Lecture 11: Box Models: Interhemispheric Flow

11.1 Preliminaries

The arguably most glaring deficiency in Stommel's box model of the THC is its confinement to a single hemisphere of a single ocean basin, ignoring that all oceans are connected. Conceptually, the simplest way of including a greater portion of the World Ocean is perhaps to use "back-to-back" Stommel models (Fig 11.1), an approach first pursued by Welander (1986, Willebrand & Anderson, Eds., NATO ASI Series, **C190**, Kluwer, 163-200). In steady state, for every pair of boxes connected by pipes, the same considerations apply as to the original Stommel model, and there are two stable steady states for the flow between every pair of adjacent boxes. Hence, if the Atlantic is viewed as two back-to-back Stommel models, one finds 4 stable equilibria, one of which corresponds to the observed "Northern Sinking" solution with a pole-to-pole circulation. In principle, this idea can be further extended to include the Pacific Ocean as well, yielding $2^4 = 16$ stable equilibria for the 4 box pairs (Marotzke, 1990).

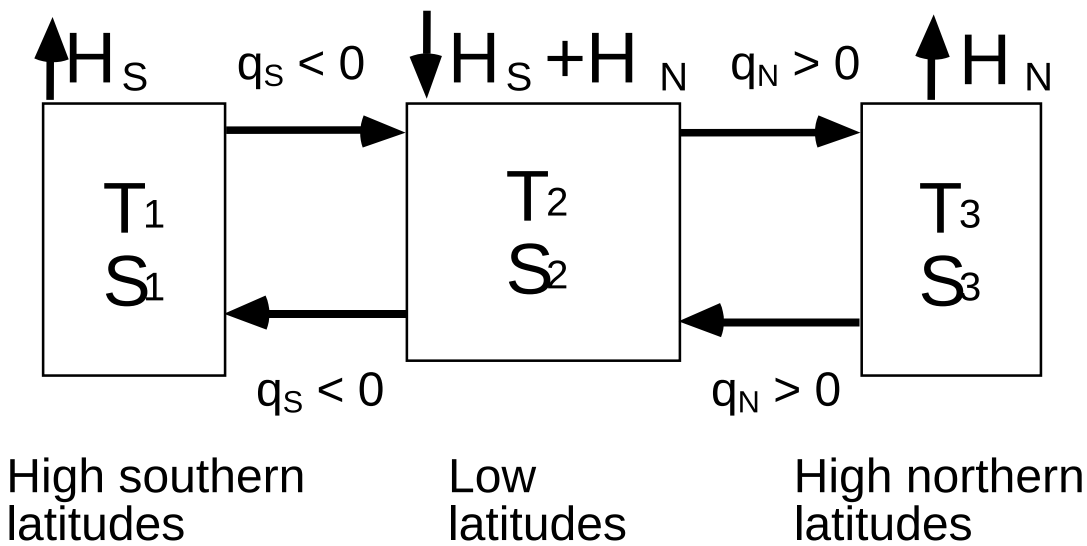

Fig. 11.1. Geometry of Welander's generalisation of Stommel's model. Shown is the "northern sinking" pole-to-pole equilibrium solution.

There is, however, a huge problem in interpreting the single Atlantic THC cell as two back-to-back Stommel models. The box model would require that the (surface) density in the Atlantic be a monotonic function of latitude – greatest in the north, intermediate at low latitudes, smallest in the south. This is in blatant contradiction to observations: Surface densities at both northern and southern high latitudes are much greater than around the equator, as indicated by the observed SST (**Levitus SST**). It appears worth its while to construct a box model that produces interhemispheric flow while maintaining the observed density minimum in the tropics. In discussing a prototype model of this kind, we encounter, along the way, an interesting example of how tortuous the path of progress in science can be.

Stommel's model was first formulated in 1961 but the paper, celebrated as it may be today, went virtually unnoticed for 25 years (see Marotzke, 1994, for a historical account of how Stommel's paper was received – or ignored! – by the community). Meanwhile, in 1982, another box model was independently proposed (Rooth, 1982; see **Fig. 11.3**), which explained how a 2-hemispheric THC symmetric about the equator might go unstable. This result inspired what is arguably the most influential study of the THC (Bryan, 1986), but faded out of the public eye owing to the "rediscovery" of Stommel (1961). Curiously, the profound difference between the dynamics in Stommel's and Rooth's models did not attract attention for a long time. As one consequence, it took more than 10 years before the model in Fig. 11.3 was extensively applied to the steady-state pole-to-pole circulation (Rahmstorf, 1996, *Clim. Dyn.* **12**, 799-811; Scott et al., 1999) – which took up only one half-sentence in Rooth (1982). We will now consider the two applications of Rooth's model – first to the instability of the equatorially symmetric state, then to the pole-to-pole steady state.

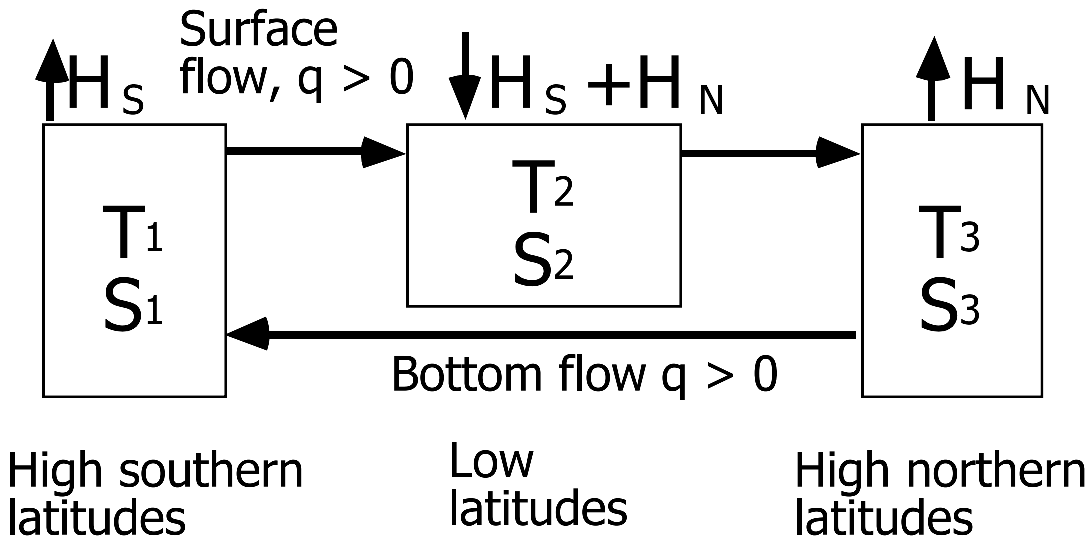

Fig. 11.3. Geometry of Rooth's model.

11.2 Rooth's model: Formulation

The model is in its physical laws equivalent to Stommel's, with the crucial difference that the flow is driven by the *pole-to-pole density difference*. In general, the equivalent surface salinity fluxes in the two hemispheres differ (H_N and H_S in northern and southern hemispheres, respectively). For simplicity, we assume that temperature is fixed and symmetric about the equator. Assuming flow directions as in Fig. 11.3, the equations are

$$\dot{S}_1 = -H_S + q\left(S_3 - S_1\right) \tag{11.1}$$

$$\dot{S}_2 = H_S + H_N - q\left(S_2 - S_1\right) \tag{11.2}$$

$$\dot{S}_3 = -H_N + q\left(S_2 - S_3\right) \tag{11.3}$$

$$q = k'\left(\rho_3 - \rho_1\right) = k'\beta\left(S_3 - S_1\right) \tag{11.4}$$

where k' is a hydraulic constant that is different from the one used in Stommel's model. Typical density differences between high latitudes are much smaller than the pole-equator density contrast, so k' must be correspondingly larger to obtain the same flow strength.

11.3 Rooth's model: Instability of the symmetric state

A symmetric state requires symmetric forcing, so we set $H_N = H_S = \phi$. Inspection of (11.1) through (11.4) shows that, starting from an isohaline state at S_0, the equations are solved by

$$\overline{S}_1 = \overline{S}_3 = -\phi t;\ \ \overline{S}_2 = 2\phi t;\ \ \overline{q} = 0\,. \tag{11.5}$$

Here, we have arbitrarily set that the initial salinities are uniformly zero. Equation (11.5) describes a rather weird reference state, obviously not an equilibrium, but with vanishing flow at all times. We can now look at the linear perturbation expansion about this reference state, by writing

$$q = \overline{q} + q' = q' = k'\beta\left(S_3' - S_1'\right), \tag{11.6}$$

$$S_{1/3} = \overline{S}_{1/3} + S_{1/3}' = -\phi t + S_{1/3}', \tag{11.7}$$

$$S_2 = \overline{S}_2 + S_2' = 2\phi t + S_2'\,. \tag{11.8}$$

We take the time derivative of the dynamical equation (11.4),

$$\dot{q} = k'\beta\left(\dot{S}_3 - \dot{S}_1\right) = k'\beta\left[-H_N + H_S + q\left(S_2 - S_3\right) - q\left(S_3 - S_1\right)\right], \tag{11.9}$$

use that the forcing is symmetric, and insert (11.5) through (11.8), to obtain to first order in primed quantities,

$$\dot{q}' = 3k'\beta\phi t q'\,. \tag{11.10}$$

This perturbation expansion requires that the no-flow solution existed long enough so that a considerable salinity difference can build up between low and high latitudes. Any small perturbation away from the reference state grows according to

$$q'(t) = q'(0)\exp\left\{\frac{3}{2}k'\beta\phi t^2\right\}, \tag{11.11}$$

so the no-flow solution is unconditionally unstable. (Thanks to Jeff Blundell for pointing out the solution.) The physical interpretation is as follows. A small salinity excess in the northern box leads to weak flow as indicated in the figure, which further increases salinity in the northern box – advecting high salinity, S_2, in and low salinity, S_3, out. In contrast, the salinity in the southern box does not change – waters advected in and out have the same salinity. As a result, the small initial excess in northern salinity over southern salinity is amplified – a positive feedback is at work. An asymmetric state develops despite the symmetric forcing – symmetry breaking.

11.4 Rooth's model: Steady states and their stability

Rooth (1982) mentioned in passing that his model also had a steady state, with flow strength proportional to $\sqrt{\phi}$. The steady-state aspects of Rooth's model were, however, not considered until Rahmstorf (1996) noticed that the steady state is as readily found in the more general case of asymmetric forcing. Insertion of (11.4) into (11.1) shows that the steady state flow strength is

$$\overline{q} = \sqrt{k'\beta H_S}\ . \tag{11.12}$$

This result has several remarkable properties. First, the steady-state THC *increases* with increased freshwater flux forcing, in stark contrast with the single-hemispheric box model but consistent with at least some steady-state 3-dimensional model results, as long as one compares (11.12) against the Atlantic component of a global model (Wang et al., 1999, *J. Climate* **12**, 71-82).

Second, and most remarkably, the Atlantic THC only depends on *Southern* Hemisphere atmospheric moisture flux (Rahmstorf, 1996). The other elements of the solution are

$$\overline{S}_3 - \overline{S}_1 = \sqrt{H_S / k'\beta}\ . \tag{11.13}$$

Insertion of the solution (11.12) for the flow into the steady-state version of (11.3), the salinity conservation equation for box 3, gives

$$\sqrt{k'\beta H_S}\left(\overline{S}_2 - \overline{S}_3\right) = H_N \tag{11.14}$$

or

$$\overline{S}_2 - \overline{S}_3 = \frac{H_N}{\sqrt{k'\beta H_S}} \tag{11.15}$$

In particular, we find that

$$\frac{\overline{S}_2 - \overline{S}_3}{\overline{S}_3 - \overline{S}_1} = \frac{H_N}{H_S} \equiv \Gamma \tag{11.16}$$

The salinity difference between equator and the southern box 1 can be inferred from (11.13) and (11.15) as

$$\begin{aligned} \overline{S}_2 - \overline{S}_1 &= \overline{S}_2 - \overline{S}_3 + \overline{S}_3 - \overline{S}_1 \\ &= H_N \Big/ \sqrt{k'\beta H_S} + \sqrt{H_S / k'\beta} \\ &= \left(H_S + H_N\right) \Big/ \sqrt{k'\beta H_S} \end{aligned} \tag{11.17}$$

Under symmetric conditions, $\overline{S}_2 - \overline{S}_1 = 2\left(\overline{S}_2 - \overline{S}_3\right) = 2\left(\overline{S}_3 - \overline{S}_1\right)$ - the salinity drop is equal between, in turn, equator, northern box, and southern box.

That the northern sinking THC strength only depends on southern hemisphere salinity forcing can be understood by the following argument. Rooth's model is essentially unidirectional pipe flow. At any given point (box), the difference between incoming and outgoing salinity is given by the ratio of salt forcing over flow strength, as follows from the kinematic steady-state condition. Box 1 is then singled out because flow strength is proportional to the difference between box 1's incoming and outgoing salinity. Both kinematic and dynamic equations for box 1 hence only include local properties (surface salt forcing, incoming salinity S_3, outgoing salinity S_1).

Everything appears quite simple now: There is a single northern-sinking state; by symmetry, there is also a single southern-sinking state, the strength of which is determined solely by northern hemisphere moisture flux. But there is more to it than meets the eye, as discovered by a graduate student at MIT, Jeff Scott (Scott et al., 1999). In numerical solutions of a more complicated version of the Rooth box model, he did not always find the putative northern sinking solution. A linear perturbation expansion about the steady state (11.12) reveals why. Notice, first, that

$$S_2' = -S_1' - S_3' \tag{11.18}$$

because global-mean salinity is a constant. Linear expansion of the perturbation equations for S_1 and S_3 gives, using (11.12), (11.13), (11.15), and (11.18),

$$\begin{aligned} \dot{S}_1' &= \overline{q}\left(S_3' - S_1'\right) + q'\left(\overline{S}_3 - \overline{S}_1\right) \\ &= 2\sqrt{k'\beta H_S}\left(S_3' - S_1'\right) \end{aligned} \tag{11.19}$$

$$\begin{aligned} \dot{S}_3' &= q'\left(\overline{S}_2 - \overline{S}_3\right) + \overline{q}\left(S_2' - S_3'\right) \\ &= k'\beta\left(S_3' - S_1'\right) H_N \Big/ \sqrt{k'\beta H_S} - \sqrt{k'\beta H_S}\left(S_1' + 2S_3'\right) \\ &= \sqrt{k'\beta H_S}\left[\frac{H_N}{H_S}\left(S_3' - S_1'\right) - \left(S_1' + 2S_3'\right)\right] \end{aligned} \tag{11.20}$$

Equations (11.19) and (11.20) can be rewritten in matrix form as

$$\begin{pmatrix} \dot{S}_1' \\ \dot{S}_3' \end{pmatrix} = \mathbf{A}\begin{pmatrix} S_1' \\ S_3' \end{pmatrix} \tag{11.21}$$

with

$$\mathbf{A} \equiv \sqrt{k'\beta H_S}\begin{pmatrix} -2 & 2 \\ -\Gamma - 1 & \Gamma - 2 \end{pmatrix}, \tag{11.22}$$

where Γ is the ratio of northern to southern hemisphere salinity forcing, defined in (11.16). The stability of the steady state is determined by the sign of the real part of the eigenvalues of **A**, which are obtained from

$$(\lambda + 2)(\lambda + 2 - \psi) + 2(\psi + 1) = \lambda^2 - (\psi - 4)\lambda + 6 = 0\,, \tag{11.23}$$

$$\lambda_{1/2} = \frac{1}{2}(\psi - 4) \pm \sqrt{\frac{1}{4}(\psi - 4)^2 - 6}\,. \tag{11.24}$$

For ψ less than about 9, the eigenvalues are complex, and their real part is negative if $\psi < 4$. For a Γ of more than 4, the northern sinking solution is unstable to infinitesimal perturbations, and the only stable equilibrium is the southern sinking one. If relatively too much freshwater is dumped into the North Atlantic, the "northern sinking" THC cannot be sustained. The **sketch** shows a phase diagram with the equilibrium solutions as a function of H_N, for given H_S.

The feedbacks present are most readily identified by using (11.19) and (11.20) to write

$$\frac{\dot{q}'}{k'\beta} = S_3' - S_1' = q'\left(\overline{S}_2 - \overline{S}_3\right) + \overline{q}\left(S_2' - S_3'\right) - \overline{q}\left(S_3' - S_1'\right) - q'\left(\overline{S}_3 - \overline{S}_1\right) \qquad (11.25)$$

With (11.16) and (11.18), this gives

$$\frac{\dot{q}'}{k'\beta} = -3S_3'\overline{q} + q'\left(\overline{S}_3 - \overline{S}_1\right)\left(\Gamma - 1\right). \qquad (11.26)$$

The first term represents the ubiquitous "mean flow feedback" (mean flow eliminates anomalies), the second term the feedback associated with anomalous flow. If $\Gamma>1$, the coefficient multiplying q' is positive, and the term contributes to exponential growth, so the salinity transport feedback is positive. The opposite applies if $\Gamma<1$.

A positive flow perturbation will cause salinity perturbations proportional to $\overline{S}_2 - \overline{S}_3$ and $\overline{S}_3 - \overline{S}_1$ in boxes 3 and 1, respectively; both perturbations are positive. If the S_3 perturbation is smaller than the S_1 perturbation (resulting from $\Gamma<1$), the flow is weakened, meaning that the salinity advection feedback is negative (stabilising). Both feedbacks are negative and the equilibrium is stable. Conversely, if $\Gamma>1$, then $\overline{S}_2 - \overline{S}_3 > \overline{S}_3 - \overline{S}_1$, the salinity advection feedback is destabilising; if $H_N > 4\,H_S$, the salinity advection feedback overcomes the negative mean flow feedback, and the steady state is unstable.

The strengths of Rooth's box model are probably not recognisable except for experts. It looks rather pathetic, owing mainly to its exclusion of equatorial upwelling (in addition to being a box model); consequently, it has not been used much over the years *except* lending the initial inspiration to Frank Bryan's experiments (making

Rooth’s paper the truly seminal one on the multiple equilibria of the thermohaline circulation). Recently, however, a number of GCM phenomena have been explained with the model; in particular, the connection between interhemispheric flow and pole-to-pole density differences has been firmly established. We will return to this point in Lecture 17 (THC Theory).