topographic Rossby waves, governed by eq.(1.6), were first considered by Robinson (1964) in the context of trying to explain anomalies in sea surface height and pressure measurements off the east Australian coast (Hamon, 1963). Robinson considered a simple rectilinear discontinuity in depth which supported a single type of topographic wave, and Mysak (1967) extended the analysis to a linear slope followed by a discontinuity, showing the possibility of an infinity of modes of waves with different numbers of nodes running parallel to depth contours. The free surface case was first considered by Longuet-Higgins (1968) who showed that Robinson's single mode becomes a double Kelvin wave in the long wavelength limit, and the same year, Buchwald and Adams (1968) found a particularly simple solution for a rigid lid case with exponential topography. All of these waves have frequencies less than f and travel with higher values of f/H to the right. Huthnance (1975) showed that these properties are generic to any such waves on rectilinear, monotonic topography, and also showed that if H_x/H is everywhere bounded, there is a maximum in the dispersion relation (for the free surface case). Mysak (1980) gives a good summary of this work.

Having looked in some detail at the equations for the depth independent case, it is now easier to understand the situation where density is allowed to vary. The argument equating torque plus advection with rate of change of local angular momentum still stands, so the linearised eq.(1.7) is unaltered except that P should be replaced by P_B. Equation (1.6), however, relies on the fact that the pressure gradient which drives the vertically averaged flow is the same as that which provides the torque. This is no longer the case, so a correction should be expected. The two pressure gradients are related, however, by the hydrostatic relation, so we should be able to express this correction in terms of density alone.

Linearising and integrating eq.(1.1) vertically gives

$$\begin{aligned}\hat{\mathbf{k}} \wedge \nabla\Psi_t - f\nabla\Psi &= -\frac{1}{\rho_0}\int_{-H}^{0} \nabla P \, dz \\ &= -\frac{1}{\rho_0}\nabla\int_{-H}^{0} P \, dz + \frac{P_B}{\rho_0}\nabla H \, .\end{aligned}$$

Taking the curl gives

$$\nabla^2\Psi_t + \beta\Psi_x = \frac{1}{\rho_0}J(P_B, H)$$

as expected. Dividing by H before taking the curl gives

$$\begin{aligned}\nabla\cdot\left(\frac{\nabla\Psi_t}{H}\right) + J(\Psi, f/H)) &= \frac{1}{\rho_0}J\left(\int_{-H}^{0} P \, dz, \frac{1}{H}\right) + \frac{1}{\rho_0 H}J(P_B, H) \\ &= \frac{1}{\rho_0}J\left(\int_{-H}^{0} P - P_B \, dz, \frac{1}{H}\right).\end{aligned}$$

But, using the hydrostatic relation, $P - P_B = -g\int_{-H}^{z}\rho \, dz$,

$$\dot{\nabla}\cdot\left(\frac{\nabla\Psi_t}{H}\right) + J(\Psi, f/H) = \frac{1}{\rho_0}J(E, 1/H) \tag{1.9}$$

where

$$E = -g\int_{-H}^{0}\int_{-H}^{z}\rho \, dz = g\int_{-H}^{0} z\rho \, dz.$$

E is here the potential energy of a fluid column measured relative to the surface and is negative definite unless ρ is defined in such a way as to allow negative values (note that an arbitrary function of z may be subtracted from ρ without altering the forcing in eq.(1.9)). The forcing term, $\frac{1}{\rho_0}J(E, 1/H)$, is commonly known as the JEBAR term (Joint Effect of Baroclinicity And Relief), and neatly expresses the interaction between density and depth variations. It is particularly useful because E depends only on density, and therefore changes on the time scale of density advection, which is rather slower than the time scale for pressure changes. There are a number of ways of interpreting this term.

Given that $-HJ(\int_{-H}^{0} P\,dz, 1/H) = J(\overline{P}, H)$ is the torque which would be felt if the flow were barotropic, and $-H^2 J(P_B, 1/H)$ is the actual torque, we see that $HJ(E, 1/H)$ is the correction to the torque due to baroclinicity.

For the steady state case, we have

$$f\hat{\mathbf{k}} \wedge \mathbf{u}_B = \frac{-1}{\rho_0}(\nabla P)_B,$$

and, taking the cross product with ∇H,

$$-f\mathbf{u}_B \cdot \nabla H = f w_B = \frac{-1}{\rho_0} J(P_B, H),$$

so

$$\frac{1}{\rho_0} J(E, 1/H) = \frac{f\tilde{w}_B}{H}$$

where $\tilde{w} = w - w_{\text{barotropic}}$. This shows how the forcing term represents the fact that the horizontal divergence due to the beta effect is balanced by the total w field and not just the barotropic w. This is an important point which will be returned to in chapter 4.

With all these different interpretations in mind, it can be seen that eq.(1.9) has a robust physical basis, and despite the fact that topographic Rossby waves are an essentially barotropic phenomenon, adding baroclinicity to the problem only alters its character over advective time scales, and not at all in the steady state, given that there would be some other sort of forcing on the right hand side of eq.(1.9) anyway. It is with this in mind, and with the model of the spin-up of a flat bottomed ocean by Rossby waves as a guide that the possible use of topographic Rossby waves in describing basin-wide flows is considered.

Chapter 2.

Topographic Basin Modes.

Before the FRAM run was started, several short, coarse resolution runs were made to check that the model was working. One of these runs started with velocities zero everywhere but the temperature and salinity (and therefore density and E) as given by Levitus (1982). In this case, two large, two-lobed structures were formed around the most prominent pieces of topography (the Kerguelan and New Zealand plateaux). These had initial flows of order 50 Sv and rotated around the topography anticlockwise with a period of order one day, growing in amplitude as they did so until the model run had to be terminated. All these features, apart from the growth with time, suggested topographic Rossby waves (shelf waves), so a simple model was designed to test this idea.

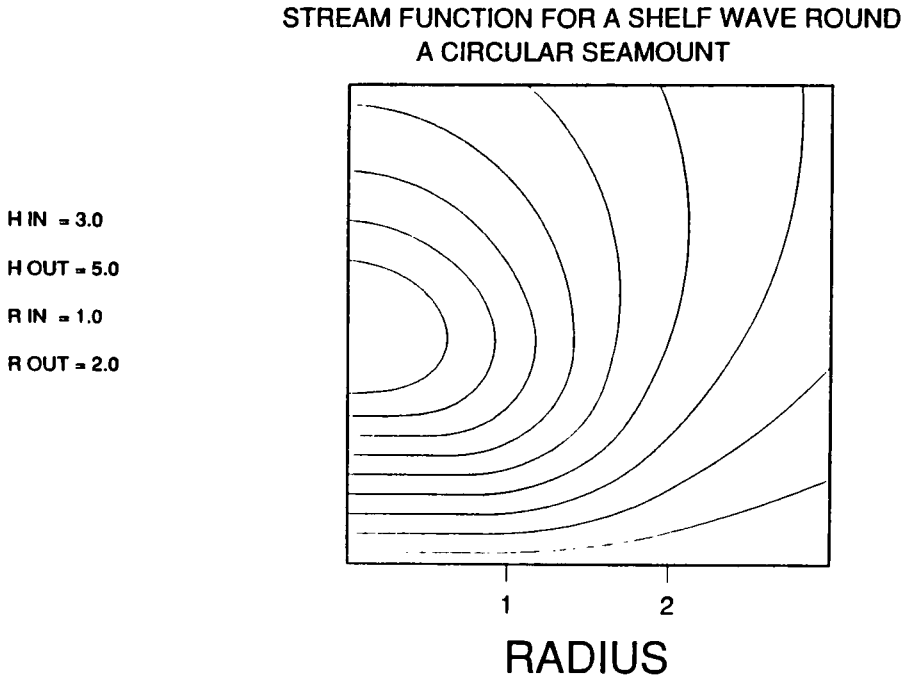
Considering a radially symmetrical topography, and a wave-like form for the stream function, $\Psi = R(r)\exp(i(n\theta - \omega_n t))$, eq.(1.9) becomes

$$r^2 R_{rr} + r R_r \left(1 - \frac{r H_r}{H}\right) - R\left(n^2 + \frac{f}{\omega_n}\frac{r H_r}{H}\right) = 0$$

for the f-plane, homogeneous case. Substituting $x = \ln r$ converts this to

$$R_{xx} + R_x \left(1 - \frac{H_x}{H}\right) - R\left(n^2 + \frac{f}{\omega_n}\frac{H_x}{H}\right) = 0,$$

which is the equivalent equation for a topography dependent on x only and stream function $\Psi = R(x)\exp(i(ny - \omega_n t))$, so the results discussed in chapter 1 for rectilinear topography can easily be extended to the case of radially symmetrical topography. In particular, the simple solution of Buchwald and Adams (1968) for exponential slopes translates into an



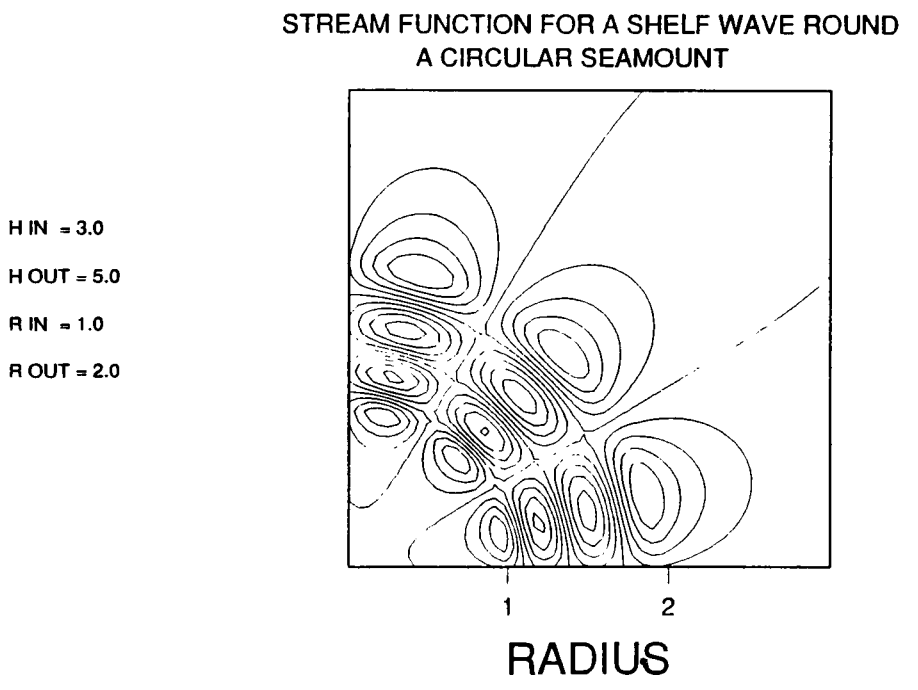


Fig.(2.1). *Two typical topographic waves around a circular seamount. The seamount has a flat region in the centre with depth H-in, a region in which depth is proportional to a power of radius between radii R-in and R-out and a flat region with depth H-out outside that. For these modes, H-in/H-out = 3/5 and R-in/R-out = 1/2. Only one quadrant is shown since the remaining quadrants are easy to reconstruct due to the symmetry of the modes. The first mode shown is the gravest mode with azimuthal wavenumber 1 and no nodes along the slope. The second mode has azimuthal wavenumber 6 and 3 nodes along the slope. The whole pattern rotates anticyclonically with a frequency given by the dispersion relation shown in fig.(2.2), and only those modes without any nodes along the slope continue to exist when the slope becomes vertical, the others converging to zero frequency.*

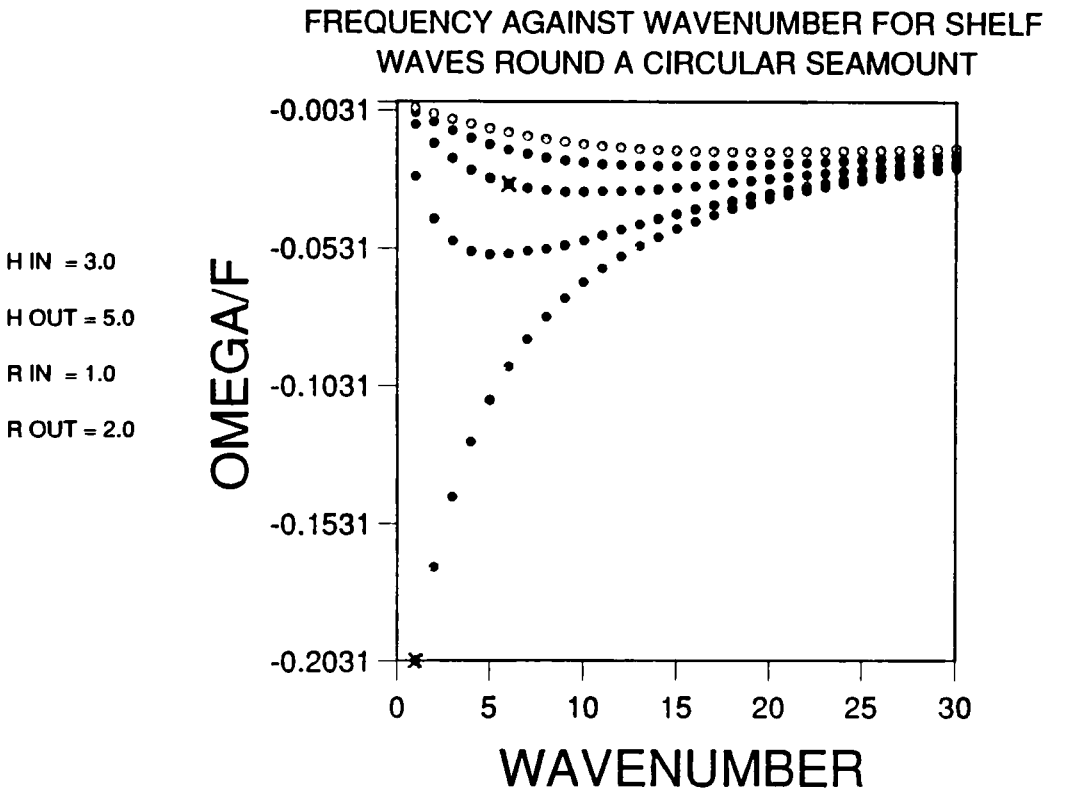


Fig.(2.2). *Dispersion relation for shelf wave modes around the circular seamount described in fig(2.1). The two modes shown in fig(2.1) are marked with crosses.*

equally simple solution for a radially symmetrical slope on which the depth is proportional to a power of r. These are sinusoidal in the azimuthal coordinate with different numbers of nodes on the sloping region at constant radius. Two examples are shown in fig.(2.1). These waves, together with an arbitrary function of depth, can easily be shown to form a complete, orthogonal set over the sloping region, so any flow pattern can be represented in sloping regions as a superposition of shelf waves plus a function of depth alone. A typical set of dispersion relations for these waves is shown in fig.(2.2). The only difference from Buchwald and Adams (1968) is that they are forced to be discrete by the periodic boundary conditions implicit in radially symmetrical geometry.

Topographic Modes.

This completeness prompted the idea of representing the solution to the forced equation as a sum of the spatial portions of the wave modes each multiplied by a function of time chosen to satisfy the forcing. For the more general situation, the basin wave modes are chosen to satisfy

$$-i\omega_n \nabla \cdot \left(\frac{\nabla \phi_n}{H} \right) + J(\phi_n, f/H) = 0 \tag{2.1}$$

and the hope is to represent the stream function as

$$\Psi = \sum_n \phi_n T_n(t)$$

so that Ψ satisfies the forced equation

$$\nabla \cdot \left(\frac{\nabla \Psi_t}{H} \right) + J(\Psi, f/H) = F. \tag{2.2}$$

Before proceeding with this analysis, however, it is necessary to specify the boundary conditions. For the case of a simple, closed basin with no islands these are obviously

$\Psi = \text{const}$, $\phi_n = \text{const}$ on the boundary, where the constant may for simplicity be taken as zero. When there are several unconnected boundaries, however, the conditions are $\Psi = \psi_\delta(t)$, $\phi_n = c_{n,\delta}$ for each boundary δ. The values of ψ_δ and $c_{n,\delta}$ must be specified by some auxiliary condition, and to obtain this condition we have to go back to the momentum equation. The reason for this is that, when there are islands, we have lost some information in taking the curl of the momentum equations. Since $\nabla \wedge \nabla P = 0$, the pressure field has dropped out of the equation, but the resulting flow must still be driven by an implied pressure field. When there are no islands, this pressure field can always be reconstructed, since any field with zero curl may be written as the gradient of some potential, but when there are islands where the field is undefined, that potential may not be single-valued as we must insist that the pressure field is. To gain the auxiliary condition, then, we must integrate $\nabla P \cdot d\mathbf{s}$ around each island and set the resultant integral to zero.

The vertical integral of eq.(1.1), linearised and with some forcing on the right hand side, gives

$$\begin{aligned}
\hat{k} \wedge \nabla\Psi_t - f\nabla\Psi &= \frac{1}{\rho_0}\int_{-H}^{0} \nabla P \, dz + \mathbf{G}' \quad \text{where} \quad \mathbf{G}' = \int_{-H}^{0} \text{forcing}\, dz \\
&= \frac{1}{\rho_0}\int_{-H}^{0} \nabla(P - P_B)\, dz + \frac{H}{\rho_0}\nabla P + \mathbf{G}' \\
&= \frac{\nabla E}{\rho_0} + \frac{H}{\rho_0}\nabla P_B + \mathbf{G}' \\
\therefore \quad \frac{\hat{k} \wedge \nabla\Psi_t}{H} - \frac{f}{H}\nabla\Psi &= \frac{\nabla P_B}{\rho_0} + \frac{(\nabla E/\rho_0 + \mathbf{G}')}{H}.
\end{aligned}$$

Integrating around a closed contour of constant Ψ (*e.g.,* a boundary) eliminates the bottom pressure giving

$$\begin{aligned}
\oint \frac{\nabla\Psi_t}{H} \wedge d\mathbf{s} = \oint \mathbf{G} \cdot d\mathbf{s} \qquad &\text{where} \qquad \mathbf{G} = \frac{(\nabla E/\rho_0 + \mathbf{G}')}{H} \\
&\text{and} \qquad \nabla \wedge \mathbf{G} = F.
\end{aligned} \tag{2.3}$$

This gives the boundary condition for Ψ. The wave modes are solutions of the unforced, barotropic equation so they must obey a similar boundary condition with zero on the right hand side, and complete boundary conditions for the wave modes are given by

$$\oint \frac{\nabla\phi_n}{H} \wedge d\mathbf{s} = 0, \qquad \phi_n = \text{const} \qquad \text{on each boundary.} \tag{2.4}$$

The constant on the second condition is defined implicitly by the necessity to obey the first condition. Conceptually, a wave mode can be found by choosing a value for the constant, solving the equation with that as a boundary condition, calculating $\oint \frac{\nabla\phi_n}{H} \wedge d\mathbf{s}$, repeating the whole process with a different constant until the first condition is satisfied.

It is now possible to demonstrate the orthogonality of the wave modes. Taking the sum of $(2.1)_n \times \phi_m^* + (2.1)_m^* \times \phi_n$ gives

$$J(\phi_m^*\phi_n, f/H) + i\omega_m^*\phi_n\nabla\cdot\left(\frac{\nabla\phi_m^*}{H}\right) - i\omega_n\phi_m^*\nabla\cdot\left(\frac{\nabla\phi_n}{H}\right) = 0.$$

Integrating over area, the Jacobian integrates to zero since $\phi_m^*\phi_n$ is constant on each boundary, and the other two terms can be written

$$i(\omega_n - \omega_m^*)\int \frac{\nabla\phi_n\cdot\nabla\phi_m^*}{H}\,dA + i\omega_m^*\int\nabla\cdot\left(\frac{\phi_n\nabla\phi_m^*}{H}\right)dA - i\omega_n\int\nabla\cdot\left(\frac{\phi_m^*\nabla\phi_n}{H}\right)dA = 0.$$

Using the divergence theorem and boundary conditions (2.4) shows that the second and third terms integrate to zero, so

$$i(\omega_n - \omega_m^*)\int \frac{\nabla\phi_n\cdot\nabla\phi_m^*}{H}\,dA = 0.$$

Putting $m = n$ shows that $\omega_n = \omega_n^*$ and therefore ω_n is real. We see that it is possible to normalise the modes so that

$$\int \frac{\nabla\phi_n\cdot\nabla\phi_m^*}{H}\,dA = \delta_{nm}. \tag{2.5}$$

Similarly, taking (2.1) $\times \phi_m^*$ and integrating gives

$$\int \phi_m^* J(\phi_n, f/H)\, dA + i\omega_n \int \frac{\nabla\phi_n \cdot \nabla\phi_m^*}{H}\, dA = 0$$

$$\therefore \quad \int \phi_m^* J(\phi_n, f/H)\, dA = -i\omega_n \delta_{mn} \tag{2.6}$$

giving an orthogonality relationship between the modes. This is not quite a self-adjoint system, but is almost as simple since the adjoint of a function is simply its complex conjugate.

If we now substitute a superposition of wave modes

$$\Psi = \sum_n \phi_n T_n(t)$$

into eq(2.2), multiply through by ϕ_m^* and integrate over area, we get

$$\int \phi_m^* \sum_n \frac{dT_n}{dt} \nabla \cdot \left(\frac{\nabla\phi_n}{H}\right) dA + \int \phi_m^* J\left(\sum_n T_n\phi_n, \frac{f}{H}\right) dA = \int \phi_m^* F\, dA.$$

Using eq.(2.5) and eq.(2.6) reduces this to

$$\frac{dT_m}{dt} + i\omega_m T_m = -\int \phi_m^* F\, dA = -Q_m \quad \text{(say)}. \tag{2.7}$$

From this equation, we can see how each component of a superposition of wave modes responds to a time dependent forcing. Equation (2.7) can be integrated to give

$$T_m = T_{m0} + e^{-i\omega_m t} \int_0^t Q_m e^{i\omega_m t'}\, dt'. \tag{2.8}$$

For the situation in the coarse resolution run, T_{m0} was zero and $\mathbf{G}'$ was zero so the forcing was given by $\frac{1}{\rho_0}\nabla \wedge \left(\frac{\nabla E}{H}\right) = \frac{1}{\rho_0} J(E, 1/H)$, which was constant after time zero as set by the Levitus data. This makes Q_m constant in time and therefore

$$T_m = \frac{(e^{-i\omega_m t} - 1)}{i\omega_m} Q_m,$$

producing a constant component and a wave-like component of equal amplitude. Choosing typical density gradients to work out the forcing of the gravest mode of the circular seamount model, with height and radius chosen to match Kerguelan, gives a flow of order 50 Sv, so the initial amplitude and character of the waves is explained quite nicely. The fact that they grew in amplitude after the initial time remains unexplained, suggesting that advection may be important in causing this instability.

In later runs, the model was started with density depending only on depth, and the density was relaxed exponentially towards Levitus values with a time constant of 180 days for the top 140 m and 540 days below that. This produces an initially linear growth in the forcing from zero at time zero, so writing $Q_m = q_m t$ in eq.(2.8) gives

$$T_m = \frac{q_m}{i\omega_m}\left(t + \frac{e^{-i\omega_m t} - 1}{i\omega_m}\right),$$

showing an initial oscillation of the same character as before, with a linear growth superimposed which becomes larger than the oscillatory part after a time $t = 1/\omega_m$. A typical time scale for a topographic mode is a few days, so the amplitude of oscillations excited by a relaxation time scale of hundreds of days is very much smaller than the equivalent for effectively instant relaxation, and in FRAM we would expect to see the sum of wave modes simply growing linearly after a few days. This is indeed what happens. The linearity of the growth is illustrated in a plot of nett transport through Drake Passage versus time (fig.2.3), and the fact that growth is everywhere linear is shown by the fact that the stream function pattern (magnitude apart) is almost identical at 10 days and at 200 days (fig 2.4). Topographic modes therefore seem to provide a promising description of the spin-up of FRAM.

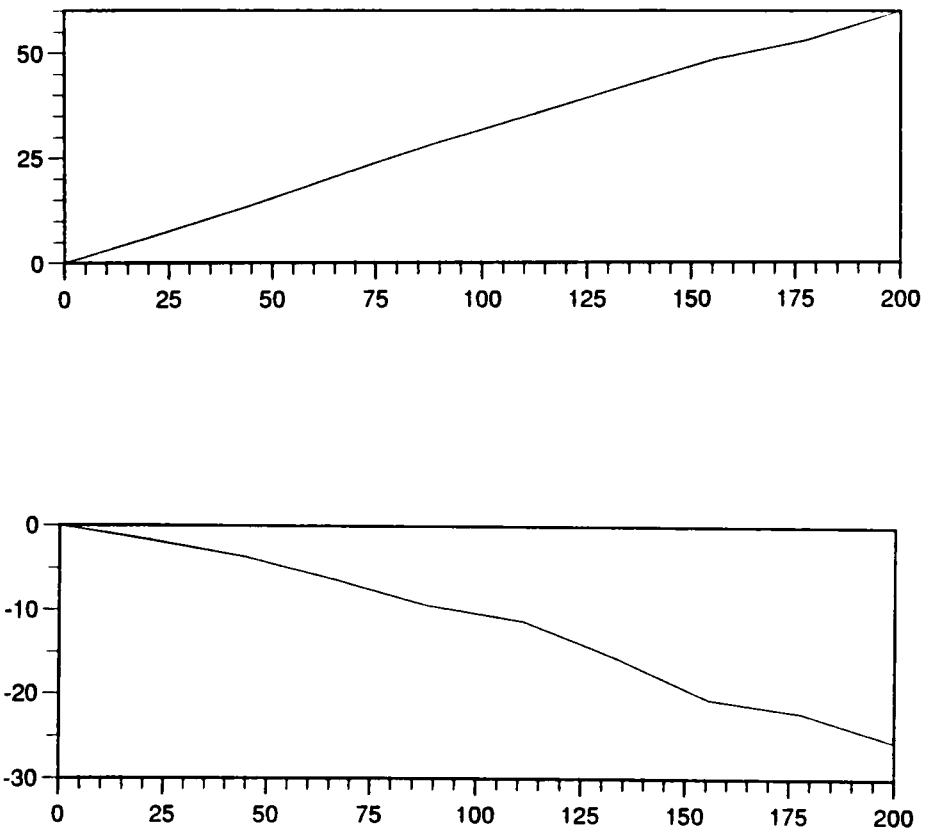
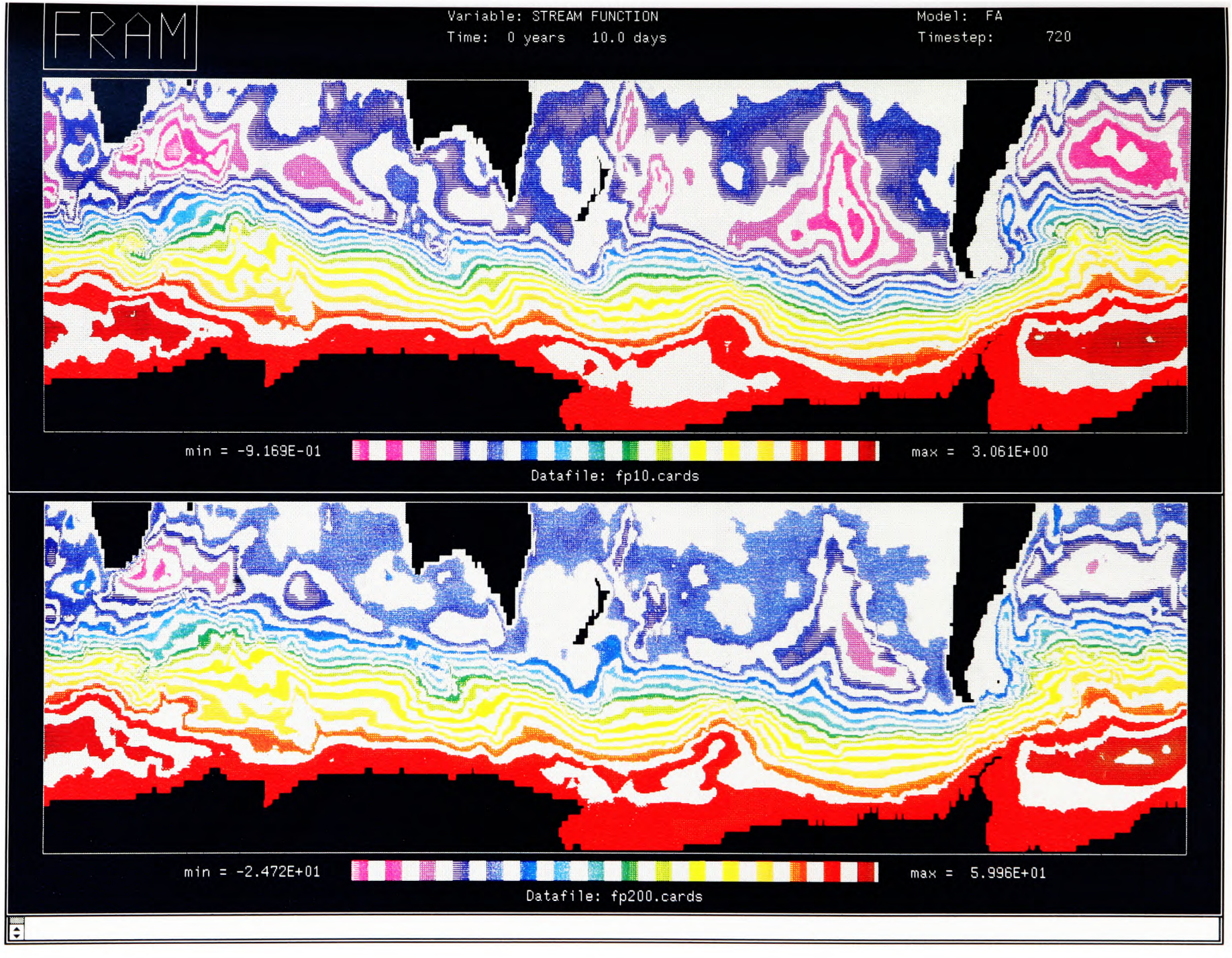


Fig.(2.3). *Graph showing how the maximum and minimum values of stream function grow with time over the first 200 days of FRAM. The stream function is in sverdrups and the time in days. Both components change remarkably linearly over this period.*

Following page: Fig(2.4). *Two plots of stream function in FRAM, after 10 days (top) and after 200 days (bottom). The scales are in sverdrups and, although the scales are different, the patterns are almost indistinguishable showing that linear growth is occurring everywhere.*


FRAM
Variable: STREAM FUNCTION
Time: 0 years 10.0 days
Model: FA
Timestep: 720
min = -9.169E-01
max = 3.061E+00
Datafile: fp10.cards
min = -2.472E+01
max = 5.996E+01
Datafile: fp200.cards

The orthogonality relationship may also be used to put a bound on the value of ω_n. Following Johnson (1988), eq.(2.6) gives us

$$\begin{aligned} i\omega_n &= -\int \phi_n^* J(\phi_n, f/H)\, dA \\ &= -\int J(\phi_n, f\phi_n^*/H)\, dA + \int \frac{f}{H} J(\phi_n, \phi_n^*)\, dA \\ &= \int \frac{f}{H} J(\phi_n, \phi_n^*)\, dA \\ &= \int \frac{f}{H} \nabla\phi_n \wedge \nabla\phi_n^*\, dA \end{aligned}$$

but, if f has amplitude everywhere less than $f_{\max}$, and is single-signed,

$$\left| \int \frac{f}{H} \nabla\phi_n \wedge \nabla\phi_n^*\, dA \right| \leq f_{\max} \int \frac{|\nabla\phi_n|^2}{H}\, dA,$$

where

$$|\nabla\phi|^2 = \nabla\phi \cdot \nabla\phi^*$$

and eq.(2.5) tells us that

$$\int \frac{|\nabla\phi_n|^2}{H}\, dA = 1$$

so we have

$$|i\omega_n| = |\omega_n| \leq f_{\max}.$$

Completeness.

Although the modes seem to provide a promising description of the spin-up of FRAM, it is simple to show that they are incomplete in their representation of barotropic potential vorticity (BPV). Integrating eq.(2.1),

$$\int \nabla \cdot \left(\frac{\nabla\phi_n}{H} \right) dA = \frac{1}{i\omega_n} \int J(\phi_n, f/H)\, dA,$$

so in any area over which the Jacobian term integrates to zero, the topographic modes have no nett BPV. This includes regions bounded by any number of closed contours of ϕ_m or of f/H, and, by the boundary condition on ϕ_m, any bounded ocean is such a region. Also included are any regions bounded by two contours of f/H and two of ϕ_m where one of the quantities takes the same value on both contours (see fig.2.5). This last constraint means that in any region where contours of f/H run out from a boundary and back into the same boundary, topographic modes cannot represent a nett BPV.

The obvious fix to this problem comes from the mode that has been overlooked, satisfying $\omega = 0,\ J(\phi, f/H) = 0$. This is simply the flow along contours of f/H which satisfies the unforced, steady state equations. If we add in a function $\phi_0(\frac{f}{H}, t)$, we can account for all the input BPV quite easily. Over an area bounded by a contour of f/H we simply put

$$\int \nabla \cdot \left(\frac{\nabla \phi_{0t}}{H} \right) dA = \oint \frac{\phi_{0nt}}{H}\, ds = \int F\, dA,$$

but, since

$$\phi_0 = \phi_0(\frac{f}{H}, t),$$

we have

$$\phi_{0n} = \left| \nabla \frac{f}{H} \right| \frac{\partial \phi_0}{\partial \frac{f}{H}}$$

$$\therefore \quad \frac{\partial}{\partial t} \frac{\partial \phi_0}{\partial \frac{f}{H}} \oint \frac{\left| \nabla \frac{f}{H} \right|}{H}\, ds = \int F\, dA,$$

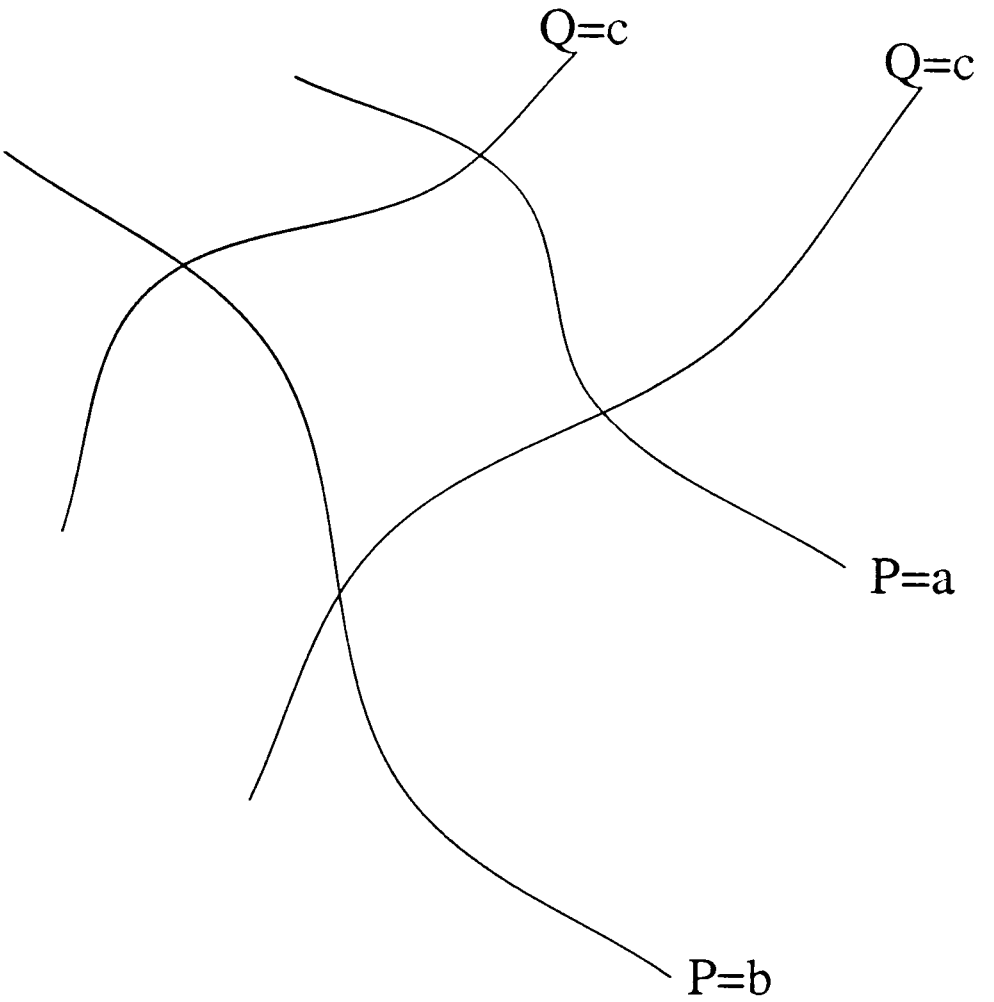


Fig.(2.5). *The most general region over which J(P,Q) integrates to zero. This is true of the region enclosed by the four contours if either P or Q takes the same value on two nonadjacent contours.*

which specifies ϕ_0 if boundary values are given. The remainder of the stream function, $\Psi' = \sum_m T_m \phi_m$, now satisfies

$$\nabla \cdot \left(\frac{\nabla \Psi'_t}{H} \right) + J(\Psi', f/H) = F - \nabla \cdot \left(\frac{\nabla \phi_{0t}}{H} \right).$$

This works as long as the boundary is one or two contours of constant f/H, as is the case in the real ocean since the boundary is where $H \to 0$. Most models, however, have vertical walls defining the boundaries and therefore allow contours of f/H to intersect the boundary. In that case the boundary conditions disallow any finite function of f/H, but an infinitesimal function, dropping to a constant value at the boundaries would be adequate to absorb any nett BPV in a delta function of BPV generated at the boundary, without involving any finite flow. This, of course, is the situation in which a western boundary flow would be generated, and for a steady state to be reached some sort of dissipation must be invoked. The continual input of BPV must be destroyed somewhere since there can be no nett advection out of the region.

A useful insight into the form of the final steady state can be gained by using an analogy due to Welander (1968). Using a Rayleigh-type approximation to Ekman friction driven by the barotropic velocities, he obtains for the steady state version of eq.(2.2),

$$J(\Psi, f/H) = F - \nabla \cdot (\chi \nabla \Psi), \tag{2.9}$$

where $\chi = Df/H^2$ and $D = \sqrt{\mu/2\rho_0 f}$ is the Ekman depth for vertical coefficient of viscosity μ. Introducing a fictitious velocity field with stream function f/H, this can be rewritten as

$$\frac{D\Psi}{Dt} - \nabla \cdot (\chi \nabla \Psi) = -F$$

where $D\Psi/Dt = J(f/H, \Psi)$, representing the advection of Ψ by the fictitious flow field. If Ψ is now thought of as temperature, χ as a coefficient of heat conduction and F as rate of heating, solutions to this equation become much easier to visualize. For example, in a flat-bottomed, beta-plane, rectangular ocean, the fictitious flow is a uniform velocity from east to west. The fluid enters with temperature given by the boundary condition and is heated and cooled as it travels across the basin. If the nett amount of heating along a particle track (f/H contour) is non-zero then the fluid will approach the western boundary at a temperature different from the boundary temperature, so a boundary layer is formed whose thickness is determined by the competing effects of diffusion of heat away from the boundary and advection of heat towards the boundary. This is an analogue of the classic Stommel western boundary layer necessary when balancing BPV input along a contour of f by advection of planetary vorticity by a north-south flow would produce a nett mass flux across the contour, so a frictional return flow is necessary.

This analogy also highlights the importance of friction in regions of closed contours of f/H. Any BPV input into these regions cannot be balanced by advection across contours of f/H without producing a nett mass flux out of the region. In the thermal analogy, closed contour regions are regions where the flow is recirculating, so any nett heating simply produces a continual increase in temperature until diffusion becomes important enough to allow the heat to escape.

The importance of critical contours separating regions of f/H contours also becomes clear. These contours pass through a saddle point where $|\nabla \frac{f}{H}| = 0$. These are points in the analogy where the fluid is stationary so the temperature is very sensitive to the forcing. Also, contours on either side of the critical contour have very different paths and so should

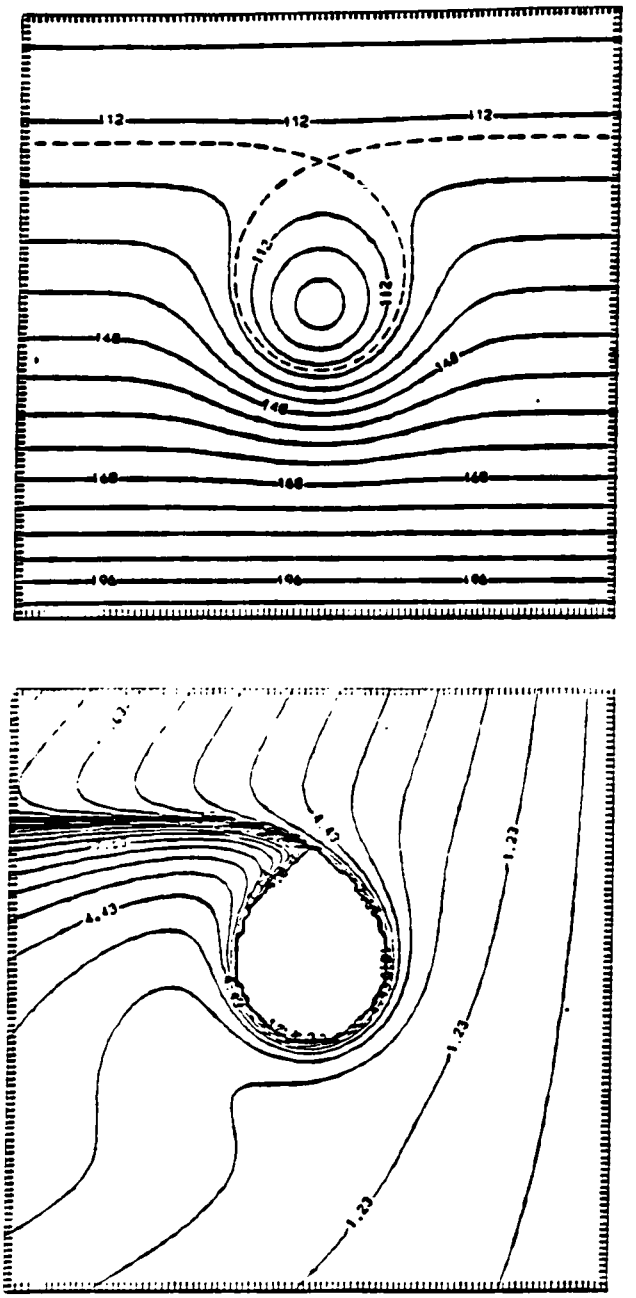


Fig.(2.6). *Top: contours of H/f chosen for a layer of fluid which is driven by an imposed vertical velocity field at the top interface. Bottom: resulting interfacial height field, showing a double jet along part of the critical contour. Contours are not shown in the region of closed contours of H/f since the flow does not reach a steady state in this region, without friction. Taken from Straub (1990).*

be expected to have very different temperatures anywhere 'downstream' of where they first separate, hence producing discontinuities (in the non-diffusive limit) of temperature at the critical contour. An example is found in Straub (1990) for the steady filling of an ocean basin. This produces an equation of the same form as eq.(2.9) without friction. Figure (2.6) shows the contours of H/f in this case, and the resulting steady state flow, where it exists. The flow in the closed contour region can never reach steady state without higher order effects (friction or non-linear terms) being taken into account. The discontinuity in stream function is only on the "downstream" side of the saddle point, as predicted.

We see, therefore, that singular behaviour (resulting in frictional boundary layers) does occur where contours of f/H intersect boundaries, but only on the western boundary or its equivalent. The infinitesimal function of f/H proposed previously produces singularities in the forcing of topographic modes at both ends of the contours, and therefore introduces small scales unnecessarily into the forcing. If topographic modes are to be used as a description of the flow, the introduction of small scale forcing is to be avoided, since it will increase the number of modes needed to adequately resolve the flow. It seems sensible, then, to try to find a way of supplying nett BPV which doesn't involve introducing these small scales.

The next simple way to absorb the nett BPV is to use a function satisfying

$$\nabla \cdot \left(\frac{\nabla \Psi_{0t}}{H} \right) = F,$$

and also satisfying the forced boundary conditions. This leaves the remainder of the solution, Ψ', to satisfy

$$\nabla \cdot \left(\frac{\nabla \Psi'}{H} \right) + J(\Psi', f/H) = -J(\Psi_0, f/H).$$

This looks a promising candidate to be expressed as a sum of wave modes, since the limitations on wave modes were due to the fact that the Jacobian term integrates to zero over a number of regions, and this time the forcing suffers from exactly the same limitations. There is a rather curious problem, though, connected with the way the boundary conditions affect the wave modes. Writing $\Psi' = \sum_n \phi_n T_n$, and

$$\nabla \cdot \left(\frac{\nabla \phi_n}{H} \right) = \frac{1}{i\omega_n} J(\phi_n, f/H)$$

gives

$$\sum_n \left(\frac{T_{nt}}{i\omega_n} + T_n \right) J(\phi_n, f/H) = -J(\Psi_0, f/H)$$

.

Along a boundary, ϕ_n and Ψ_0 are constant, so we may rewrite the Jacobian terms giving

$$\sum_n \left(\frac{T_{nt}}{i\omega_n} + T_n \right) \phi_{np} \left(\frac{f}{H} \right)_s = \Psi_0 \left(\frac{f}{H} \right)_s$$

where s is a coordinate measured along the boundary, and p is perpendicular to s. Wherever $(f/H)_s$ is non-zero, we have

$$\sum_n \left(\frac{T_{nt}}{i\omega_n} + T_n \right) \frac{\phi_{np}}{H} = \frac{\Psi_{0p}}{H},$$

but Ψ_0 was chosen to obey the boundary condition

$$\oint \frac{\nabla \Psi_{0t}}{H} \wedge d\mathbf{s} = \oint \frac{\Psi_{0pt}}{H}\, ds = \oint \mathbf{G} \cdot d\mathbf{s},$$

and ϕ_n to obey

$$\oint \frac{\phi_{np}}{H}\, ds = 0,$$

so either Ψ' cannot be represented as a sum of wave modes at the boundary or there must be some sort of singularity where $(f/H)_s = 0$. If the boundary runs parallel to f/H contours over some finite distance, the problem disappears.

Beta-Plane models.

In order to examine these problems in more detail, two simple models are considered. These models take advantage of a result due to Longuet-Higgins (1964) that the basin modes for a closed, flat-bottomed, beta-plane ocean consist simply of westward propagating plane waves multiplied by envelope functions which are the normal modes of a stretched membrane held fast at the same boundary as the ocean boundary. The mode equation is

$$-i\omega_n \nabla^2 \phi_n + \beta \phi_{nx} = 0.$$

Substituting $\phi_n = \psi_n e^{ik_n x}$, $k_n = -\beta/2\omega_n$ gives

$$\nabla^2 \psi_n + \frac{\beta^2}{4\omega_n^2} \psi_n = 0,$$

which can easily be solved for ψ for a number of regular geometries.

To investigate the problem of topographic modes failing to represent the nett input of BPV, a rectangular ocean is considered with boundaries running north-south and east-west. Modes for this geometry have the form

$$\phi_n = a_n \sin\left(\frac{m\pi x}{K}\right) \sin\left(\frac{n\pi y}{L}\right) e^{ik_n x}$$

if the basin runs from $x = 0$ to K, $y = 0$ to L. If we choose a forcing of the form $F = F_0 \sin(\pi y/L)$ so that we need not consider modes with $n \neq 1$, then the forced equation may be written

$$(\Psi_{XX} - \Lambda\Psi)_T + \Psi_X = 1, \tag{2.10}$$

and the mode equation becomes

$$-i\Omega_n(\phi_{nXX} - \Lambda\phi_n) + \phi_{nX} = 0, \tag{2.11}$$

where $X = x/K$, $T = t\beta K$, $\Omega_n = \omega_n/\beta K$, $F_0 = \beta/K$, $\Lambda = (K\pi/L)^2$ and all functions are of X only, to be multiplied by $\sin(\pi y/L)$ to get the complete spatial dependence.

Eq.(2.10) is the same as eq.(6.4) of Anderson and Gill (1975)(AG hereafter), which is a representation of a very similar problem; the spin-up of an ocean by Rossby waves, but in their case with a free surface. The solution given by AG for $\Lambda = 20$ (corresponding to $\Lambda = 80$ in eq.(2.10) since AG have boundaries at $X = \pm 1$ instead of $X = 0$ and $X = 1$) will be taken as a reference against which to compare sums of topographic modes. This reference solution is recalculated using a Fourier decomposition of Ψ,

$$\Psi = \sum_n a_n(T)\sin(n\pi X)$$

which leads to a simple pair of equations for the coefficients a_n:

$$\text{For even } n, \qquad a_{nT}\left[\frac{n^2\pi^2 + \Lambda}{2}\right] = \sum_{\text{odd } m} 2a_m \frac{mn}{n^2 - m^2}.$$

$$\text{For odd } n, \qquad a_{nT}\left[\frac{n^2\pi^2 + \Lambda}{2}\right] = \sum_{\text{even } m} 2a_m \frac{mn}{n^2 - m^2} \quad - \frac{2}{n\pi}.$$

This can be efficiently timestepped by a leapfrog method since the rate of change of even coefficients depends only on the odd coefficients and vice-versa. The resulting reference solution is shown in fig.(2.7).

The normalised wave modes from eq.(2.11) are

$$\phi_n = \frac{1}{\sigma_n}\sin(n\pi X)e^{i\sigma_n X}$$

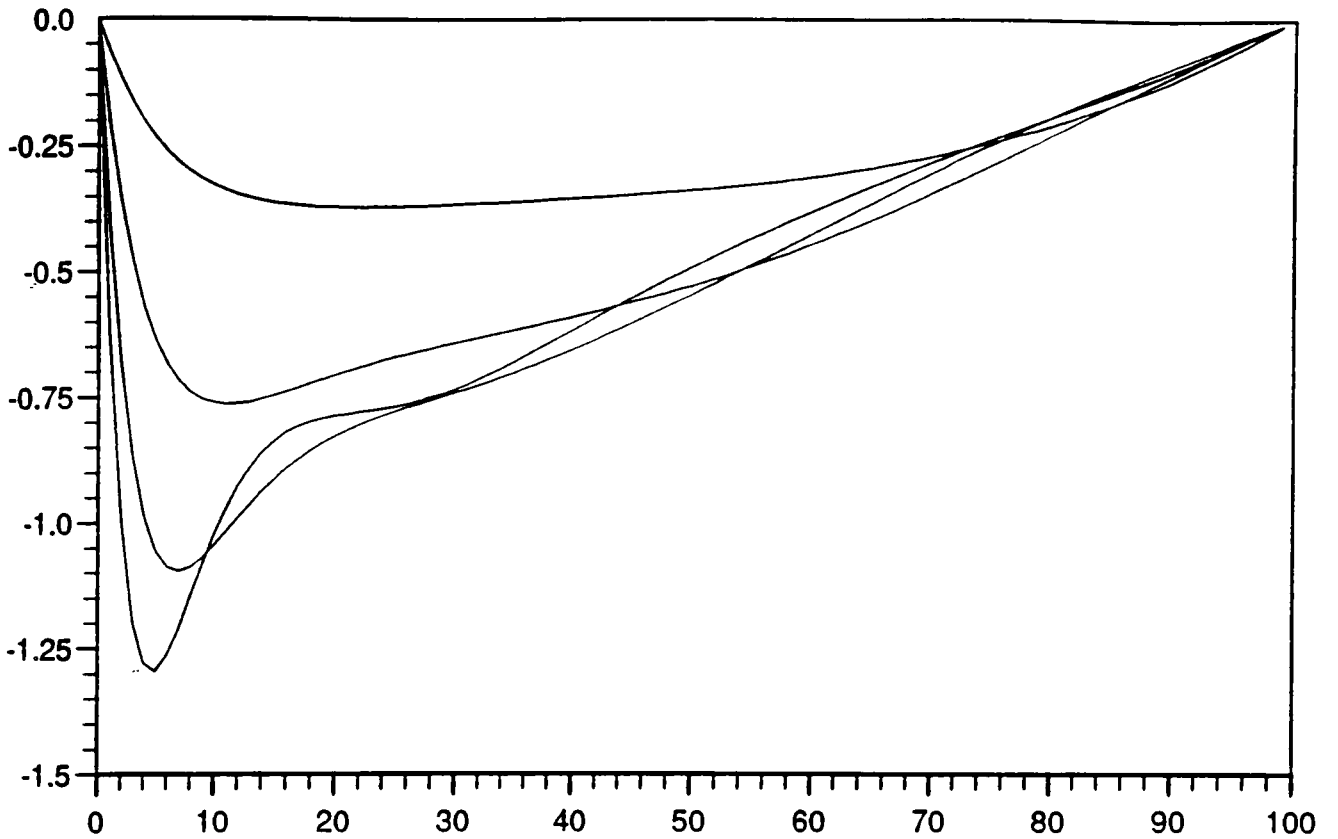


Fig.(2.7). *Reference solution for spin-up of a rectangular ocean, showing the x-variation of stream function only. The solution is the same as in Anderson and Gill (1975), except that time and amplitude are different by a factor of two due to a slightly different non-dimensionalisation. The solution is shown at dimensionless times 30, 60, 90 and 120. The vertical axis is stream function and the horizontal axis is 100 times the dimensionless distance across the basin.*

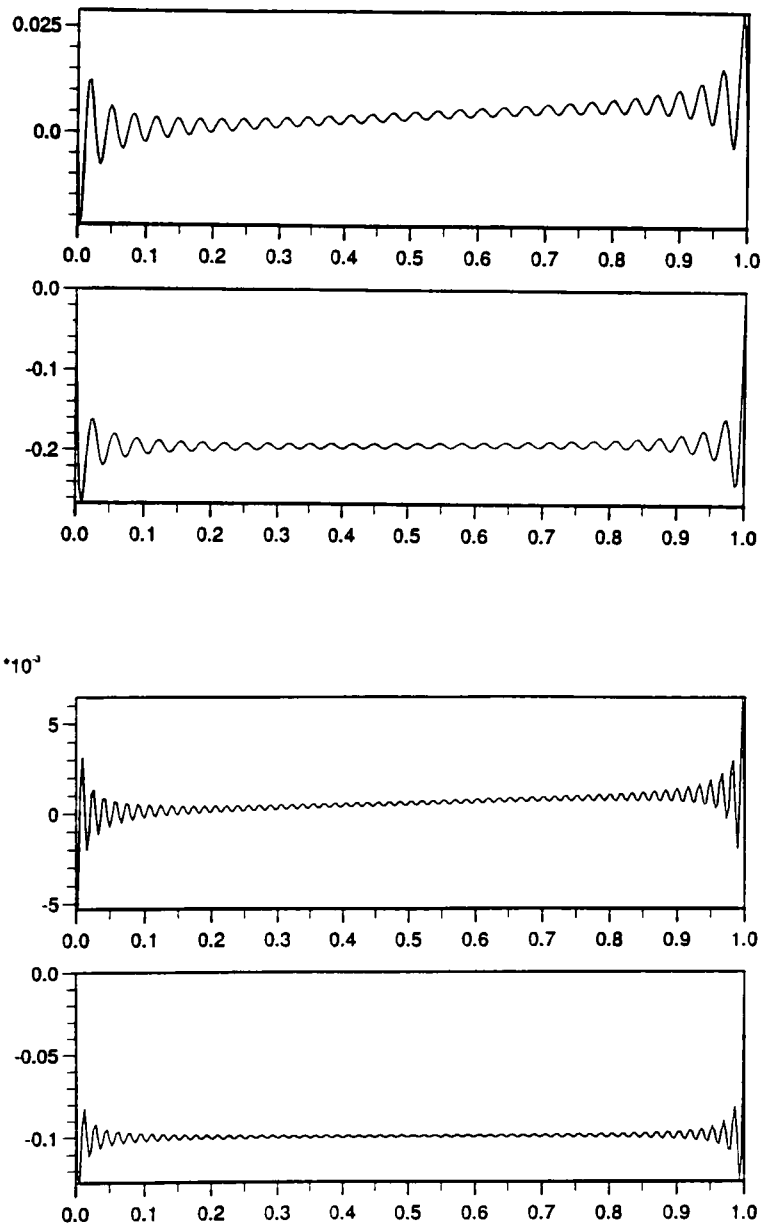


Fig.(2.8). *Differences between stream functions calculated as a sum of basin modes and an accurate solution found by time stepping fourier modes. The top pair uses 30 basin modes and the bottom pair uses 60. The upper graph of each pair is a solution which considers the nett input of BPV separately and the lower graphs are calculated purely as sums of basin modes. Note the different vertical scales.*

where $\sigma = \sqrt{\Lambda + n^2\pi^2}$ and $\Omega_n = -1/2\sigma_n$. From these, Ψ may be calculated in two different ways. First, assuming that the whole solution may be represented as a sum of wave modes, we put $\Psi = \sum_n \phi_n \Theta_n(T)$ and obtain

$$\Theta_n = \frac{n\pi}{\Lambda\sigma_n}\left(\frac{1 - e^{-i\Omega_n T}}{i\Omega_n}\right)(1 \pm e^{-i\sigma_n})$$

where the positive sign is taken when n is even, the negative sign when n is odd. The second method is to absorb the BPV into a function Ψ_0 satisfying

$$(\Psi_{0XX} - \Lambda\Psi_0)_T = 1,$$

and to write the remainder of the solution as $\Psi' = \sum_n \phi_n \Theta_n(T)$, which gives

$$\Psi_0 = T\left[\frac{\cosh\left[(X - \frac{1}{2})\sqrt{\Lambda}\right]}{\Lambda\cosh(\frac{1}{2}\sqrt{\Lambda})} - \frac{1}{\Lambda}\right],$$
$$\Theta_n = \frac{n\pi}{\Lambda\sigma_n}\left(\frac{1 - e^{-i\Omega_n T}}{i\Omega_n} - T\right)(1 \pm e^{-i\sigma_n}).$$

The ability of these two solutions to represent the true solution is tested by calculating them for $T = 120$ and subtracting them from the Fourier solution to see how the residuals behave. Two such cases are shown in fig.(2.8), the first for a sum of 30 wave modes and the second for a sum of 60 wave modes. Both cases show the same general features. The residual in all cases oscillates with the wavenumber of the first uncalculated mode. In addition, the residual for the sum of wave modes only, has in each case a constant offset from zero, which is inversely proportional to the number of modes. This constant offset represents the nett BPV, concentrated in a region whose extension away from the boundary is limited to the length scale of the first uncalculated mode. In the limit of an infinite number of wave modes, this may be interpreted as the infinitesimal function of f/H mentioned earlier. In addition to avoiding this constant offset, the solution in which

the nett BPV input is handled separately has the advantage that the oscillatory part of the residual is smaller in amplitude (and increasingly so as the number of modes is increased), than the oscillatory part of the residual in the pure wave mode case. In short, when the wave modes alone are used to represent the solution, they suffer from something akin to a Gibbs' phenomenon which increases the number of wave modes needed for a given accuracy. This problem is avoided if the nett BPV is treated separately and the wave modes are used only to represent the difference between the solution for a non-rotating ocean and that for a rotating ocean.

Another way of looking at this is that the singularities introduced by using a description in terms of a sum of wave modes plus a function of f/H produce a small scale input to the forcing of the wave modes and thus increase the number of modes needed to resolve the solution. The function of f/H was not included explicitly in this calculation, but it only forces the topographic modes with a delta function at the boundary, and since the modes all go to zero at the boundary, this forcing is not cast onto the modes at all. Assuming the solution can be expressed purely as a sum of wave modes is therefore equivalent to adopting an infinitesimal function of f/H to absorb the nett BPV.

We cannot use this model to look at the boundary condition problem since the rectangular ocean has northern and southern boundaries which run parallel to contours of f/H. We must therefore move to a slightly more complex basin shape: a square with boundaries at 45° to the north-south direction. If the x and y axes are chosen as in fig.(2.9) and the basin has sides of length K, then, non-dimensionalising so that $X = x/K$, $Y = y/K$, $T = t\beta K/\sqrt{2}$, and choosing a simple, doubly sinusoidal forcing produces the equation

$$(\Psi_{XX} + \Psi_{YY})_T + \Psi_X + \Psi_Y = \sin(\pi X)\sin(\pi Y),$$

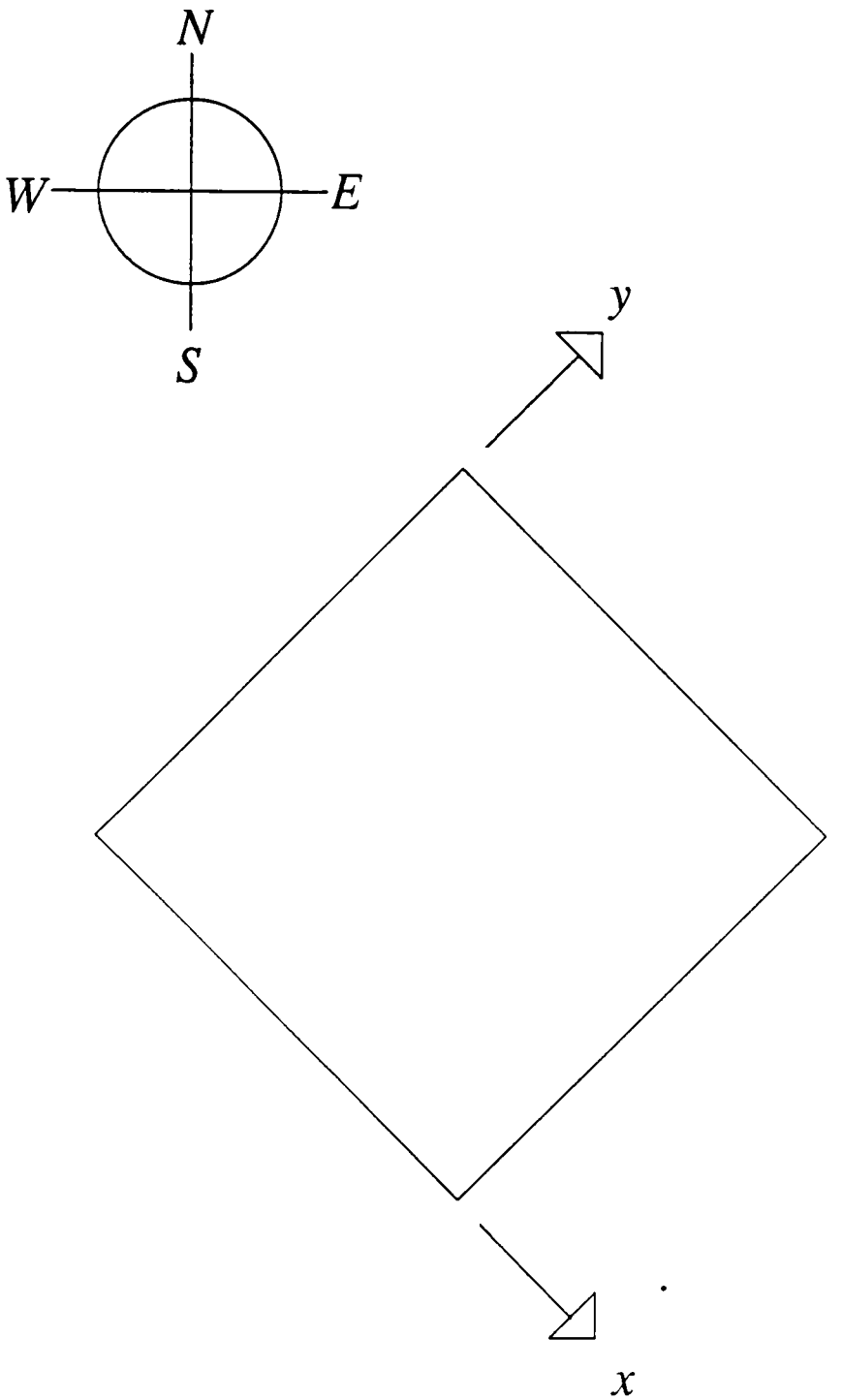


Fig.(2.9). *Diagram to show the geometry of the basin and the orientation of axes used in further investigations of the completeness of basin modes (see text) .*

and for the wave modes,

$$-i\Omega(\phi_{XX} + \phi_{YY}) + \phi_X + \phi_Y = 0.$$

The normalised wave modes for this equation are

$$\phi_{mn} = \frac{1}{\sigma_{mn}} \sin(m\pi X) \sin(n\pi Y) e^{i\sigma(X+Y)},$$

with $\sigma_{mn} = \pi\sqrt{(m^2+n^2)/2}\,, \quad \Omega_{mn} = -1/2\sigma_{mn}.$

The problem may be simplified further by concentrating only on the component of the solution which is proportional to T^2. Expanding Ψ in powers of T:

$$\Psi = \Psi_1 T + \Psi_2 T^2 + \Psi_3 T^3 + \ldots$$

produces a series of equations

$$\nabla^2 \Psi_1 = \sin(\pi X) \sin(\pi Y)$$

$$2\nabla^2 \Psi_2 = -(\Psi_{1X} + \Psi_{1Y})$$

$$3\nabla^2 \Psi_3 = -(\Psi_{2X} + \Psi_{2Y})$$

$$\ldots \text{etc.}$$

The first may be solved immediately giving

$$\Psi_1 = \frac{-1}{2\pi^2} \sin(\pi X) \sin(\pi Y).$$

This gives us

$$\nabla^2 \Psi_2 = -\frac{1}{2}(\Psi_{1X} + \Psi_{1Y}) = \frac{-1}{4\pi} \left[\cos(\pi X) \sin(\pi Y) + \sin(\pi X) \cos(\pi Y)\right].$$

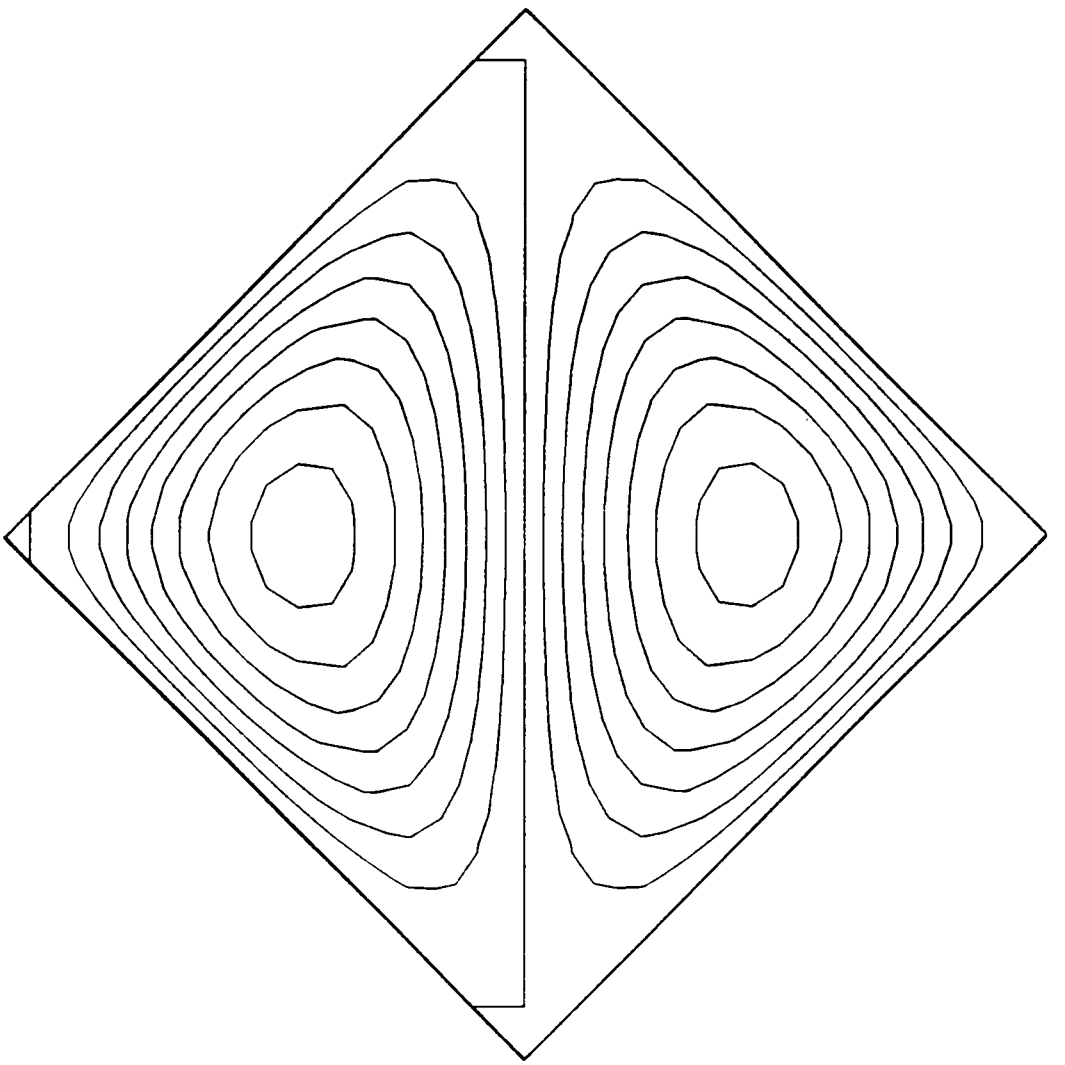


Fig.(2.10). *The structure of the component of stream function which grows initially as time squared, calculated as a sum of fourier modes, although a similar plot calculated for a sum of basin modes is not visibly different. The structure is antisymmetric about the north-south axis.*

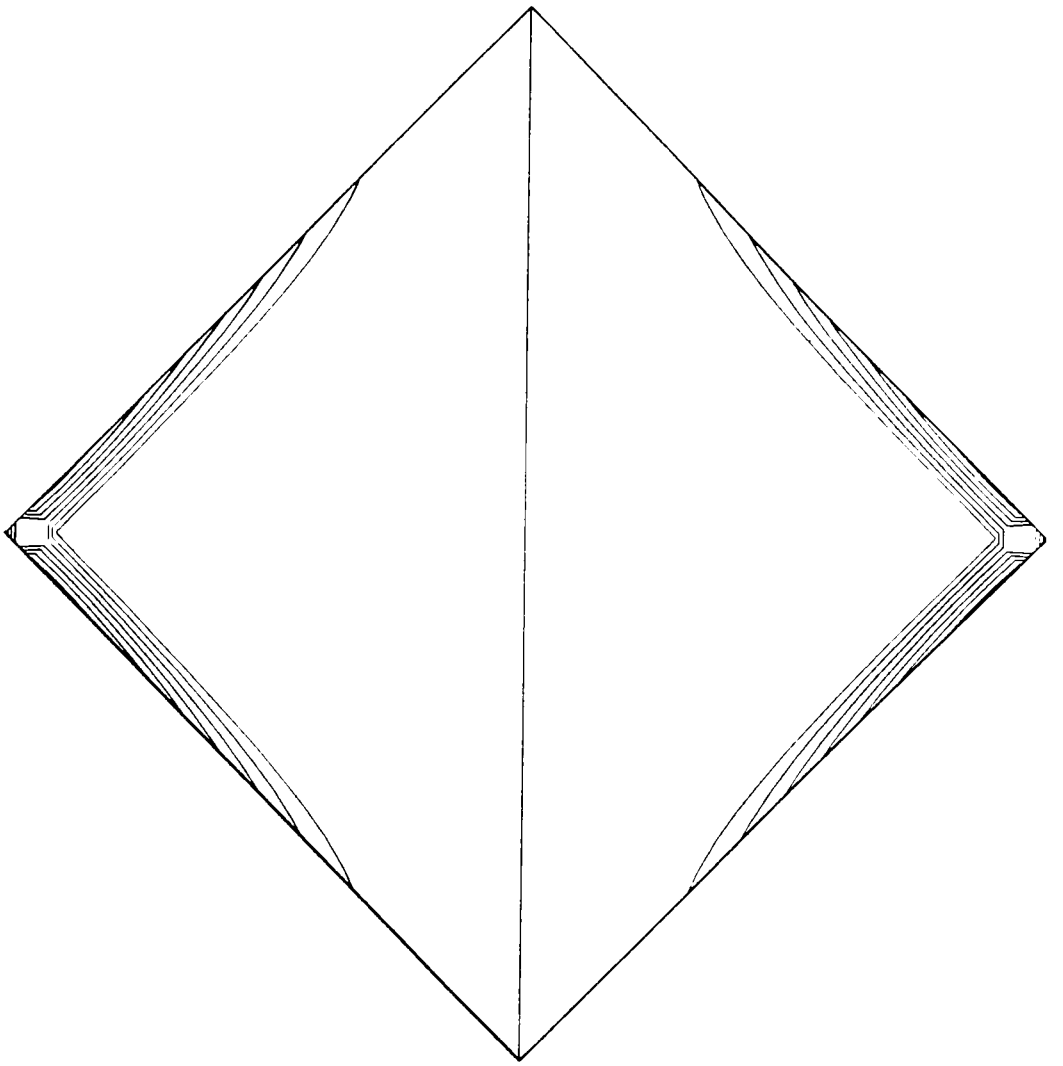


Fig.(2.11). *The difference between analytical and basin mode calculations of the BPV of the stream function shown in fig.(2.10). Again antisymmetric about the north-south axis, this shows that the small inaccuracy of the basin modes velocity field is due to misrepresentation of BPV at the boundary only. The basin modes solution uses 19 modes in each direction, hence the boundary layer is as narrow as it can be with this resolution. The range is ±0.224.*

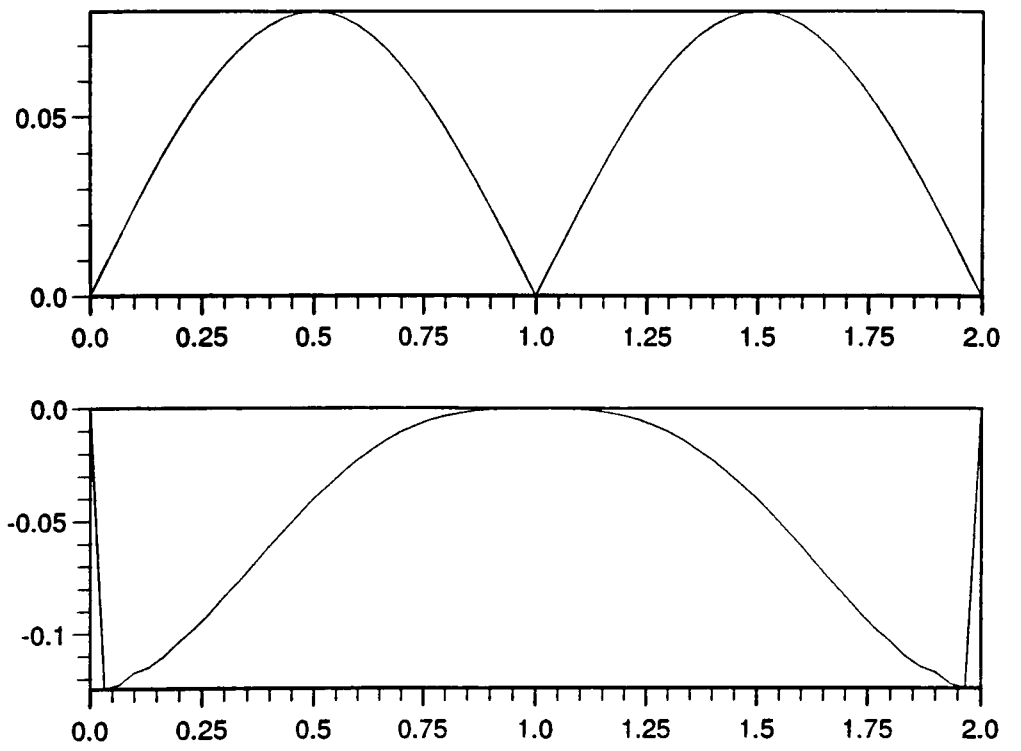


Fig.(2.12). *Similar to fig.(2.11). The top plot shows the analytical BPV along the boundary, and the bottom plot shows the difference between analytical and basin modes calculations of BPV. In this case there are 100 basin modes in each direction and it can be seen that the size of the anomaly at the boundary remains finite as resolution increases. The horizontal coordinate is non-dimensional distance along the western boundary, with the southern corner at 0, the western corner at 1 and the northern corner at 2.*

Writing the cosine terms as a fourier series of sines, the ∇^2 operator may be inverted again giving

$$\Psi_2 = \frac{1}{\pi^4} \sum_{\text{even } m} \frac{m}{(m^2-1)^2} \left[\sin(\pi Y)\sin(m\pi X) + \sin(\pi x)\sin(m\pi Y)\right], \tag{2.12}$$

which, happily, turns out to be a sum over only one index. Ψ_2 may be used to exemplify the boundary condition problem since it contains no nett BPV but, if it were written as a sum of wave modes, $\nabla^2\Psi_2$ could be written as a sum of Jacobians and integrating around the boundary would produce the same contradiction as that derived in the previous section.

Writing Ψ_2 as a sum of wave modes produces, after a fair amount of algebra,

$$\Psi_2 = \frac{1}{2\pi^4} \sum_{m,n} Re\Bigg(\frac{i\alpha\alpha_2}{\sigma_{mn}} \frac{(1 \pm e^{-i\sigma_{mn}})(1 \pm e^{-i\sigma_{mn}})}{(\alpha^2-\beta^2)(\alpha_2^2-\beta_2^2)} \\ \times \quad \sin(m\pi X)\sin(n\pi Y) e^{i\sigma_{mn}(X+Y)} \Bigg), \tag{2.13}$$

where $\alpha = 2m$, $\alpha_2 = 2n$, $\beta = 1 + (m^2-n^2)/2$, $\beta_2 = 1 + (n^2-m^2)/2$ and a + sign is taken for each of m and n which is even.

Figure (2.10) shows Ψ_2 calculated from eq.(2.12) and the difference between Ψ_2 from eq.(2.12) and Ψ_2 from eq.(2.13). A problem near the east and west corners is apparent, although the size of the residual is actually quite small.

The difference between $\nabla^2\Psi_2$ from eq.(2.13) and the analytical form

$$\begin{aligned} \nabla^2\Psi_2 &= \frac{-1}{4\pi} \left[\cos(\pi X)\sin(\pi Y) + \sin(\pi X)\cos(\pi Y)\right) \\ &= \frac{-1}{4\pi} \sin\left[\pi(X+Y)\right], \end{aligned}$$

shown in fig.(2.11) shows that the topographic modes are representing BPV well except along two regions of the boundary. The size of this BPV anomaly remains finite as the number of modes increases, while the area to which it is confined shrinks, so in the limit it

must have no effect on the velocity field. Figure (2.12) shows a plot of the BPV anomaly along the northern section of the boundary for a much higher resolution, demonstrating that the amplitude does remain finite despite the compression of scale.

With this result as a guide, we can see how to generalise to a basin of any shape with topography. We have

$$\nabla \cdot \left(\frac{\nabla \Psi_t}{H} \right) + J(\Psi, f/H) = F \,,$$
$$\oint \frac{\nabla \Psi_t}{H} \wedge d\mathbf{s} = \oint \mathbf{G} \cdot d\mathbf{s} \,.$$

We now write $\Psi = \Psi_0 + \Psi'$, where

$$\nabla \cdot \left(\frac{\nabla \Psi_{0t}}{H} \right) = F + \delta F, \qquad \oint \frac{\nabla \Psi_t}{H} \wedge d\mathbf{s} = \oint \mathbf{G} \cdot d\mathbf{s},$$
$$\nabla \cdot \left(\frac{\nabla \Psi'}{H} \right) + J(\Psi', f/H) = -J(\Psi_0, f/H) - \delta F.$$

If δF is chosen to have a finite value at the boundary, so as to eliminate the problem of representing Ψ' as a sum of wave modes, but to be zero elsewhere so that

$$\int \delta F \, dA = 0 \qquad \text{over any area,}$$

then δF produces no change in the coefficients of the wave modes since it must be integrated to produce them, and similarly no change in Ψ_0. The only difference is a finite change in $\nabla \cdot \left(\frac{\nabla \Psi_0}{H} \right)$ right at the boundary, of amplitude $\int_0^t \delta F \, dt'$, which can only alter the flow in the inviscid case, as $t \to \infty$.

In conclusion, then, it appears that the deficiency of the wave modes in representing nett BPV input may be remedied in either of two ways. Adding a function of f/H (which may be interpreted as a set of wave modes with zero frequency) works, but at

the expense of introducing small scales to the problem and increasing the number of wave modes necessary to resolve the solution to any given accuracy. The alternative—using the wave modes to represent the difference between solutions to the non-rotating and the rotating problem—works more effectively, and the apparent difficulty with this description, connected with boundary conditions, is only actually a problem when considering either BPV at the boundary or the behaviour of the solution as $t \to \infty$.

Friction.

All of these inviscid equations display solutions which continue to change with time as long as the forcing is still there. This is inevitable because there is a nett input of BPV which is not being destroyed anywhere. Clearly, for a steady state to be reached, either the nett input of BPV by the forcing must become zero (which is a possibility in a stratified fluid since it is possible that the fluid flow adjusts the JEBAR term so as to cancel the input of BPV due to wind stress) or there must be some form of friction capable of destroying BPV. An example is the approximation to Ekman bottom friction used by Welander to give eq.(2.9). If this were added in to the forcing we would obtain

$$\nabla \cdot \left(\chi \nabla \Psi + \frac{\nabla \Psi_t}{H} \right) + J(\Psi, f/H) = F, \tag{2.14}$$

where $\chi = Df/H^2 = \frac{1}{H^2}\sqrt{\mu f/2\rho_0}$.

This was derived on the premise that the Ekman layer is driven by the barotropic velocity. In fact it is driven by the velocity near to the bottom which was shown in chapter 1 to differ from the barotropic velocity by an amount determined by the JEBAR term. A more accurate form of eq.(2.14) when F includes a JEBAR term may therefore be calculated, with a correction which depends only on E and the bottom density (Mertz and

Wright, 1992). Whatever the precise form of the friction, however, its chief function in this context is to destroy BPV, which it can only do to a significant extent in regions of strong shear or steep bottom gradient. In practice this means that friction is only important in boundary layers or in regions of closed f/H contours where it limits the steady growth of circulatory flow and fixes the precise distribution of the recirculation velocities. The form of the solution outside these regions (and the general character inside) is unaffected by the form chosen for friction (unless non-linear terms are important or the friction is so large as to dominate the whole solution), and for this reason it is felt to be justified to consider only the most conveniently simple form for the friction. This is obtained by writing $\chi = r/H$, which does not violate the form of bottom friction too drastically, and converts eq.(2.14) to

$$\left(r + \frac{\partial}{\partial t}\right) \nabla \cdot \left(\frac{\nabla \Psi}{H}\right) + J(\Psi, f/H) = F, \tag{2.15}$$

with corresponding wave mode equation

$$(r - i\omega_n) \nabla \cdot \left(\frac{\nabla \phi_n}{H}\right) + J(\phi_n, f/H) = 0.$$

This has no effect on the spatial form of the wave modes, and only changes their eigenvalues in the manner

$$\omega_n(\text{viscous}) = \omega_n(\text{inviscid}) - ir.$$

No change is necessary to the methods of solution, and the only real changes to a wave modes description of the flow are: The oscillations excited at the beginning are now damped out after a time of the order $1/r$; The steady component of wave modes with $|\omega| \leq (\sim r)$ is diminished so that boundary layers and other features which are singularities in the long

time limit of the inviscid case are restricted to finite width. That a steady state can now be achieved is shown by the fact that the non-wave-modes component of the solution now satisfies

$$\left(r + \frac{\partial}{\partial t}\right) \nabla \cdot \left(\frac{\nabla \Psi_0}{H}\right) = F,$$

so that $\Psi_0 = \Psi_0^1(1 - e^{-rt})$ where $\nabla \cdot \left(\frac{\nabla \Psi_0^1}{H}\right) = F/r$.

Friction thus operates in the regions where a purely wave mode description of the solution has difficulties. In a region of closed contours of f/H it acts to limit the eventual circulation in these regions with an exponential decay in growth typified by the above formula. The other common kind of region where wave mode descriptions have problems is the western (or equivalent-western) boundary layer. In these regions, light can be shed on the rôle of friction by going back to the rectangular beta-plane ocean. If friction is included, eq.(2.10) becomes

$$\left(R + \frac{\partial}{\partial T}\right) (\Psi_{XX} - \Lambda \Psi) + \Psi_X = 1,$$

where R is a dimensionless friction. This can easily be solved for the steady-state solution, giving

$$\Psi = \frac{-1}{R\Lambda} \left(1 - \frac{\left[(1 - e^{\lambda_2})e^{\lambda_1 X} + (1 - e^{\lambda_1})e^{\lambda_2 (X-1)}\right]}{(1 - e^{\lambda_1 - \lambda_2})}\right)$$

where

$$\lambda_1 = -\frac{1}{2R} - \sqrt{\frac{1}{4R^2} + \Lambda}\,, \qquad \lambda_2 = -\frac{1}{2R} + \sqrt{\frac{1}{4R^2} + \Lambda}\,.$$

This has the form of two exponentials decaying away from the boundaries at $X = 0$ and $X = 1$. Of the exponents, λ_1 is always negative and larger than λ_2, the difference in size increasing as R decreases, so the decay away from the western boundary ($X = 0$) is

always faster than the decay away from the east ($X = 1$). As $R \to 0$ this solution tends to an inviscid limit

$$\Psi = X - 1$$

which is valid everywhere except $X = 0$, where $\Psi = 0$. Near this limit, the length scale of the western boundary region is given by $|\lambda_1| \approx 1/R$. This is precisely the length scale of the wave mode for which $|\Omega| \approx R$, so the effect of friction in spreading the singularity out away from the western boundary can be seen to follow through in the wave modes description since modes with characteristic length scale smaller than about $1/R$ have reduced amplitudes compared with the inviscid case, and the jump at the western boundary is therefore smoothed out. If friction is included, then, and enough wave modes are used to resolve the frictional boundary layer, a reasonable solution with a simple, passive western boundary layer will result.

Chapter 3.

Numerical Determination of Modes.

The aim in this determination is to use the same finite difference scheme as was used in FRAM and CRAM, *i.e.,* the barotropic component of the Cox code (Cox, 1984). This is somewhat buried due to the semi-implicit treatment designed to stop instabilities from developing when using the code with a time step comparable with the period of an inertial oscillation, but the forms of the del-squared term and the Jacobian term are easy to pick out and may be written as

$$\nabla \cdot \left(\frac{\nabla \Psi}{H} \right) \Rightarrow \delta_\lambda \left(\frac{m}{\overline{H}^{\phi} a^2} \delta_\lambda \Psi \right) + \delta_\phi \left(\frac{1}{m \overline{H}^{\lambda} a^2} \delta_\phi \Psi \right)$$

$$J_{xy}(\Psi, f/H) \Rightarrow \frac{m}{a^2} \delta_\phi \overline{\left(\frac{f}{H} \delta_\lambda \Psi \right)}^{\lambda} - \frac{m}{a^2} \delta_\lambda \overline{\left(\frac{f}{H} \delta_\phi \Psi \right)}^{\phi}$$

where a is the radius of the earth, $m = \sec \phi$ where ϕ is latitude. The symbols δ and overline represent differentiation and averaging with respect to the indicated coordinate (λ =longitude or ϕ =latitude) in finite difference terms, *i.e.,*

$$\delta_\lambda(\mu) = \frac{\mu_{+\frac{1}{2}} - \mu_{-\frac{1}{2}}}{\Delta \lambda}$$

$$\overline{\mu}^{\lambda} = \frac{\mu_{+\frac{1}{2}} + \mu_{-\frac{1}{2}}}{2}$$

where the subscripts indicate whether the value used is to be taken at the current position plus or minus half a grid point in the λ direction.

The main difference between these terms and the terms in the theory of chapter 1 is that here we are working in spherical polar coordinates. The theory works just as well in spherical polars, but is rather more messy to work with.

The finite difference representations above are both five point representations, representing

$$m\left[m\left(\frac{\Psi_\lambda}{H}\right)_\lambda+\left(\frac{\Psi_\phi}{Hm}\right)_\phi\right],$$

the del-squared term, and

$$mJ_{\lambda\phi}(\Psi,f/H),$$

the Jacobian term, evaluated at a particular grid point, in terms of H, f, m at various grid point and half grid point positions, and Ψ at the grid point being considered and the four adjacent grid points. Interestingly, the contribution from the grid point itself to the Jacobian term is zero, thus reducing this part to a four point approximation.

The topographic mode equation has the form

$$i\omega(\text{del}-\text{squared term})=(\text{Jacobian term}),$$

so in the finite difference scheme this becomes a set of simultaneous equations, one for each grid point, for ψ_m. These may be written in matrix form as

$$\mathcal{J}\psi_m=i\omega\mathcal{D}\psi_m \tag{3.1}$$

where $\psi_{\mathbf{m}}$ is a vector of length N, $\mathcal{J}$ is an $N\times N$ matrix representing the Jacobian term, and $\mathcal{D}$ is an $N\times N$ matrix representing the del-squared term (N is the total number of grid points.). This is a generalised eigenvalue problem and is soluble directly using the QZ algorithm. This is the method used to find the topographic modes of the Southern Ocean. It is interesting though, before looking at the results of this calculation, to look at some of the effects of the discretisation on the eigenvalues and vectors produced.

Errors.

It is informative to look at one simple situation for which an analytical solution can be found for both the complete and the finite difference equations. This is the case of an exponential depth profile with no variation in the x-direction, solved for an f-plane in cartesian coordinates. Consider a periodic channel of length K in the x-direction and width L in the y-direction. The depth in the channel is given by $H = he^{by}$ and the topographic modes equation reduces to

$$-i\omega(\psi_{xx} + \psi_{yy} - b\psi_y) - bf\psi_x = 0,$$

the solution to which is (Buchwald and Adams, 1968)

$$\psi = e^{y(\frac{b}{2} + \frac{in\pi}{L})} e^{ix\frac{2\pi m}{K}},$$

which results in the dispersion relation

$$\frac{\omega}{f} = \frac{2\pi mb/K}{\frac{4\pi^2 m^2}{K^2} + \frac{\pi^2 n^2}{L^2} + \frac{b^2}{4}}\,.$$

The cartesian finite difference terms may be obtained from the spherical terms by replacing λ with x, ϕ with y, and m and a with 1. For this depth and the f-plane, several of the terms cancel and the topographic modes equation becomes

$$\begin{aligned}\frac{-2i\omega}{(\delta x)^2(H_{+\frac{1}{2}} + H_{-\frac{1}{2}})}\,(\psi_{+1,0} - 2\psi_{0,0} + \psi_{-1,0}) \\ - \frac{i\omega}{(\delta y)^2}\left(\frac{\psi_{0,1}}{H_{+\frac{1}{2}}} - \frac{\psi_{0,0}}{H_{+\frac{1}{2}}} + \frac{\psi_{0,-1}}{H_{-\frac{1}{2}}} - \frac{\psi_{0,0}}{H_{-\frac{1}{2}}}\right) \\ + \frac{f}{2\delta x\delta y}\left[\frac{1}{H_{-\frac{1}{2}} - H_{+\frac{1}{2}}}\right](\psi_{1,0} - \psi_{-1,0}) = 0\,.\end{aligned}$$

Substituting for ψ the function

$$e^{i(2\pi mi\delta x/K + \pi nj\delta y/L)} e^{bj\delta y/2}$$

where i and j are indices for the grid points counting in the x and y directions respectively, satisfies these equations giving a dispersion relation

$$\frac{\omega}{f} = \frac{\sinh\delta \sin\Delta/\delta x\delta y}{\frac{2\sin^2(\frac{\Delta}{2})}{(\delta x)^2\cosh\delta} + \frac{\cosh\delta - \cos D}{(\delta y)^2}}$$

where $\delta = b\delta y/2$, $\Delta = 2\pi m\delta x/K$, and $D = \pi n\delta y/L$; m and n being integers.

This pair of solutions is curious in that the eigenvectors are completely unaffected by the discretisation, but the eigenvalues are altered. It is more usual for the eigenvalues to be well approximated and the vectors less well approximated.

The two equations for ω allow us to estimate how well a topographic mode need be resolved to get a good approximation to the eigenvalue. The modes in this case are sinusoidal in the x direction and in the y direction are sinusoids modulated by the square root of the depth. Choosing m or n to be anything other than 1 gives the same eigenvalues as for a channel of length K/m and width L/n, so the only parameters that affect ω are the aspect ratio of the channel, the length scale $1/b$ and the numbers of grid points across the channel in each direction. We can thus produce contour plots of error in ω due to discretisation against resolution in both directions, for various aspect ratios and depth contrasts (the depth contrast being the ratio of depths at $y = 0$ and at $y = K$, determined by Kb). A set of these plots is presented in fig.(3.1). The most striking aspect of all the plots is a line on which the error is zero. Usually an almost straight line through the origin, it shows that when a particular wave is resolved approximately equally well in both directions, the errors tend to cancel, giving a better eigenvalue than might be expected. The similarity between the plots shows that the errors are fairly insensitive to aspect ratio and depth contrast except when the waves are very elongated along the slope when varying the depth contrast can make a relatively large difference. Even in this situation, though, a

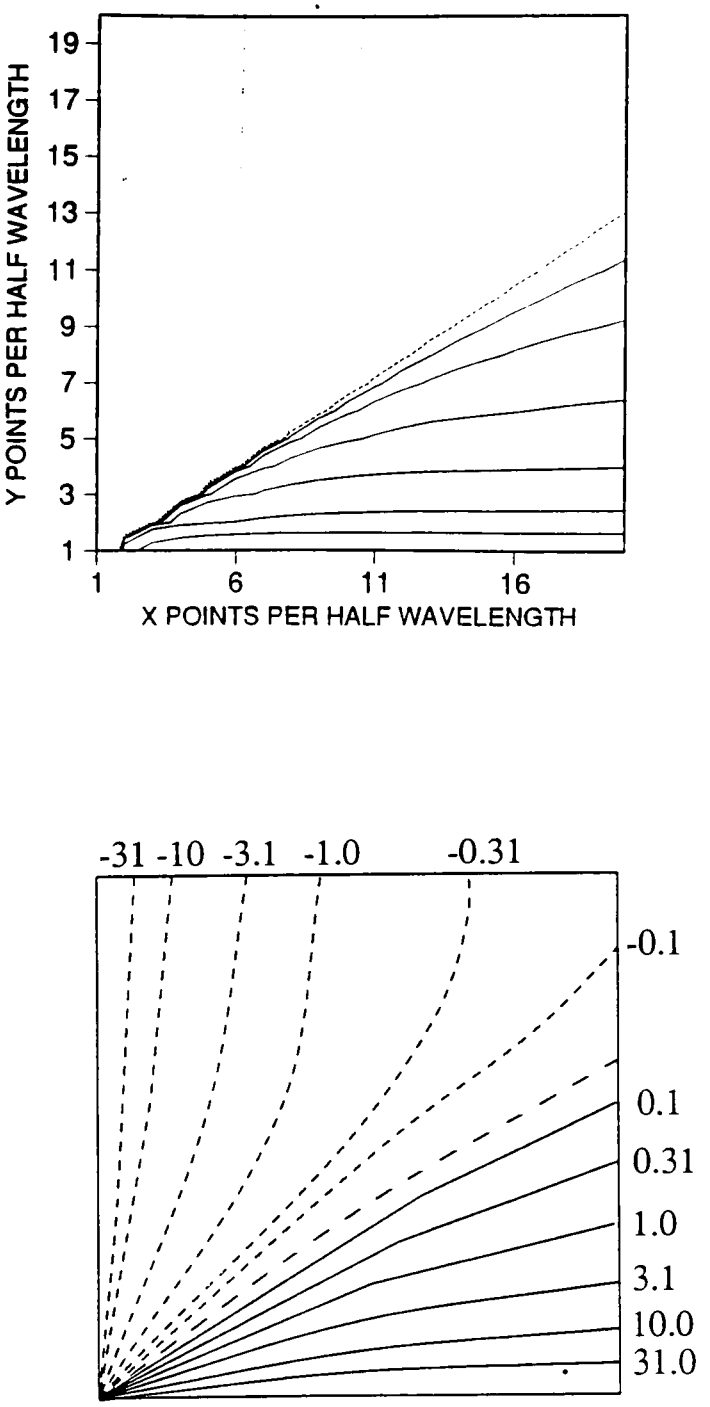


Fig.(3.1) a) *Plot of errors due to resolution for a topographic wave with aspect ratio 1 on a sloping region with ratio of largest to smallest depth of 10. The schematic is to show the values attached to the contours, the values being percentage error in the eigenvalue when calculated at the resolution given along the axes.*

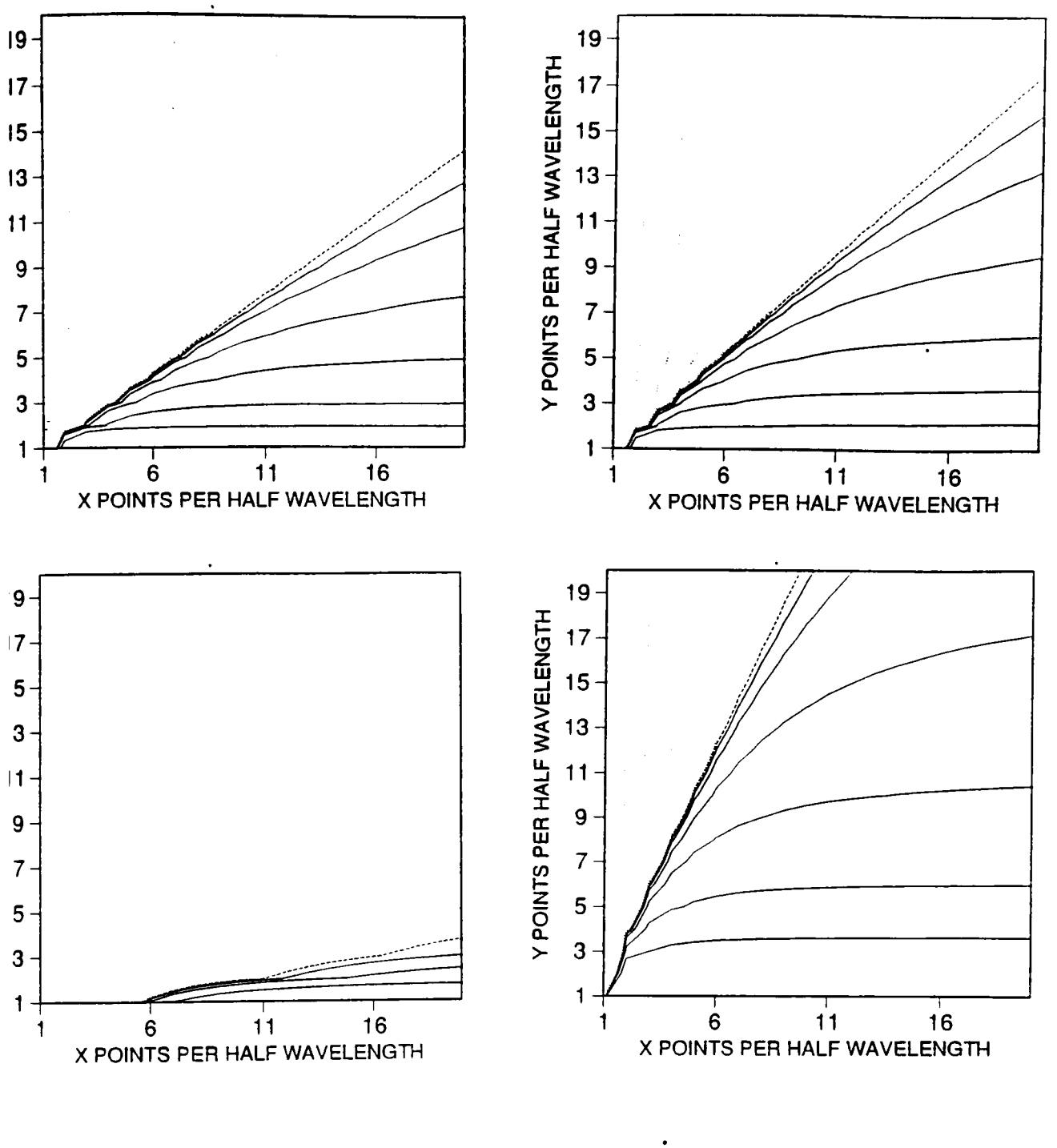


Fig.(3.1) b) *As fig.(3.1) a but with different aspect ratios and depth ratios. Aspect ratios are 100 at the top and 0.01 at the bottom. depth ratios 100 on the left and 1.5 on the right.*

rule of thumb applies for the accuracy of the eigenvalue. That is that ω is accurate to 3% or better if there are 5 grid points between nodes in each direction, 1% or better if there are 10 grid points.

Representation of topography.

Another way of looking at errors is to compare a high resolution solution with a low resolution solution. This is difficult to do with a two dimensional topography because of the size of the calculation necessary for high resolution, but choosing a one dimensional topography and taking a solution that is sinusoidal in the along-slope direction, the eigenvalue and the down-slope structure can be solved for at a high resolution. Two questions arise in this situation. Given a high resolution topography to be cast onto a low resolution grid, what is the best way to average a number of depths to give a depth at the grid point, and given a model that only allows depths to be at the bottom of a given number of levels, what is the best way to distribute these levels in the vertical?

In an attempt to answer these questions, two slices of real topography from the Southern Ocean were taken from the DBDB5 data set used to generate the topography for FRAM. The topographies were chosen using the criterion that they must give a monotonic slope for 200 points (16° 40′). The only region where such slopes occur in a convenient north-south or east-west direction is around the Kerguelan Plateau, so two slopes were chosen, one going east from Kerguelan and the other going west. The first topography, known as Kerg A, starts at 58° 45′ S, 77° 20′ E and stretches to the west of Kerguelan through approximately 960 km. The second (Kerg B) starts at 48° 40′ S, 69° 15′ E and stretches to the east through 1220 km. Neither are quite monotonic, but Kerg B was close

enough for calculation purposes while Kerg A required the reversal of a small patch of ten points to facilitate calculation. The topographies actually used are illustrated in fig.(3.2), and the most obvious difference is that Kerg B penetrates much closer to the surface. Its shallowest point is actually 3 m deep.

Three ways of averaging the depths are considered. The most obvious is to take the mean, but it may be better to take a median depth or, since the equation shows dependence on $\log H$, to take a geometrical mean. The results of applying these averaging techniques to blocks of ten points from the two topographies are shown in fig.(3.3). The different techniques produce very similar results. To compare the effects on the topographic modes, finite difference calculations were performed on the different topographies and compared with a finite difference calculation of the first few modes of the full 200 point topography. This was done for wavelengths of 1000 km and 250 km, and the comparison was made by comparing the eigenvalues and the positions of the nodes in the eigenvectors. Two numbers were produced for each mode: the percentage error of eigenvalue and the root mean square value of

$$\frac{\text{error in nodal position}}{\text{mean inter-nodal distance}}$$

where the mean inter-nodal distance is half the distance between the node before and the node after the node being considered.

Graphs of these two quantities for solutions with the different topographies are shown in fig.(3.4). The results are not discernibly better for any one averaging procedure, so any averaging procedure seems as good as any other. It should be borne in mind though that this is only for a monotonic topography in which the procedures produce almost identical results, and in other situations there may be a best procedure.

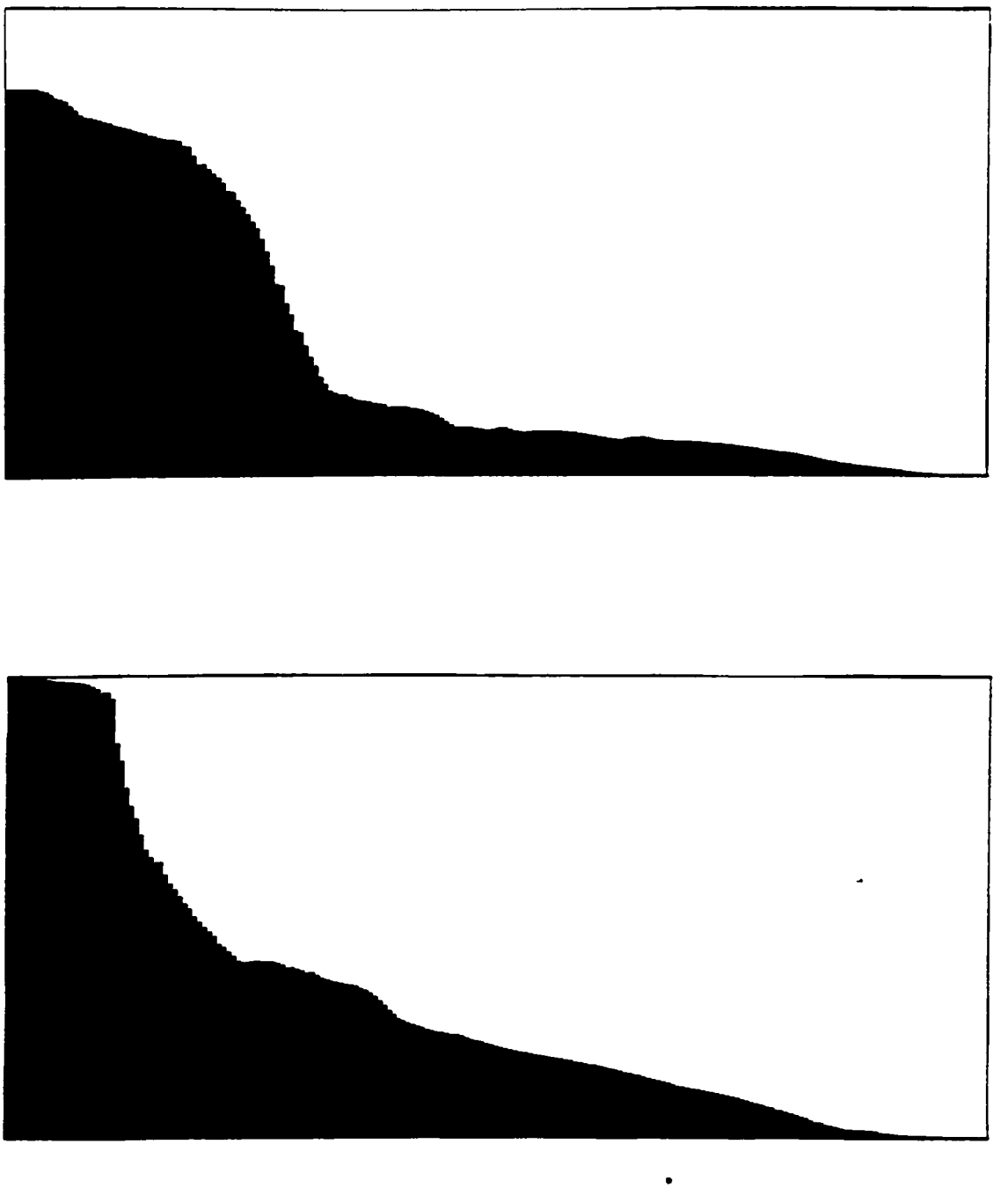


Fig.(3.2). *The topographies Kerg A (top) and Kerg B (bottom) as described in the text.*

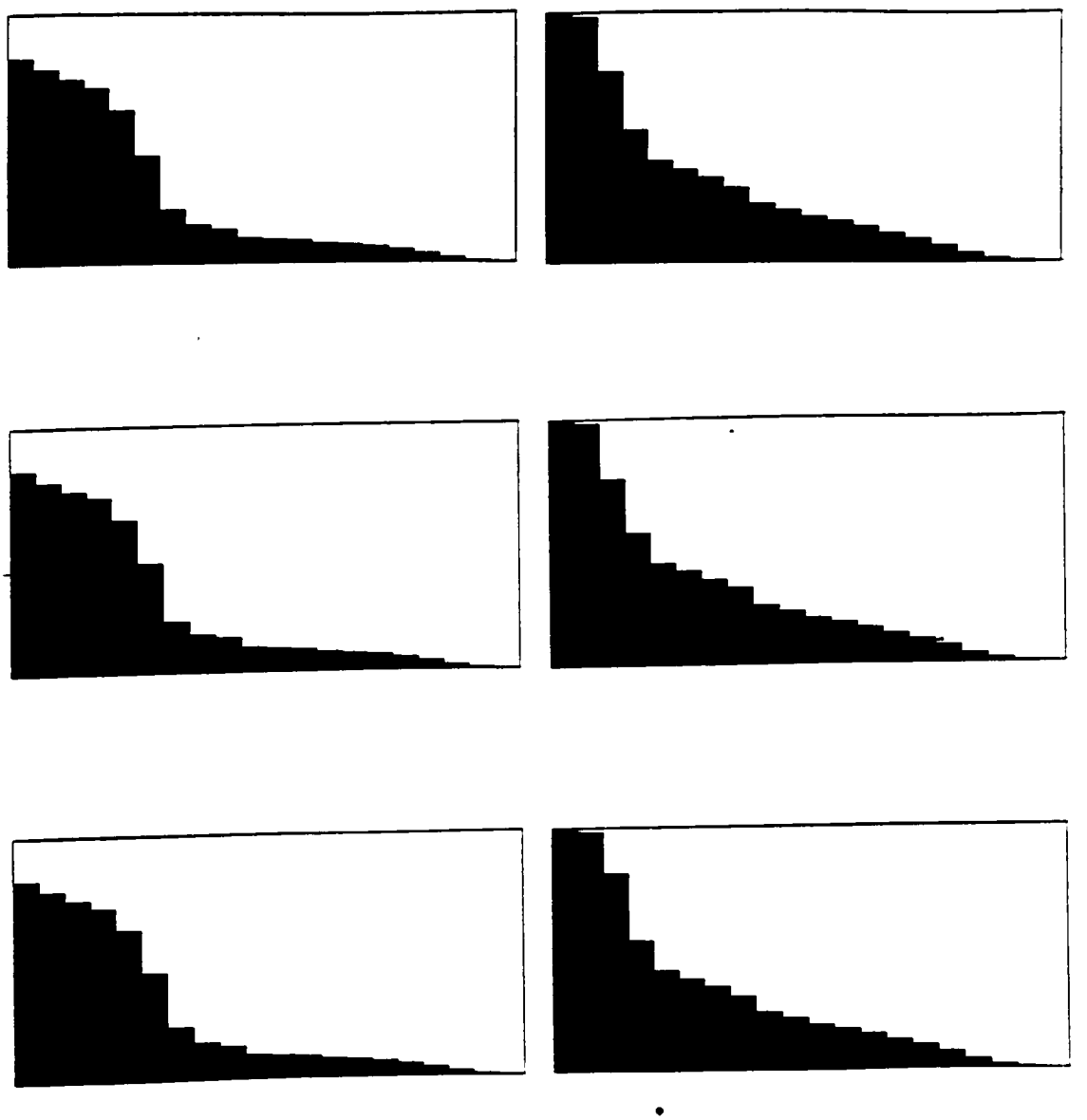


Fig.(3.3). *Various approximations to the topographies Kerg A (left) and Kerg B (right) shown in fig.(3.2). Groups of ten points were taken together and an average depth was found in three different ways. Top: the arithmetic mean was taken. Centre: the median value was chosen. Bottom: the geometric mean was taken.*

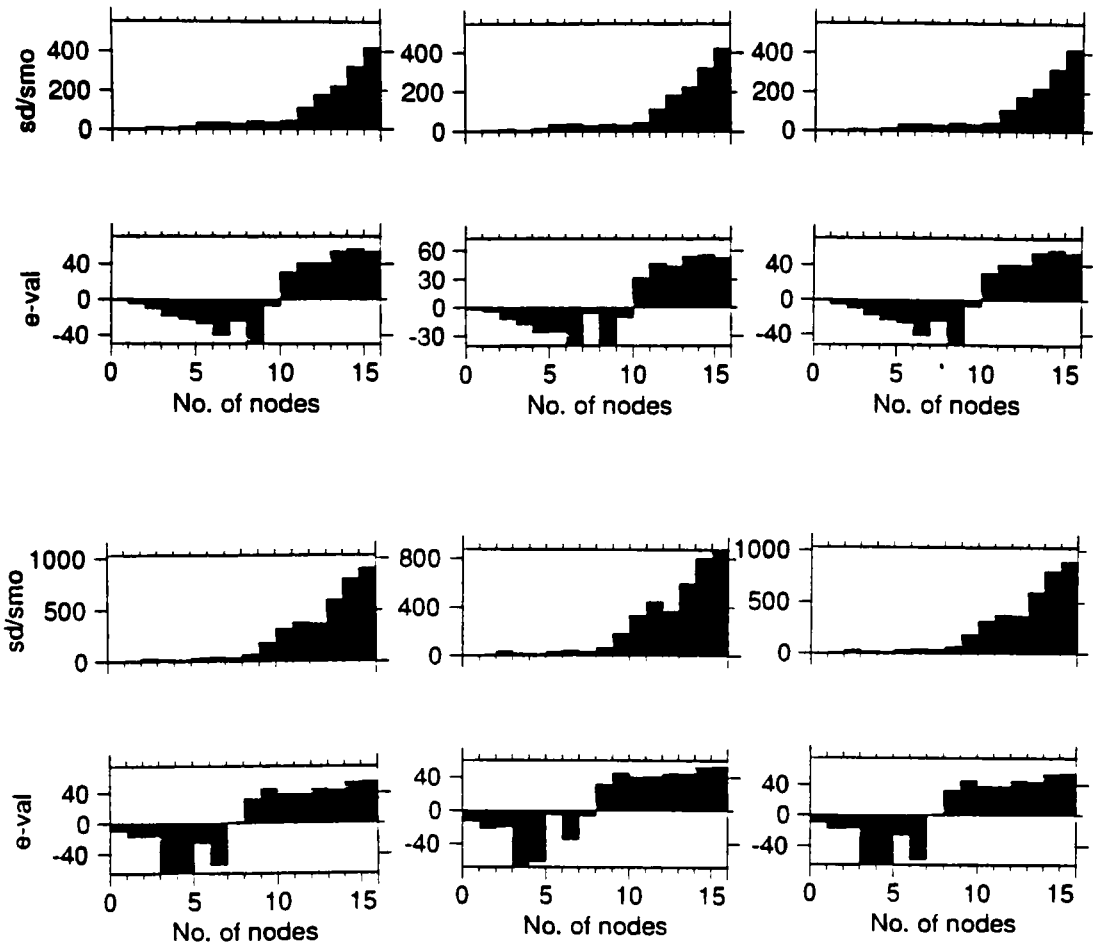


Fig.(3.4). *Histograms of errors in topographic modes calculated for (top) Kerg A and (bottom) Kerg B topographies. Comparison of errors due to the three different methods of averaging; arithmetic mean (left), median (centre) and geometric mean (right) shows no one method to be noticeably better than the others. The quantities shown are the root mean square value of the ratio of error in position of node to mean inter-nodal distance (upper graph of each pair, expressed as a percentage), and percentage error in eigenvalue.*

The other question of how to distribute levels in the vertical is considered by generating more sets of topographies. This time each point of the Kerg A and Kerg B topographies is moved to the nearest of 32 depths specified by one of three distributions. EVEN has 32 levels each of thickness 168.20m going down to 5382.5m. LOG has 32 levels whose thickness increases geometrically from 20.7m at the top to give a bottom at 5382.5m, and FRAM has 32 levels that are concentrated toward the top (the first layer being 20.7m thick again), but become evenly spaced lower down to give a bottom at 5499m (the discrepancy between bottom depths is a mistake).

Topographic modes for the six topographies were again calculated and measurements of the errors are shown in fig.(3.5). This time some fairly consistent answers were obtained. Not all of the 32 levels were used in any of the topographies, so the number of levels used affected the effective resolution and this showed up quite strongly in the results.

For Kerg A, with no very shallow water, EVEN (using 26 depths) was best, followed by FRAM (20 depths) and then LOG (11 depths). Comparing FRAM (20 depths) with one of the solutions with 20 grid points shows that the errors are very similar, showing that no information is lost in FRAM apart from what is implicit in having only 20 grid points resolution.

For Kerg B, with a very shallow region, LOG (22 depths) produced the most accurate vectors but FRAM (23 depths) produced the best eigenvalues, with EVEN (21 depths) a poor third in both cases. In this case, the 20 grid point solutions are clearly better than the EVEN solutions, so EVEN has lost more information than is implicit in 20 (or 21) grid point resolution. This is due to its failing to resolve the important structure in the shallow regions.

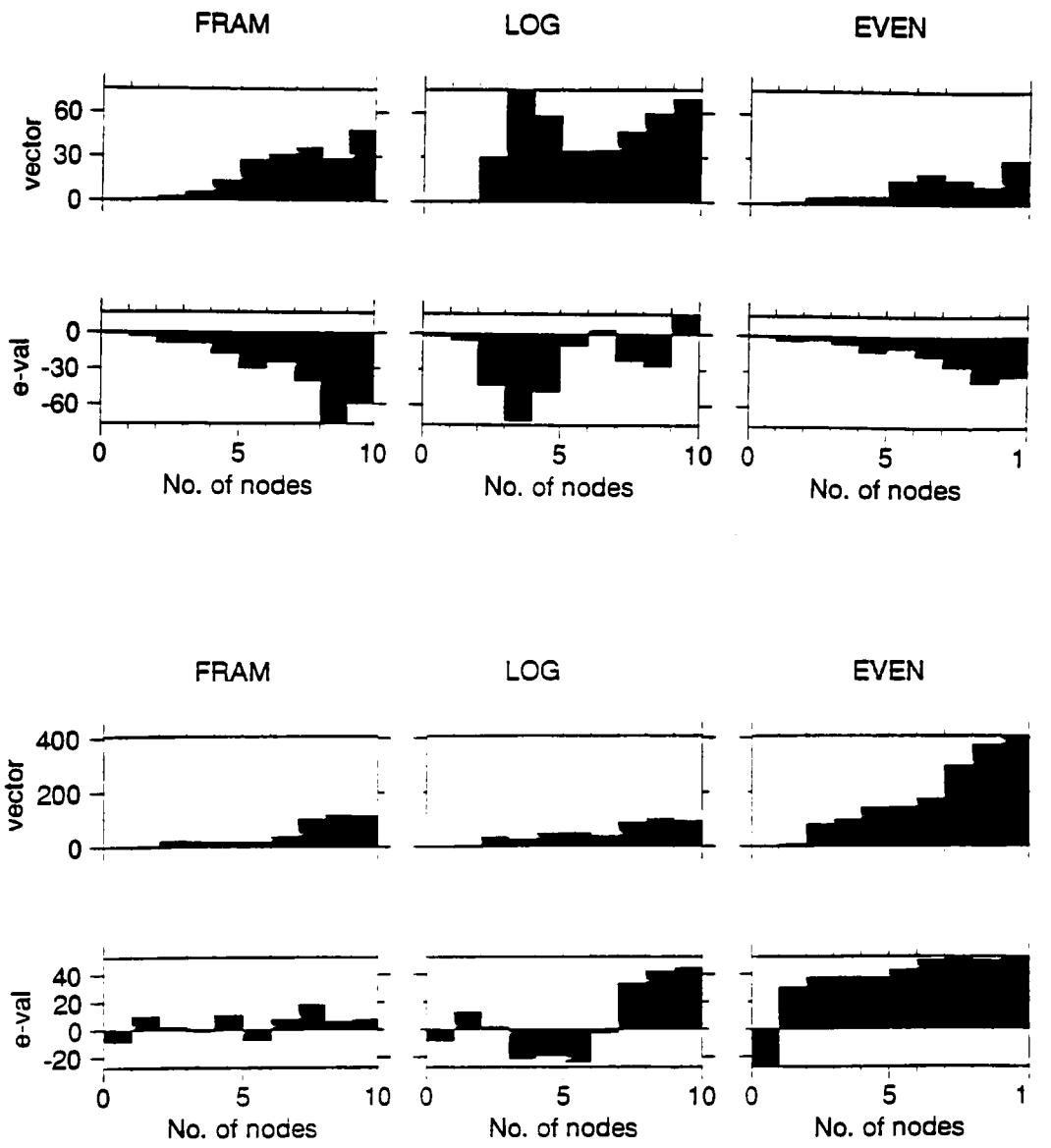


Fig.(3.5). *The same as fig.(3.4) except the comparison is between three different methods of distributing levels in the vertical; using the FRAM levels (left), a logarithmic distribution concentrated near the surface (centre) and a simple regular distribution (right).*

Overall, then, FRAM scores best, scoring over EVEN when shallow water regions are important, and scoring over LOG in two ways; by not squandering its resources entirely in the upper layers when the topography does not reach into them, and by resolving the deep water structure better in the large, deep regions where the difference between some slope and none at all can be crucial. The strong dependence of the results on number of levels used indicates the importance of tailoring the vertical distribution to maximise the number of levels used in the kinds of slope actually incorporated in the model, bearing in mind the relative importance of shallow regions, where they occur.

A rough criterion for optimizing the distribution of depth levels is therefore to maximise the number of pairs of adjacent horizontal grid points which differ in depth by one vertical grid point. If there is no difference in depth between adjacent grid points then horizontal resolution is effectively wasted, and jumps of more than one vertical grid point are also a problem as is shown in chapter 4. This optimized distribution takes no account of the extra importance of resolving depths near the surface except in so far as there are a large number of small depth changes over the continental shelf, and it ignores the need to resolve the thermocline, so it may be advisable to 'tack on' a region of increased resolution in the top few hundred metres to account for these requirements, and to perform the optimization only on depths below this.

Wave modes in the Southern Ocean.

Equation (3.1) represents the wave mode equation only at internal points, and must be augmented by boundary conditions. In the case of a closed basin, the boundary condition $\Psi = 0$ is appropriate, and is easy to accommodate in eq.(3.1) by shrinking the matrices

so as to exclude boundary points, which has the effect of implicitly setting the stream function on the boundary to zero. For an ocean with more than one coastline, the more complicated boundary condition is impossible to incorporate into the generalised eigenvalue problem form of eq.(3.1) without making both $\mathcal{J}$ and $\mathcal{D}$ singular and thus making solution more complicated. For this reason the modes of the Southern ocean were calculated using the wrong boundary condition, $\Psi = 0$, thus producing modes which were incapable of modelling any nett flow around Antarctica. The aim in the first instance was not to form a description of the flow in the Southern Ocean as a superposition of the calculated modes, but to get a picture of the general character of basin modes with a realistic topography, and this aim should not be greatly affected by the change in boundary conditions. Even with this simplification, solving eq.(3.1) using the QZ algorithm was slightly problematic. In several cases one or more eigenvalues were returned as indeterminate, showing that $\mathcal{J} - i\omega\mathcal{D}$ is a 'singular pencil', in other words $\det(\mathcal{J} - i\omega\mathcal{D}) = 0$ for some ω (see Wilkinson, 1979). When this occurs, although it frequently happens that the other eigenvectors are correct, it is quite possible that they are all wrong, so only those solutions which included no indeterminate eigenvalues may be treated with confidence.

The domain for solution was taken to be the same as in FRAM, and depths were taken as the median of the FRAM depths falling within a region bounded by half grid point values of latitude and longitude. The highest resolution used (limited by a combination of memory space and cpu time requirements) was 72 points of longitude, separated by 5° intervals, by 17 points of latitude separated by 3.25° intervals. The northern boundary, open in FRAM, was taken as a solid wall at 24° S.

Three particular solutions, all at the same resolution, will be considered. The first is a basic run in which depth and land points are calculated as described above. The second includes a two grid point representation of New Zealand as land, since New Zealand is too small to be resolved in the basic solution. The third has no New Zealand but has an artificial barrier placed across the Drake Passage.

The interaction between these changes and the artificial boundary condition is interesting. Blocking the Drake Passage does not stop any nett circulation around Antarctica, since the boundary condition ensures that there can be no circulation anyway, so the effect of the barrier should tell us about the importance of communication between ocean basins, separately from the question of allowing or disallowing circulation. The presence of New Zealand has almost the opposite effect. It hardly alters the topography of the region, and doesn't stop communication between one basin and another. However, the boundary condition—stream function is zero on all land points—means that the presence of New Zealand stops there from being any nett circulation around the island. The run with New Zealand, then, should tell us about the importance of the boundary condition.

The results of these three runs are summarised in table (3.1), in which the 48 shortest period modes of the standard run are matched with corresponding modes of the other runs. The fact that, in all cases but one, a match can be made shows that these apparently quite drastic alterations have no global impact at all, and relatively little local impact. This is illustrated in fig.(3.6), which shows mode 10 for the three cases. Despite occupying almost the whole Pacific region, the effects of New Zealand and the Drake Passage on this mode are very small, the largest effect being a 1.5% change in period due to the inclusion of New Zealand. Mode 10 is the largest scale mode observed, and demonstrates

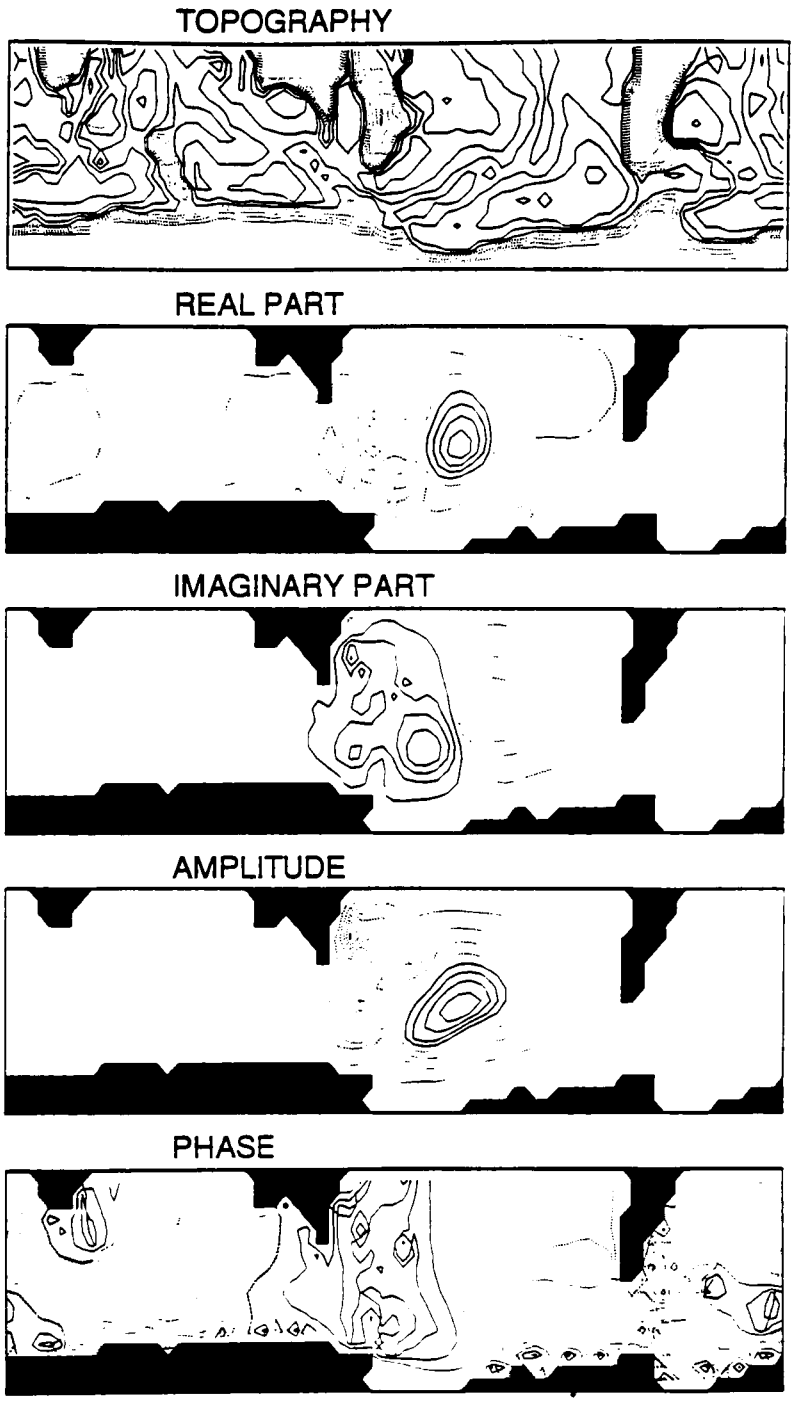


Fig.(3.6) a) *The largest scale mode found in a calculation of basin modes for the Southern Ocean, located above the Pacific Antarctic rise. The resolution of the calculation is 72 grid points east-west by 17 grid points north-south. This mode has a period of 75.46 hours.*

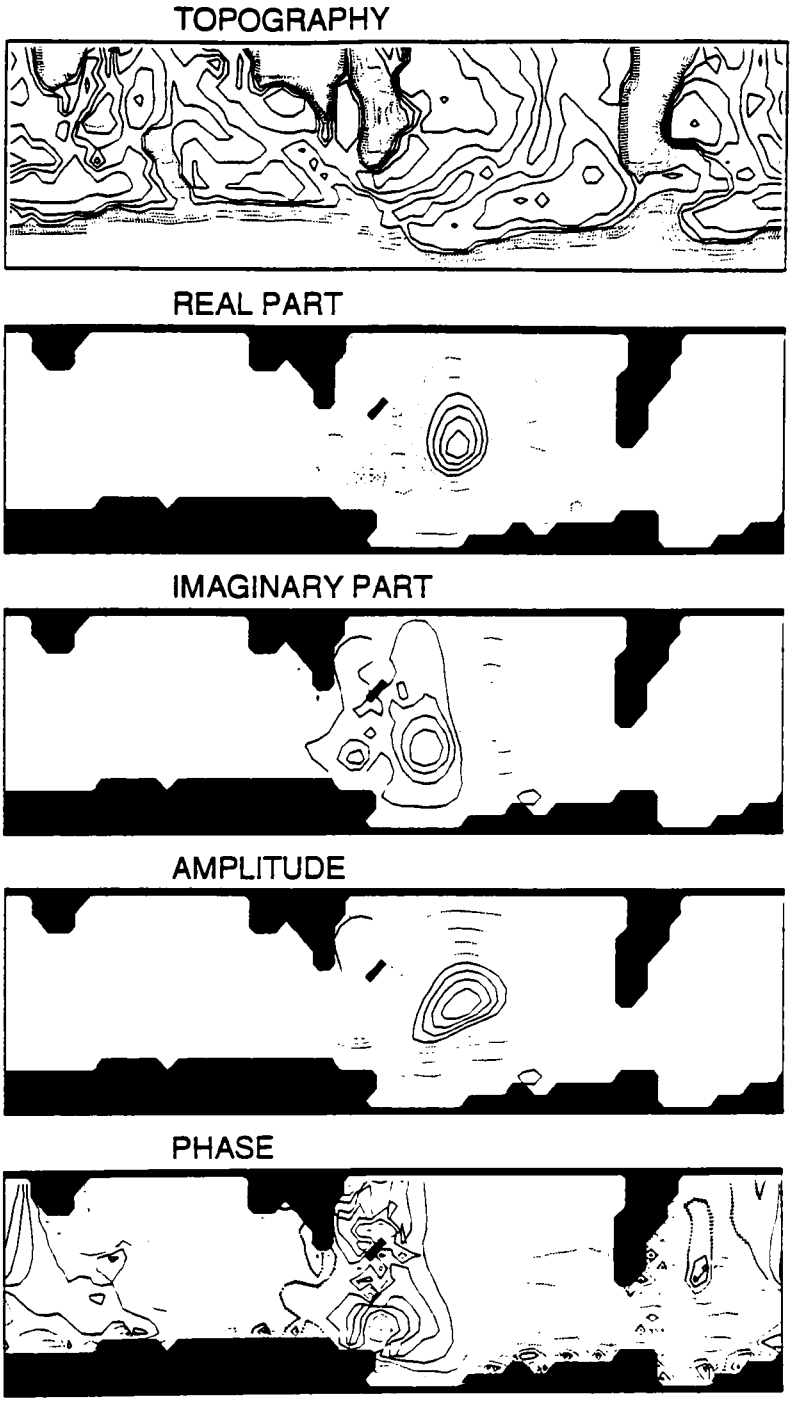


Fig.(3.6) b) *As for fig(3.6) a but with New Zealand added. The period is altered to 76.61 hours.*

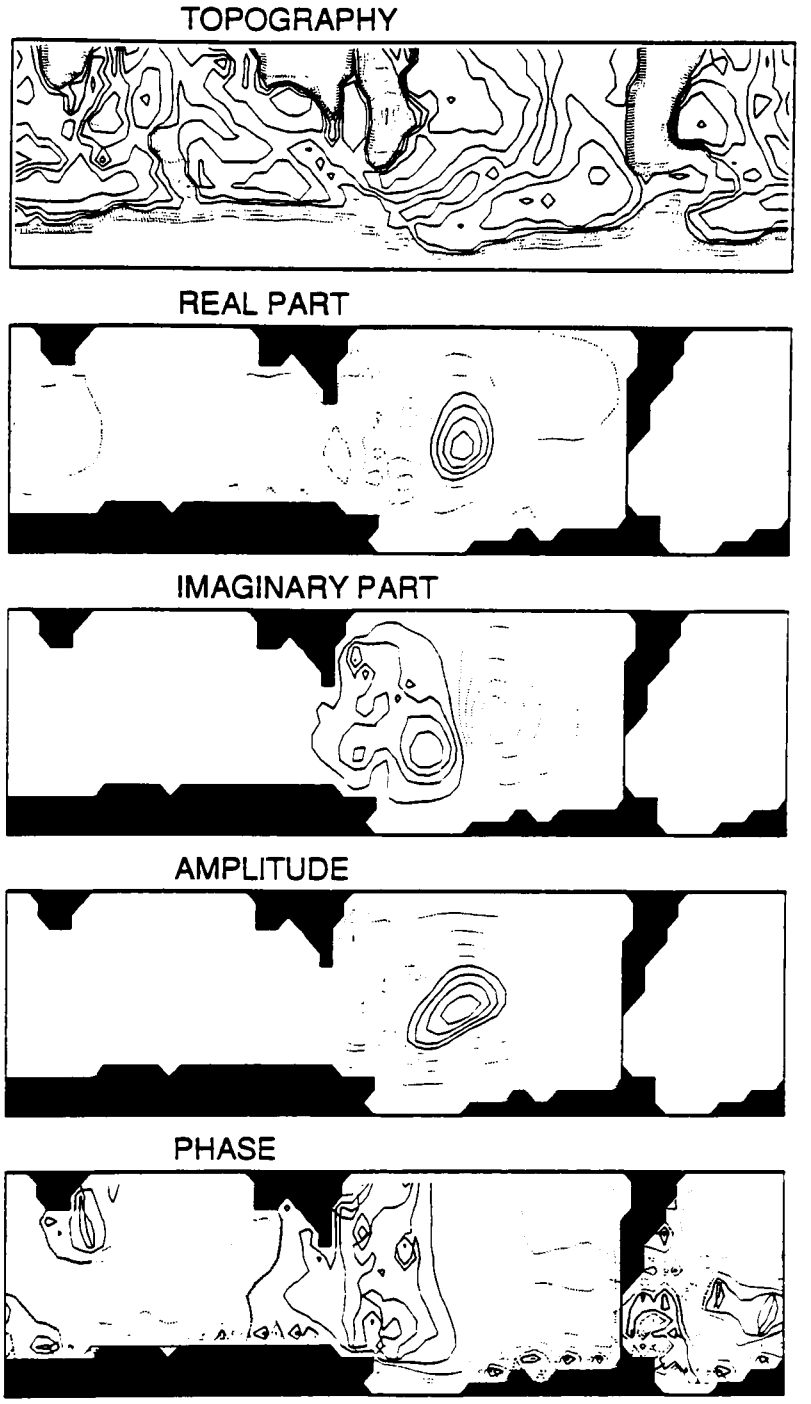


Fig.(3.6) c) *As for fig.(3.6) a but with the Drake Passage closed off. The period is altered to 75.47 hours.*

Table (3.1)

Description	Period (hours)	New Zealand period	Drake Passage period
2-lobed, W of Kerguelan	94.43	94.43	94.44
Wavenumber 2 on Pacific ridge	91.92	91.98	91.92
Small structures in Atlantic	90.95	90.96	91.15
Small structures in Atlantic	89.61	89.61	89.89
Small structures in Pacific	86.73	87.85	86.73
Wavenumber 3 on Pacific ridge	83.62	**84.65**	83.63
Small structures in Atlantic and S Indian Ocean	81.34	81.34	81.35
Small structures in Atlantic	81.20	81.20	81.21
Wavenumber 1 in S Drake Passage	15.24	15.24	**15.94**
Small structures in Atlantic and to E of Africa	78.61	78.62	78.65
Large mode, wavenumber 1 on Pacific ridge	75.46	**76.61**	75.47
Large mode on ridge to S of Africa	75.74	75.75	75.77
Small structures in Drake Passage	21.29	21.29	**21.65**
Small structures in Drake Passage with phase showing wavenumber 1	21.91	21.91	**22.05**
Shelf wave in Ross Sea, wavenumber 1	25.45	25.46	25.45
Small structure on shelf at longitude 0	24.80	24.80	24.80
Shelf wave in Weddell Sea, wavenumber 1	25.84	25.84	25.85
Small structures in Drake Passage	26.88	26.88	**31.44**?
Large, 2-lobed structure covering New Zealand Plateau	28.02	**29.24**	28.02
Large, 2-lobed structure covering New Zealand Plateau	28.53	**32.66**	28.54

Table (3.1) continued

Description	Period (hours)	New Zealand period	Drake Passage period
Shelf wave, wavenumber 1, to W of Drake Passage	28.40	28.40	28.48
Shelf wave, wavenumber 1 in Ross Sea	36.63	36.63	36.63
4-lobed structure at N end of New Zealand Plateau	35.11	**37.11**?	35.11
Wavenumber 1.5, S of Australia	35.29	35.29	35.29
Shelf wave between Drake Passage and Ross Sea. Wavenumber 1	38.03	38.03	38.03
Wave radiating outwards from Argentuna	38.48	38.48	**43.03**?
Complex structure on N of New Zealand Plateau	40.01	**40.67**?	40.01
Large, 2 or 4-lobed structure on Kerguelan Plateau	41.83	41.84	41.84
Shelf wave to W of Ross Sea, wavenumber 1.5	44.00	44.01	44.01
Small structures in Weddell Sea and Drake Passage	43.83	43.83	**44.50**
Shelf wave from Ross Sea to Drake Passage, wavenumber 1.5	47.62	47.63	47.66
Shelf wave from Ross Sea to Drake Passage, wavenumber 2	48.34	48.34	48.37
Small structures along coast from Ross Sea to Weddell Sea	48.54	48.55	**48.96**
2-lobed structure with amplitude highest at ends of Kerguelan Plateau	50.16	50.16	50.16
Small structures over New Zealand Plateau	52.27	**55.79**?	52.27
Shelf wave in and to E of Ross Sea	52.72	52.72	52.72
Large structure moving SE to NW in Pacific	71.19	**63.42**?	71.20
Small structures in Weddell Sea and near Falklands	54.45	54.45	**65.07**?
2- or 4-lobed structure over Kerguelan Plateau	54.83	54.84	54.84
Small structures to E of and in Drake Passage	55.06	55.06	**54.64**?

Table (3.1) continued

Description	Period (hours)	New Zealand period	Drake Passage period
Shelf wave on Antarctica S of Africa, wavenumber 1	70.37	70.37	70.25
Large structure, wavenumber 2 along ridge in Indian Ocean	67.85	**67.44**	67.86
Large structure in Atlantic, moving towards Argentina	68.46	68.46	**62.44**?
Large structure in Pacific, concentrated on S coast	66.60	66.48	66.65
Shelf wave on Antarctica, S of Africa. Wavenumber 1	64.79	64.79	64.79
Large structure on idge between Ross Sea and Australia. Wavenumber 1	64.32	**None**	64.33
Small structures on New Zealand Plateau	60.59	**70.77**?	60.59
Wavenumber 1.5 to S of Australia	61.39	61.44	61.39
Small structures on Kerguelan Plateau	62.78	62.78	62.78

Table (3.1). *the periods of the 49 shortest-period basin modes in a 72 by 17 grid point model of the Southern Ocean, along with the equivalent modes in two variants of the model, one with New Zealand added and the other with the Drake Passage closed. Question marks indicate modes which were difficult to identify visually, and bold type marks modes which would be expected to be affected by the change to the model.*

the typical pattern of being confined to one particular region, and being easily identified with a significant piece of topography—in this case the Pacific–Antarctic Rise. There are virtually no inter-basin connections, and most of the modes are much smaller than mode 10, being confined to a smaller topographic feature or shelf region. These 48 modes are all clearly topographically controlled, and all the longer period motions that were looked at were very small scale, so beta clearly has very little effect on the barotropic modes.

The effects of New Zealand and the Drake Passage can be seen from table (3.1). Identification of the modes was made by visual comparison, and there is a question mark next to the modes for which identification was unsure. Note that the question marks are always next to numbers in bold script. Bold script is used for modes which have significant amplitude in the region of altered land. The fact that question marks occur next to emboldened numbers shows that the influence of the change is very local. Note also that many of the numbers in bold have no question marks by them and match fairly well with the period derived from the standard run. This is most dramatically demonstrated with mode 8 (fig.(3.7)), which is pressed right up against the Drake Passage barrier, but was easily visually identified and has a period altered by less than 5%. Figure (3.8) demonstrates a similar result for mode 18 over New Zealand. The boundary condition forces a rather larger change in structure of the mode, but the eigenvalue again changes by less than 5%. Apparently large topological changes, then, seem to produce effects only very locally, and even locally the effects on a particular mode can be quite small. The only large effect is on the ability of a sum of modes to represent a nett circulation around a given region.

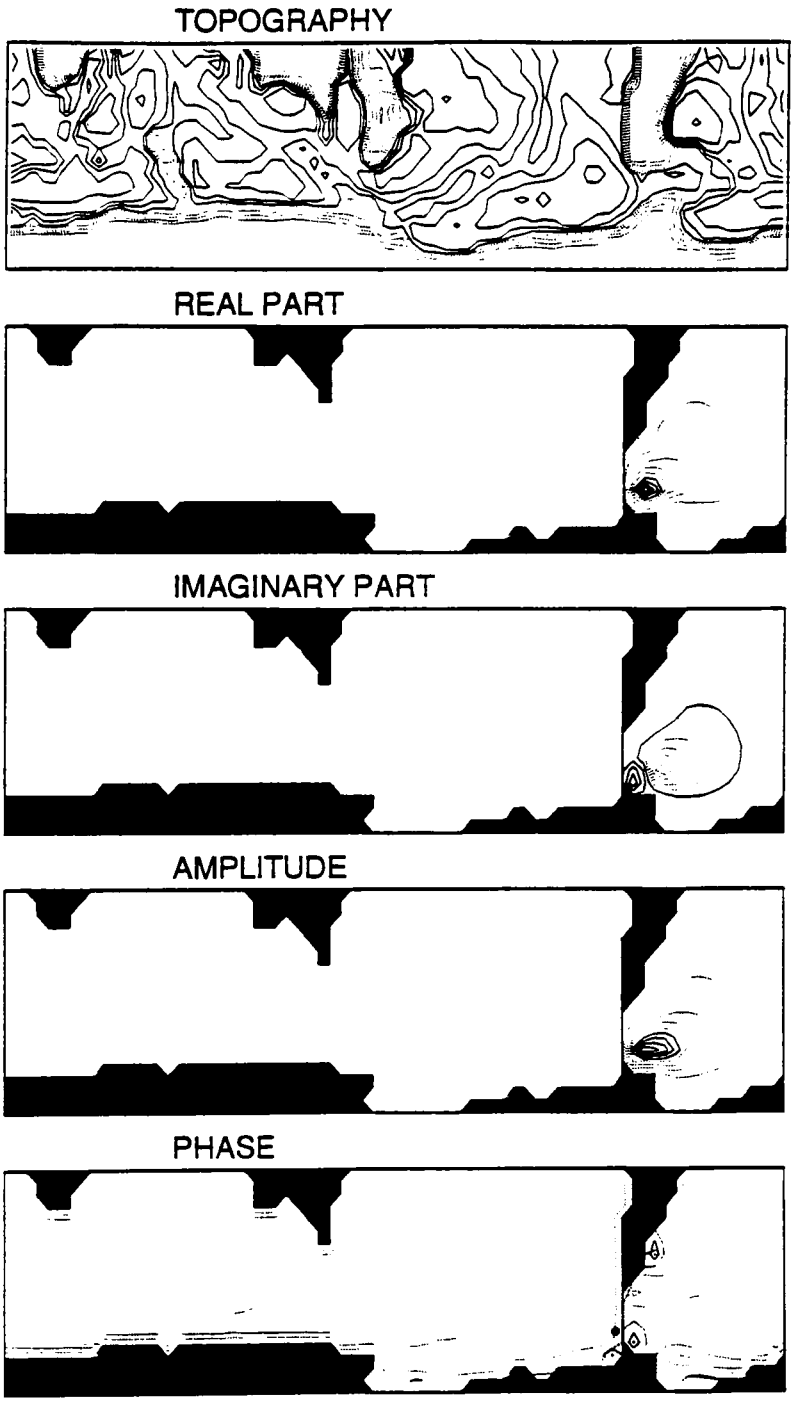


Fig.(3.7) *A basin mode as in fig.(3.6), located in the Drake Passage, whose period is only altered slightly by the closure of the Drake Passage. The period is 15.94 hours, compared with 15.24 hours with the Drake Passage open.*

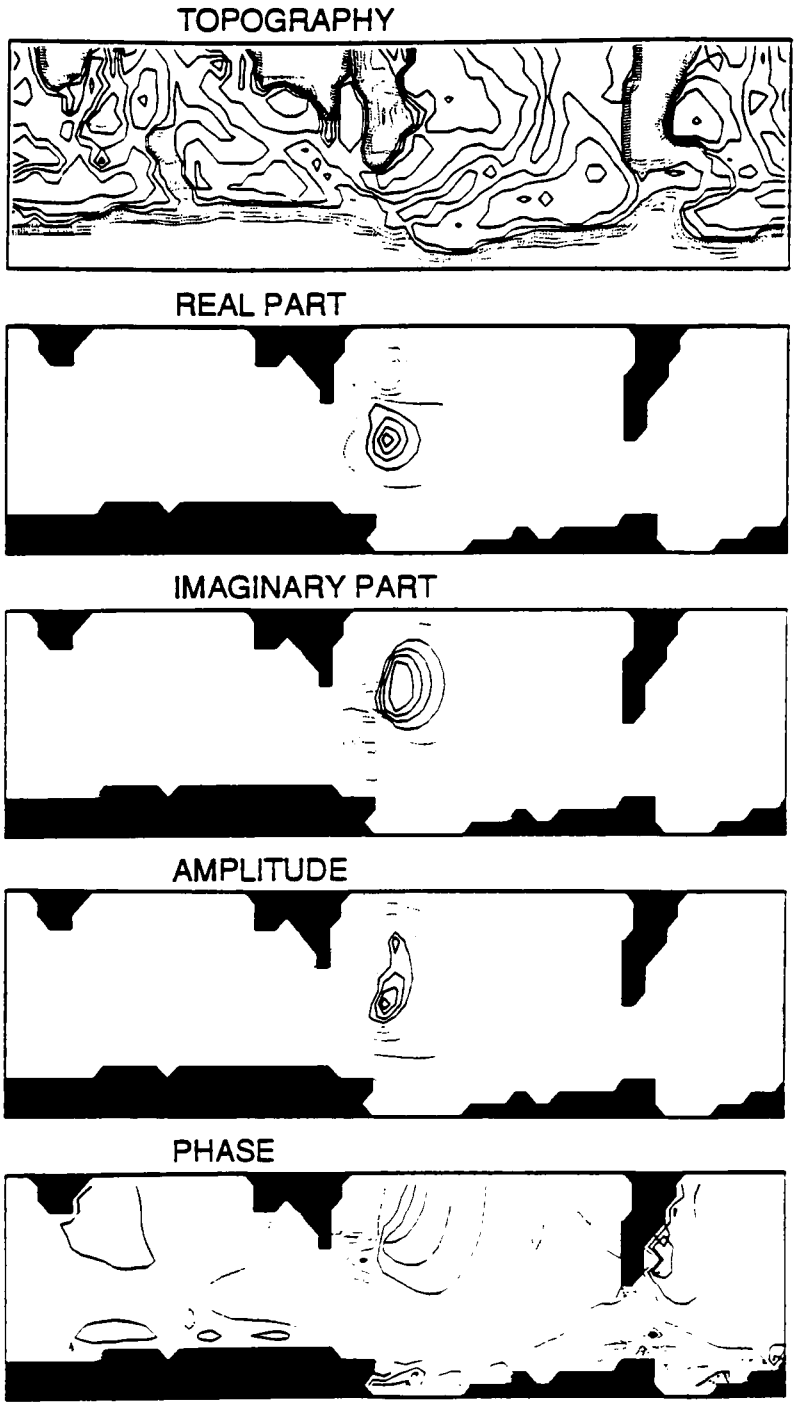


Fig.(3.8) a) *A basin mode as in fig.(3.6), located on the New Zealand Plateau. This mode strongly resembles one of the structures seen in the coarse resolution run of FRAM which became unstable. It has a period of 28.02 hours.*

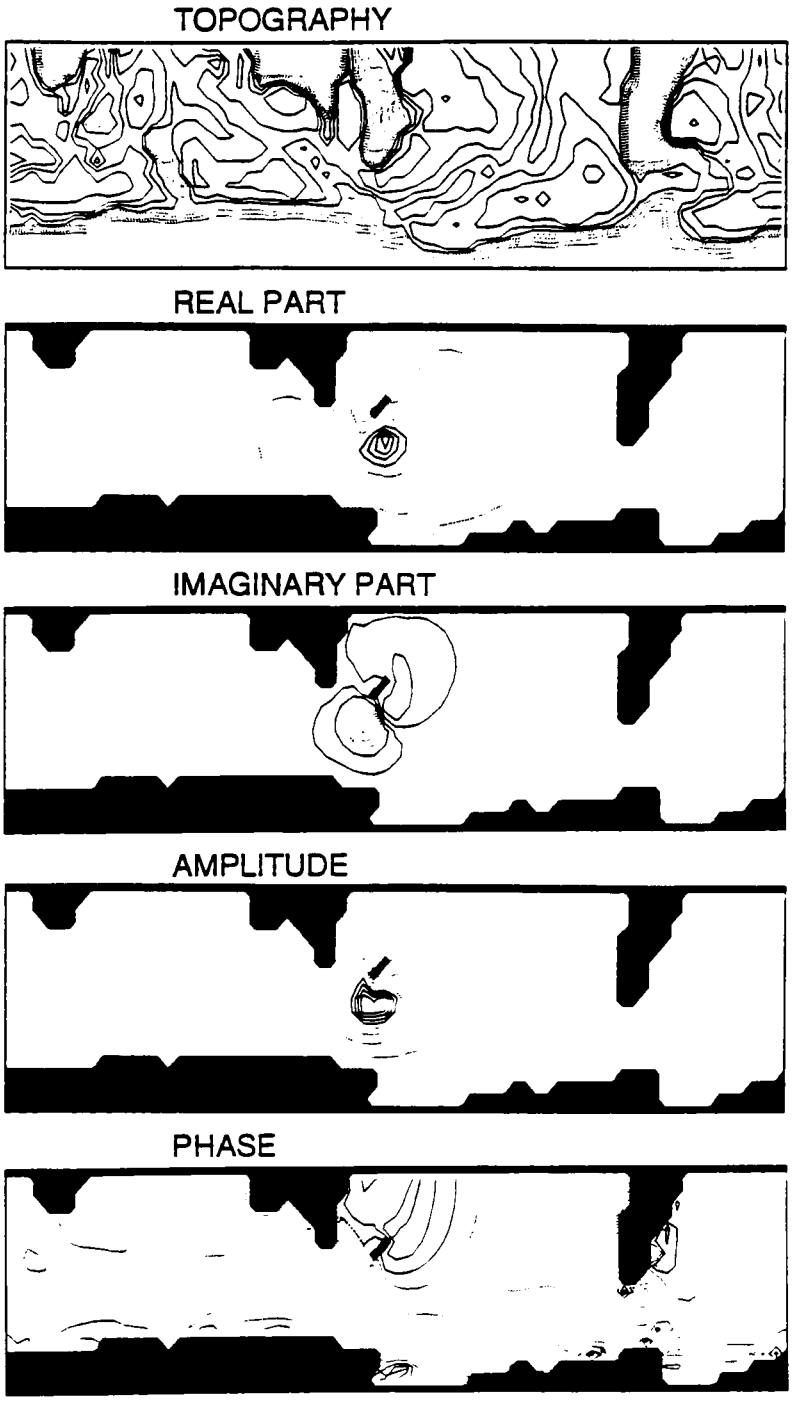


Fig.(3.8) b) *As fig.(3.8) a but with New Zealand included. The inclusion of the island makes surprisingly little difference to the form of the mode and only alters the period slightly to 29.24 hours.*

Summary.

An analysis of errors due to lack of horizontal resolution produces the rule of thumb that the eigenvalue of a topographic wave will be correct to within 3% if there are 5 grid points per half wavelength in both directions, 1% if there are 10 grid points. Results of an investigation of vertical resolution suggest that the way in which an "average" depth is chosen for a grid point is relatively unimportant. Far more important is the vertical distribution of depths used if the depth must be approximated by one of a number of given values. A regular spacing of depths does very badly in regions of shallow topography, as the governing equation suggests, but the logarithmic sequence of depths which is suggested by the governing equation is actually not as good as the distribution of depths used in FRAM, which has a logarithmic concentration near the surface but a more regular spacing in deep regions. The reason for this is found to be that the actual distribution of depths in the ocean is such that many of the depths in the logarithmic distribution are not used, whereas more depths are used from the FRAM distribution. There are two constraints defining the best vertical distribution of depths. The depths must be concentrated towards the top, and must be arranged so as to maximise the number of depths used along a "typical" slope. This latter constraint depends on the horizontal resolution that is chosen.

Topographic modes calculated for a fairly coarse resolution Southern Ocean are found to be very localised, with effects of boundary conditions and factors distant from the region in which the mode has large amplitude producing only small alterations in the structure of the modes. These small alterations can, however, be important in allowing or disallowing a circulation around a given region, suggesting that even the higher, smallest scale modes may be important in satisfactorily representing such a circulation. The modes were found

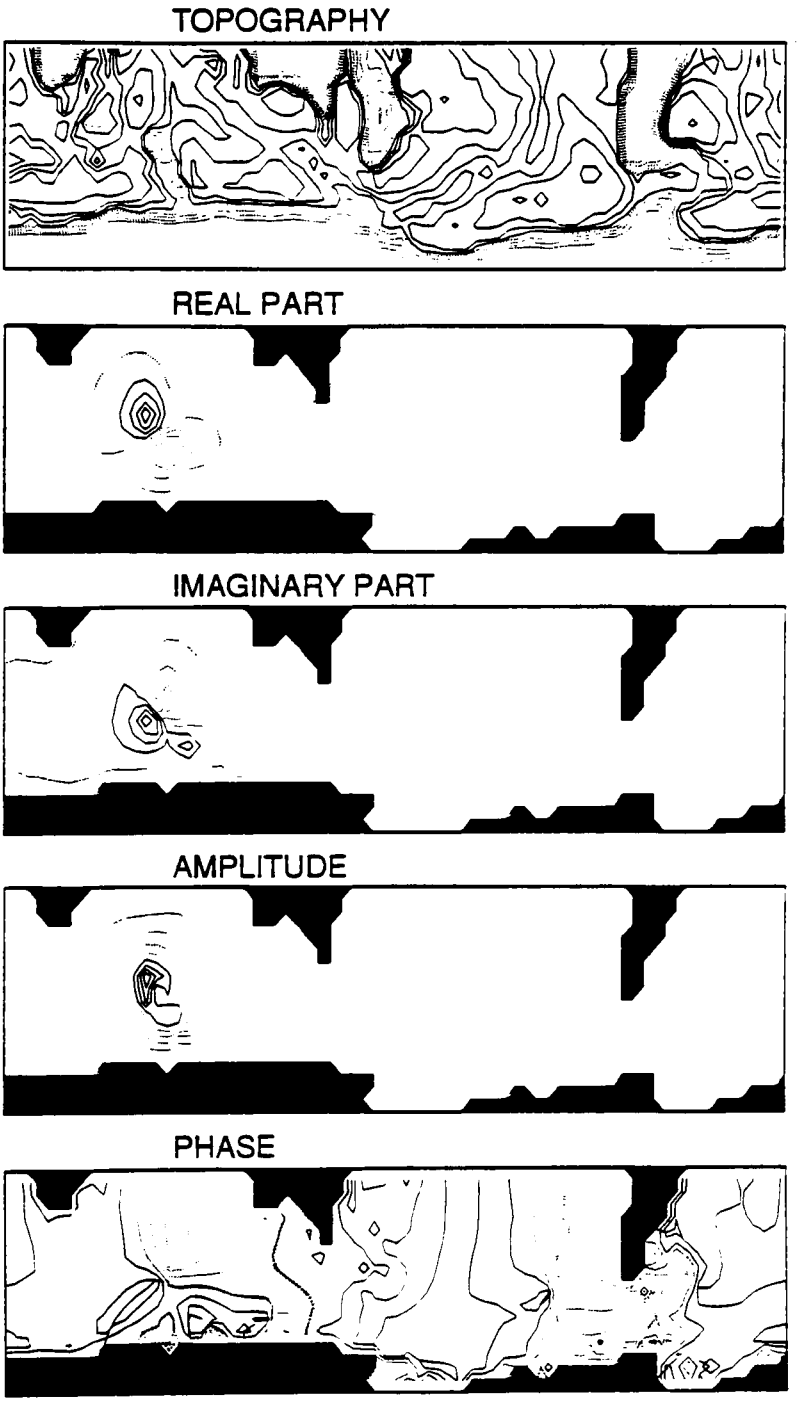


Fig.(3.9). *A basin mode as in fig.(3.6), located on the Kerguelan Plateau. This mode strongly resembles the other structure seen in the unstable, coarse resolution run of FRAM. It has a period of 41.83 hours.*

to be topographically controlled, with beta unimportant, and the modes which caused problems in CRAM were easily identified(fig.(3.8) and fig.(3.9)).

The conclusion, then, has to be that although topographic modes are an interesting theoretical tool, useful for describing spin-up or the response of the ocean to a change in forcing, their use in a real situation to describe basin-wide flows is impractical due to the scale of typical modes, and hence to the number necessary to provide reasonable resolution. Nonetheless, in regions where contours of f/H are very wiggly, they do have an advantage over integration along characteristics as a means of finding the steady state solution in that they avoid the problem of accumulation of errors resulting from integrating along the long path length of a contour of f/H. They may also be useful in limited sub-regions of the ocean basin, where the topography may vary rather more smoothly over the whole region. The problem of their being basin modes and therefore needing to be calculated over the whole basin might be circumvented by using the variational method described by Johnson (1989) to calculate modes in open domains.

Chapter 4.

Effects of Stratification.

Discussion of Previous Work.

A number of models of the Southern Ocean were discussed in the introduction, and several of the numerical models included some representation of stratification, but the innate complexity of the interactions between stratification and topography means that there have been few analytical studies which have produced any useful results.

Hide (1971) and Hogg (1973), among others, looked at the behaviour of the Taylor column in a stratified fluid, and concluded that it would decay away from the topography over a characteristic height scale $H = fL/N$, where N is the Brunt-Väisälä frequency. Both therefore agreed on the importance of the parameter $S = N^2H^2/f^2L^2$ as a measure of the importance of stratification. Taking from Levitus (1982) a value of $\partial\rho/\partial z = -5 \times 10^{-4}\,\mathrm{kg\,m^{-4}}$ for the Southern Ocean, and choosing $H = 5000\,\mathrm{m}$, $f = 10^{-4}\,\mathrm{s^{-1}}$ gives $S = 1$ at $L \approx 100\,\mathrm{km}$, so over length scales larger than about 100 km, topographic effects in the Southern Ocean should be expected to influence the whole fluid column.

When a flat bottom is assumed, stratification is simple to take into account in terms of vertical normal modes (see Gill, 1982), but the inclusion of topography makes separation of variables impossible. Rhines (1970) tackled this problem by looking at waves decaying exponentially away from a boundary at a small angle to the horizontal, and found that short waves tend to be trapped more strongly to the bottom, obeying the scaling $S \approx 1$

where H and L are taken as characteristic scales for the wave. He then examined the low frequency limit for constant slope on a beta plane, and found that stratification causes topographic waves to split up into a set of vertical modes. One of these modes is like the barotropic Rossby wave, only slightly modified so as to have a node near the bottom thus diminishing the topographic influence. Other modes either have a node near the bottom and resemble baroclinic Rossby waves, or are bottom trapped and thus strongly influenced by topography. Stratification is thus seen to separate the influences of β and of topography. The separation is far from complete though, and when the waves get long enough so that $S \ll 1$ the waves become essentially barotropic and contours of f/H become important (Straub, 1990).

A number of models of the effects of stratification on shelf waves have been produced. Mysak (1967) showed that , when the stratification is significant only below the continental shelf break, a significant increase in the phase speed of the waves can result, but later studies (*e.g.,* Wang, 1975) showed that this is not true when the stratification extends over the shallow sloping region as well. Other studies—an analytical model by Allen (1975) and a numerical simulation by Wang and Mooers (1976)—produced similar results, but general results other than those given in Rhines (1970) were few and far between until Huthnance (1978) published a thorough investigation of the problem for a straight coastline and monotonic depth profile on an f-plane. Looking at the problem in terms of differential equations, variational methods, finite difference analysis and finally computer simulation, this paper showed that the direction of phase propagation on an f-plane is unaltered by stratification, and that the relationship between the barotropic shelf wave and the double Kelvin wave discovered by Longuet-Higgins (1968) can be extended to encompass shelf waves with along-slope nodes as well. As the stratification increases so that S becomes

large for a given wave, the nodal surfaces become inclined, tending towards the horizontal plane, and the shelf waves become internal Kelvin waves, trapped against the slope which now acts like a vertical wall, since it is "steep" as defined in terms of S.

All these investigations rely on simplifying the equations by allowing the topography to vary only in one dimension. The BPV equation provides us with another approach to the problem, in which vertical integration simplifies the equations by making them two dimensional in the horizontal plane, with the effects of stratification simply represented by a two dimensional field. Advection of density, of course, is due to the full three dimensional fields, so some approximations have to be made to close the problem in two dimensions. In this form, the problem consists of an interesting feedback loop. The JEBAR term forces a current to flow, which in turn advects density thus causing a change in the JEBAR term. Shaw and Csanady (1983) used this description to show how density perturbations on the continental slope could induce a flow which causes them to be carried along the slope. It is also possible to rewrite a two-layer model so that, instead of separate equations for the two layers, the two equations are the BPV equation and an equation for the evolution of the JEBAR term. More recently, Olbers and Wübber (1991) have proposed a simple model that is two dimensional and involves only the stream function and a form of potential energy. It is to this description, and variations on it, that this chapter is devoted.

Advection by barotropic flow.

Equation (1.9):

$$\nabla \cdot \left(\frac{\nabla \Psi_t}{H} \right) + J(\Psi, f/H) = \frac{1}{\rho_0} J(E, 1/H),$$

shows that (in the absence of wind, friction and non-linear terms,) the barotropic flow can be considered to be driven by a single quantity involving only the distribution of density, $E = g\int_{-H}^{0} z\rho\, dz$. In order to close the feedback loop described above, all that is needed is to calculate how advection of density causes E to change.

Olbers and Wübber (1991) looked at a particularly simple way to close this loop, by considering advection only by the barotropic component of the flow. Since this is fully specified in terms of the stream function, it is simple to calculate and turns out to have a particularly neat form.

By continuity of volume, a barotropic flow with horizontal velocity $\mathbf{U}/H$ has an associated vertical velocity

$$w = \frac{z}{H^2}\mathbf{U}\cdot\nabla H.$$

We also know that

$$\rho_t + \nabla_3\cdot(\rho\mathbf{U}_3) = 0,$$

so

$$\rho_t = -\nabla\cdot\left(\frac{\rho\mathbf{U}}{H}\right) - \frac{\partial}{\partial z}(\rho w),$$

giving

$$E_t = g\int_{-H}^{0} z\rho_t\, dz = -g\int_{-H}^{0} z\nabla\cdot\left(\frac{\rho\mathbf{U}}{H}\right) dz - g\int_{-H}^{0} z\frac{\mathbf{U}\cdot\nabla H}{H^2}\frac{\partial}{\partial z}(\rho z)\, dz.$$

Using $\nabla\cdot\mathbf{U} = 0$,

$$\begin{aligned}-g\int_{-H}^{0} z\nabla\cdot\left(\frac{\rho\mathbf{U}}{H}\right) dz &= -g\frac{\mathbf{U}\cdot}{H}\int_{-H}^{0} z\nabla\rho\, dz - g\mathbf{U}\cdot\nabla\frac{1}{H}\int_{-H}^{0} z\rho\, dz,\\ &= -g\frac{\mathbf{U}\cdot\nabla}{H}\int_{-H}^{0} z\rho\, dz - g\rho_B\mathbf{U}\cdot\nabla H - g\mathbf{U}\cdot\nabla\frac{1}{H}\int_{-H}^{0} z\rho\, dz.\end{aligned}$$

Integrating by parts;

$$-g\int_{-H}^{0} z\frac{\mathbf{U}\cdot\nabla H}{H^2}\frac{\partial}{\partial z}(\rho z)\,dz = g\rho_B\mathbf{U}\cdot\nabla H + g\frac{\mathbf{U}\cdot\nabla H}{H^2}\int_{-H}^{0}\rho z\,dz.$$

Putting these together produces

$$\begin{aligned} E_t &= \frac{-\mathbf{U}\cdot\nabla E}{H} - 2E\mathbf{U}\cdot\nabla\frac{1}{H} \\ &= \frac{-1}{H}J(\Psi,E) - 2EJ(\Psi,1/H) \\ &= -HJ(\Psi,E/H^2), \\ \therefore \quad \frac{E_t}{H^2} &= \frac{-1}{H}J(\Psi,E/H^2), \end{aligned} \tag{4.1}$$

showing that the quantity E/H^2 behaves almost as though it is advected by the two dimensional mean flow $\overline{\mathbf{u}} = \mathbf{U}/H$. The analogy is not perfect though since, although eq.(4.1) can be rewritten as

$$e_t + \overline{\mathbf{u}}\cdot\nabla e = 0, \qquad e = E/H^2,$$

this is the conservation equation for a horizontally non-divergent flow, and $\overline{\mathbf{u}}$ is generally divergent, since $\frac{1}{H}\nabla\cdot\overline{\mathbf{U}} = \nabla\cdot\overline{\mathbf{u}} + \overline{\mathbf{u}}\cdot\nabla H = 0$.

If eq.(4.1) is accepted then we have a closed system with two equations, (1.9) and (4.1), for two variables, Ψ and E. There are, however, a number of approximations that have gone into the derivation of eq.(4.1).

Different Densities.

The density used to derive eq.(1.9) is important for its capacity to create changes with depth in the horizontal pressure gradient. That means it must be the in-situ density plus an arbitrary function of depth alone. The density used to derive eq.(4.1), however, is

important for its quality of being conserved when advected. For that to be the case when vertical advection is considered, the density must be some form of potential density.

It would be possible to reconcile these two different densities if we could write

$$\rho_{\text{in situ}} = \rho_{\text{potential}} + A(z),$$

but that relies on the compressibility of seawater being independent of potential temperature and salinity. In fact it varies by about 3%, depending most strongly on temperature, for typical ocean values. This means that, where a 10°C potential temperature difference causes a typical density difference of about $1\,\mathrm{kg\,m^{-3}}$ for two water parcels at the same depth, the different compressibilities of the two parcels mean that moving them both up or down through 1000 m would produce an additional density difference of about $0.2\,\mathrm{kg\,m^{-3}}$ (using compressibilities and thermal expansion coefficients from Gill (1982)). Something like a fifth of the density variation of the ocean is intrinsically due to interactions between pressure and potential temperature and cannot be divided into individual contributions from the two properties. Nonetheless, as a qualitative description of interactions between flow and stratification, eq.(4.1) is certainly valuable for its simplicity, and in the much more common circumstances where vertical excursions of water parcels are limited to much less than 1000 m it should be quite accurate.

Baroclinic Dynamics.

The pair of equations (1.9) and (4.1) present a tempting model of the feedback between stratification and barotropic flow, but there is one rather odd feature of the system. It was shown in chapter 1 that the forcing term $J(E, 1/H)$ can be interpreted as the difference between the geostrophic bottom velocity across isobaths and the barotropic cross-isobath

flow, the difference in the steady, geostrophic case being due to stratification and the thermal wind relationship. If we consider the bottom velocity to be different from the barotropic velocity, as we must in order to have any forcing at all in eq.(1.9), then surely it is not consistent to allow advection of the bottom density by the barotropic flow, as was assumed in the derivation of eq.(4.1). On the other hand, in a two-layer flow, the additional baroclinic flow is along contours of interfacial depth and therefore produces no change in density and makes no contribution to changes in E. In order to throw some light on this apparent contradiction it is necessary to consider the vertical velocities rather carefully.

The vertical velocity used in deriving eq.(4.1) was that which was necessary to satisfy mass continuity with a purely barotropic flow, but vertical velocities are notoriously difficult to produce in a near-geostrophic system, and it was shown in chapter 1 that the forcing $J(E, 1/H)$ can be seen as being born of the necessity of matching the geostrophic divergence with the bottom w. Values of w produced by geostrophic balance are typically much smaller than $\overline{\mathbf{u}} \cdot \nabla H$, so the forcing can be seen as the thermal wind acting to reduce the cross-isobath flow near the bottom, so that the barotropic component of the flow is shielded from the topography. Although the horizontal component of the baroclinic flow produces little change in E, the vertical component (as derived from mass continuity) must be such as to almost cancel the vertical component of the barotropic velocity in order to produce a total w which is in near-geostrophic balance.

Another complication, which arises when wind stress is present, is that the barotropic velocity no longer represents the average flow through the bulk of the fluid, since a significant part of it consists of the Ekman transport confined to a surface layer. When

considering density advection, therefore, it is necessary to subtract off this Ekman contribution to the vertically averaged flow.

In order to quantify these effects, a more complete equation for the variation of E is derived. First of all, considering the effect due to the horizontal component of the barotropic flow,

$$\begin{aligned} E_t^{(1)}(\overline{\mathbf{u}}) &= -g\int_{-H}^{0} z\overline{\mathbf{u}}\cdot\nabla\rho\,dz \\ &= -g\overline{\mathbf{u}}\cdot\int_{-H}^{0} z\nabla\rho\,dz \\ &= -g\overline{\mathbf{u}}\cdot\nabla\int_{-H}^{0} z\rho\,dz - gH\rho_B\overline{\mathbf{u}}\cdot\nabla H \\ &= -\overline{\mathbf{u}}\cdot\nabla E - gH\rho_B\overline{\mathbf{u}}\cdot\nabla H, \end{aligned}$$

so

$$E_t^{(1)} = -\overline{\mathbf{u}}\cdot(\nabla E + gH\rho_B\nabla H). \tag{4.2}$$

The next contribution to be calculated is from the horizontal component of the baroclinic velocity $\mathbf{u}'$. We might expect this to be small since the argument given above shows that the geostrophic contribution in the two-layer case is zero. Rewriting eq.(1.1) to include vertical friction,

$$\mathbf{u}_t + f\hat{\mathbf{k}}\wedge\mathbf{u} = -\frac{\nabla P}{\rho_0} + \frac{\mathbf{T}_z}{\rho_0} + \mathbf{NL}, \tag{4.3}$$

where $\mathbf{NL}$ represents the non-linear terms and $\mathbf{T}$ is the horizontal stress vector.

Averaging over depth gives

$$\begin{aligned} \overline{\mathbf{u}}_t + f\hat{\mathbf{k}}\wedge\overline{\mathbf{u}} &= \overline{\mathbf{NL}} + \frac{\mathbf{T}_T - \mathbf{T}_B}{\rho_0 H} - \frac{1}{\rho_0 H}\int_{-H}^{0}\nabla P\,dz \\ &= \overline{\mathbf{NL}} + \frac{\mathbf{T}_T - \mathbf{T}_B}{\rho_0 H} - \frac{\nabla P_B}{\rho_0} - \frac{\nabla E}{\rho_0 H} \end{aligned} \tag{4.4}$$

where the hydrostatic relation has been used to derive the second stage.

Taking the difference, (4.3)-(4.4) gives

$$\mathbf{u}'_t + f\hat{\mathbf{k}} \wedge \mathbf{u}' = \mathbf{NL} - \overline{\mathbf{NL}} + \frac{1}{\rho_0}\left(\mathbf{T}_z + \frac{\mathbf{T}_B - \mathbf{T}_T}{H}\right) - \frac{1}{\rho_0}\left(\nabla(P - P_B) - \frac{\nabla E}{H}\right) \tag{4.5}$$

where $\mathbf{u}' = \mathbf{u} - \overline{\mathbf{u}}$. Using the hydrostatic relation again, this may be rewritten as

$$\begin{aligned}\mathbf{u}'_t + f\hat{\mathbf{k}} \wedge \mathbf{u}' = \mathbf{NL} - \overline{\mathbf{NL}} + \frac{1}{\rho_0}\left(\mathbf{T}_z + \frac{\mathbf{T}_B - \mathbf{T}_T}{H}\right) \\ + \frac{g}{\rho_0}\int_{-H}^{z}\left(1 + \frac{z}{H}\right)\nabla\rho\, dz + \frac{g}{\rho_0}\int_{z}^{0}\frac{z}{H}\nabla\rho\, dz.\end{aligned} \tag{4.6}$$

If we assume that the right hand side varies slowly due to the slow advection of density, the left hand side can be written simply as $f\hat{\mathbf{k}} \wedge \mathbf{u}'$, giving

$$\begin{aligned}\mathbf{u}' = \frac{-\hat{\mathbf{k}}\wedge}{f}\Bigg[\mathbf{NL} - \overline{\mathbf{NL}} + \frac{1}{\rho_0}\left(\mathbf{T}_z + \frac{\mathbf{T}_B - \mathbf{T}_T}{H}\right) \\ + \frac{g}{\rho_0}\int_{-H}^{z}\left(1 + \frac{z}{H}\right)\nabla\rho\, dz + \frac{g}{\rho_0}\int_{z}^{0}\frac{z}{H}\nabla\rho\, dz\Bigg].\end{aligned}$$

Changes in E due to this component of the velocity can be written as

$$E_t^{(2)}(\mathbf{u}') = -g\int_{-H}^{0} z\mathbf{u}' \cdot \nabla\rho\, dz,$$

so, substituting for $\mathbf{u}'$ gives

$$\begin{aligned}E_t^{(2)} =& \frac{g}{\rho_0 f}\frac{(\mathbf{T}_B - \mathbf{T}_T)}{H} \wedge \int_{-H}^{0} z\nabla\rho\, dz + \frac{g}{\rho_0 f}\int_{-H}^{0} z\mathbf{T}_z \wedge \nabla\rho\, dz \\ &- \frac{g^2}{\rho_0 f}\int_{-H}^{0} z\nabla\rho \wedge \left[\int_{-H}^{z}\left(1 + \frac{z'}{H}\right)\nabla\rho\, dz' + \int_{z}^{0}\frac{z'}{H}\nabla\rho\, dz'\right] dz \\ &+ \frac{g}{f}\int_{-H}^{0} z(\mathbf{NL} - \overline{\mathbf{NL}}) \wedge \nabla\rho\, dz.\end{aligned} \tag{4.7}$$

The first term can be rewritten as

$$\frac{g}{\rho_0 f}\frac{(\mathbf{T}_B - \mathbf{T}_T)}{H} \wedge \int_{-H}^{0} z\nabla\rho\, dz = \frac{(\mathbf{T}_B - \mathbf{T}_T)}{\rho_0 f H} \wedge (\nabla E - gH\rho_B\nabla H),$$

and represents the correction to the advection by barotropic flow due to the fact that part of the barotropic flow is concentrated in Ekman layers and should not be taken to advect density within the bulk of the fluid.

The second term represents changes in E due to horizontal advection of density within the Ekman layers. In the limit of thin Ekman layers (low vertical viscosity), the contribution due to the shear layer at the top tends to zero thanks to the factor of z within the integral. At the bottom things are more complicated, and vertical Ekman velocities have to be considered also, the final result depending on the detailed structure of the Ekman layer and its interaction with the surrounding flow. Since this term is so complicated, and the effects of bottom friction are generally found to be small in models with topography, the contribution to the second term from bottom friction will also be taken to be zero.

The third term I have dubbed the "spiral term", because it is zero unless $\nabla\rho$ changes direction with depth. In the two-layer limit it is therefore identically zero. In fact, in FRAM, density contours on a horizontal surface are everywhere closely parallel to the barotropic flow, creating an equivalent barotropic mode for most of the region (Killworth, 1992). A simple explanation, given by Killworth, is in terms of conservation of potential vorticity on an f-plane. This requires fluid parcels to move with a constant vertical thickness, so flow in a slab of constant density must be along contours of thickness. Since this applies at all depths, there can be no turning of the horizontal density gradient with depth and the flow direction must be independent of depth. Applying this argument all the way to the bottom, of course, means that the bottom velocities, and therefore velocities everywhere, must be along isobaths. This clearly is not the case, the reason being the alternative possibility that bottom velocities are zero. Again, bottom velocities are not zero, but this is due to departures from f-plane geostrophic behaviour, which are exactly the processes which provide exceptions to the simple argument given above. Cross-isobath flow is therefore possible due to a combination of shielding of the barotropic flow from the

topography by horizontal density gradients reducing the velocities near the bottom, and beta and ageostrophic effects allowing small bottom velocities.

This argument only applies to the long term steady state, in which no advection of density is taking place. If we are interested in interactions between density advection and barotropic flows, we cannot expect the flow to be parallel to density contours, so a different form of scaling must be applied.

The simplest form of density gradient which gives a non-zero value for the spiral term is one with constant amplitude, but direction spiralling with depth, represented by $\underline{A}e^{iz/L}$, where $\underline{A}$ is a complex number $\underline{A} = a + ib$ representing the vector (a, b). Inserting this form for $\nabla\rho$ into the spiral term gives a value for the spiral term of

$$\text{spiral term} = -\frac{g^2|\nabla\rho|^2H^3}{\rho_0 f}\left[\frac{L}{2H} - \frac{L^3}{H^3}(1-\cos(H/L))\right].$$

The term in square brackets has a maximum value of about 0.1 at $H/L \approx 4$. The spiral term therefore obeys the relation

$$|\text{spiral term}| \lesssim \frac{g^2|\nabla\rho|^2H^3}{10\rho_0 f}.$$

In order to compare this with the barotropic advection term, a scaling must be made for that term under similar circumstances. Assuming no wind or non-linear terms, a constant value of $\nabla\rho$ with depth would produce a flow across contours of f/H of order $gH\nabla\rho/2f\rho_0$ (using the curl of eq. 4.4). Since this is only the cross contour flow, we have

$$|\overline{\mathbf{u}}| \gtrsim \frac{gH\nabla\rho}{2f\rho_0}.$$

This gives for $E_t^{(1)}$ a value of about $g^2H^3|\nabla\rho|^2/4f\rho_0$, but this may be larger if the flow along f/H contours is much stronger than that across contours, or smaller if the flow

happens to be along density contours. This latter possibility was necessarily ignored in the derivation of scaling for the spiral term, so we have

$$E_t^{(1)} \gtrsim \frac{g^2|\nabla\rho|^2H^3}{4f\rho_0},$$
$$\text{spiral term} \lesssim \frac{g^2|\nabla\rho|^2H^3}{10f\rho_0}.$$

The spiral term is thus typically smaller, though not by a large margin, than the horizontal barotropic advection.

The final contribution to changes in E comes from density advection by the vertical velocity field w. The vertical distribution of w, because it is innately dependent on beta or ageostrophic terms, does depend on the full details of the flow field, so it is impossible to find a simple description of the effect of the full w field. One approximation, however, is particularly neat, and that is assuming a w which varies linearly between top and bottom. That is the form which would result from a barotropic flow, but this approximation does not use the barotropic w but a linear interpolation between the actual top and bottom w values (actually, between w at the bottom of the surface Ekman layer and w at the top of the bottom Ekman layer, if such a layer is included). The weak dependence of flow with depth provides some justification for this, and the barotropic character of certain time-dependent phenomena in particular makes it a useful approximation, as will be seen in a later section. In any case, it should be a useful guide to the scale of effects due to vertical density advection.

For such a velocity field, the change in E is given by

$$E_t^{(3)}(w^{(1)}) = -g\int_{-H}^{0} w^{(1)} z\rho_z\, dz.$$

Writing $w^{(1)} = w_T + (w_T - w_B)z/H$ gives

$$\begin{aligned} E_t^{(3)} &= -gw_T \int_{-H}^{0} z\rho_z \, dz - \frac{g}{H}(w_T - w_B) \int_{-H}^{0} z^2 \rho_z \, dz \\ &= gw_T \int_{-H}^{0} \rho - \rho_B \, dz + \frac{(w_T - w_B)}{H}[2E + \rho_B g H^2]. \end{aligned} \tag{4.8}$$

Applying eq.(4.5) at the bottom, ignoring the bottom Ekman layer, and assuming the non-linear terms are negligible here, gives

$$\mathbf{u}'_{Bt} + f\hat{\mathbf{k}} \wedge \mathbf{u}' = \frac{1}{\rho_0 H}(\nabla E - \mathbf{T}_T) + g\rho_B \nabla H - \overline{\mathbf{NL}}.$$

Assuming slow variations in ρ and $\mathbf{T}$, and taking the cross product with ∇H gives

$$\mathbf{u}'_B \cdot \nabla H = -\frac{1}{\rho_0 f H} J(E, H) + \frac{1}{\rho_0 f H} \mathbf{T}_T \wedge \nabla H + \frac{1}{f} \overline{\mathbf{NL}} \wedge \nabla H,$$

but

$$w_B = -\mathbf{u}_B \cdot \nabla H = -\overline{\mathbf{u}} \cdot \nabla H - \mathbf{u}'_B \cdot \nabla H,$$

and

$$w_T = w_{\text{Ekman}} = \frac{1}{\rho_0} \nabla \wedge \left(\frac{\mathbf{T}_T}{f} \right) \approx \frac{1}{\rho_0 f} \nabla \wedge \mathbf{T}_T,$$

so

$$\begin{aligned} w_T - w_B &= \frac{1}{\rho_0 f} \nabla \wedge \mathbf{T}_T + \overline{\mathbf{u}} \cdot \nabla H \\ &\quad - \frac{1}{\rho_0 f H} J(E, H) + \frac{1}{\rho_0 f H} \mathbf{T}_T \wedge \nabla H + \frac{1}{f} \overline{\mathbf{NL}} \wedge \nabla H \\ &= \frac{H}{\rho_0 f} \nabla \wedge \left(\frac{\mathbf{T}_T}{H} \right) - \frac{1}{\rho_0 f H} J(E, H) \\ &\quad + \overline{\mathbf{u}} \cdot \nabla H + \frac{1}{f} \overline{\mathbf{NL}} \wedge \nabla H \end{aligned} \tag{4.9}$$

Also, taking the curl of eq.(4.4), we have

$$\nabla \cdot \left(\frac{\nabla \Psi_t}{H} \right) + J(\Psi, f/H) = \nabla \wedge \overline{\mathbf{NL}} + \frac{1}{\rho_0} \nabla \wedge \left(\frac{\mathbf{T}_T}{H} \right) + \frac{1}{\rho_0} J(E, 1/H), \tag{4.10}$$

which can be written as

$$\begin{aligned}\overline{\mathbf{u}}\cdot\nabla H &= -HJ(\Psi,1/H)\\ &= -\frac{H}{f}\left[-\nabla\cdot\left(\frac{\nabla\Psi_t}{H}\right)+\frac{1}{\rho_0}\nabla\wedge\left(\frac{\mathbf{T}_T}{H}\right)+\frac{1}{\rho_0}J(E,1/H)+\nabla\wedge\overline{\mathbf{NL}}\right]+\frac{\beta\Psi_x}{f}.\end{aligned}$$

So, substituting for $\overline{\mathbf{u}}\cdot\nabla H$ in eq.(4.9), we have

$$w_T - w_B = \frac{H}{f}\nabla\cdot\left(\frac{\nabla\Psi_t}{H}\right)+\frac{\beta\Psi_x}{f}-\frac{1}{f}\nabla\wedge\int_{-H}^{0}\mathbf{NL}\,dz. \tag{4.11}$$

Substituting (4.10) in (4.8) gives

$$\begin{aligned}E_t^{(3)} = \frac{1}{f}(2E+\rho_B g H^2)\times&\\ \left[\nabla\cdot\left(\frac{\nabla\Psi_t}{H}\right)+\frac{\beta\Psi_x}{H}-\frac{1}{H}\nabla\wedge\int_{-H}^{0}\mathbf{NL}\,dz\right]&\\ +\frac{g}{\rho_0 f}\nabla\wedge\mathbf{T}_T\int_{-H}^{0}\rho-\rho_B\,dz.&\end{aligned} \tag{4.12}$$

So, pulling together all the contributions, we have

$$\begin{aligned}E_t = &-\frac{1}{H}\left(\nabla\Psi+\frac{\mathbf{T}_T}{\rho_0 f}\right)\wedge(\nabla E+gH\rho_B\nabla H)\\ &+\text{spiral term}\\ &+\text{baroclinic non}-\text{linear term}\\ &+\frac{1}{f}(2E+\rho_B g H^2)\left[\nabla\cdot\left(\frac{\nabla\Psi_t}{H}\right)+\frac{\beta\Psi_x}{H}-\frac{1}{H}\nabla\wedge\int_{-H}^{0}\mathbf{NL}\,dz\right]\\ &+\frac{g}{\rho_0 f}\nabla\wedge\mathbf{T}_T\int_{-H}^{0}\rho-\rho_B\,dz\\ &+\text{advection due to remaining } w \text{ term}\\ &\quad\text{(due to baroclinic flow).}\end{aligned} \tag{4.13}$$

To find a scale for the fourth term, we note that

$$2E+\rho_B g H^2 = 2g\int_{-H}^{0} z(\rho-\rho_B)\,dz,$$

and using a constant value of ρ_z, this gives a value of

$$2E + \rho_B g H^2 = -\rho_z \frac{H^3 g}{3}.$$

Choosing the same scaling for $\overline{\mathbf{u}}$ as before—

$$\overline{\mathbf{u}} = gH\nabla\rho/2f\rho_0$$

—gives for the beta term a value of

$$\begin{aligned} &- \frac{\rho_z g H}{3} \frac{g H^3 \beta \nabla \rho}{2 f^2 \rho_0} \\ &= -\frac{2}{3} \frac{\beta H}{f} \frac{\rho_z}{|\nabla \rho|} \times \text{scaling for} E_t^{(1)}. \end{aligned}$$

With this scaling we can see that, for w due to the beta effect to have a significant effect on E_t, density contours must be tilted at such a shallow angle that they would not reach from the bottom to the top of the ocean even over distances comparable with the radius of the earth. In the Southern Ocean this is certainly not the case, with some density surfaces stretching from top to bottom over distances of order 1000 km (see *e.g.,* Nowlin, Whitworth and Pilsbury, 1977).

The processes contributing to changes in E have thus been divided into a number of separate terms, all of which are "small" in some sense. The first and second terms - horizontal advection by the vertically averaged velocity - we simply observed to be small in the sense that the mean flow in the Southern Ocean is along contours of constant density, and the flow direction is almost independent of depth. This can be thought of as the state to which the ocean is brought by density advection. Obviously, in the steady state, densities do not change, so density advection must be zero. Were the density to change, the horizontal advection terms are the ones which would be most important in bringing

about a new state of equilibrium (with one particular exception to be discussed later). There is no a priori reason to believe that the spiral term would play an insignificant rôle in this process. The scaling arguments presented show that the spiral term is potentially of the same order as the horizontal barotropic term, suggesting an important deficiency in two-layer models which are unable to represent it. In the steady state, though, Killworth's argument from conservation of potential vorticity shows that it is not just the combination of horizontal advection terms that becomes small, but each is individually small.

The remaining terms are all due to ageostrophic or beta plane processes, and are thus much smaller than many naïve scalings would suggest. For example, the fourth term in eq(4.13) has a term in square brackets

$$\nabla \cdot \left(\frac{\nabla \Psi_t}{H} \right) + \frac{\beta \Psi_x}{H} - \frac{1}{H} \nabla \wedge \int_{-H}^{0} \mathbf{NL} \, dz.$$

Ignoring wind and non-linear terms, substitution from eq. 4.10 shows that this can be rewritten as

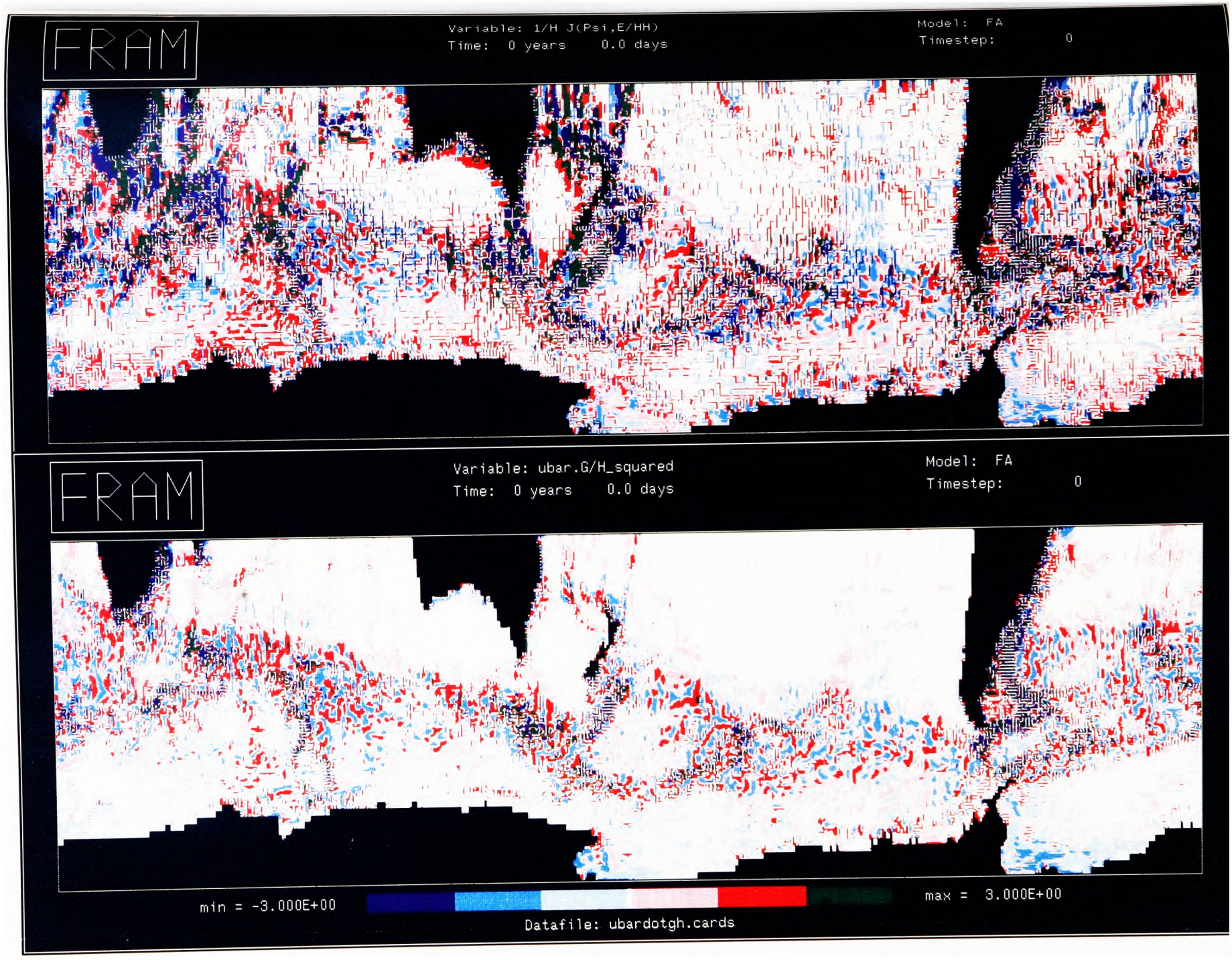
$$f J(\Psi, 1/H) - \frac{1}{\rho_0} J(E, 1/H);$$

two terms, each of which are much larger than their difference. Interestingly, this allows an assessment of the accuracy of considering only the complete barotropic velocity field for advection. As was shown at the beginning of this chapter, that assumption leads to an equation for E_t:

$$E_t^{(\mathrm{bt})} = -H J(\Psi, E/H^2).$$

Continuing to ignore wind and non-linear terms, the first and fourth terms of eq.(4.13) can be combined, using eq.(4.10), to give an equation for E_t assuming only linear dynamics

Following page: Fig.(4.1). *Two calculations of the rate of change of the quantity E/(H×H) due to advection of density. Both calculations use the six year mean values of temperature, salinity and stream function. The top figure assumes advection by the full barotropic component of the flow, including a vertical velocity component necessary to make the flow non-divergent. The bottom figure assumes advection by the same velocity field but with vertical velocity set to zero. In order to show structure on as many scales as possible, the scale used is quite complicated. The fields have been scaled in the following way: the absolute value is taken, and if less than some base value, the variable is set to that base value. The field is then divided by the base value and the logarithm is taken, giving a number which is non-negative. Finally, the sign is reattached, so the effect is that any value smaller than the base value is set to zero, and other values are plotted on a logarithmic scale while retaining their sign. Values below the below the base value are then plotted as white, while the other colours each cover a range of a factor of ten. Red and green represent positive values, and blue and purple represent negative values. Bearing in mind the logarithmic nature of the scaling, it is obvious that values in the top picture are much higher than those in the bottom picture, showing that leaving out the vertical velocity gives a better approximation to the advection. The base value used here is 3×10^-13 in S.I. units. The strange linear features in the top picture are due to the topography, which is mostly at a lower resolution than the model permits, in order to facilitate comparison with the coarse resolution model. This has the effect of producing gradients only in limited regions, often forming lines of steep gradient surrounded by plateaux.*


FRAM
Variable: 1/H J(Psi,E/HH)
Time: 0 years 0.0 days
Model: FA
Timestep: 0
FRAM
Variable: ubar.G/H_squared
Time: 0 years 0.0 days
Model: FA
Timestep: 0
min = -3.000E+00
max = 3.000E+00
Datafile: ubardotgh.cards

and no spiral term:

$$\begin{aligned} E_t^{(1,3)} &= -HJ(\Psi, E/H^2) + \frac{1}{\rho_0 f} J(E, 1/H)(2E + \rho_B g H^2) \\ &= -\frac{1}{H}\nabla\Psi \wedge \mathbf{F} + \frac{1}{f}\left[fJ(\Psi, 1/H) - \frac{1}{\rho_0} J(E, 1/H) \right] (2E + \rho_B g H^2), \end{aligned} \tag{4.14}$$

where $\mathbf{F} = \nabla E + g\rho_B H \nabla H = g\int_{-H}^{0} z\nabla\rho\, dz$.

This shows that, when ageostrophic terms are small, a better approximation to E_t is given by advection of $\mathbf{F}$ by the mean flow rather than advection of E/H^2. This is good news for models since $\mathbf{F}$ is totally dependent on density variations and is zero where there is no gradient, whereas E/H^2 is very sensitive to different representations of topography. The other interesting point about the term $E_t^{(1)} = -(1/H)\nabla\Psi \wedge \mathbf{F}$ is that, for a mean flow, any deviation from zero must show where ageostrophic and beta terms are playing an important part in density advection.

Comparison with FRAM.

In order to test these ideas out, the two quantities $-E_t^{(\mathrm{bt})}/H^2$ and $-E_t^{(1)}/H^2$ were calculated using FRAM data from a six year mean data set taken from after the relaxation to Levitus was turned off in all but the top level. Density was calculated as potential density referenced to 2000 m depth. Plots of these quantities are shown in fig.(4.1). It is clearly visible that, in regions of topographic slope, $-E_t^{(\mathrm{bt})}/H^2$ has much larger values than $-E_t^{(1)}/H^2$ (note that where there is no slope the two quantities are, by definition, equal). The distribution of $-E_t^{(1)}/H^2$ is also interesting. It reaches peak values particularly on the downstream (eastern) side of topographic features crossed by the circumpolar current. These include the southern end of the Madagascar Ridge, the east of the Kerguelen Plateau, the region between the Pacific Antarctic Ridge and the Campbell Plateau, the south-

eastern edge of the Campbell Plateau, other parts of the Pacific Antarctic Ridge especially in the Udintsev fracture zone, the eastern edge of the Falklands Plateau, the South Scotia Ridge and the Mid Atlantic Ridge. In these regions the field shows very small scale (often one grid point) fluctuations between large positive and large negative values, which must make their interpretation somewhat suspect, especially given the relative smoothness of the smaller amplitude features visible everywhere else where there is a strong eastward flow.

It is apparent, then, that the zero order balance for density advection in the Southern Ocean must be taken as $-(1/H)\nabla\Psi \wedge \mathbf{F} = 0$, with exceptions to this occurring in the form of ageostrophic processes operating throughout the circumpolar current, possibly with particularly strong ageostrophic processes over certain pieces of topography. This zero order balance corresponds to a flow in which the vertical velocity is negligible, so the barotropic flow is determined by the thermal wind relation and the constraint that the cross-isobath flow at the bottom, $\mathbf{u}_B \cdot \nabla H$, must be small (much smaller than $\overline{\mathbf{u}} \cdot \nabla H$). This, if true, approximately determines the flow field for a given density field. Slow changes in the density field would then occur as a result of the ageostrophic processes, so it must be the balance of these processes which determines the long-term structure of the density field.

Barotropic Potential Vorticity in FRAM.

The scaling arguments in the previous section resulted in a scenario in which the main effect of stratification is effectively to shield the barotropic flow from the effects of bottom

topography. This scenario has an interpretation in terms of the balance of terms in the BPV equation. Rewriting eq.(4.10) to include friction gives

$$\nabla \cdot \left(\frac{\nabla \Psi_t}{H} \right) + J(\Psi, f/H) = \nabla \wedge \overline{\mathbf{NL}} + \frac{1}{\rho_0} \nabla \wedge \left(\frac{\mathbf{T}_T}{H} \right) + \frac{1}{\rho_0} J(E, 1/H) + \nabla \wedge \overline{\mathbf{R}} \tag{4.14}$$

where $\mathbf{R}$ represents friction. If the ageostrophic and beta terms are small this reduces to

$$f J(\Psi, 1/H) \approx \frac{1}{\rho_0} J(E, 1/H)$$

which, if the equality were exact and f constant, would imply $\Psi = E + \Psi_1(H)$ (with a different Ψ_1 for each closed region of f/H contours). Again, the residual $f J(\Psi, 1/H) - \frac{1}{\rho_0} J(E, 1, H)$ should show up regions where ageostrophic and beta plane effects are important to the balance of BPV. In order to assess how well this breakdown of terms works, all the terms in eq.(4.14) (except friction which is calculated simply as "the remainder") have been calculated using FRAM data. Before presenting the results, however, it is worth pointing out some difficulties in calculating these terms from FRAM output.

All the terms in eq.(4.14) are represented in the Cox code (Cox,1984) by finite difference terms which are analogous in form, with the exception of $\frac{1}{\rho_0} J(E, 1/H)$ which is represented by a finite difference analog of $\frac{1}{\rho_0} \nabla \wedge \overline{\mathbf{G}}$, where

$$\overline{\mathbf{G}} = \frac{g}{H} \int_{-H}^{0} \nabla \int_{z}^{0} \rho \, dz' \, dz.$$

For purposes of calculating the term balance, it is obviously best to use exactly the representation that was used in the code, but for other purposes, E itself, or more precisely, E/H^2, is worth calculating. If the advection of density is due to the complete barotropic velocity field alone (and there are important circumstances where this may be the case, as

will be discussed later), it is the quantity E/H^2 which can be considered to be advected by the stream function.

If E/H^2 is to be calculated in a manner consistent with the dynamics of the model, a value for E must be found for which the finite difference analog of $J(E, 1/H)$ produces the same value as the finite difference analog of $\nabla \wedge \overline{\mathbf{G}}$. It turns out that it is possible to achieve this for all grid points except where the change in depth across one horizontal grid point corresponds to a jump of more than one vertical grid point.

The manner in which the finite difference analogy breaks down depends on the derivation of the equivalence of the two representations of the forcing term. In order to pinpoint the breakdown as clearly as possible, the derivation will be given first in the continuous case and then, following by analogy as far as possible, in the finite difference case.

For the continuous case:

$$\begin{aligned}
\overline{\mathbf{G}} &= \frac{g}{H}\int_{-H}^{0} \nabla \int_{z}^{0} \rho \, dz' \, dz \\
&= \frac{g}{H}\int_{-H}^{0} \nabla \int_{-H}^{0} \rho \, dz' \, dz + \frac{g}{H}\int_{-H}^{0} \nabla \int_{z}^{-H} \rho \, dz' \, dz \\
&= g\nabla \int_{-H}^{0} \rho \, dz + \frac{g}{H}\nabla \int_{-H}^{0}\int_{z}^{-H} \rho \, dz' \, dz + \frac{g}{H}\int_{-H}^{-H} \rho \, dz \nabla H \\
&= \nabla D + \frac{\nabla E}{H} \qquad \text{where} \qquad D = g\int_{-H}^{0} \rho \, dz.
\end{aligned}$$

Hence,

$$\nabla \wedge \overline{\mathbf{G}} = \nabla \wedge \left(\frac{\nabla E}{H}\right) = J(E, 1/H).$$

For the finite difference case, it is useful first to derive an equation for the equivalent to taking the gradient outside the integral sign.

$$\sum_{k=1}^{K_0} \delta_x A \Delta_k = \frac{1}{\delta x}\left(\sum_{k=1}^{K_0} A_{\frac{1}{2}} \Delta_k - \sum_{k=1}^{K_0} A_{-\frac{1}{2}} \Delta_k\right)$$

$$= \frac{1}{\delta x}\left(\sum_{k=1}^{K_{\frac{1}{2}}} A_{\frac{1}{2}}\Delta_k - \sum_{k=1}^{K_{-\frac{1}{2}}} A_{-\frac{1}{2}}\Delta_k - \sum_{k=K_0+1}^{K_{\frac{1}{2}}} A_{\frac{1}{2}}\Delta_k + \sum_{k=K_0+1}^{K_{-\frac{1}{2}}} A_{-\frac{1}{2}}\Delta_k\right)$$

where K_0 is the number of vertical grid points at the horizontal point x_0, $K_{\frac{1}{2}}$ is the number at $x_0 + \frac{1}{2}\delta x$, $K_{-\frac{1}{2}}$ at $x_0 - \frac{1}{2}\delta x$. If x_0 is at a u, v point then from the way the Cox code handles topography,$K_{\frac{1}{2}} \geq K_0$ and $K_{-\frac{1}{2}} \geq K_0$ since depths at u, v points are the minimum of the depths of the surrounding T, S points.

Writing $\overline{A} = \frac{1}{H}\sum_{k=1}^{K} A_k\Delta_k$, we have

$$H\overline{(\delta_x A)} = \delta_x(H\overline{A}) + \frac{1}{\delta x}\left(-\sum_{k=K_0+1}^{K_{\frac{1}{2}}} A_{\frac{1}{2}}\Delta_k + \sum_{k=K_0+1}^{K_{-\frac{1}{2}}} A_{-\frac{1}{2}}\Delta_k\right)$$

or, averaging over y,

$$H\overline{(\overline{\delta_x A}^{y})} = \overline{\delta_x(H\overline{A})}^{y} +$$
$$\frac{1}{\delta x}\left(-\sum_{k=K_0+1}^{K_{\frac{1}{2}\frac{1}{2}}} A_{\frac{1}{2}\frac{1}{2}}\Delta_k + \sum_{k=K_0+1}^{K_{-\frac{1}{2}\frac{1}{2}}} A_{-\frac{1}{2}\frac{1}{2}}\Delta_k - \sum_{k=K_0+1}^{K_{\frac{1}{2}-\frac{1}{2}}} A_{\frac{1}{2}-\frac{1}{2}}\Delta_k + \sum_{k=K_0+1}^{K_{-\frac{1}{2}-\frac{1}{2}}} A_{-\frac{1}{2}-\frac{1}{2}}\Delta_k\right) \tag{415}$$

where K_0 =min$(K_{\frac{1}{2}\frac{1}{2}}, K_{\frac{1}{2}-\frac{1}{2}}, K_{-\frac{1}{2}\frac{1}{2}}, K_{-\frac{1}{2}-\frac{1}{2}})$.

This is the finite difference equivalent of

$$\int_{-H}^{0} A_x\, dz = \frac{\partial}{\partial x}\int_{-H}^{0} A\, dz + A_{(-H)}H_x\,.$$

In the Cox code, the u component of $\overline{\mathbf{G}}$ is defined as

$$\overline{G}_u = \frac{1}{H}\sum_{k=1}^{K} \overline{\delta_x P'_k}^{y}\Delta_k \qquad \text{where} \qquad P'_k = g\sum_{k'=\frac{1}{2}}^{k-\frac{1}{2}} \overline{\rho_{k'}}^{z}\overline{\Delta_{k'}}^{z}\,.$$

Writing $P'_k = D - P''_k$, where

$$D = g\sum_{k'=\frac{1}{2}}^{K-\frac{1}{2}} \overline{\rho_{k'}}^{z}\overline{\Delta_{k'}}^{z}\,,\; P''_k = g\sum_{k'=k+\frac{1}{2}}^{K-\frac{1}{2}} \overline{\rho_{k'}}^{z}\overline{\Delta_{k'}}^{z}$$

gives

$$H\overline{G}_u = \sum_{k=1}^{K} \overline{\delta_x D}^y \Delta_k \; - \; \sum_{k=1}^{K} \overline{\delta_x P''_k} \Delta_k$$
$$= H\overline{\delta_x D}^y - H\overline{(\overline{\delta_x P''}^y)}.$$

Using eq.(4.15), with P'' substituted for A gives

$$H\overline{G_{\rho u}} = H\overline{\delta_x D}^y - \overline{\delta_x (H\overline{P''})}^y -$$

$$\frac{1}{\delta x}\left[- \sum_{k=K_0+1}^{K_{\frac{1}{2}\frac{1}{2}}} P''_{\frac{1}{2}\frac{1}{2}} \Delta_k + \sum_{k=K_0+1}^{K_{-\frac{1}{2}\frac{1}{2}}} P''_{-\frac{1}{2}\frac{1}{2}} \Delta_k - \sum_{k=K_0+1}^{K_{\frac{1}{2}-\frac{1}{2}}} P''_{\frac{1}{2}-\frac{1}{2}} \Delta_k + \sum_{k=K_0+1}^{K_{-\frac{1}{2}-\frac{1}{2}}} P''_{-\frac{1}{2}-\frac{1}{2}} \Delta_k \right].$$

From the definition of P'', however, we know that $P''_K = 0$ *i.e.,* that $P'' = 0$ at the bottom, so for the terms in square brackets: if $K_{\pm\frac{1}{2}\pm\frac{1}{2}} = K_0$ we get zero since the sum is over zero points; if $K_{\pm\frac{1}{2}\pm\frac{1}{2}} - K_0 = 1$ we get zero since $P''_K = 0$; if $K_{\pm\frac{1}{2}\pm\frac{1}{2}} \geq 2$ we do not get zero. As long, therefore, as the jump in K between a T, S point and the adjacent u, v points is restricted to 0 or 1 we can define $E = H\overline{P''}$ and we have

$$\overline{G}_u = \overline{\delta_x D}^y + \frac{1}{H}\overline{\delta_x E}^y \, .$$

Similarly

$$\overline{G}_v = \overline{\delta_y D}^x + \frac{1}{H}\overline{\delta_y E}^x \, ,$$

and the forcing is given by

$$\overline{\delta_x \overline{G}_v}^y - \overline{\delta_y \overline{G}_u}^x = \overline{\delta_x \left(\frac{\overline{\delta_y E}^x}{H} \right)}^y - \overline{\delta_y \left(\frac{\overline{\delta_x E}^y}{H} \right)}^x \, .$$

This peculiarity represents a potential problem in the representation of flow over topography in the Cox code. The fact that, for barotropic advection, it is E/H^2 which is advected by the stream function and also E/H^2 which forces the stream function via the BPV equation produces a potentially strong feedback when such flows occur. In the

Cox code, first of all, the density is advected in a manner which does not conserve E/H^2 (although, again, the approximation is better when vertical changes of more than one grid point are disallowed), and secondly, in certain regions, E/H^2 cannot be considered to be the quantity which is producing the forcing. This feedback loop is therefore being quite drastically tampered with.

A second problem of finite differencing is that it is impossible to separate $J(\Psi, f/H)$ into individual beta and topographic terms. The term is therefore calculated complete and an indication of the beta term is found by using the same finite difference formulation to find $J(\Psi, f)$, a term which never actually occurs in the model.

The first calculation was performed using the six year mean for which the time dependent term drops out. A potential problem with using this data set is that the (in situ) densities calculated from the mean values of temperature and salinity may differ from the mean values of the density, and also the non-linear terms calculated from the mean flow may differ from the mean value of the non-linear terms. In practice, these effects seem to be small (in either amplitude or scale) and are ignored in calculating the balance of terms.

The quantities $J(\Psi, f/H)$ and $J(\Psi, f/H) - \frac{1}{\rho_0}\nabla \wedge \overline{\mathbf{G}}$ are shown for the mean flow in fig.(4.2). The smoothness of the difference term shows that it must be predominantly due to the beta effect, since $J(\Psi, 1/H)$ and $\frac{1}{\rho_0}\nabla \wedge \overline{\mathbf{G}}$ are both non-zero only where changes in depth occur - hence the stripy look of the plot of $J(\Psi, f/H)$. The reason for the relatively large number of grid points where no depth change occurs is that, for consistency, and ease of comparison, the topography has been kept the same as that used in the coarse resolution ($1^\circ \times 1^\circ$) model except in regions where that was so steep as to cause problems with instability (Killworth, 1987).

The difference plot shows similar features to the plot of $(1/H^3)\nabla\Psi \wedge \mathbf{F}$ shown in fig.(4.1), supporting the idea that both show the effects of ageostrophic flows. High values are found over the same regions of topography, with a similar problem of fluctuation between positive and negative over grid point scales, and the remaining structure is much smoother and is distributed over the same regions of eastward flow.

Most of this structure, though, is due to the beta term (see fig.(4.3)) and would not have been there if the originally intended quantity

$$fJ(\Psi, 1/H) - \frac{1}{\rho_0}\nabla \wedge \overline{\mathbf{G}}$$

had been calculated, suggesting that the ageostrophic and beta terms are balancing each other without having a strong influence on the geostrophic balance. This conjecture is borne out by the other plot in fig.(4.3) showing the effect of non-linear terms. Clearly this is in close balance with the beta term. The other two terms, wind and "the remainder" (representing in this case friction plus the non-linear time dependent effects mentioned earlier), shown in fig.(4.4) are smaller again, so the time mean circulation can be divided into three sections: a geostrophic balance between $fJ(\Psi, 1/H)$ and $\frac{1}{\rho_0}\nabla \wedge \overline{\mathbf{G}}$; a balance between time-averaged non-linear terms and beta effect; and smaller effects due to wind, friction and time dependent processes, with some regions over topographic features where all terms are important over very small scales.

Another noticeable property of the beta and non-linear terms is their regularity. Over many sections, fluctuations between positive and negative occur with a regular zonal wavelength. These waves appear to be strongest near a number of "western boundary" regions, notably around the Falklands Plateau and the eastern side of the Kerguelen Plateau. This

regularity and the balancing of beta with non-linear terms suggest that these may be stationary Rossby waves, propagating with westward phase velocity relative to the current and so standing still, but with eastward group velocity (which is the case always for the rigid lid approximation, but would be the case anyway for waves of such short wavelength ... at 60°and in 4000 m deep water the maximum in the Rossby wave dispersion relation is at approximately 10^4 km).

A stationary Rossby wave was proposed by Moore,(1963) as a feature of a gyre circulation at intermediate Reynolds numbers (higher than those needed for a frictional balance to hold, but lower than is the case for the purely inertial recirculation of Fofonoff (1954). It has since been shown that the presence of a northern boundary stops the formation of the Rossby wave train, producing instead a strong recirculation, but numerical experiments by Cessi et al (1990) have shown that the Rossby wave reappears if the boundary layer separates well away from any northern boundary. In a careful examination of the dynamics of this flow, Cessi et al show that the westernmost "loop", which looks simply like the first oscillation of the Rossby wave, is in fact a separate dynamical regime in which beta, inertial and viscous terms are all important. It is in this region that almost all of the Lagrangian change in potential vorticity takes place, and viscous forces in the subsequent Rossby wave train are much less important. Although the situation in the Southern Ocean is not that of a classical ocean gyre with western boundary layer, there are many similarities between the observed wave trains and those predicted by this theory. The bulk of the wave train consists of a balance between beta and non-linear terms, with friction playing a much less important rôle, but the first (strongest) oscillation of each wave train has a term balance in which friction is very important. The currents are clearly separating from the boundaries (generally topographic features rather than coastlines) far from any southern or northern

boundary, and there is a hint of damping to the east of the source regions. A significant difference from the theory is the importance of the JEBAR and $fJ(\Psi, 1/H)$ terms in the western boundary regions - an effect which could not have been predicted by the model since it assumes completely barotropic dynamics with no topography.

The wavelength of the waves is seen to vary with latitude. One contribution to the apparent variation is the map projection which stretches out features that are further south, but part of the variation is genuine, with typical wavelengths at 44° S being 290 km and at 65° S being 340 km, corresponding to Rossby wave speeds of $3.5\,\mathrm{cm\,s^{-1}}$ and $4.0\,\mathrm{cm\,s^{-1}}$ respectively. Given the approximate nature of this calculation, and the sensitivity of the calculated speed to wavelength ($c \sim 1/\lambda^2$), these two speeds are surprisingly similar, suggesting that the flow against which the waves are propagating is quite uniform. This is consistent with the mean stream function (fig. 0.3) which shows that the waves predominate in regions where contours of stream function, although quite variable on the small scale, are strikingly uniformly spaced on the large scale. The typical variation of about 160 Sv over 15°gives, for a depth of 4000 m a mean current speed of $2.42\,\mathrm{cm\,s^{-1}}$; a good agreement, given that the dispersion relation used was for barotropic Rossby waves. The actual structure of the Rossby waves must be far from barotropic or the effect of topography would predominate over beta and the waves would not be zonally orientated. Rhines (1970) showed that at wavelengths where stratification begins to become important, shown at the beginning of this chapter to be of order 100 km in the Southern Ocean, a form of Rossby wave can develop which is decoupled from the topography by having a node near the bottom of the ocean. The Rossby waves seen here must have that sort of form, although they will be different in detail due to the vertical shear of the flow against which they are propagating.

Another interesting point is that many of the Rossby wave features can also be identified in the plot of $(1/H^3)\nabla\Psi \wedge \mathbf{F}$ (fig.(4.1)). At first sight this is not surprising, since both are supposed to represent ageostrophic and beta effects, but the balance found in a stationary Rossby wave is such that there is no vertical velocity; the non-linear and beta contributions to w cancel out, so the waves make no contribution to the fourth term of eq.(4.13). Any reflection of the waves in fig.(4.1) must therefore be due to horizontal advection, but all the barotropic contribution to horizontal advection has been taken into account (that is what the plot represents), so this must be balanced by baroclinic terms - the non-linear baroclinic advection and, possibly, the spiral term. The fact that the waves show up in this plot as well is further confirmation of their baroclinicity.

Time Dependence.

Although the balance of terms for the six year mean showed relatively little contribution from the non-linear time dependent terms (lumped in with friction as "the remainder"), stream function plots from different times show significant variations over small spatial scales. These would be expected to make a contribution to the BPV balance. As an example, the BPV term balance in FRAM was calculated for a particular snapshot at 12 years, part of the period over which the mean data set was calculated. As shown in fig.(4.5), the term $J(\Psi, f/H)$ looks similar to the time mean, and again the largest values are accounted for by geostrophic balance with $\frac{1}{\rho_0}\nabla \wedge \overline{\mathbf{G}}$. The ageostrophic component again shows the regular trains of Rossby waves seen in the mean flow, but this time (fig.(4.6)) the non-linear terms have a more random and smaller scale structure superimposed, particularly over topographic features and in a region to the South East of the Southern tip of Africa. The time dependent term (also shown in fig.(4.6)) matches this

very small scale structure and also contains features of somewhat larger scale in the Rossby wave train, perhaps suggesting a temporary propagation of the Rossby waves. Fig.(4.7) shows that, again, the wind stress makes only a small contribution in almost all regions, and friction operates on very small scales, and particularly in the Rossby wave band and over topographic features.

The small scale of the time dependent perturbations, and the smaller scale of the frictional effects associated with them is temptingly reminiscent of the enstrophy cascade found in geostrophic turbulence, and the fact that the perturbations occur in the same region as standing Rossby waves raises the intriguing idea that they may be caused by Rossby wave instabilities, but no real conclusions can be reached about these speculations without much further work.

To sum up, the balance of terms in the BPV equation seems to split into a number of separate paired balances. First of all, the geostrophic flow across depth contours, $fJ(\Psi, 1/H)$, is closely balanced by the JEBAR term $\frac{1}{\rho_0}\nabla \wedge \overline{\mathbf{G}}$. Secondly, the beta term is closely balanced by the non-linear terms due to the mean flow. Thirdly, the non-linear terms due to fluctuations from the mean are balanced by the time dependent terms, and also partly by friction (a balance which points to a purely local instability), and on top of this there is a small effect due to wind stress. That is the balance over the bulk of the flow. The main exceptions occur in the "western boundary" regions which appear to be the source of the Rossby waves, in which all the terms except wind make important contributions to the overall balance.

Conclusion.

In this chapter, the possibility has been explored of representing the effects of stratification on the barotropic flow by using the JEBAR term to force the BPV equation, and calculating the effect of advection of density on the quantity E, which in turn determines the JEBAR term. The primary conclusion is that, for flows that are in near-geostrophic balance, the first approximation to changes in E should be found using the horizontal component of barotropic flow with zero vertical velocity rather than including the vertical velocity which would make the barotropic flow non-divergent. It was also noted that a baroclinic effect, the "spiral term", is also potentially important, and that this effect is missed by models with only two layers. This first order balance was calculated from FRAM data, and it was shown that the balance is indeed much better if the vertical velocity is ignored. Remaining imbalances were taken to represent ageostrophic and beta effects.

To see whether this balance was reflected in the BPV equation, the various terms in this equation were calculated from FRAM data. Again, the primary balance which would result from zero w field (f-plane geostrophic flow) was confirmed, and ageostrophic and beta effects were found which matched those seen in the truncated density advection balance. These were identified as stationary Rossby waves emanating from western boundary layers, closely matching those described in Cessi et al (1990), and some form of local instability in which non-linear effects balance time dependence and friction. In showing up these features beneath the basic geostrophic balance, the value of the BPV equation and the truncated density advection equation as diagnostics has been established.

Following pages: Figures (4.2) to (4.7). *Terms in the BPV equation, calculated from FRAM data. Figures (4.2) to (4.4) are values calculated from the six year mean data set, and figures (4.5) to (4.7) are calculated from a 'snapshot' at 12 years. The scale is as for fig.(4.1), with a base value of 10^-14 s^-2.*

Fig.(4.2). *Advection of planetary potential vorticity (top) and the difference between this and the forcing due to horizontal density gradients (bottom), from the mean data set.*

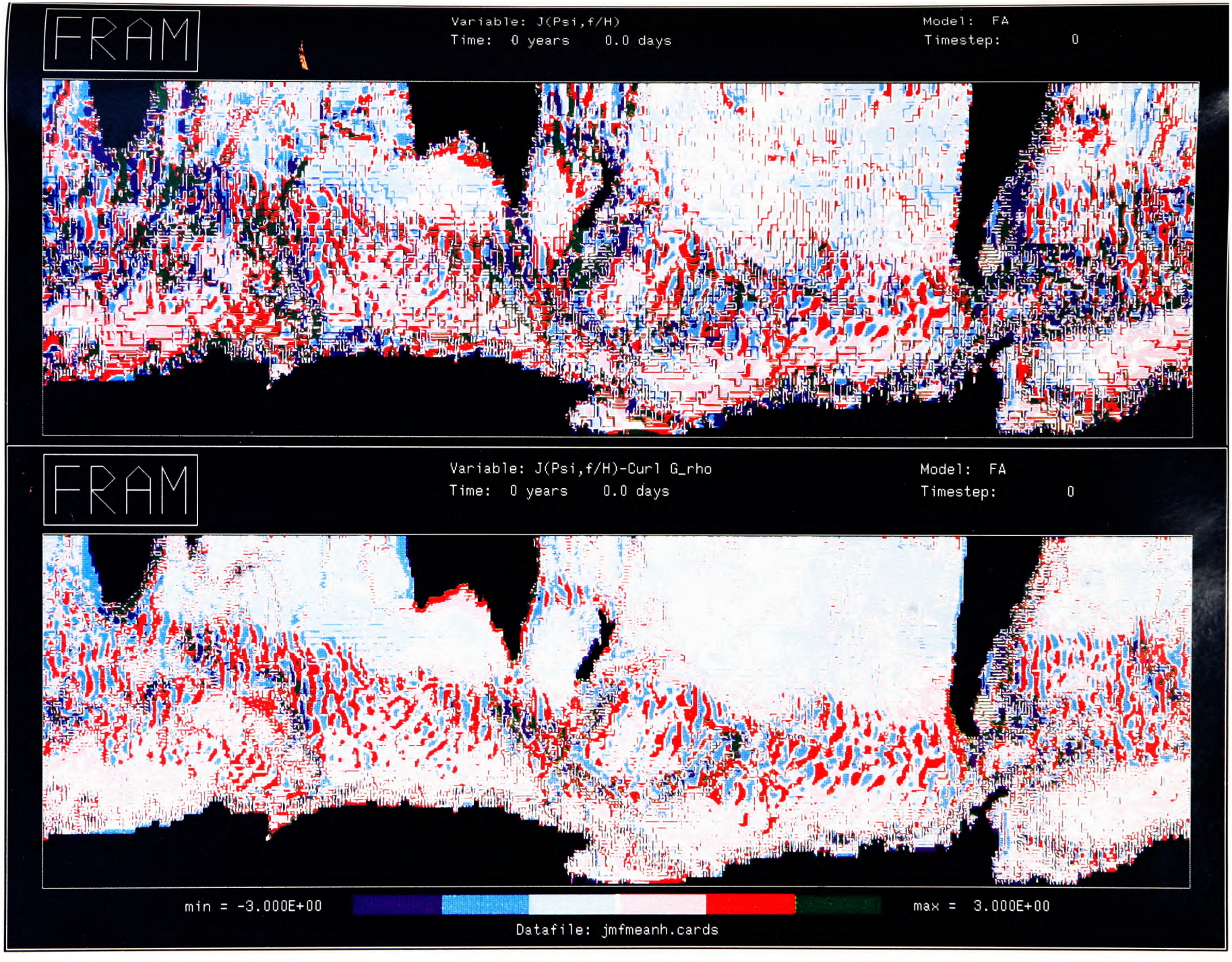
Fig.(4.3). *The beta contribution to fig.(4.2) (top) and the non-linear terms due to the mean flow.*

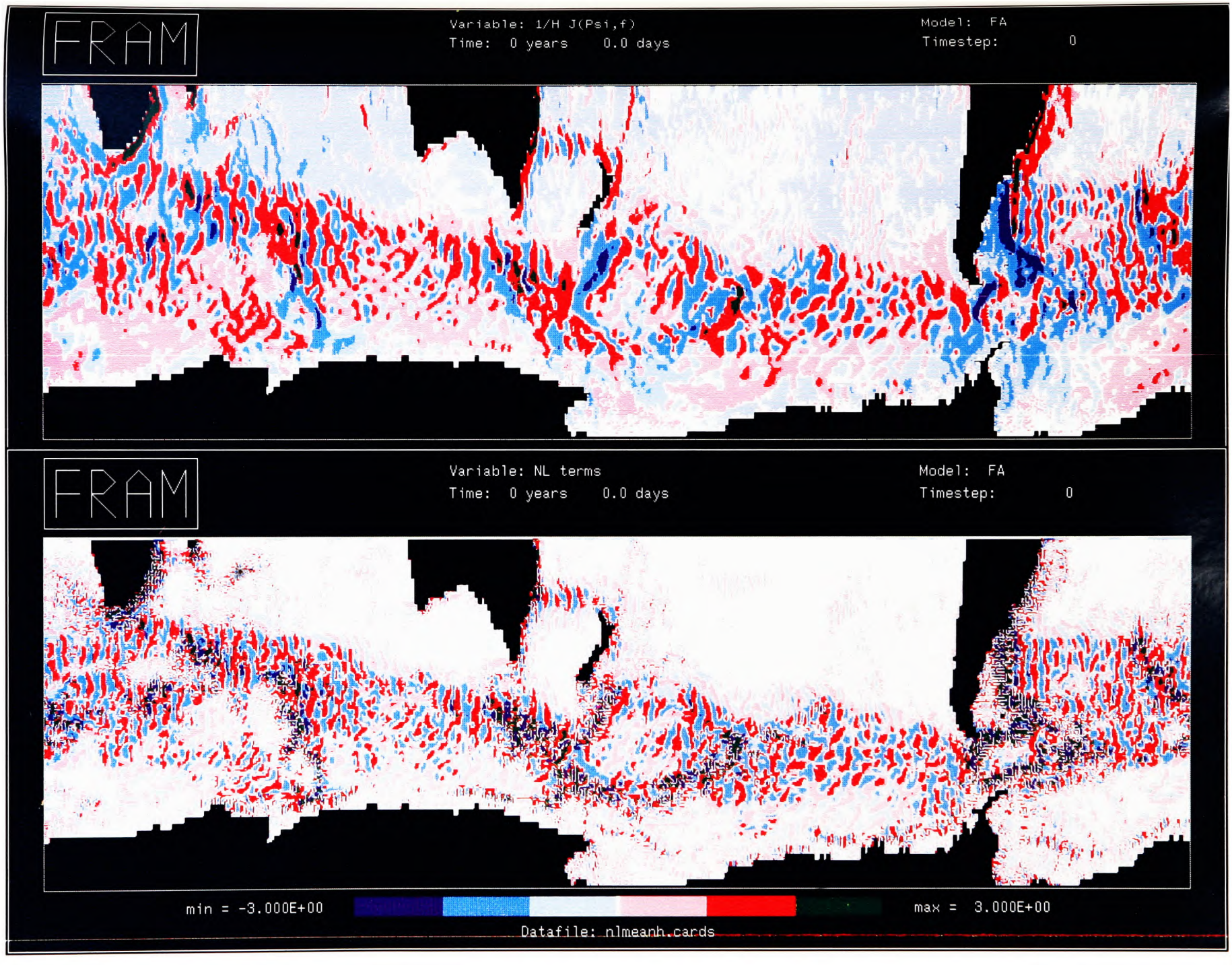
Fig.(4.4). *The forcing due to wind stress (top) and the remaining terms (bottom), in this case representing friction and time-dependent non-linear terms.*

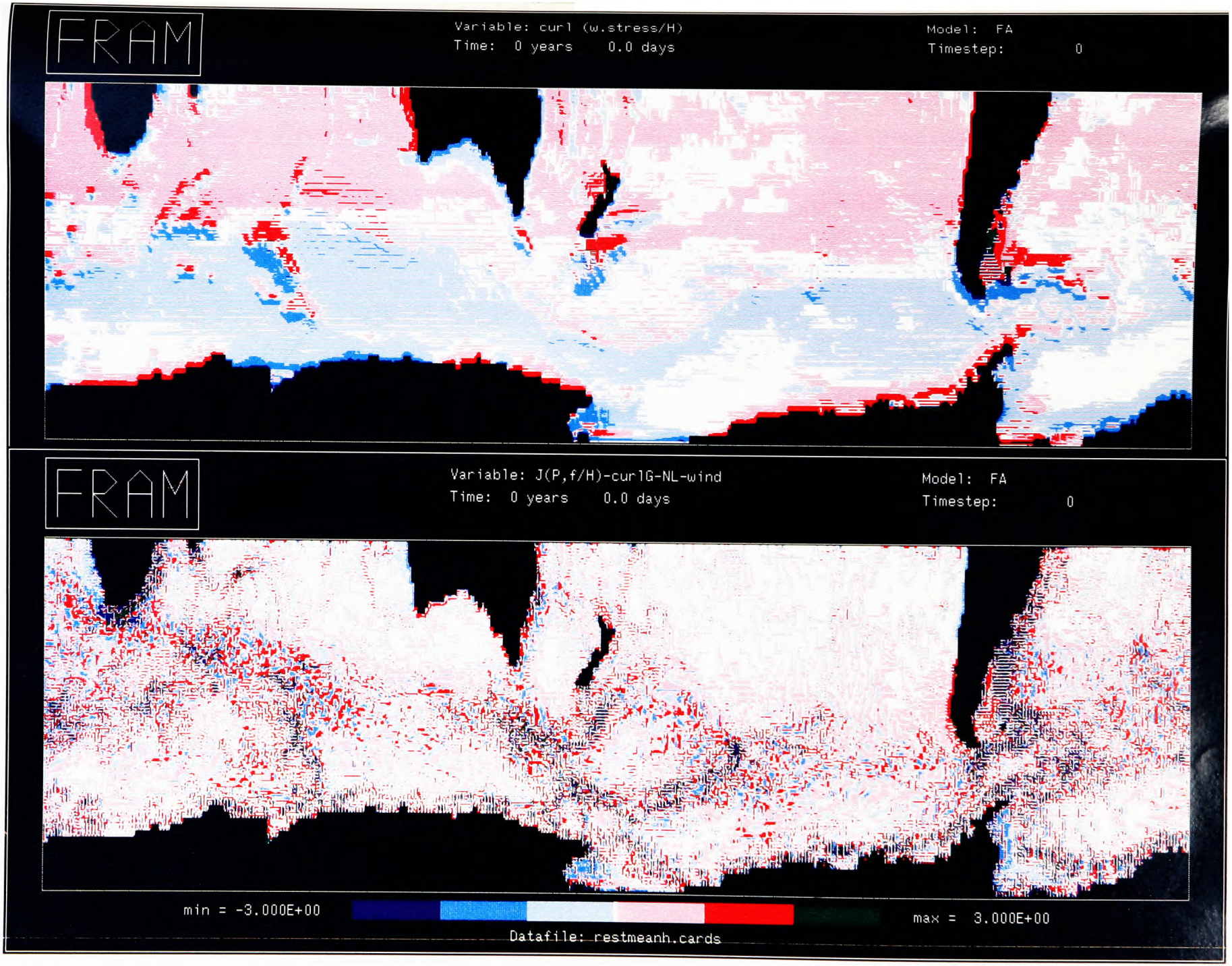
Fig.(4.5). *As fig.(4.2), but calculated from the 12 year 'snapshot'.*

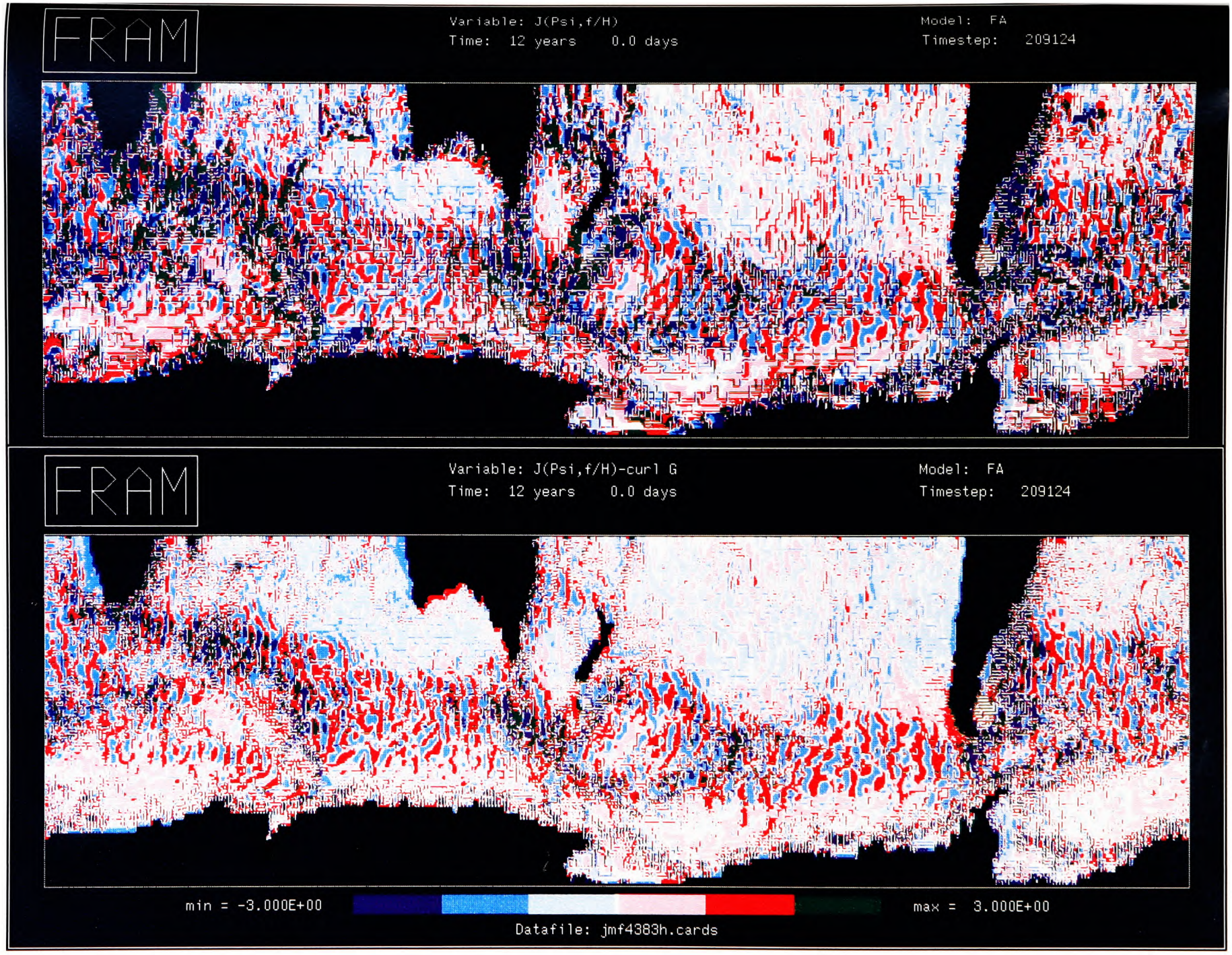
Fig.(4.6). *Non-linear (top) and time-dependent (bottom) terms at 12 years.*

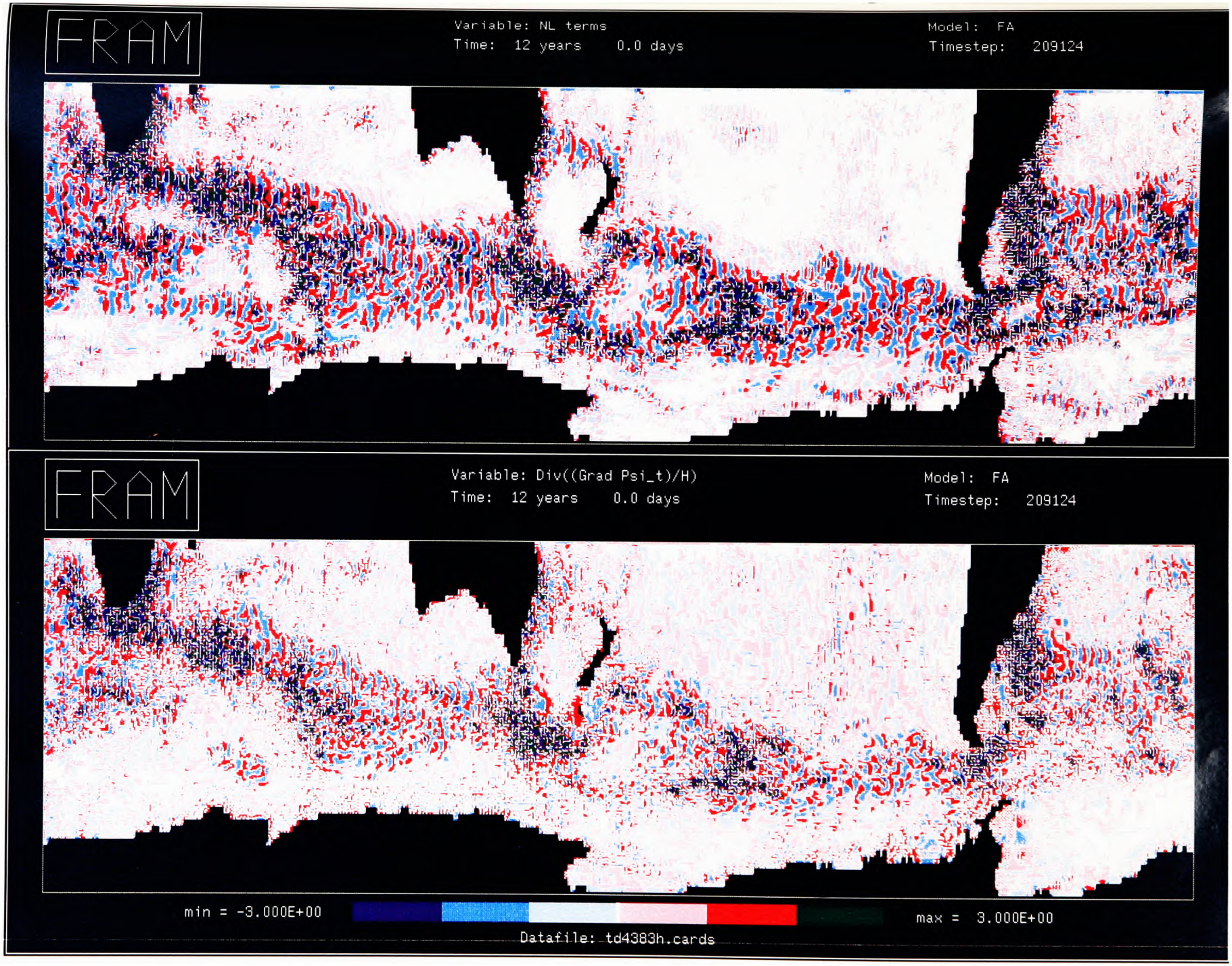
Fig.(4.7). *Wind stress forcing (top) at 12 years and the remaining terms (bottom), in this case representing only friction.*

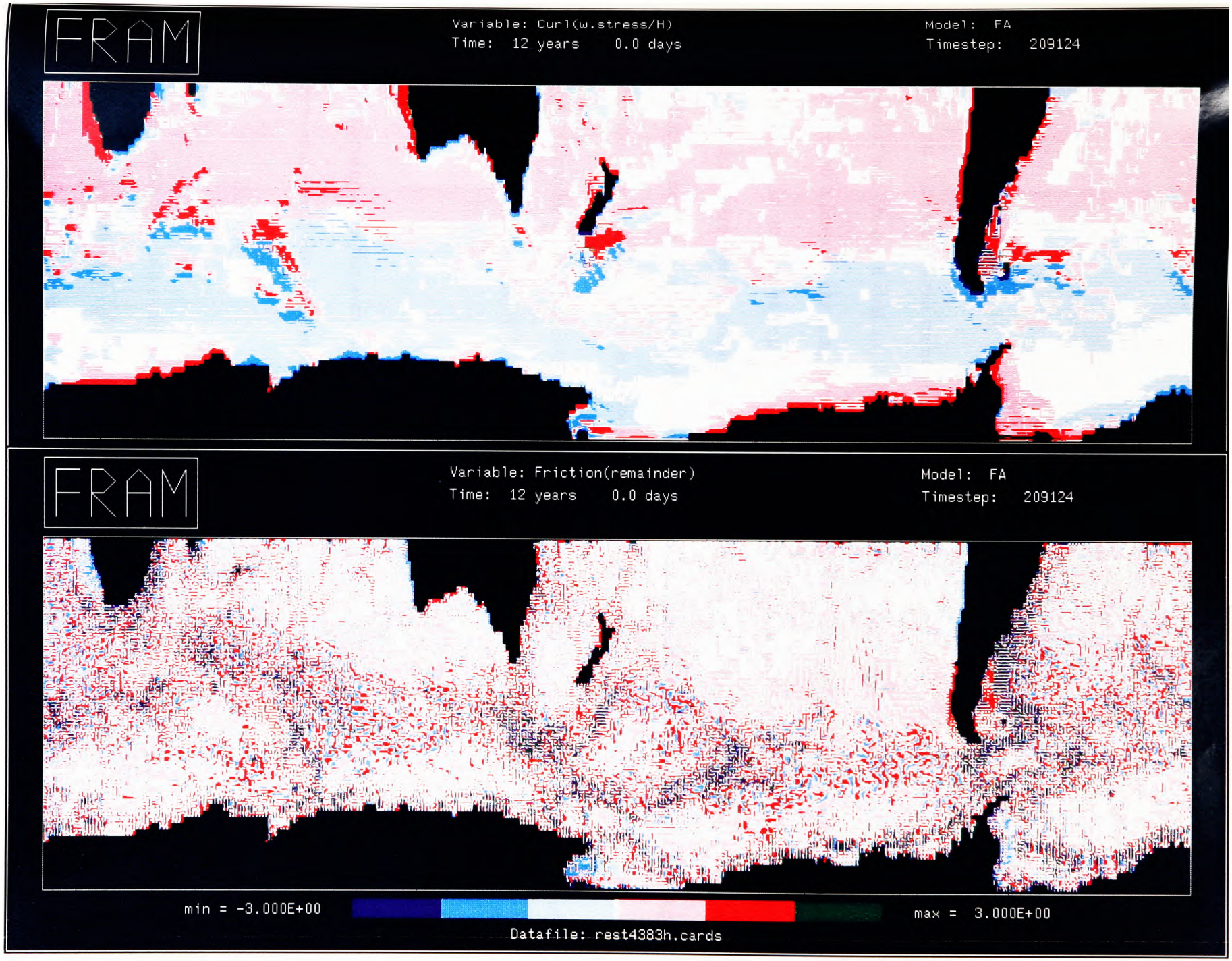

FRAM
Variable: J(Psi,f/H)
Time: 0 years 0.0 days
Model: FA
Timestep: 0
FRAM
Variable: J(Psi,f/H)-Curl G_rho
Time: 0 years 0.0 days
Model: FA
Timestep: 0
min = -3.000E+00
max = 3.000E+00
Datafile: jmfmeanh.cards


FRAM
Variable: 1/H J(Psi,f)
Time: 0 years 0.0 days
Model: FA
Timestep: 0
FRAM
Variable: NL terms
Time: 0 years 0.0 days
Model: FA
Timestep: 0
min = -3.000E+00
max = 3.000E+00
Datafile: nlmeanh.cards


FRAM
Variable: curl (w.stress/H)
Time: 0 years 0.0 days
Model: FA
Timestep: 0
FRAM
Variable: J(P,f/H)-curlG-NL-wind
Time: 0 years 0.0 days
Model: FA
Timestep: 0
min = -3.000E+00
max = 3.000E+00
Datafile: restmeanh.cards


FRAM
Variable: J(Psi,f/H)
Time: 12 years 0.0 days
Model: FA
Timestep: 209124
FRAM
Variable: J(Psi,f/H)-curl G
Time: 12 years 0.0 days
Model: FA
Timestep: 209124
min = -3.000E+00
max = 3.000E+00
Datafile: jmf4383h.cards


FRAM
Variable: NL terms
Time: 12 years 0.0 days
Model: FA
Timestep: 209124
FRAM
Variable: Div((Grad Psi_t)/H)
Time: 12 years 0.0 days
Model: FA
Timestep: 209124
min = -3.000E+00
max = 3.000E+00
Datafile: td4383h.cards


FRAM
Variable: Curl(w.stress/H)
Time: 12 years 0.0 days
Model: FA
Timestep: 209124
FRAM
Variable: Friction(remainder)
Time: 12 years 0.0 days
Model: FA
Timestep: 209124
min = -3.000E+00
max = 3.000E+00
Datafile: rest4383h.cards

Chapter 5.

The Spin-up of FRAM.

In Chapter 2, brief mention was made of the linearity of the spin-up seen in FRAM, and it was observed that such linearity could be explained by the response of topographic modes to a linearly changing density field. It was, however, shown that the full response could not in general be represented by topographic modes, and some extra function was needed if there were regions of closed f/H contours. There are such closed contours in FRAM, and this causes a problem because the flow around such contours should, according to the BPV equation, grow in proportion to t^2.

Consider the BPV equation as it would apply at the start of the FRAM run: linearised and forced only by the JEBAR term.

$$\nabla \cdot \left(\frac{\nabla \Psi_t}{H} \right) + J(\Psi, f/H) = \frac{1}{\rho_0} J(E, 1/H). \tag{5.1}$$

Integrating over the area enclosed by a contour of f/H, the Jacobian term integrates to zero and use of the divergence theorem gives

$$\frac{\partial}{\partial t} \oint \frac{\nabla \Psi}{H} \wedge d\mathbf{s} = \oint \frac{\nabla E}{H} \cdot d\mathbf{s}.$$

If the density is increasing linearly with time, then $E \propto t$ and this equation predicts that the circulation $\oint \frac{\nabla \Psi}{H} \wedge d\mathbf{s}$ must increase in proportion to t^2. The fact that the total transport increases linearly with t (fig.(2.2)) and that the pattern of the stream function remains almost perfectly constant between about days 10 and 200 (fig.(2.3)), shows that this is not happening.

A possible reason for this is that the coefficient of the t^2 term may be small enough that no significant flow is set up before non-linear and frictional effects start to become important. This view is particularly plausible given that, on an f-plane, the forcing is identically zero. In this case, contours of f/H are also contours of $1/H$ and integrating $J(E, 1/H)$ over an area enclosed by such a contour gives zero, so this forcing of a t^2 component can be seen to be an effect of variation of f, and will only be significant for contours of f/H which traverse a large enough range of latitudes.

On the other side of the argument, there are regions, most noticeably the Kerguelen Plateau and the New Zealand – Campbell Plateau region, where closed contours of f/H do traverse ranges of order 20°of latitude, and the term balance of the steady state shows that beta and non-linear terms are likely to eventually balance out in these regions. The JEBAR term will probably, therefore, give a non-zero forcing. Assuming this is reflected reasonably accurately in the original Levitus data, the spin-up should see this driving a circulation until the non-linear terms become strong enough to balance it.

Assuming the JEBAR term is strong enough to drive a significant flow, there is one other possible explanation of why the quadratic behaviour is not observed, an explanation that requires the insights gained in the analysis in chapter 4 to be fully appreciated.

The near-geostrophic flow scenario developed in chapter 4 shows that the stratification is in near equilibrium with a baroclinic flow driven by the thermal wind relation, and a barotropic flow which is such as to make cross-isobath flow at the bottom much smaller than a naïve scaling would predict. This determines the cross-isobath component of the barotropic flow, and the along-isobath component is then determined by mass continuity in the same way that the zonal flow is determined in a Sverdrup flow calculation, and with the

same proviso that an equivalent to the western boundary layer must be invoked to close the circulation. With the "spiral term" small in the Southern Ocean, this equilibrium specifies a flow which produces little advection of density, so, to the extent that the Levitus data represents the true situation in the Southern Ocean, the spin-up of FRAM produces a flow which does not advect density very strongly, so the relaxation to Levitus can be taken as the primary driving term for a long time (at least 200 days).

There is an exception to this equilibrium, though, in regions of closed contours of f/H. Here, the small imbalance due to the lack of non-linear terms produces a flow around these contours which is barotropic and bears no directional relation to the contours of constant density. The flow which would grow in proportion to t^2, were relaxation to Levitus the only driving force, is exactly the flow which produces significant advection of density and therefore changes the driving force. It is possible that this produces a feedback between flow and forcing which limits the flow to only linear growth.

The fact that this additional flow is barotropic allows the change in E due to advection of density to be modelled by the simple eq.(4.1) formulated by Olbers and Wübber to represent flow due to the complete barotropic field. This barotropic flow has the complete barotropic w field associated with it since it is a result of the time dependent term in eq.(5.1), rather than $J(\Psi, f/H)$, balancing the forcing. A simple model then results for the coupling of the pair of equations:-

$$\nabla \cdot \left(\frac{\nabla \Psi'_t}{H} \right) = \frac{1}{\rho_0} J(E, 1/H)$$
$$\frac{E_t}{H^2} = -\frac{1}{H} J(\Psi, E/H^2) + Q$$

where Q is a source term for E/H^2. To test these ideas, the flow around Kerguelen was calculated from the FRAM spin-up. The Kerguelen Plateau was chosen because it

is the most extensive and clear cut region of closed contours of f/H in the Southern Ocean, it covers a large range of latitudes, and it contains (in the approximation of FRAM topography) no islands, which can be a complicating factor. The circulation was calculated for the contour of $f/H = -5 \times 10^{-8}\,\mathrm{m^{-1}\,s^{-1}}$ by calculating the finite difference equivalent of $\nabla \cdot \left(\frac{\nabla \Psi}{H}\right)$ and integrating over all points on the Kerguelen Plateau with values of f/H less than this. Figure (5.1) shows this quantity plotted at intervals of 10 days for the first 400 days of FRAM. For the first 250 days, the plot is clearly much closer to linear than quadratic, although there seems to be some slight flattening with time and there is a hint of possible quadratic behaviour within the first 10 days. (A point at 1 day is also included but has no visible effect.) The length of this contour is about 3,000 km, so the value at 200 days of $-12,000\,\mathrm{m^2\,s^{-1}}$ for the circulation corresponds to a mean velocity of $1.2\,\mathrm{cm\,s^{-1}}$. A fluctuation of this order from the actual flow would have been clearly visible as a change of pattern in the stream function before 200 days. The argument that the forcing is simply weak enough to show no visible effect is therefore proven false by the observation that the circulation after 1 day is approximately $-5.2\,\mathrm{m^2\,s^{-1}}$. At one day, virtually no advection can have taken place, so the circulation must still be at the stage where it is increasing as t^2. Extrapolating this quadratic increase to 200 days would produce a circulation of $208,000\,\mathrm{m^2\,s^{-1}}$, more than seventeen times the observed circulation. Clearly some process is at work which is acting to stop the circulation from increasing as t^2.

After 250 days the circulation starts to vary rather less regularly. This is the time when visible changes in the stream function pattern begin to appear, and so must be when other advection or non-linear effects begin to be significant. This regime is therefore beyond the range of the simple model.

To make as simple a model as possible, consider a region in which contours of f/H are circular, and f varies linearly (beta plane) so that the variation of $1/H$ along a contour of f/H is sinusoidal. Also, assume that the circulation is due to a velocity which has no variation along the contour (the remaining components of the velocity fields are supposed to be in the near- geostrophic balance given by $J(\Psi, f/H) = \frac{1}{\rho_0} J(E, 1/H)$). If this velocity is written as u, the circulation is given by

$$C = \int \nabla \cdot \left(\frac{\nabla \Psi'}{H} \right) dA = \oint \frac{\nabla \Psi'}{H} \wedge d\mathbf{s} = 2\pi r u.$$

The variation of circulation is then given by

$$C_t = \frac{1}{\rho_0} \int J(E, 1/H)\, dA = \frac{1}{\rho_0} \oint E \nabla \frac{1}{H} \cdot d\mathbf{s},$$

but $\nabla \frac{1}{H} = \frac{1}{f} \nabla \frac{f}{H} - \beta \hat{\jmath}/fH$, so $\nabla \frac{1}{H} \cdot d\mathbf{s} = -\beta dy/fH$, giving

$$C_t = -\frac{\beta H}{f \rho_0} \oint \frac{E}{H^2}\, dy. \tag{5.1}$$

For the variation of E/H^2, we have

$$\frac{E_t}{H^2} = -\frac{1}{H} J(\Psi', E/H^2) + \frac{1}{\Theta} \left(L - \frac{E}{H^2} \right),$$

where L is the Levitus equilibrium value of E/H^2 to which the density field is being relaxed with a time scale Θ. Since the flow is along circular contours,

$$\frac{-1}{H} J(\Psi', E/H^2) = -u \frac{\partial}{\partial s} (E/H^2),$$

where s is a coordinate along the contour.

If we represent E/H^2 and L as sums of Fourier modes,

$$\frac{E}{H^2} = \sum_n A_n \sin(n\phi/2\pi) + B_n \cos(n\phi/2\pi)$$
$$L = \sum_n a_n \sin(n\phi/2\pi) + b_n \cos(n\phi/2\pi),$$

where ϕ is the angle along the contour, measured anticlockwise with zero in the east, then notice that eq.(5.1)picks out only one fourier coefficient

$$\oint \frac{E}{H^2}\, dy = \oint \frac{E}{H^2} r\cos\phi\, d\phi = \frac{rB_1}{2}.$$

The second equation is

$$\frac{E_t}{H^2} + u\frac{\partial}{\partial s}(E/H^2) = \frac{1}{\Theta}\left(L - \frac{E}{H^2}\right),$$

in which changes in bB_1 depend only on u, A_1, B_1, a_1 and b_1, and changes in A_1 depend on the same variables, so we can write

$$\begin{aligned} C_t &= -\frac{\beta H r}{2f\rho_0} B \\ A_t &= \frac{CB}{2\pi r^2} + \frac{a-A}{\Theta} \\ B_t &= -\frac{CA}{2\pi r^2} + \frac{b-B}{\Theta} \end{aligned}$$

where subscripts have been dropped for simplicity.

Non-dimensionalising on time scale τ and on forcing $L_0 = \sqrt{a^2+b^2}$,

$$X = \frac{A\Theta}{L_0\tau}, \qquad Y = \frac{B\Theta}{L_0\tau}, \qquad Z = \frac{C\tau}{2\pi r^2}, \qquad T = \frac{t}{\tau},$$

gives

$$X_T = YZ + \cos\delta - \tau X/\Theta$$

$$Y_T = -XZ + \sin\delta - \tau Y/\Theta$$

$$Z_T = -\frac{\beta H L_0 \tau^3}{4\pi r f \rho_0 \Theta} Y,$$

where $\sin\delta = b/L_0$, $\cos\delta = a/L_0$.

Choosing the time scale τ to satisfy

$$\frac{\tau^3}{\Theta} = -\frac{4\pi r\rho_0 f}{L_0 H\beta} \qquad \text{(positive in the south)}$$

gives $Z_T = Y$, and writing $\underline{X} = X + iY$ we have the coupled pair of equations

$$\begin{aligned} \underline{X}_T &= -i\underline{X}Z + e^{i\delta} - \underline{X}\tau/\Theta \\ Z_T &= Im(\underline{X}), \end{aligned} \tag{5.2}$$

which have two free parameters; δ, which is the angle at which the Levitus forcing acts, and τ/Θ, the ratio of the "natural" time scale of the equations, τ, to the time scale for relaxation to Levitus, Θ. This latter quantity is a measure of the "suddenness" of relaxation to Levitus. To give a physical interpretation to these equations, Z represents the circulation, so the second equation shows how the circulation is forced by the torque due to the component of the density which is correlated with the gradient of depth along a contour of f/H. The horizontal density gradient is represented by the complex number X, whose amplitude gives the amplitude of the gradient and whose phase gives the direction. The first equation, then, shows that the change in this gradient consists of a rotation due to advection by the circulating flow, $-i\underline{X}Z$, an increase in the component that represents the Levitus data, $+e^{i\delta}$, and a decrease otherwise, $-\underline{X}\tau/\Theta$.

This pair of equations, similar in some ways to the Lorenz equations, was integrated using a simple Runge-Kutta method for a variety of values of δ and τ/Θ. A selection of the resulting plots of circulation, Z, versus time T are shown in fig.(5.2). Two interesting types of behaviour can be identified. The long-term behaviour is governed by the parameter δ. If $\delta = 0$, then nothing ever happens, since Z is initially zero, and so $\underline{X}$ remains real for all time and Z remains zero. Physically speaking, the Levitus density is oriented in such a way as to produce no forcing of the circulation, so the density simply relaxes to the Levitus value and nothing else happens. This corresponds to a precisely north-south density gradient, the normal situation being to have more dense water to the south in the southern hemisphere. If this density gradient is twisted slightly anticlockwise, $(0 < \delta < 90^\circ)$ then the circulation

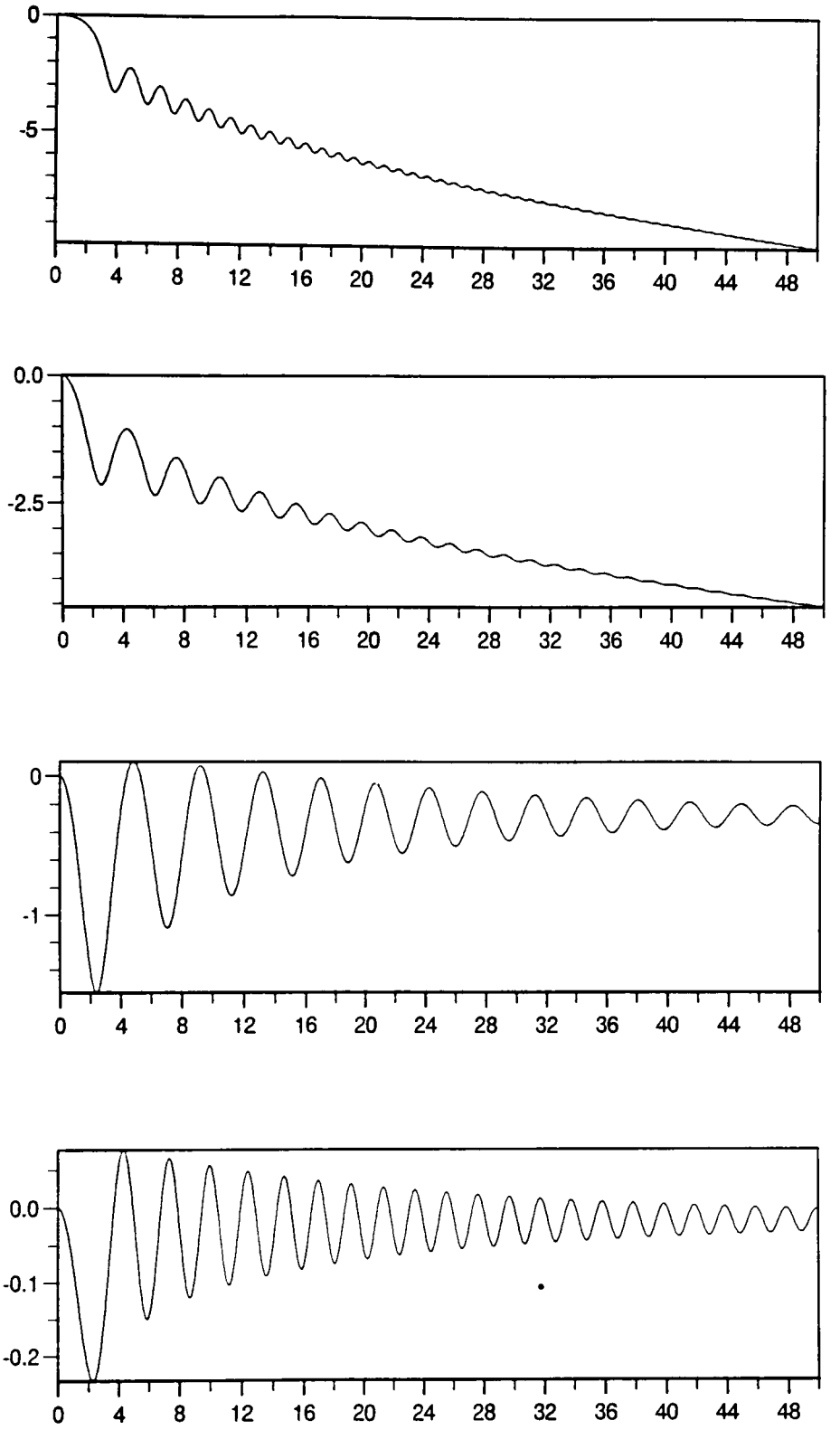


Fig.(5.2) a) *Dimensionless circulation in a simple model of flow around a circular piece of topography, forced by relaxation towards a given density field. The suddenness of the relaxation (see text for definition) is 0.1, and the principal density gradient is at an angle anticlockwise from East of 10, 80, 110 and 170 degrees reading from top to bottom. The x-axis represents dimensionless time.*

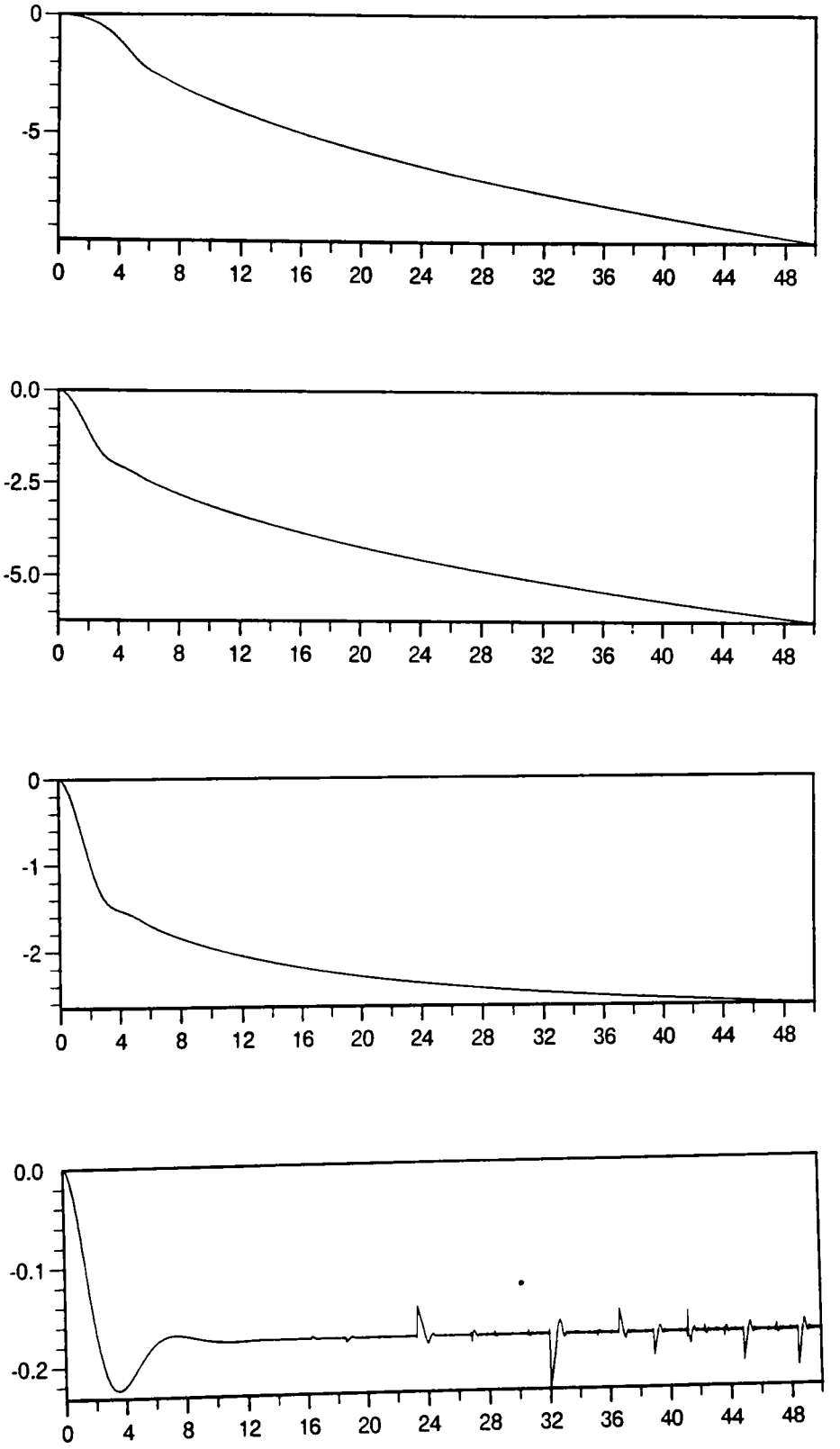


Fig.(5.2) b) *As fig.(5.2) a but with a suddenness of 1. The fluctuations in the bottom figure appear to be due to a numerical instability.*

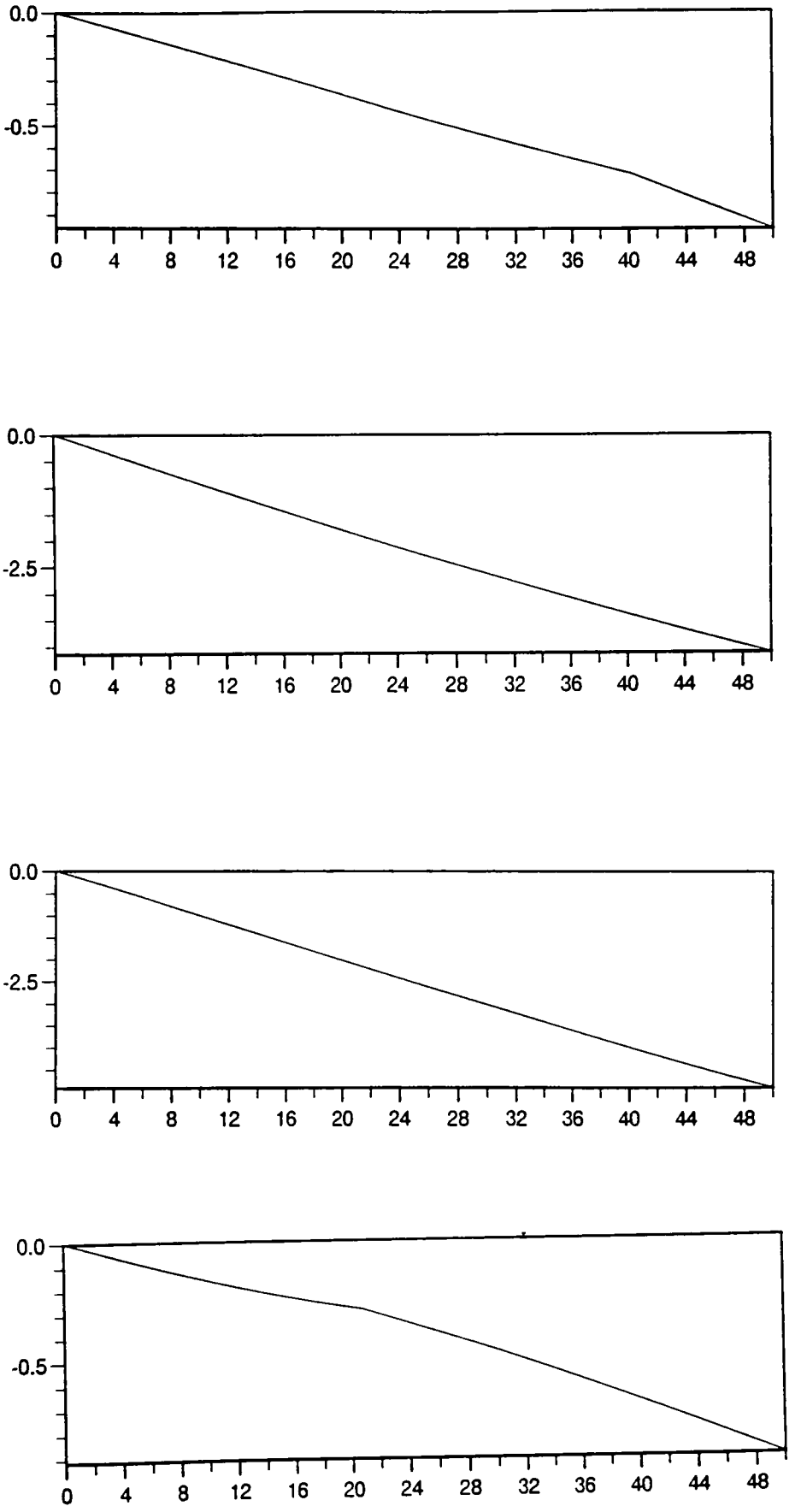


Fig.(5.2) c) *As fig.(5.2) a but with a suddenness of 10.*

increases roughly as $t^{\frac{1}{2}}$, with superimposed oscillations governed by the "suddenness". Around $\delta = 90°$there is a complicated region in which the situation changes, so that for $90° < \delta < 180°$there is no long-term trend in the circulation, and only the oscillations are present. The other two quadrants are simply the same as this but with a negative sign, as can be seen by taking the complex conjugate of eq.(5.2). The reason for this difference is the change in the effect of advection between the two situation. At the boundary between the two quadrants, the phase of the density change is such as to produce the maximum forcing. Either side of this boundary, the circulation initially induced is in the same direction, but in one case is moving the density around towards the point where it produces forcing most efficiently, and in the other case it is moving it away, producing drastically different results.

The oscillations depend largely on the "suddenness", although the frequency depends to some extent on δ. Values of τ/Θ greater than about 1 allow no oscillations and simply smooth the curve out, making it linear in the limit of high τ/Θ whereas low values allow an increasing number of oscillations. For low enough τ/Θ it can be seen that the frequency of oscillation increases with time, but is generally of order $1/\tau$.

The best approximation to the observed behaviour at Kerguelen is given by $\delta \approx 10°$, $\tau/\Theta \approx 2$, but with $\Theta \approx 400$ days (somewhere between 180 and 540 days) for FRAM, the time scales are clearly incompatible by a factor of about 100. This factor could reasonably be explained, though, as being due to the simplicity of the model. τ is a measure of the time necessary for advection effects to become important, and in the model that involves flow around a significant fraction of the complete circle. In reality, there will be higher modes than the first fourier mode contributing to the forcing, since real topography is far from circular, thus reducing the length scale over which advection has to act. In the

case of FRAM, this argument is even stronger, since FRAM topography consists of jumps across one grid point and so the relevant length scale is one grid point or about 27 km, whereas the overall length scale for Kerguelen is more like 3000 km. Another argument based on the simplicity of the model is that, in the presence of more than one fourier mode, oscillations would tend to interfere with each other and would be far less obvious in the time series. If the simple model can be thought of as a bell, the Kerguelen Plateau is more like a piece of scrap metal; likely to ring much less noticeably.

The value of τ/θ used above was derived simply by choosing a rough fit of the model with the observed behaviour. The non-dimensionalisation, however, is based on the scaling

$$\tau^3 = -\frac{4\pi r \rho_0 f \Theta}{L_0 H \beta},$$

which should, in principle, give us an independent way to calculate τ. In practice, it is quite difficult to give a "typical" value of L_0, but since τ depends only on the cube root of this, a rough value ought to be adequate.

From calculations of E/H^2 in FRAM performed in conjunction with the analysis in Chapter 4, a sensible value of L_0 would be about $4\,\mathrm{kg\,m^{-2}\,s^{-2}}$. Taking a time scale of 400 days, a depth of 2000 m, radius of 5×10^5 m and f and β calculated at 55° S, gives

$$\tau \approx 73\,\text{days}, \qquad \tau/\Theta \approx 0.2,$$

giving support to the idea that the actual value of τ for the fundamental mode is somewhat less than the value which gives the apparent best fit.

Conclusion

A curiosity in the spin-up of FRAM, namely the absence of any flow features varying in proportion to t^2, was pointed out. Two possible reasons for this absence were put forward; the absence of any significant forcing of such flow features or a feedback between density advection and the forcing of a barotropic flow. A graph of the circulation around Kerguelen Plateau during the first 400 days of FRAM was shown, and values from this graph were sufficiently large to dispense with the idea that there is no significant forcing in this region.

A simple model of the feedback loop was presented, which produced an interesting variety of behaviours. None of these behaviours, however, allowed flows proportional to t^2 to persist beyond some natural time scale τ, and one integration in particular gave results quite similar to those seen in FRAM. A problem with the simple model was found in that the timescales it suggested were too long to account for the behaviour seen in FRAM, but it was argued that the simplifications made in the model were of a form which would be expected to cause this kind of discrepancy. Partial support for this argument was furnished by showing that the natural time scale predicted from FRAM data was shorter than that predicted by choosing the "best fit" parameters of the model.

A plausible reason for the lack of any quadratic behaviour in the spin-up of FRAM has thus been put forward, demonstrating again, and in quite a different context, the special nature of regions of closed contours of f/H.

References

Allen, J. S. 1975. Coastal trapped waves in a stratified ocean. *Journal of Physical Oceanography* 5: 300–25

Anderson, D. L. T. and Gill, A. E. 1975. Spin-up of a stratified ocean, with applications to upwelling. *Deep-Sea Research* 22: 583–596

Bryan, K. and Cox, M.D. 1972. The circulation of the World Ocean: a numerical study. Part I, a homogeneous model. *Journal of Physical Oceanography* 2: 319–335

Buchwald, V. T. and Adams, J. K. 1968. The propagation of continental shelf waves. *Proceedings of the Royal Society of London, Series A* 305: 235–50

Cessi, P., Condie, R. V. and Young, W. R. 1990. Dissipative dynamics of western boundary currents. *Journal of Marine Sciences* 48: 677–700

Cox, M.D. 1984. A primitive equation, three-dimensional model of the ocean. *GFDL Ocean Group Technical Report No.1* (Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, N.J.08542, U.S.A.).

Fofonoff, N. P. 1954. Steady flow in a frictionless, homogeneous ocean. *Journal of Marine Sciences* 13: 254–262

Gill, A. E. 1968. A linear model of the Antarctic Circumpolar Current.*Journal of Fluid Mechanics* 32: 465–488

Gill, A. E. 1982. Atmosphere-Ocean Dynamics. *Academic Press, 662pp* a) Chapter 6. b) Appendix 3.

Gill, A. E. and Bryan K. 1971. Effects of geometry on the circulation of a three-dimensional Southern Hemisphere model. *Deep-Sea Research* 18: 685–721

Green 1970. Transfer properties of the large scale eddies in the general circulation of the atmosphere. *Quarterly Journal of the Royal Meteorological Society* 96: 157–185

Hamon, B. V. 1962. The spectrums of mean sea level at Sydney, Coff's Harbour, and Lord Howe Island. *Journal of Geophysical Research* 67: 5147–55

Hamon, B. V. 1963. Correction to "The spectrums of mean sea level at Sydney, Coff's Harbour, and Lord Howe Island". *Journal of Geophysical Research* 68: 4635

Hidaka, K. and Tsuchiya, M. 1953. On the Antarctic Circumpolar Current. *Journal of Marine Sciences* 12: 214–222

Hide, R. 1971. On geostrophic motion of a non-homogeneous fluid. *Journal of Fluid Mechanics* 49 part 4: 745–751

Hogg, N. G. 1973. On the stratified Taylor column. *Journal of Fluid Mechanics* 58 part 3: 517–537

Huthnance, J. M. 1975. On trapped waves over a continental shelf. *Journal of Fluid Mechanics* 69: 689–704

Huthnance, J. M. 1978. On coastal trapped waves: analysis and numerical calculation by inverse iteration. *Journal of Physical Oceanography* 8: 74–92

Johnson, E. R. 1988. Topographic waves in open domains. Part 1. Boundary conditions and frequency estimates. *Journal of Fluid Mechanics* 200: 69–76

Johnson, G. C. and Bryden, H. L. 1989. On the size of the Antarctic Circumpolar Current. *Deep-Sea Research* 36 No.1: 39–53

Killworth, P. D. 1987. Topographic instabilities in level model OGCMs. *Ocean Modelling: Unpublished Manuscript* 75: 9–12

Killworth, P. D. 1992. An equivalent barotropic mode in FRAM. *Journal of Physical Oceanography* in press.

Levitus, S. 1982. Climatological atlas of the world ocean. *NOAA Professional Paper No.13.* (Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, N.J.08542, U.S.A.).

Longuet-Higgins, M. S. 1964. Planetary waves on a rotating sphere. *Proceedings of the Royal Society of London, Series A* 279: 446–473

Longuet-Higgins, M. S. 1968. On the trapping of waves along a discontinuity of depth in a rotating ocean. *Journal of Fluid Mechanics* 31: 417–434

McWilliams, J. C., Holland, W. R. and Chow, J. H. S. 1978. A description of numerical Antarctic Circumpolar Currents. *Dynamics of Atmospheres and Oceans* 2: 213–291

McWilliams, J. C. and Chow, J. H. S. 1981. Equilibrium geostrophic turbulence 1: A reference solution in a β-plane channel. *Journal of Physical Oceanography* 11: 921–949

Moore, D. W. 1963. Rossby waves in ocean circulation. *Deep-Sea Research* 10: 735–747

Munk, W. H. and Palmén, E. 1951. Note on the dynamics of the Antarctic Circumpolar Current. *Tellus* 3: 53–55

Mysak, L. A. 1967. On the theory of continental shelf waves. *Journal of Marine Sciences* 25: 205–227

Mysak, L. A. 1980. Topographically trapped waves. *Annual Reviews of Fluid Mechanics* 12: 45–76

Nowlin, W.D. Jr., and Klinck, J. M. 1986. The physics of the Antarctic Circumpolar Current. *Reviews of Geophysics and Space Physics* 24: 469–491

Nowlin, W. D. Jr., T. Whitworth III and Pilsbury, R. D. 1977. Structure and transport of the Antarctic Circumpolar Current at Drake Passage from short-term measurements. *Journal of Physical Oceanography* 7: 788–802

Olbers, D. and Wübber, Chr. 1991. The Role of Wind and Buoyancy Forcing of the Antarctic Circumpolar Current. *Alfred-Wegener-Institut Technical note in the* "Berichte aus dem Fachbereich Physic" *series.*

Proudman, J. 1916. *Proceedings of the Royal Society of London, Series A* 92: 408

Rhines, P. B. 1970. Edge-, bottom-, and Rossby waves in a rotating stratified fluid. *Geophysical Fluid Dynamics* 1: 273–302

Robinson, A. R. 1964. Continental shelf waves and the response of the sea level to weather systems. *Journal of Geophysical Research* 69: 367–368

Schulman, E. E. 1970. The Antarctic Circumpolar Current. *As referred to by Nowlin and Klinck, 1986.*

Semtner, A. J. 1974. An oceanic general circulation model with bottom topography. *UCLA Dept. of Meteorology Technical Report No.9,* 99pp.

Shaw, P. T. and Csanady, G. T. 1983. Self-advection of density perturbations on a sloping continental shelf. *Journal of Physical Oceanography* 13: 769–782

Stone, P. H. 1974. The meridional variation of the eddy heat fluxes by baroclinic waves and their parameterisation. *Journal of the Atmospheric Sciences* 29: 405–418

Straub, D. N. 1990. Some effects of Large Scale Topography in a Baroclinic Ocean. *University of Washington, doctoral thesis.* 162pp, chapter 1.

Taylor, G. I. 1923. Experiments on the motion of solid bodies in rotating fluids. *Proceedings of the Royal Society of London, Series A* 104: 213–218

Treguier, A. M. and McWilliams, J. C. 1990. Topographic influences on wind-driven, stratified flow in a β-plane channel: an idealised model for the Antarctic Circumpolar Current. *Journal of Physical Oceanography* 20: 321–343

Wang, D. -P. 1975. Coastal trapped waves in a baroclinic ocean. *Journal of Physical Oceanography* 5: 326–333

Wang, D. -P. and Mooers, C. N. K. 1976. Coastal trapped waves in a continuously stratified ocean. *Journal of Physical Oceanography* 6: 853–863

Webb, D. J., Killworth, P. D., Coward, A. and Thompson, S. 1991. The FRAM atlas of the Southern Ocean. *Natural Environment Research Council, Swindon, U.K..*

Welander, P. 1968. Wind-driven circulation in one- and two-layer oceans of variable depth. *Tellus* 20: 1–15

Wilkinson, J. H. 1979. Krönecker's canonical form and the QZ algorithm. *Linear Algebra and Appl.* 28: 285–303

Wolff, J. -O., Maier-Reimer, E. and Olbers, D. J. 1991. Wind-driven Flow over Topography in a Zonal β-Plane Channel: A Quasi-geostrophic Model of the Antarctic Circumpolar Current. *Journal of Physical Oceanography* 21: 236–264