4.6

The Antarctic Circumpolar Current System

Stephen R. Rintoul, Chris Hughes and Dirk Olbers

4.6.1 Flow in the zonally unbounded ocean

The absence of land barriers in the latitude band of Drake Passage has a profound influence on the dynamics of currents in the Southern Ocean and, more generally, on the earth's climate. Within this band, the strong eastward flow of the Antarctic Circumpolar Current (ACC) connects each of the ocean basins. Sverdrup dynamics in their usual form cannot be applied to flows within a zonally unbounded ocean, and as a consequence the dynamics of the ACC have long been a topic of debate. Eddy fluxes are believed to play a more central role in both the dynamical and thermodynamical balances of the Southern Ocean than in other areas of the world ocean. The interbasin connection providerchd by the ACC permits a global overturning circulation to exist; the overturning circulation, in turn, dominates the global transport of heat, fresh water and other properties that influence climate (see Gordon, Chapter 4.7; Bryden, Chapter 6.1; and Wijffels, Chapter 6.2). The vigorous interbasin exchange accomplished by the ACC also admits the possibility of oceanic teleconnections, where anomalies formed in one basin may be carried around the globe to influence climate at remote locations (e.g. White and Peterson, 1996). The fact that no net meridional geostrophic flow can exist across the unblocked latitudes isolates the Antarctic continent from the warmer waters at lower latitudes to some extent, contributing to the glacial climate of Antarctica; what heat does get carried poleward to balance the heat lost to the atmosphere must be carried by eddies.

Energetic interactions between the atmosphere, ocean and sea ice result in the formation of water masses that play an important role in the global overturning circulation, and ventilate a substantial fraction of the volume of the global ocean. As a result of these unique aspects, many characteristics of the present-day ocean circulation and climate reflect the influence of the Southern Ocean.

The major currents of the southern hemisphere oceans are shown in Fig. 4.6.1. The ACC is the dominant feature in terms of transport, carrying a mean transport of 134 ± 13 Sv through Drake Passage (Whitworth, 1983; Whitworth and Peterson, 1985). The ACC consists of a number of circumpolar fronts, which correspond to water mass boundaries as well as deep-reaching jets of eastward flow (Orsi *et al.*, 1995). The two main fronts, the Subantarctic and Polar Fronts, are shown in Fig. 4.6.1. Poleward of the ACC, cyclonic gyres are found in the embayments of the Weddell and Ross seas. Westward flow near the continental margin of Antarctica (the Antarctic Slope Front) is found at many locations around the continent (Jacobs, 1991; Whitworth *et al.*, 1998). Equatorward of the ACC, the circulation consists of westward-intensified subtropical gyres in each basin. Strong poleward flow in the western boundary currents is balanced by weaker equatorward flows in the interior of the basins. The main focus of this chapter is on the ACC itself, but exchanges between the ACC and the subtropical and subpolar regimes are also an important part of the story.

While the horizontal flows shown in Fig. 4.6.1 are the dominant circulation features, the weaker
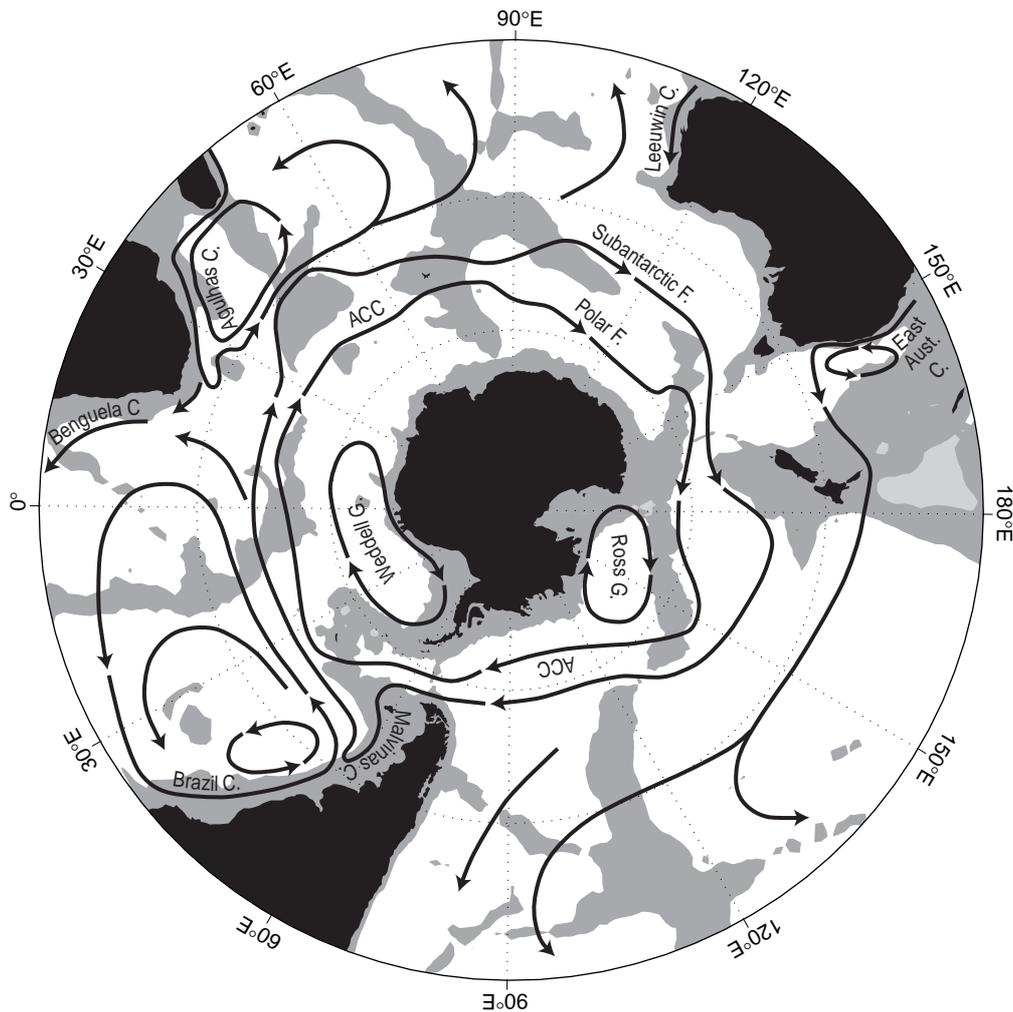
Fig. 4.6.1   Schematic map of major currents in the southern hemisphere oceans south of 20°S. Depths shallower than 3500 m are shaded. The two major cores of the Antarctic Circumpolar Current (ACC) are shown, the Subantarctic Front and Polar Front. Other abbreviations used are F for front, C for Current and G for gyre.

flows in the meridional-vertical plane are also significant. By following property extrema such as an oxygen minimum or salinity maximum, early investigators (e.g. Sverdrup, 1933) inferred the overturning circulation shown schematically in Fig. 4.6.2. Deep water spreads poleward and upward across the ACC and is balanced by equatorward flow in lighter and denser layers. This pattern is driven at least in part by the wind stress acting on the sea surface: south of the westerly wind stress maximum (which generally lies near the axis of the ACC), the Ekman transport is divergent and deep water upwells into the surface layer; north of the westerly wind maximum, the Ekman transport is convergent and surface waters are downwelled into the ocean interior. The water masses exported from the Southern Ocean to lower latitudes as part of

this overturning circulation are responsible for renewing the intermediate and abyssal depths of the southern hemisphere oceans. However, although the general pattern and significance of the meridional circulation of the Southern Ocean has been recognized for many decades, until recently no attempt had been made to quantify the flow paths and water mass conversions implied by Fig. 4.6.2.

The ACC is also unique in the extent to which eddy fluxes contribute to the dynamical and thermodynamical balances. For example, in the absence of mean meridional geostrophic flow across the Drake Passage gap, eddies must carry a significant poleward heat flux to balance the heat lost to the atmosphere at high latitudes and the heat carried equatorward in the Ekman layer (de Szoeke and Levine, 1981). A significant eddy
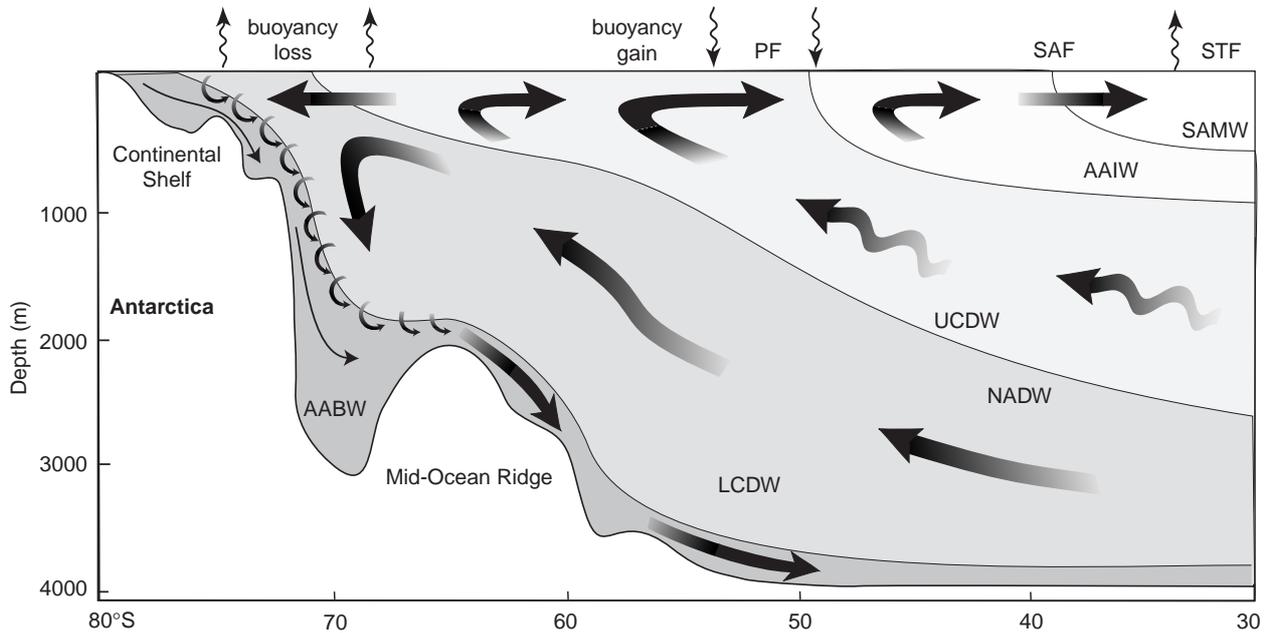
Fig. 4.6.2   A schematic view of the meridional overturning circulation in the Southern Ocean (from Speer *et al.*, 2000). An upper cell is formed primarily by northward Ekman transport beneath the strong westerly winds and southward eddy transport in the UCDW layer. A lower cell is driven primarily by formation of dense AABW near the Antarctic continent. PF, Polar Front; SAF, Subantarctic Front; STF, Subtropical Front; AAIW, Antarctic Intermediate Water; UCDW, Upper Circumpolar Deep Water; NADW, North Atlantic Deep Water; LCDW, Lower Circumpolar Deep Water; AABW, Antarctic Bottom Water.

heat flux is consistent with observations in Drake Passage that the fronts satisfy the linear criteria for baroclinic instability (Bryden, 1979; Wright, 1981). Extrapolation of eddy heat flux measurements in Drake Passage and southeast of New Zealand to the entire circumpolar belt suggested the eddy fluxes were large enough to close the heat balance (Bryden, 1979; Bryden and Heath, 1985), but concern remained that these observations were not representative of the ACC as a whole.

In terms of theory and modelling of the ACC, the state of the art prior to the WOCE (World Ocean Circulation Experiment) period can be briefly summarized as follows. Early wind-driven models of the ACC gave huge transport values unless very large friction was included. Munk and Palmén (1951) proposed that bottom form stress balanced the wind stress. High-resolution quasi-geostrophic models tended to confirm this balance (e.g. McWilliams *et al.*, 1978). At the same time, Sverdrup theory appeared to do a reasonable job of predicting the transport and path of the ACC (Stommel, 1957; Baker, 1982). However, prior to the WOCE era no high-resolution primitive equation models with realistic geometry, stratification,

and full thermodynamics had been run, and many dynamical questions remained open.

Here we review what has been learnt about the Antarctic Circumpolar Current system over the last decade, largely as a result of the WOCE programme. By the 'ACC system' we mean not only the zonal flow of the ACC itself, but also the meridional overturning circulation and water masses of the Southern Ocean. Nowlin and Klinck (1986) provide a comprehensive review of ACC physics as understood at that time, primarily on the basis of measurements in Drake Passage and earlier circumpolar hydrographic surveys. We assume the reader has some familiarity with the dynamics of ocean or atmosphere circulation. Readers interested in more background on topics such as Sverdrup balance, Ekman layers, and potential vorticity are referred to textbooks such as Gill (1982).

We first describe recent observations of the structure and transport of the ACC (Section 4.6.2). Numerical and analytical models have led to substantial advances with regard to the theory and dynamics of the ACC, and its links to the meridional circulation (Section 4.6.3). New observations

and analysis techniques have allowed the circulation, formation and modification of Southern Ocean water masses to be quantified for the first time, as discussed in Section 4.6.4. Observations, theory and modelling over the last decade have led to new insights into the meridional circulation of the Southern Ocean and its link to the global overturning circulation (Section 4.6.5). Finally, we identify a number of important questions that remain open, despite the substantial progress made during the WOCE era (Section 4.6.6). These open questions will provide the challenges for the next generation of field programmes and modelling efforts to follow WOCE.

4.6.2 Observations of the Antarctic Circumpolar Current

A decade ago our knowledge of the structure of the ACC was largely built on detailed measurements in Drake Passage and coarse-resolution hydrographic sections at other longitudes. The increase in the number of high-quality, high-resolution sections and advances in remote sensing have revealed a number of new features of the ACC.

4.6.2.1 Fronts of the ACC

An example of the high-quality sections spanning the ACC collected during WOCE is shown in Fig. 4.6.3. This winter section between Australia and Antarctica near 140°E (WOCE repeat section SR3) illustrates a number of features that are characteristic of the ACC system. As noted first by Deacon (1937), isopleths of all properties generally slope upward to the south across the ACC in a series of steps, or fronts. The locations of the major fronts are indicated above the plots in Fig. 4.6.3. (For the criteria used to define the fronts at SR3, see Rintoul and Bullister (1999).) A number of prominent property extrema in Fig. 4.6.3 define the well-known water masses of the Southern Ocean, as discussed in Section 4.6.4.

Orsi *et al.* (1995) and Belkin and Gordon (1996) have described the circumpolar path and characteristics of the major fronts of the ACC based on careful analysis of a large number of hydrographic sections across the ACC (e.g. **Fig. 4.6.4**, see Plate XX). In addition to the two main fronts of the ACC, the Subantarctic Front (SAF) and Polar Front (PF), Orsi *et al.* (1995) identified two other fronts that were circumpolar in extent,

the 'southern front' and 'southern boundary' of the ACC (**Fig. 4.6.4**, see Plate XX). They also showed that while in Drake Passage the fronts are almost always distinct features separating zones of quieter flow and uniform water properties, at other longitudes the fronts of the ACC merge or split. The merging of fronts is particularly dramatic in the southwest Indian Ocean (**Fig. 4.6.1**), where the Subantarctic and Polar Fronts of the ACC and the Subtropical Front and Agulhas Return Current are all in close proximity and together produce some of the largest temperature and salinity gradients in the world ocean (Olbers *et al.*, 1992; Park *et al.*, 1993; Read and Pollard, 1993; Belkin and Gordon, 1996; Sparrow *et al.*, 1996). The SR3 section shown in Fig. 4.6.3 provides another example of a frontal structure more complex than the classical description based on Drake Passage experience: both the SAF and PF are split into two branches, and the southern SAF and northern PF have merged near 53°S.

While studies like those above have shown that at most longitudes it is possible to identify particular fronts in hydrographic sections, and so verify their circumpolar extent, advances in remote sensing and modelling of the ACC have provided a new view of the rich structure of the current. By representing the SAF and PF as meandering Gaussian-shaped jets, Gille (1994) was able to map the full circumpolar path of the fronts from GEOSAT altimeter data and illustrate the extent to which the fronts were steered by topography. Stream-function maps from eddy-resolving numerical models reveal a more complex, filamented structure to the ACC than that generally inferred from hydrographic climatologies (e.g. Maltrud *et al.*, 1998). Mean gradients of sea-surface temperature (**Fig. 4.6.4**, see Plate XX) also show a complex, filamented structure similar to that seen in the models (Hughes and Ash, 2000).

4.6.2.2 ACC transport

The need to measure the transport of the ACC was a key motivation for the International Southern Ocean Studies (ISOS) experiment in the 1970s. Early attempts to use a small number of current meters to provide a reference for geostrophic calculations were frustrated by the banded nature of the flow: the resulting transport estimates varied dramatically depending on whether a particular instrument was in a front or not. During ISOS, the
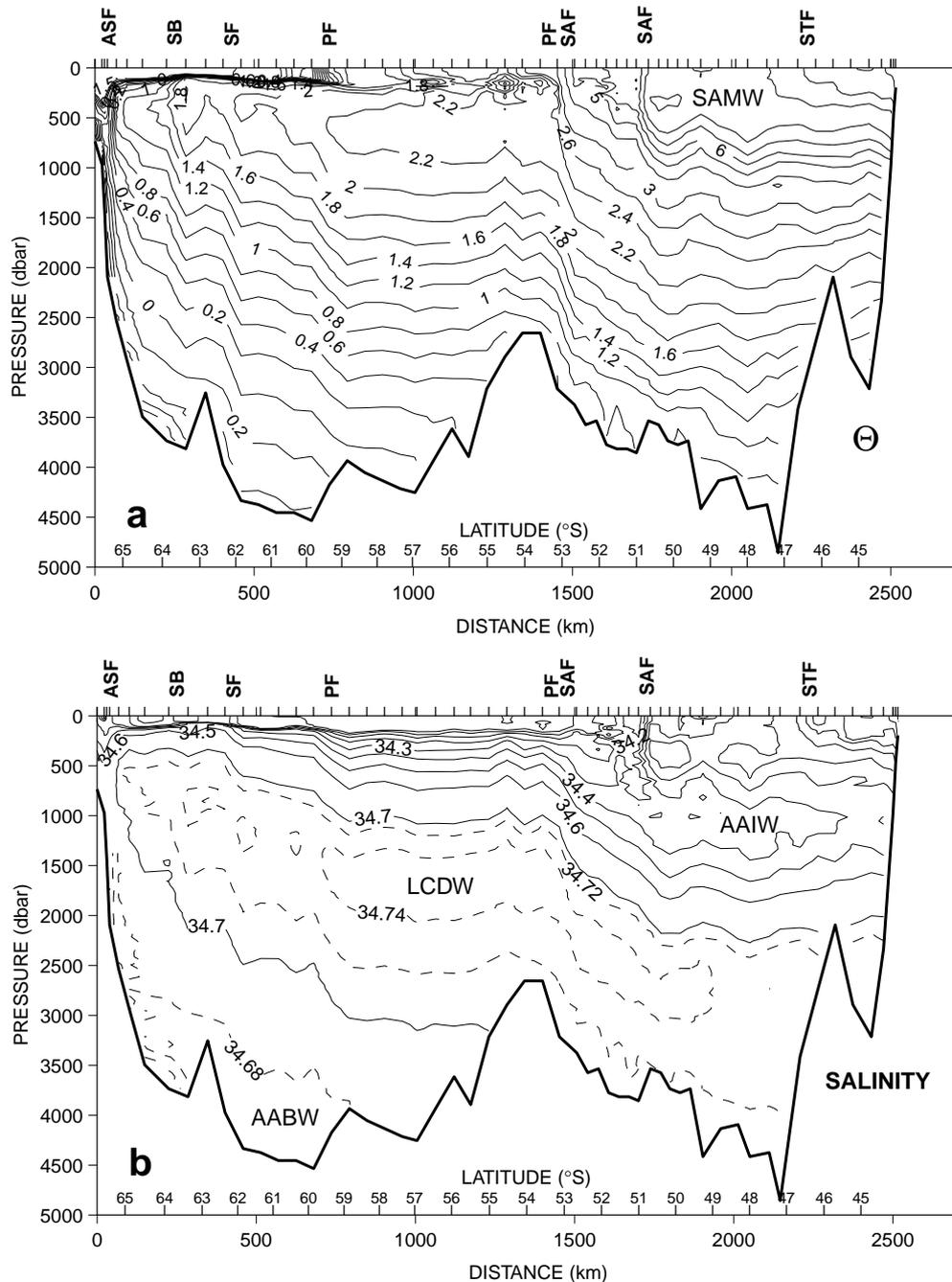
Fig. 4.6.3 Properties versus pressure along the WOCE SR3 repeat section between Australia and Antarctica ($\approx140°E$): (a) potential temperature (°C; contour interval is 1° for $\tau > 3°C$, and 0.2° for $\tau < 2.6°$); (b) salinity (on the practical salinity scale, contour interval is 0.1 for solid contours, and 0.02 for dashed contours); (c) oxygen (μmol kg^{-1}, contour interval is 20); and (d) neutral density γ_n (kg m^{-3}, contour interval is 0.1). The section was occupied in winter (September) 1996. The major fronts are indicated above the plots: STF, Subtropical Front; SAF, Subantarctic Front; PF, Polar Front; SF, southern ACC front; SB, southern boundary of the ACC; ASF, Antarctic Slope Front. The major Southern Ocean water masses are also indicated: SAMW, Subantarctic Mode Water; AAIW, Antarctic Intermediate Water; L(U)CDW, Lower (Upper) Circumpolar Deep Water; AABW, Antarctic Bottom Water (see also Section 4.6.4). Bold contours in (d) denote the isopycnals used to define the layers in Fig. 4.6.14.

mean transport of the ACC was estimated to be 134 Sv, with an uncertainty in the mean of 13 Sv, and a range of 98 to 154 Sv (Whitworth, 1983; Whitworth and Peterson, 1985). The ISOS transport time series was based on pressure gauges at either side of the passage, with an average speed
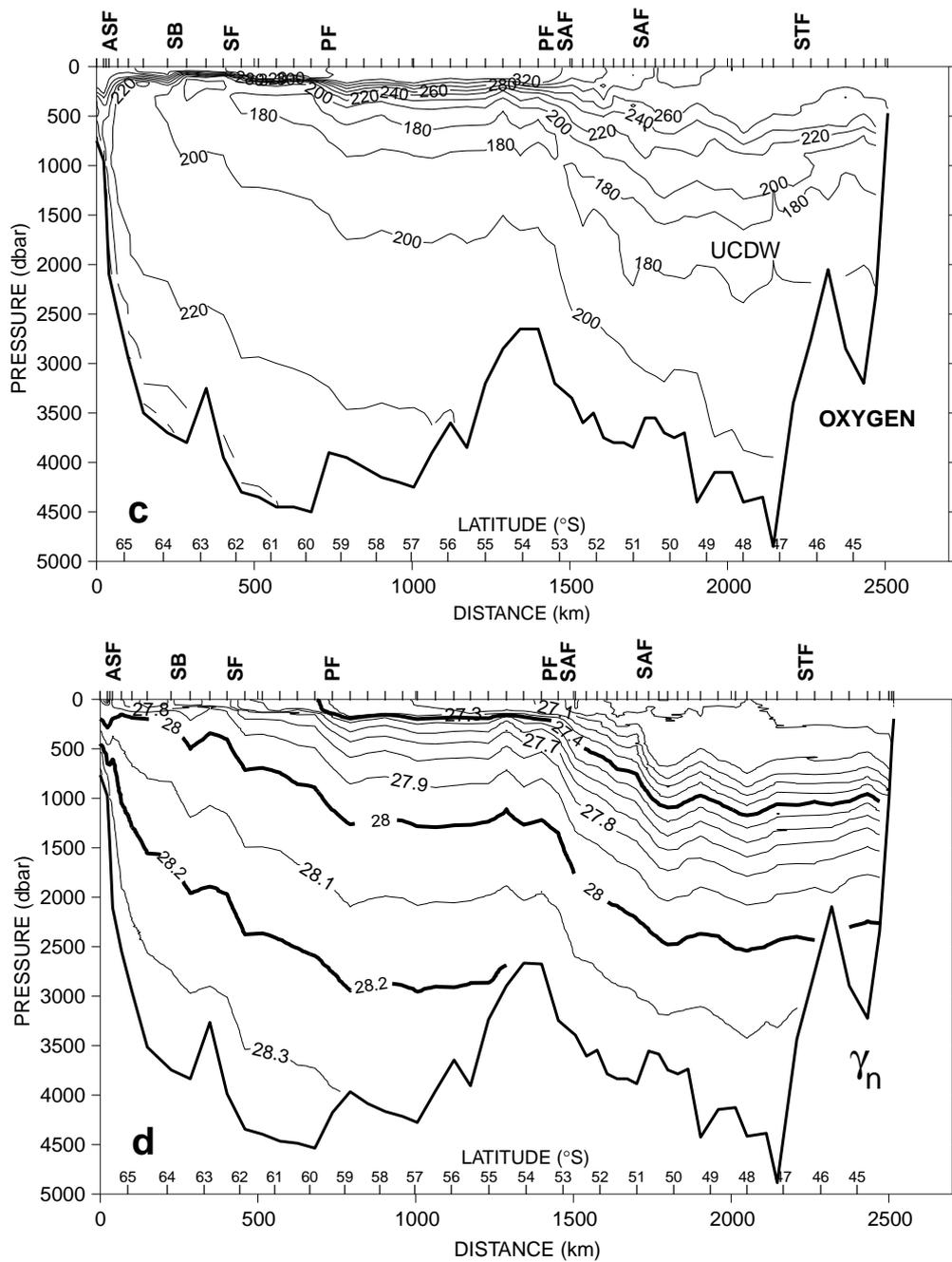
Fig 4.6.3.   Continued

calculated from three hydrographic sections referenced to current meters used to 'level' the gauges and so determine the absolute transport (Whitworth *et al.*, 1982). By resolving the narrow fronts of the ACC, the ISOS measurements significantly reduced the uncertainty in transport estimates of the ACC. Nevertheless, the loss of two moorings in the vicinity of the Subantarctic and Polar Fronts, the discrepancy between direct and hydrographic estimates of the vertical shear, and the fact that the current meter moorings were not coherent, mean that the transports may be subject to somewhat more uncertainty than the error bar quoted above would indicate.

Whitworth *et al.* (1982) and Whitworth (1983) showed from ISOS measurements that the barotropic variability was higher in frequency and larger in magnitude than the baroclinic variability. This observation, together with the similarity between repeat sections across Drake Passage

(Reid and Nowlin, 1971), has sometimes been interpreted to mean that the variability at Drake Passage is entirely barotropic. This is not supported by the ISOS measurements: the range of observed baroclinic transport in the upper 2500 m (70–100 Sv) is comparable to the range of absolute transport (105–140 Sv), and the respective standard deviations are 5.5 Sv and 8.5 Sv. The barotropic variability is larger and of higher frequency, but the baroclinic variability is also significant.

Prior to WOCE, there were no measurements of either baroclinic or barotropic variability of the ACC at other locations. Repeats of WOCE section SR3 south of Australia (140°E) show that the baroclinic variability there is similar in magnitude to that measured by dynamic height moorings deployed for 1 year in Drake Passage during ISOS. The variability is dominated by changes in dynamic height at the northern end, as also found during ISOS. The SR3 section south of Australia extends both further south (into colder water, with lower dynamic height) and north (warmer water, higher dynamic height) than the Drake Passage section, and so the dynamic height difference and baroclinic transport is larger at SR3. For example, the mean transport in Drake Passage above and relative to 2500 m is 87 Sv (Nowlin and Clifford, 1982; Whitworth, 1983); at SR3, the mean of six CTD (Conductivity-Temperature-Depth) sections is 107 Sv (Rintoul and Sokolov, 2000), while the mean based on 36 summer XBT (eXpendable Bathy Thermograph) sections is 109 Sv (Rintoul *et al.*, 2000). The mean baroclinic transport south of Australia (relative to a 'best guess' reference level: at the bottom except near the Antarctic margin, where a shallower level is used consistent with westward flow over the continental slope and rise; see Rintoul and Sokolov, 2000) is 147 ± 10 Sv (mean ± 1 standard deviation), about 13 Sv larger than the ISOS estimate of absolute transport through Drake Passage. The transport south of Australia must be larger than that at Drake Passage to balance the Indonesian Throughflow, which is believed to be of O(10 Sv) (Gordon, Chapter 4.7; Cresswell *et al.*, 1993; Meyers *et al.*, 1995). However, given the large remaining uncertainty in the barotropic flow at both chokepoints, the agreement is likely to be fortuitous.

Rintoul and Sokolov's (2000) estimates of the baroclinic transport variability south of Tasmania show that the ACC itself is surprisingly steady with time (**Fig. 4.6.5**). Variations in net transport largely reflect variations in westward flow across the northern end of the section, rather than changes in transport of the ACC fronts. Because the water flowing to the west south of Tasmania is warm relative to the rest of the section, the changes in this current branch have a relatively large impact on the net interbasin exchange of heat. In other words, while the ACC is undoubtedly the primary means of interbasin exchange, *variations* in the transport between the Indian and Pacific basins are dominated by changes in the flow north of the ACC.

While our picture of the circumpolar baroclinic structure of the ACC has become more complete in recent years, progress in determining the barotropic flow, and hence improving our estimates of the absolute transport, has been slower. The WOCE strategy to determine the transport of the ACC relied on several elements: repeat hydrographic sections across each of the Southern Ocean 'chokepoints', pairs of deep pressure gauges spanning each chokepoint, and direct velocity measurements from shipboard and lowered ADCPs (Acoustic Doppler Current Profilers). Because of the width of the sections, directly monitoring the absolute transport with a coherent array of traditional moored instruments is not feasible.

Shipboard and lowered ADCPs combined with much more accurate navigation and heading measurements are likely to provide valuable constraints on the barotropic component of the ACC. For example, Donohue *et al.* (2000) use ADCP observations in the Pacific to infer that the flow at the bottom beneath the SAF is significant, and in the same direction as the near-surface flow, thus enhancing the transport of the SAF over that estimated from the thermal wind alone. At least at the present time, however, both SADCP (Shipboard ADCP) and LADCP (Lowered ADCP) measurements are subject to uncertainties that are large enough to prevent their direct use as a reference velocity for estimating transports across long sections (Donohue *et al.*, 2000).

However, it is feasible to monitor a portion of the current directly, and such a strategy was adopted at the Australian chokepoint, where a 425-km-long coherent array of current meters, Inverted Echo Sounders (IESs), and seafloor electrometers (HEMs) was deployed across the main
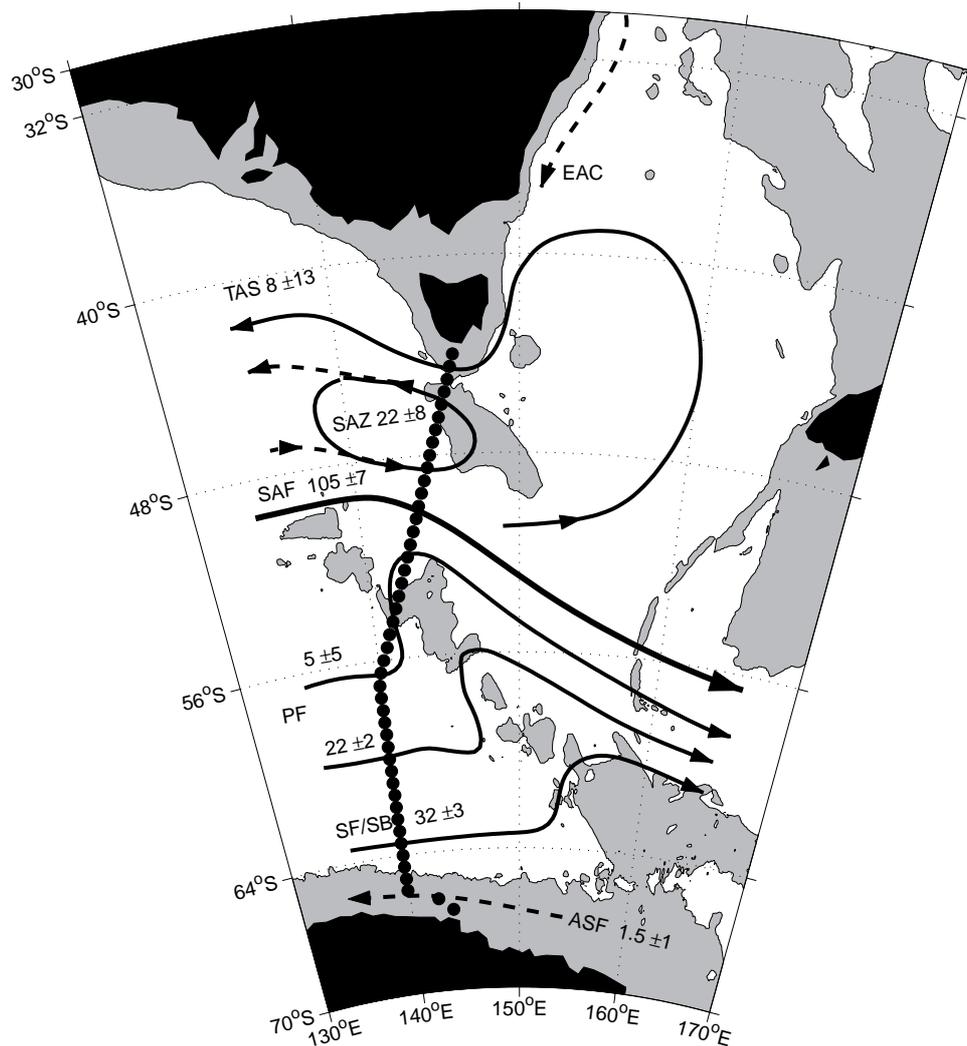
Fig. 4.6.5   Schematic summary of the main circulation features south of Tasmania, based on six repeats of WOCE repeat line SR3. The numbers represent top-to-bottom transports (mean ± 1 standard deviation). TAS, ouflow of Tasman Sea water; SAZ, anticyclonic recirculation in the Subantarctic Zone; SAF, Subantarctic Front; PF, two branches of the Polar Front; SF/SB, southern front and southern boundary of the ACC; ASF, Antarctic Slope Front. See Rintoul and Sokolov (2000) for details.

axis of the ACC to measure both absolute transport and dynamics of the current (the Subantarctic Flux and Dynamics Experiment, SAFDE; Luther *et al*., 1997). The array spanned the SAF, the main core of the ACC at this longitude, along the line of WOCE repeat section SR3 (between roughly 49 and 53°S in **Fig. 4.6.3**). Preliminary analysis of the baroclinic (from the IESs) and absolute (from the HEMs) transport time series shows that while substantial barotropic flows occur in some parts of the array at some times, the 701-day mean baroclinic (relative to the bottom) and absolute transports through the central 200 km portion of the array are similar (54.4 ± 3.1 Sv and 50.4 ± 4.2 Sv, respectively) (Luther *et al*., 1998).

Pairs of deep (1000 m) pressure gauges were deployed across a number of passages to monitor transport variability during WOCE. The idea, based on ISOS experience, is that pressure difference across the passage is proportional to changes in absolute transport. Meredith *et al*. (1996) found that the variability of pressure difference measured across Drake Passage for 4 years during WOCE was somewhat smaller than measured during the ISOS experiment (corresponding to a transport standard deviation of 8 Sv in WOCE compared with 10 Sv during ISOS). The greatest difference between the two records was the absence during WOCE of any events like two observed during ISOS where transport changed by 40% in a few

weeks. Hughes *et al.* (1999) used results from two eddy-permitting numerical models to show that transport correlated better with pressure measured on the south side of Drake Passage than with pressure difference across the passage. Pressure to the south was also highly coherent around the coast of Antarctica. The model transport variations were well correlated with zonally averaged wind stress (with a lag of less than 3 days) near the south of Drake Passage, and occurred in currents that are strongly steered by *f/H* contours, rather than following the path of the ACC. The circumpolar coherence of pressure at the Antarctic continental margin is also observed in the WOCE pressure records, as is the relationship (also noted from the ISOS measurements) between bottom pressure and wind stress, for semiannual and shorter periods. The pressure record at the northern side of the passage is dominated by local effects, resulting in the relatively weak correlation between pressure difference and transport.

While much has been learnt about the circumpolar structure of the ACC in the last decade, we have not yet made much progress in refining our estimate of the mean absolute transport of the ACC. Improved estimates of absolute transport are likely to come from inverse models capable of synthesizing the complete suite of WOCE observations (hydrography, Eulerian and Lagrangian velocity measurements, and remote sensing) with dynamical constraints. Development of such models is an active research area. Several recent models give absolute transport estimates that are similar to geostrophic estimates relative to the bottom (e.g. Macdonald, 1998; Sloyan and Rintoul, 2000a; Yaremchuk *et al.*, 2000). However, given that each of these models start with a first guess of zero barotropic flow, and no such calculations have yet included the full WOCE data set including direct velocity measurements, it is inappropriate to conclude that the barotropic contribution to the mean absolute transport of the ACC is small.

4.6.2.3 Antarctic Circumpolar Wave
The ACC is of interest in part because it allows communication between the ocean basins. One phenomenon that depends on the oceanic teleconnection provided by the ACC is the Antarctic Circumpolar Wave (ACW) identified by White and Peterson (1996). The ACW consists of anomalies in sea-surface temperature, sea-level pressure, and sea-ice extent that propagate eastward around the Southern Ocean. The patterns have zonal wavenumber 2 and circle the globe in about 8–9 years, so the apparent period at any location is about 4 years.

The discovery of the ACW has sparked considerable interest. Part of this interest lies in the potential predictability offered by the ACW. Two recent studies suggest that the ACW has a substantial impact on rainfall in Australia and New Zealand, and may provide some predictive skill (White and Cherry, 1998; White, 2000). The physics of the ACW, in particular the extent to which it represents a coupled mode of the ocean–atmosphere system, has also been a topic of active debate. The initiation of the ACW may be the result of atmospheric teleconnections related to the El Niño-Southern Oscillation (ENSO) (Peterson and White, 1998; Baines and Cai, 2000). Other studies suggest the ACW arises from, or is at least maintained by, atmosphere–ocean coupling within the Southern Ocean (Qiu and Jin, 1997; White *et al.*, 1998; Goodman and Marshall, 1999; Talley, 1999; Baines and Cai, 2000). Several recent model experiments suggest, on the other hand, that the ACW is a passive ocean response to atmospheric forcing, and not a true coupled mode. These studies themselves differ as to the nature of the atmospheric forcing that drives the ACW, with ACW-like oscillations resulting from stochastic forcing (Weisse *et al.*, 1999), standing patterns in the atmosphere (Christoph *et al.*, 1998; Cai *et al.*, 1999), or ECMWF (European Centre for Medium Range Weather Forecasts) re-analysis fluxes (Bonekamp *et al.*, 1999). In summary, a variety of dynamical hypotheses have been proposed to explain the ACW, each of which succeeds in explaining at least some of the characteristics of the ACW. Longer time series of observations (including subsurface ocean measurements) and further modelling studies will likely be required to improve our understanding of the mechanism of the ACW.

4.6.2.4 Eddy fluxes of heat and momentum
The large-scale heat budget of the area south of the ACC implies a significant poleward eddy heat flux across the current (de Szoeke and Levine, 1981), and observed fluxes in Drake Passage (Bryden, 1979; Nowlin *et al.*, 1985) and southeast

of New Zealand (Bryden and Heath, 1985) are of the right sign and sufficient magnitude to close the heat budget if extrapolated around the circumpolar belt. But the reliability of such an extrapolation is obviously open to question, given the length and heterogeneity of the ACC. With regard to eddy momentum fluxes, both the Drake Passage and New Zealand measurements suggest the momentum flux carried by the lateral Reynolds stresses is small relative to the wind stress. The primary significance of eddies in the momentum budget of the Southern Ocean lies in their ability to transfer momentum downward across density surfaces, rather than horizontally, as described below.

Over the last decade, our understanding of the eddy field and its influence on the ACC has improved as a result of several advances: satellite altimeter observations of the ACC as a whole, a limited number of additional current meter measurements, and numerical models capable of resolving (or at least 'permitting') eddies.

Measurements of sea-surface height variability from satellite altimeters has permitted the eddy energy distribution around the entire ACC to be mapped for the first time (Wunsch and Stammer, 1995). High eddy energy is found where the ACC interacts with topography or with poleward extensions of the subtropical western boundary currents (e.g. the Malvinas–Brazil Current Confluence). Morrow *et al.* (1994) showed that the lateral Reynolds stresses were generally small, but on average tended to transfer momentum into the jets of the ACC, accelerating the mean flow, although more recent results suggest that the eddies act to decelerate some of the strongest jets (Hughes and Ash, 2000).

There have been only a few *in-situ* measurements of eddy fluxes in the ACC during WOCE. South of Australia, an array of four tall current meter moorings was maintained for 2 years (Phillips and Rintoul, 2000). The array was deployed at the Subantarctic Front along the WOCE SR3 line (centred on $\approx 50.7°$ S in **Fig. 4.6.3**), in a region that altimetry suggests is one of moderate eddy activity, with the eddy energy increasing rapidly downstream. Although the eddy heat flux varies across the array, the mean values show poleward eddy heat fluxes at all depths between 300 and 2500 m that are significant at the 95% level. The eddy heat fluxes are larger in magnitude than the two previous such measurements, in Drake Passage

and southeast of New Zealand. If extrapolated to the circumpolar belt, the eddy heat flux south of Australia would carry 0.9 PW of heat poleward (40-h to 90-day band-passed data, 'poleward' defined as normal to the direction of daily shear), more than sufficient to balance the heat loss to the atmosphere and the export of heat in the Ekman layer. (Note that this estimate of the eddy heat flux contains both the divergent, dynamically active part of the eddy heat flux and the non-divergent part, see Marshall and Shutts (1981).) The eddy heat flux scaled by the mean vertical temperature gradient gives the vertical momentum flux (e.g. Johnson and Bryden, 1989). South of Australia, fluctuations in the 'eddy band' (40-h to 90-day periods) carry momentum downward at a rate of about $0.2\,\mathrm{N\,m^{-2}}$ (2 dyne cm^{-2}) at all depths (i.e. at about the same rate as momentum is supplied by the wind stress).

4.6.3 Dynamics of the ACC

The absence of continental barriers in the latitude band of Drake Passage makes the dynamics of the ACC distinctly different in character from those of currents at other latitudes. At levels where no topography exists to support zonal pressure gradients, there can be no mean meridional geostrophic flow. The vertically integrated vorticity balance in the Sverdrup approximation, which at least qualitatively succeeds in describing the wind-driven circulation in the interior of closed basins, cannot be used to infer zonal flows in the zonally unbounded Southern Ocean. Even the concept of a wind-driven circulation in the Southern Ocean is inappropriate, as the wind- and buoyancy-forced circulations are inextricably linked. The unique dynamics of the zonal and meridional circulation of the Southern Ocean have attracted the attention of theoreticians for many years. Recent work has led to substantial progress in understanding the heat, momentum and vorticity budgets of the ACC, although some questions remain a source of controversy.

4.6.3.1 Sverdrup balance arguments applied to the ACC

Sverdrup balance holds in the interior of the subtropical gyres because the wind-driven circulation does not penetrate deep enough to interact with bottom topography. In spin-up calculations, the deep circulation is effectively cut off (in the

absence of thermohaline forcing) by the westward propagation of baroclinic Rossby waves generated at the eastern boundary and on topographic features (Anderson and Gill, 1975; Young, 1981). In fact, as long as the mean flow is slow enough that Rossby wave propagation is minimally affected, a succession of Rossby waves of increasing vertical mode number acts to confine the circulation to an ever shallower depth. This only stops when the flow speed becomes comparable to the Rossby wave speed, for the wave mode with the same depth scale as the current.

In the Southern Ocean, observation confirms that the flow at all depths is strongly influenced by bottom topography. From the above argument, this would imply a flow speed comparable to the Rossby wave mode with a vertical scale of 2000 m: the first baroclinic mode. This is consistent with the observation that mesoscale features seen in temperature and sea-surface height propagate eastwards in the ACC, compared with westward propagation elsewhere (Hughes *et al.*, 1998). More importantly, it implies that the Sverdrup balance must be upset by interactions with bottom topography.

Nevertheless, several attempts have been made to apply Sverdrup theory to the ACC by assuming that various topographic features act as 'effective continents', blocking the flow. Stommel (1957), for example, suggested that the Scotia Island arc, east of Drake Passage, effectively extended the Antarctic Peninsula across the Drake Passage gap. Southward interior flow in Sverdrup balance with the wind stress curl was returned in an unusual arrangement of boundary currents: a western boundary current against South America, and an eastern boundary current along the coast of the Antarctic Peninsula and Scotia Island arc, the two currents being joined in some unspecified way by flow through Drake Passage. The transport is then given by the zonally integrated wind stress curl at the southernmost latitude of South America (which is also the northernmost latitude of the Scotia Island arc – there is no overlap; any overlap would complicate this, since it would not then be clear at which latitude the wind stress curl was relevant). Baker (1982) found some support for this argument in a comparison of wind stress curl at 55°S with baroclinic transport through Drake Passage from hydrography.

Webb (1993) suggested the Kerguelen Plateau was a sufficient barrier effectively to block the flow. Webb's model is a highly idealized source–sink flow in a homogeneous, flat-bottomed ocean, but it can also be recognized as an application of Godfrey's (1989) 'island rule' to the geometry of Antarctica, which immediately generalizes it to the case of a stratified ocean obeying Sverdrup dynamics except in specified western boundary regions. The non-Sverdrup flow all occurs in western boundary currents off the eastern coasts of South America and Kerguelan Plateau (and possibly the Antarctic Peninsula), resulting in a flow around Antarctica that is proportional to an integral of the wind stress along a line encircling the continent. This flow becomes infinite in the limit where the northernmost latitude of Kerguelan Plateau is equal to the southernmost latitude of South America, with no overlap. With only Sverdrup balance and western boundary currents involved, Webb's model is the most natural extension of wind-driven gyre dynamics to the Southern Ocean.

While the Sverdrup models of Stommel (1957) and Webb (1993) give reasonable values for the ACC transport (about 120 Sv) when combined with climatological wind stress estimates, the theories are incomplete. There are latitudes at which neither of the proposed 'effective' continental boundaries is shallower than 2000 m, so the assumption that these block the flow must at least be contingent upon some assumption of the weakness of stratification at these latitudes. The dynamics allowing the eastern and western boundary currents to join in Stommel's model are not clear. Perhaps most importantly, these flat-bottomed Sverdrup models ignore interaction between the flow and the bottom topography.

The observation that the ACC penetrates to great depth suggests this assumption is not justified. Topographic interactions link the horizontal and meridional circulations, as can be seen most clearly from the barotropic vorticity equation:

$$\rho_0 \beta \Psi_x = \mathbf{k} \cdot \nabla p_b \times \nabla H + \mathbf{k} \cdot \nabla \times \boldsymbol{\tau} + \mathbf{k} \cdot \nabla \times \mathbf{F} \quad (4.6.1)$$

where Ψ is the barotropic streamfunction, p_b is bottom pressure, H is ocean depth, $\boldsymbol{\tau}$ is wind stress, and $\mathbf{F}$ represents frictional and non-linear terms and $\mathbf{k}$ is the unit vector in the local vertical (upwards). Integrating (4.6.1) over a zonal band enclosed by two latitude lines ϕ_1 and ϕ_2, there is no net northward transport so the left-hand-side

integrates to zero, and Stokes' theorem can be applied to turn the right-hand-side into two line integrals along the bounding latitude lines, giving

$$\oint_{\phi 1}[p_b H_x + \tau^x + F^x]\,dx - \oint_{\phi 2}[p_b Hx + \delta^x + F^x]\,dx = 0$$

(4.6.2)

In fact, the zonal momentum balance tells us more than this, since the integral over depth and longitude of this balance is precisely

$$\oint[p_b H_x + \tau^x + F^x]\,dx = 0 \qquad (4.6.3)$$

at any latitude.

It has been clearly established (Gille, 1997; Stevens and Ivchenko, 1997) that this balance is dominated by the first two terms: the northward Ekman flux is balanced by a geostrophic southward return flow at depth, with a very small contribution from F^x. This means that F^x can also be neglected in equation (4.6.2). There are some complications due to the different meridional scales of terms in (4.6.2), but in practice this is true when ϕ_1 and ϕ_2 are separated by more than about 3–5 degrees of latitude.

The implication of the above is that, for the area integral between these latitudes, $\nabla \times \boldsymbol{\tau}$ is almost entirely balanced by $\nabla p_b \times \nabla H$. Returning to the barotropic vorticity balance, this means that the southward flow driven by the wind stress curl (as in a flat-bottomed Sverdrup balance) returns north in a flow balanced not by viscous terms as in a Munk or Stommel boundary current, but by the bottom pressure torques. (If the two latitudes are separated by less than about 3°, the dominant balance is between bottom pressure torques and nonlinear terms, as found by Wells and de Cuevas (1995).)

The role of topographic torques is graphically illustrated in Fig. 4.6.6 (see Plate XX). This shows the barotropic streamfunction from the Southern Ocean of the global eddy-permitting model OCCAM (Ocean Circulation and Climate Advanced Modelling Project) (Coward, 1996), superimposed on the bottom pressure torque (the first term on the right-hand-side of (4.6.1). Both quantities have been smoothed by 4.25° longitude by 3.25° latitude averaging to reduce the effect of non-linear terms. It is clear from Fig. 4.6.6 that northward flows are associated with positive torques, and southward flows with negative torques, as in equation (4.6.1).

Physically, the curvature of the earth means that a small circle drawn around a point on the earth's surface has its poleward extremity closer to the earth's axis than its centre, and its equatorward extremity further away, but not by the same amount. Water flowing into the circle from the polar side carries with it more azimuthal velocity relative to the circle's centre, as a result of planetary rotation, than water leaving the circle at the equatorial side. Hence (since f changes sign at the equator), a northward flow represents a removal of anticlockwise angular momentum from the circle, which must be balanced by a positive torque, such as bottom pressure torque or wind stress curl. For comparison with Fig. 4.6.6, a typical wind stress curl over this region gives a torque of 10^{-7} N m^{-3}. The fact that the bottom pressure torques are large relative to the wind stress curl suggests that the agreement with Sverdrup balance found by Baker (1982) was fortuitous.

4.6.3.2 Meridional circulation

Much discussion of ACC dynamics has centred on the meridional overturning circulation. In particular, with an eastward wind driving a northward Ekman flux, how does the southward return flow cross the ACC? Integrating the zonal momentum equation over depth and longitude at a particular latitude, the fact that there must be a return flow means that the integrated Coriolis force is zero, and the question can be rephrased as, what zonal force balances the zonal wind stress? This is the question addressed almost 50 years ago by Munk and Palmén (1951), who concluded that Reynolds stresses and viscous terms were probably too small, and that the wind stress is likely balanced by a bottom form stress due to pressure differences across major topographic features.

This viewpoint is now well established, with *in-situ* and satellite altimeter measurements (Bryden and Heath, 1985; Morrow *et al.*, 1994; Phillips and Rintoul, 2000) confirming the smallness of lateral Reynolds stresses, and eddy-resolving primitive equation models (Gille, 1997; Stevens and Ivchenko, 1997) clearly demonstrating a balance between wind stress and bottom form stress. Eddy-resolving Quasi-Geostrophic (QG) models must (by construction) also show such a balance (McWilliams *et al.*, 1978; Treguier and McWilliams, 1990; Wolff *et al.*, 1991), but are useful for illustrating how the balance is established. A summary

of the momentum balance in the different model types is given in Olbers (1998).

In terms of the meridional overturning (the zonally integrated flow at fixed depths), this balance is equivalent to saying that the northward Ekman flux returns in a geostrophic southward flow, supported by zonal pressure differences across topography. The only difference from the situation in other oceans is that, at the latitudes of Drake Passage, the topography does not reach the surface, so the return flow must occur at depths below about 2000 m, as in Fig. 4.6.7.

At first sight, such a deep overturning cell would seem to require large diapycnal fluxes; 10–15 Sv sinking below 1000 m at around 40°S would be comparable to the effect of deep convection in the North Atlantic, but in a region where deep convection does not occur. However, there is an important difference between flows averaged at constant depth and at constant density. In the Fine Resolution Antarctic Model (FRAM), Döös and Webb (1994) have shown that a large fraction of this overturning, labelled the Deacon Cell, occurs with very little associated change in density, as in Fig. 4.6.8. Fluid at a given density flows north, just east of Drake Passage, at one depth, and southwards elsewhere at a slightly deeper level. Seen at a given depth, denser fluid flows north, and lighter fluid flows south, with no integrated flow except at the top (the Ekman layer) and at depths blocked by topography. The flow integrated at constant depth then shows a cell penetrating to great depth, but a fluid particle will pass northward and southward across a given latitude at depths separated by only a few hundred metres, and with no appreciable change of density. The possibility of such circulations decouples the meridional overturning integrated at constant depth from that integrated at constant density. Analogous circulations are also seen in the tropospheric Ferrel cells (McIntosh and McDougall, 1996; Karoly et al., 1997).

This observation means that the meridional overturning averaged on potential (or neutral) density surfaces is worth looking at in more detail. For purposes of discussion, it is useful to consider two extreme possibilities.

- *Case I*. The northward Ekman flux at some latitude is all returned to the south in density layers that do not intersect topography. The most extreme (and least realistic) version of this
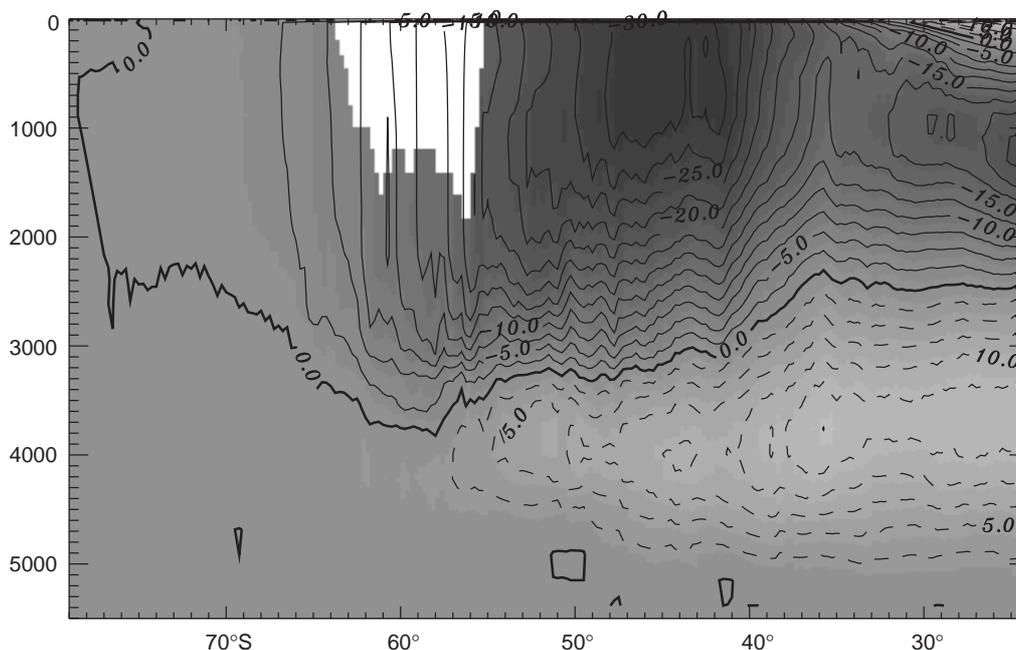
Fig. 4.6.7 The overturning streamfunction from a 6-year mean of the Fine Resolution Antarctic Model, calculated by integrating meridional velocity at constant depth, and then integrating vertically. Flow is anticlockwise around highs (pale shading) and clockwise around lows (dark), with a contour interval of 2.5 Sv. The unshaded region shows the range of latitudes and depths (Drake Passage latitudes), which are unblocked by topography at any longitude. See FRAM Group (1991) and Döös and Webb (1994) for details of the FRAM model.
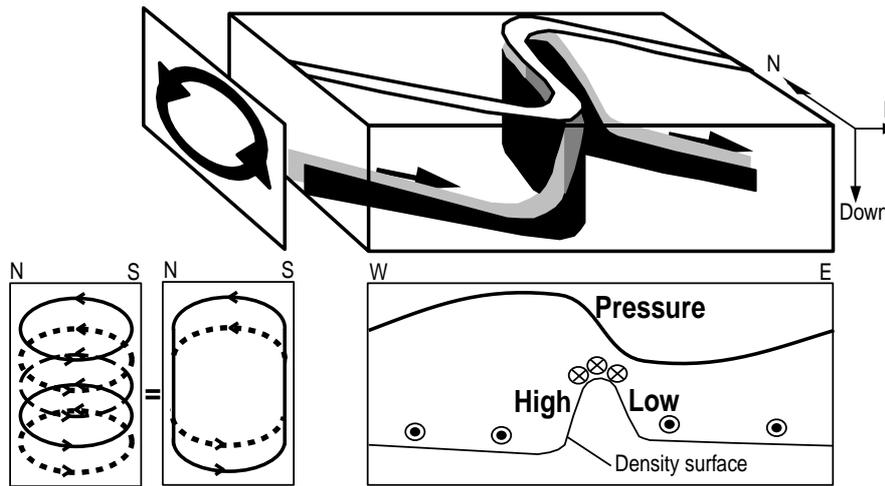
Fig. 4.6.8   Schematic showing an idealized trajectory of a water particle in the ACC moving on a density surface. The trajectory is shown in three dimensions, and projected onto the horizontal plane (top), a constant longitude plane (left), and a constant latitude plane (lower right). The resulting circulation integrated at constant latitude and depth, for this density surface, is an overturning cell with a vertical extent of a few hundred metres. Deeper density surfaces show similar overturning cells, with northward branches at the same depth as the southward branch of the cell related to lighter water, so the zonally integrated cell including all density classes represents a meridional overturning penetrating to great depth, without a need for any water particles to traverse such a large depth range (lower left). Note that this circulation implies higher pressure where the density surface is rising to the east compared with where it is deepening to the east. This results in an eastward pressure force (interfacial form stress) on the water below. This is related to the fact that the northward flow occurs where the vertical thickness of water above the density surface is small, and southward flow where the thickness is large, so there is a net southward mass flux at lighter densities due to the geostrophic flow. This partly balances the northward surface Ekman flux, since the interfacial form stress partly balances the eastward surface wind stress. The same kind of pressure force acting on the sloping bottom topography leads to the bottom form stress, which closely balances the zonally integrated zonal wind stress, since the zonal and depth integral of northward transport is very small.

scenario is the QG picture with only a few density layers, as in the models referred to above. With no flux between layers, all the Ekman flux returns to the south in the top layer, and isopycnal averaging shows no overturning, whereas level averaging shows an overturning cell reaching the bottom layer, which is the only layer containing topography.

- *Case II*. The southward return flow is all in density surfaces that intersect topography. This scenario requires a feedback mechanism between the wind stress and thermohaline forcing, so that the diapycnal flux into and out of these deep layers to the north and south of the chosen latitude can balance the northward Ekman flux.

Consider the balance of zonal momentum, integrated zonally and over two layers (which may be stratified), separated by an isopycnal. The upper layer of thickness h includes the Ekman layer, the lower one reaches from $z = -h$ to the ocean bottom. Writing the depth-integrated northward

volume flux in each layer as V_i, $i = 1, 2$, the steady-state balances read

$$-\rho_0 f \overline{V_1} = -\overline{h' p'_x} + \overline{\tau^x} \qquad (4.6.4)$$

$$-\rho_0 f \overline{V_2} = -\overline{h' p'_x} - \overline{H p_{bx}} \qquad (4.6.5)$$

where the overbar denotes time and zonal mean (see **Fig. 4.6.9** for a schematic showing the relationship between geostrophic meridional flow and interfacial form stresses on an arbitrary layer). In case I, with the Ekman flux all returning above the isopycnal at $z = -h$, we have $\overline{V_1} = \overline{V_2} = 0$. Friction and Reynolds stress are generally negligible and the only remaining terms are wind stress and the pressure forces on the boundaries: interfacial form stress and bottom stress. Thus, for case I, the wind stress and bottom stress are in balance, and are equal to the interfacial form stress at any isopyc-nal below the Ekman return flow and above the bottom. Some dynamical mechanism, such as stationary waves excited by topography in an eastward current, or
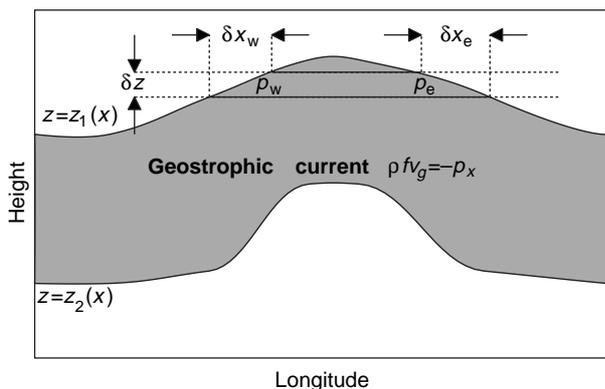
Fig. 4.6.9 Schematic demonstrating the meaning of interfacial form stress for an arbitrary (not necessarily constant density) layer of water (shaded). The net eastward force on the layer is given by $-\int p_x \, dx \, dz$, which is related to the net northward geostrophic mass transport in the layer by $f \int \int \rho v_g \, dx dz = -\int p_x \, dx \, dz$. The contribution to this area integral from the vertical portion δz is $(p_w - p_e) \, \delta z$, where p_w and p_e are pressures at the upper boundary of the layer. This can be written as $p_w z_{1x} \delta x_w + p_e z_{1x} \, \delta x_e$. Performing the vertical integral then gives $-\int p_x \, dx \, dz = \oint (p_1 z_{1x} - p_2 z_{2x}) dx$, where $p_{1,2}$ is pressure at $z = z_{1,2}$. This is the difference between the eastward pressure force on the top interface from above, and that on the lower interface. The layer considered may be bounded by isopycnals, in which case these boundary forces are interfacial form stresses, or the lower interface may be the ocean floor, in which case the corresponding boundary stress is the bottom form stress. In the limit of a small density difference $\delta \rho$ between upper and lower surface, the difference in boundary stresses becomes the interfacial form stress divergence (times $\delta \rho$).

baroclinic instability (again requiring zonal flows comparable to the baroclinic Rossby wave speed), is necessary to maintain the structure of correlated pressure gradients and isopycnal heights that produces interfacial form stresses at depth (see Section 4.6.3.3).

In case II, if we identify $z = -h$ with the position of any isopycnal that does not intersect the bottom or the Ekman layer at the latitude under consideration, then the northward Ekman flow all returns to the south beneath this level, giving $\overline{V_1} = -\overline{V_2} = \overline{\tau^x}/\rho_0 f$. The Coriolis force in the lower layer then exactly balances the form stress, and there is no interfacial form stress on density layers above topography. The difficulty here is in supplying the lower layer with the sources and sinks of water necessary to maintain this flow, without inducing

flows in the intermediate layers. A change in wind stress, for example, would upset the balance, and cause some layers to start filling and others emptying. If this change in configuration could then change the buoyancy forcing by some mechanism such as that proposed by Gnanadesikan and Hallberg (2000) (a feedback between buoyancy forcing and interface height), then an equilibrium might be attainable. Understanding the feedback mechanism could then lead to a prediction for the density structure and therefore the baroclinic flow.

Taken together, these two cases make plain the intimate relationship between wind forcing and buoyancy forcing. Models can produce circumpolar currents when forced by wind alone or by buoyancy alone. The steady state requires a balance for both, and cannot be said to be driven by one or the other. Whether the real Southern Ocean is closer to case I or case II is discussed in detail in Section 4.6.5.

4.6.3.3 Theoretical predictions of ACC transport

A complete theory capable of predicting the absolute transport of the ACC is a formidable challenge. Such a theory would need to account for both wind and buoyancy forcing, stratification, the effect of eddy fluxes in the momentum and buoyancy budget, and for interactions between the strong deep currents and bottom topography. While a complete theory requires elaborate mathematics, some insight can be gained into the factors controlling the transport of the ACC by appealing to a variety of simpler models.

Estimates from eddy flux parameterizations
Simple estimates of the baroclinic transport can be derived from the above considerations of momentum transfer in the ACC. These estimates rest on the assumption that the transfer is mainly downward and carried by the interfacial form stress of the transient eddies. In the extreme case when the stress is transferred undiminished through the water column down to depth (and taken up there by topographic form stress) its magnitude is set by the surface wind stress τ^x. With $h' = \rho'/\rho_z$ and $p'_x = \rho_0 f v'$ the interfacial stress $\overline{h' p'_x}$ turns into the lateral buoyancy flux and the momentum balance in the water column below the Ekman layer becomes

$$\tau^x = \overline{h' p'_x} \approx -\frac{fg}{N^2} \overline{v' \rho'} \qquad (4.6.6)$$

Following Green (1970) and Stone (1972) the lateral buoyancy transport of eddies, growing in the instability process, can be parameterized in term of the gradient of the mean flow, $\overline{v'\rho'} = -\kappa\rho_y = -\kappa\rho_0 f u_z/g$. The idea to combine (4.6.6) with parameterizations of the buoyancy flux for inferring the transport in the form

$$\kappa\frac{f^2}{N^2}u_z = \tau^x/\rho_0 \qquad (4.6.7)$$

was first pursued by Johnson and Bryden (1989). They used Green's form of the diffusivity $\kappa = \alpha|f|l^2/\sqrt{Ri}$ obtained for a baroclinically unstable flow, where $Ri = N^2/(u_z)^2$ is the local Richardson number, l is a measure of the eddy transfer scale and the constant α measures the level of correlation between v' and ρ' in the buoyancy flux ($\alpha = 0.015 \pm 0.005$ according to Visbeck et al. (1997)). The shear of the zonal flow and wind stress are then related by

$$\alpha\frac{|f|^3}{N^3}l^2 u_z^2 = \tau^x/\rho_0 \qquad (4.6.8)$$

Johnson and Bryden's results are obtained by equating the turbulence scale l with the baroclinic Rossby radius λ. For $l = \pi^2\lambda$, with $\lambda = NH/(|f|\pi)$, we get their estimate of the shear

$$u_z = \frac{N}{|f|}\left(\frac{\tau^x/\rho_0}{\pi^3\alpha\lambda H}\right)^{1/2} = \left(\frac{\tau^x/\rho_0}{\pi^3\alpha H^2}\cdot\frac{N(z)}{|f|}\right)^{1/2} \quad (4.6.9)$$

The first relation was used by Johnson and Bryden (1989), with λ taken to be a measure of the bulk Rossby radius, and shows the shear is proportional to the local Brunt–Väisälä frequency $N(z)$. More importantly, the shear is proportional to the root of the wind stress amplitude τ^x. In the following we use a local Rossby radius and an exponential Brunt–Väisälä frequency profile, $N(z) = N_0 \exp(z/2d)$. With $\tau^x = 0.2\,\mathrm{N\,m^{-2}}$, $H = 3500\,\mathrm{m}$, $N_0 = 1.4 \times 10^{-3}\,\mathrm{s^{-1}}$, $d = 2500\,\mathrm{m}$, and a width $B = 600\,\mathrm{km}$ of the ACC, integration of (4.6.9) yields a transport of 82 Sv relative to the bottom.

Visbeck et al. (1997) suggest that in the presence of differential rotation the eddy transfer may be restricted by the Rhines scale $\sqrt{u/\beta}$ rather than the Rossby radius. With $l = \sqrt{u/\beta}$ we find a cubic relation between τ^x and the velocity,

$$uu_z^2 = \frac{\tau^x\beta}{\rho_0\alpha}\frac{N^3(z)}{|f|^3} \qquad (4.6.10)$$

For the exponential $N(z)$ this is easily integrated. A transport of 67 Sv relative to the bottom and a total transport of 124 Sv is obtained for the above set of parameters. In this model the transport would only mildly increase with the magnitude of the wind stress, as $(\tau^x)^{1/3}$.

The action of eddies is not only manifested in the interfacial form stress, it also implies an eddy transport of potential vorticity. A formulation of the momentum balance which is more precise than (4.6.6) is expressed as a balance between the eddy Potential Vorticity (PV) flux and the vertical divergence of the frictional stress (Marshall et al., 1993),

$$-\frac{\partial}{\partial y}\overline{u'v'} + f\frac{\partial}{\partial z}\frac{\overline{v'p'}}{\rho_z} = \overline{v'q'} = -(\tau^x/\rho_0)_z \quad (4.6.11)$$

This balance holds above the depth level where topographic blocking sets in. The eddy PV flux consists of the lateral Reynolds stress divergence and the vertical divergence of the interfacial form stress. Equation (4.6.6) is in fact the consequence of (4.6.11) if the Reynolds stress divergence is small and significant frictional effects are absent below the Ekman layer. If eddy mixing of PV is down the mean PV gradient, $\overline{v'q'} = -kq_y$, vanishing of the eddy PV flux implies homogeneous mean PV. Observations indeed show that isopycnal vorticity gradients are small in and north of the Antarctic Current regime (Marshall et al., 1993). Furthermore, a linear relation was found to exist between the large-scale PV and density, $f\rho_z = a + b\rho$, with $d = f/b$, the e-folding scale of the density field. This implies an exponential $N(z)$, as assumed before, and it also imposes a constraint on the current shear,

$$u_{zz} - \frac{u_z}{d} = \beta\frac{N^2}{f^2} \qquad (4.6.12)$$

obtained by taking the meridional derivative of $f\rho_z = a + b\rho$. Vertical integration leads immediately to the velocity profile and the transport, expressed in term of the shear at some level z_0, or the corresponding density gradient, or the parameters of Green's parameterization green at the level z_0. A more meaningful interpretation is found if (4.6.12) is reformulated as constraint on the vertical profile of the diffusivity κ by inserting (4.6.7),

$$\frac{\partial}{\partial z}\frac{N^2}{\kappa} - \frac{N^2}{\kappa d} = \beta\frac{\rho_0 N^2}{\tau^x} \qquad (4.6.13)$$

Apparently, the assumption of a homogeneous PV state sets the vertical profile of the lateral diffusivity of buoyancy. Since N^2 decays exponentially with scale d in this model, we find $(1/\kappa)_z = \rho_0\beta/\tau^x = $ constant, and thus

$$\kappa(z) = \frac{\kappa_0}{1 + (\rho_0\beta\kappa_0/\tau^x)(z - z_0)} \quad (4.6.14)$$

where $\kappa_0 = \kappa(z_0)$. In this model the shear consists of two parts,

$$u_z = \frac{N^2}{f^2}\left[\frac{\tau^x}{\rho_0\kappa_0} + \beta(z - z_0)\right] \quad (4.6.15)$$

The first contribution is directly wind-driven. The second contribution is driven by the eddies that homogenize the associated PV. The transport (relative to the bottom) of this latter is fairly small and westward ≈ -2 Sv) whereas the first part contributes 39 Sv for our standard values and a diffusivity $\kappa_0 = 1000\,\mathrm{m^2\,s^{-1}}$ at $z_0 = -1000$ m. Following kap then increases to $1200\,\mathrm{m^2\,s^{-1}}$ at depth 3500 m, and (4.6.7) then implies $\rho_y(z_0)/\rho_0 \approx -2.1\times10^{-10}\,\mathrm{m^{-1}}$, in good agreement with observations.

Wind-driven flow in a two-layer QG channel yields very sluggish flow in the deep layer when its topography is arranged such that the geostrophic contours are blocked by the walls (see e.g. Wolff et al., 1991). Straub (1993) found a regime where the baroclinic instability arrests the shear at its critical level, and with the assumption that the deep flow vanishes, the transport becomes $BH\beta\lambda^2$ (notice that this correspond to the second term in equation (4.6.15)). This is only a few Sv, and Straub argues that this contribution would add as a 'channel component of the Southern Ocean' to the values obtained from Sverdrup-type estimates. Though neat as a concept, the baroclinically arrested state seems not to occur in more realistic models like FRAM, nor in the real ocean: here the current is highly supercritical with respect to the baroclinic Rossby wave propagation (Hughes et al., 1998).

All these concepts determine the transport relative to the bottom velocity. Evidently, with a bottom velocity of only $1\,\mathrm{cm\,s^{-1}}$ (this is the typical size of bottom velocities obtained with inverse models, see e.g. Olbers and Wenzel, 1989) and a depth of 3500 m we gain a contribution of 21 Sv for a current width of 600 km. How good is the assumption of zero bottom velocity? The component of the bottom velocity that is normal to the height contours is constrained by the kinematic condition of no flow through the bottom, $w + \mathbf{u}\cdot\nabla H = 0$ at $z = -H$. An estimate of the vertical velocity at the bottom may be obtained by integration of the planetary vorticity equation, $fw_z = \beta v$, from below the surface Ekman layer (with depth D) to the bottom. One finds

$$\mathbf{k}\cdot\nabla\times\boldsymbol{\tau}/(\rho_0 f) - w(-H) = \frac{\beta}{f}\int_{-H}^{D} v\,\mathrm{d}z \quad (4.6.16)$$

In the zonal mean the transport below the Ekman layer is returned in the Ekman layer, then $w(-H) = -u(-H)\cdot\nabla H \approx -(\partial\tau^x/\partial y)/(\rho_0 f)$ and thus $u(-H) \approx \tau^x/(\rho_0 f\delta H)$ where δH is the height of the topography. Values of the order of a few $\mathrm{mm\,s^{-1}}$ are obtained. In view of the fact that the cancellation between the geostrophic flow and the Ekman transport certainly does not occur locally, and the fact that this constraint applies only to the component of $\mathbf{u}$ normal to the bathymetry, the estimate of the bottom velocity must be considered as a lower bound.

As is evident from (4.6.9), (4.6.10) and (4.6.15), the dependence of the baroclinic transport on the amplitude of the wind stress and the Brunt–Väisälä frequency is generally governed by the degree of non-linearity of the eddy flux parameterization. It should be kept in mind that in these parameterizations only transient eddy effects are taken into account. As shown below (and in all analyses of the zonally averaged momentum balance of numerical models), vertical transfer of momentum is also established by standing eddies.

The barotropic formstress mechanism

Estimates of the transport from a more complete theory, which includes the barotropic component of the flow, are difficult to obtain without elaborate mathematics and extreme simplifications. The flat-bottom case, with the usual frictional parameterizations of the bottom or Reynolds stress, is certainly an unrealistic oversimplification. For a flat-bottomed channel with constant wind stress, the total transport is $B\tau^x/(\rho_0 R)$ or $B^3\tau^x/(12\rho_0 A)$, where B is the channel width, R the coefficient of linear bottom friction and A the lateral eddy viscosity. But this model leads to extremely large transports for reasonable choices of the frictional parameters (Hidaka's dilemma). This dilemma is

somewhat relieved when partial barriers are introduced representing continents and leaving smaller gaps (Drake Passage) for the current to pass through (Gill, 1968).

The flat-bottom case gives unrealistic results because it does not allow for the bottom form stress to work in the overall momentum balance (4.6.3), repeated here as

$$\tau^x - \tau_b^x - \overline{Hp_{bx}} = 0 \qquad (4.6.17)$$

This balance has been shown to hold in all more or less realistic numerical models (τ_b^x is the frictional bottom stress, put to the linear from $\rho_0 RHu$ below). The relevance of equation (4.6.17) to the transport becomes clear when the relation of the bottom form stress to the physical mechanisms responsible for establishment of the bottom pressure field are considered.

The simplest of such models are barotropic with simple topography. For Charney and DeVore's (1979) barotropic model of QG flow over sinusoidal terrain (with wavelength $2\pi/k$), the bottom form stress is evaluated as

$$\overline{Hp_{bx}/\rho 0} = \frac{1}{2}(f\delta)^2 \frac{RHu}{R^2 + \kappa^2(u - c_R)^2} \qquad (4.6.18)$$

where $c_R = \beta/\delta^2$ is the speed of barotropic Rossby waves and δ is the amplitude of the topography relative to the mean depth. From the form of (4.6.18) it is obvious that the form stress is most effective if the current speed equals the speed of the Rossby wave, a situation termed 'topographic resonance'. Adapted to ACC conditions, (4.6.17) only yields the subcritical solution ($R^2 \ll (\kappa c_R)^2$ and $u \ll c_R$) and the transport per unit width becomes

$$Hu = \frac{\tau^x/(\rho_0 R)}{1 + (1/2)(\delta ak)^2} \qquad (4.6.19)$$

with $a = |f|/\beta$. If the flow is constricted in a channel this relation still applies (Olbers and Wübber, 1991), but if the topography gets sufficiently high so that blocking of the geostrophic contours by the walls occurs, i.e. $\delta > \delta_c \sim B/a$, the flow switches to a different regime with transport

$$Hu = \frac{L\tau^x}{\rho_0 B \pi f} \frac{\delta_c}{\delta(\delta - \delta_c)} \qquad (4.6.20)$$

as shown by Krupitsky and Cane (1994) for $R/|f| \leqslant O(\delta^3)$, $\delta > \delta_c$, and in similar form by Wang and Huang (1995). The barotropic pressure form stress reflected in these expressions is seen to act as a drag on the flow that considerably reduces the transport compared to the flat-bottom value $B\tau^x(\rho_0 R)$, so that transports of only 10–20 Sv are easily achieved. In the blocked state with transport (4.6.20), the current runs through the channel entirely in boundary layers at the southern and northern walls, connected by an internal boundary-layer current following the blocked geostrophic contours. Krupitsky *et al*. (1996) use an heuristic equivalent barotropic model (see also Ivchenko *et al*., 1999) to show that stratification can relieve this unrealistic behaviour by modifying the geostrophic contours. In an unblocked channel – a Charney–DeVore model with topographic perturbations approaching zero at the walls – the current is allowed to cross the geostrophic contour by frictional processes at all values of topography height and only friction processes allow for a component of the pressure that is out-of-phase with respect to the topography.

Baroclinic mechanisms

The reaction of the zonal barotropic pressure force on the topography leads to a strong reduction in the transport in wind-driven barotropic models. In numerical General Circulation Models (GCMs), it is found that baroclinicity increases the transport from the small values of the barotropic topographic state to realistic values in the range of the observed transport of the ACC. This appears both in coarse-resolution models, for example the early experiments by Bryan and Cox (1972), Cox (1975), and more recently by Olbers and Wübber (1991) and Cai and Baines (1996), and in models with eddy resolution, for example the FRAM experiment (FRAM Group, 1991) and Gille (1997). Analysis of the momentum balance (4.6.17) in FRAM shows that the barotropic and baroclinic bottom form stress components exceed the wind stress by two orders of magnitude (Stevens and Ivchenko, 1997), with eastward acceleration by the barotropic pressure field and a corresponding deceleration by the baroclinic pressure field largely cancelling, such that the wind stress is almost balanced by the residual and the momentum balance (4.6.17) works essentially without friction. Notice that in these baroclinic
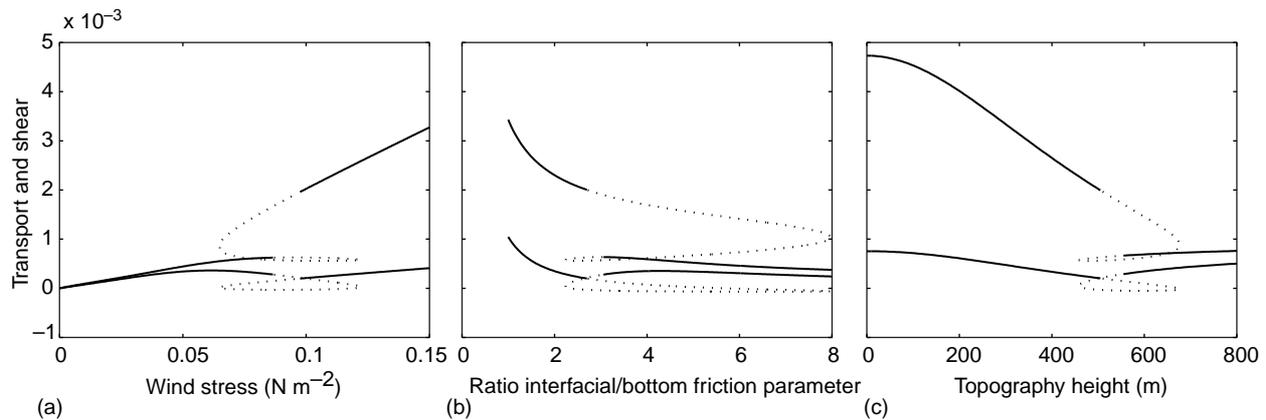
Fig. 4.6.10   Sensitivity of the zonal barotropic transport and shear in the low-order two-layer QG model of Völker (1999) to (a) wind stress amplitude, (b) interfacial and bottom friction, and (c) height of the topography. The flow is forced by sinusoidal wind stress τ^x in a zonal β-plane channel with sinusoidal topography elevation (periodic in the zonal direction, one half sine in meridional direction and vanishing on the walls), friction between the layers and at the bottom is linear. The lower curve in each panel is the shear, the upper curve the transport (both are scaled). Bold lines indicate stable solutions, dotted lines indicate unstable solutions. Notice that there is a small window in the parameter space where only unstable solutions exist. Values of the parameters (if not varied): interfacial friction parameter 2.9×10^{-7} bottom friction parameter $1.1 \times 10^{-7}\,\mathrm{s}^{-1}$, topography height 500 m, wind stress amplitude 0.1 N m^{-2}.

conditions the barotropic pressure does not act as a drag as in homogeneous models. Coarse models have an equally strong effect of the baroclinic pressure field but an unrealistically large contribution from lateral friction.

Baroclinic pressure gradients are established either by thermohaline forcing changing the stratification by water mass conversion, or simply by adiabatic rearrangement of a prescribed layering of stratified mass being lifted over the topography. The latter mechanism is operating in baroclinic adiabatic models (case I in the terminology of Section 4.6.3.2). As shown in Olbers and Völker (1996) and Völker (1999), the waves that produce the topographic resonance of Charney and DeVore (1979) are now baroclinic. These are generated in resonance with the topography and become stationary when the barotropic current speed equals the baroclinic Rossby wave speed. The transport decreases strongly with increasing topography height, starting from a frictionally controlled state at low heights with a transition to a complex resonant regime with multiple equilibria at intermediate heights, and further to a state controlled by barotropic and baroclinic bottom form stresses at high topography. The dependence of the transport and shear of this model on the topography height and other system parameters (friction and forcing) are displayed in Fig. 4.6.10. Stable solutions exist where the curves are bold. Solutions with dotted

curves are unstable (in the window of unstable solutions homoclinic orbits and chaotic behaviour is found; this disappears when increasing the number of resolved modes). Though the momentum balance (4.6.17) in this latter solution seems to operate without friction, it should be pointed out that the barotropic and baroclinic bottom stresses are due to phase shifts of the topographically induced pressure gradients with respect to the topographic undulations, which in turn are proportional to the coefficients of bottom and interfacial friction of the model, again in correspondence to the barotropic model. It is particularly interesting that eddy effects do not appear explicitly in (4.6.17), but transient eddies or bottom friction are needed to produce the phase shift in standing eddies, which is necessary to produce bottom form stress. The baroclinic topographic resonance theory determines the transport in adiabatic models in a manner similar to the barotropic Charney–DeVore mechanism: the bottom form stress is a complicated resonance function of the barotropic and baroclinic velocities and the transport follows from (4.6.17) and a corresponding balance for the baroclinic momentum. The structural properties of this low-order model are preserved when the degrees of freedom are increased from the simplest non-trivial model with 11 modes to a number representing a moderately resolved coarse model (with 75 modes).

The crucial role of transient eddies or non-ideal fluid behaviour beneath the mixed layer is also clear in realistic conditions. Consider a contour of constant time-averaged PV on some density surface beneath the mixed layer but above topography. Since PV and potential density are both materially conserved for an ideal fluid, any steady-state flow can have no component across the PV contour. Thus if there is a return flow at this density, since it must cross this PV contour (assumed to close around Antarctica), either friction or transient eddies are required to permit this.

Analytical theories of the ACC in which the baroclinic pressure field, and thus the bottom stress, is established by thermohaline forcing are still lacking. Several numerical studies (e.g. Olbers and Wübber, 1991; Cai and Baines, 1996; Gent *et al.*, 2000; Gnanadesikan and Hallberg, 2000) with coarse-resolution models have recently investigated the dependence of transport on the buoyancy forcing at the surface. The state of the ACC in these models is generally intermediate between the extreme cases I and II described in Section 4.6.3.2. There is conversion of deep to lighter water masses to allow for a deeper (than Ekman layer) reaching meridional cell but there is also a parameterized interfacial stress of some kind. It is likely that the barotropic and baroclinic form stresses are not solely created by thermohaline processes because the topographic resonance mechanism should be operating as well.

An increase of the ACC transport by an increase of the buoyancy loss by increased brine release off the Antarctic shelf was documented in PE (Potential Energy) models by Olbers and Wübber (1991) and clearly described in Gent *et al.* (2000). Using restoring boundary conditions for heat and salt and different wind fields, Gnanadesikan and Hallberg (2000) found a similar strong increase of the ACC transport with strengthening of the overturning circulation (linked to a deeper thermocline and increased water mass transformation in the northern hemisphere). As shown by Cai and Baines (1996) and Gent *et al.* (2000), parameterizations of sub-grid mixing play an essential role: the ACC transport and the overturning transport are larger in the presence of a larger vertical diffusivity and a smaller isopycnal diffusivity. The latter is in qualitative agreement with the baroclinic transport models (4.6.7) and (4.6.15). It also

agrees with the baroclinic Charney–DeVore model described above, where the transport strongly decreases with friction between the layers (see **Fig. 4.6.10b**).

Topographic steering

The important role of submarine topography in the dynamics of the ACC was discussed in the preceding sections. The topography also acts to steer the current, as noted very early on by Sverdrup *et al.* (1942, pp. 468; 606–7) and described by Gordon *et al.* (1978) (see also the pressure maps in Webb *et al.*, 1991, and the steric height maps in Olbers *et al.*, 1992). Steering by bathymetry has also been detected in the sea surface topography obtained from altimeter data, as reported by Chelton *et al.* (1990) and Gille (1994). A laboratory model of homogeneous and linearly stratified flow over realistic topography is reported by Boyer *et al.* (1993). The resulting flow is in fair agreement with observations regarding steering by the major ridges and troughs, but it also shows significant discrepancies that are traced back to inadequate representation of small-scale passages and fracture zones, unrealistic forcing (which was simulated by sources and sinks of mass) and neglect of the planetary β effect. Early models of topographic effects on the ACC were homogeneous (Kamenkovich, 1962; Johnson and Hill, 1975), so that f/H contours inevitably dominated the flow pattern. The breaking of f/H or bathymetry control by stratification is demonstrated in many simple numerical models (e.g. Klinck, 1993) and the coarse- and high-resolution models of the circumpolar circulation discussed above.

Theory suggests the flow may be steered along bathymetry contours, latitude circles or the geostrophic contours f/H. The conditions under which one of these effects will dominate can be clarified by use of the balance (4.6.1) of integrated vorticity. If the deep ocean is motionless, $\rho_0 f \, \mathbf{k} \times \mathbf{u}_\mathrm{b} = -(\nabla p)_\mathrm{b} = 0$, the bottom torque term in (4.6.1) vanishes because $(\nabla p)_\mathrm{b} = \nabla p_\mathrm{b} - g\rho_\mathrm{b}\nabla H$. This allows a 'free mode' in the transport streamfunction following f contours; in case of locally weak stress curl the flow would follow such a path. However, a motionless abyss requires strong stratification to shield the flow from the influence of topography. The ACC is far from such a state but the converse condition of a homogeneous water

column is inappropriate as well. The influence of stratification is visible in other forms of the vorticity balance

$$(\mathbf{k} \times \nabla \Psi) \cdot \nabla \frac{f}{H} + \frac{f}{H^2} \mathbf{U}^g \cdot \nabla H = \mathbf{k} \cdot \nabla \times (\boldsymbol{\tau}/(\rho_0 H))$$

$$(4.6.21)$$

or

$$H^2 \mathbf{u_b} \cdot \nabla \frac{f}{H} + \mathbf{U}^g \cdot \nabla f = f \mathbf{k} \cdot \nabla \times (\tau/\rho_0 f) = f w_E \quad (4.6.22)$$

obtained from (4.6.1) using $\mathbf{k} \times \nabla \Psi = H \mathbf{u_b} + \mathbf{U}^g - \mathbf{k} \times (\tau/f)$ (ignoring lateral stresses and bottom friction for simplicity). Here, $\mathbf{U}^g$ is the baroclinic (thermal wind) transport relative to the bottom. Obviously, in the case of weak stratification when the baroclinic transport term in the above balances could be ignored, the transport streamfunction and the bottom velocity both would follow f/H contours where the corresponding stress curls are weak. If, in addition, the variation of the planetary vorticity f along the path of the flow is small the bathymetry contours act as characteristics (the topographic $\beta_{T=} f \Delta H/(H \Delta L)$ is in fact generally larger than the planetary β).

A more detailed consideration of stratification effects in models would obviously be required to distinguish between these different possibilities. An intelligent shortcut has been pursued by Marshall (1995a,b) using the homogeneous potential vorticity model of Marshall *et al.* (1993). With a functional dependence $f\rho_z = Q(\rho)$, the density field ρ and the baroclinic transport $\mathbf{U}^g$ is determined by a boundary value, say $\rho(\mathbf{x}, z=0) = \rho_s(\mathbf{x})$. Also the bottom density ρ_b is determined by ρ_s. Furthermore, the bottom is a material surface and – since Montgomery potential, $M = p + g\rho z$, and density are conserved along the three-dimensional flow in adiabatic conditions – we have a functional dependence $M_b = M_b(\rho_b)$. Assuming no friction in the abyss, the bottom velocity is geostrophic, i.e. $f_{ub} = \mathbf{k} \times (\nabla M_b + gH \nabla \rho_b)$. Inserting these relations into (4.6.22) we find (after some manipulation)

$$A(\mathbf{k} \times \nabla \rho_s) \cdot \nabla \frac{f}{H} + B(\mathbf{k} \times \nabla \rho_s) \cdot \nabla f = \frac{f^2}{g} Q_s w_E$$

$$(4.6.23)$$

with

$$A = H^2(H - H_{ref})Q_b, \qquad B = \int_{-H}^{0} Qz \, dz \quad (4.6.24)$$

This vorticity equation determines the surface density from the Ekman pumping w_E, provided the functional relations $H_{ref}(\rho_b) = (1/g)\partial M_b/\partial \rho_b$ and $Q(\rho)$ are specified. The extreme cases with their characteristics $f/H = $ constant and $f = $ constant appear again in (4.6.23) – corresponding to a homogeneous ocean with large A (because here $H_{ref} \gg H$) and motionless abyss with large B (because here $Q \gg Qb$). For constant H_{ref} and linear $Q(\rho) = a + b\rho$ (as found by Marshall *et al.*, 1993) the characteristic problem (4.6.23) is linear and characteristics are easily computed. Marshall (1995b) presents a solution for the transport streamfunction and the interior circulation for an unforced case ($w_E = 0$), prescribing a surface density section between Antarctica and Australia (**Fig. 4.6.11**). The solution transports 160 Sv, and streamlines are largely zonal and quite indifferent to topography in large portions of the domain. Nevertheless, topography clearly steers the current over the major topographic features such as the Atlantic Ridges, the South East Indian and Macquarie Ridges, the Pacific–Antarctic Ridge and the East Pacific Rise. These patterns are enhanced for larger values of the reference depth H_{ref}, which leads to an increase of the deep currents. Outstanding in the solution is an exaggerated northward deflection over the Kerguelen Plateau, whereas the observed equatorward displacement of the current behind Drake Passage and the gradual poleward migration in the rest of the Southern Ocean is not reproduced, presumably due to neglect of the wind forcing.

4.6.4 Water mass formation and conversion

Isopycnals shoal steeply to the south across the Southern Ocean, reflecting the baroclinicity of the ACC (**Fig. 4.6.3**). Water spanning a wide range of density is thus directly forced by exchange of momentum, heat and fresh water with the overlying atmosphere and sea ice. The air–sea–ice interactions strongly modify the physical and chemical properties of outcropping layers and transform water from one density class to another. The water masses formed in this way ventilate a substantial fraction of the world ocean volume and are a key link in the global overturning circulation (Section 4.6.5).
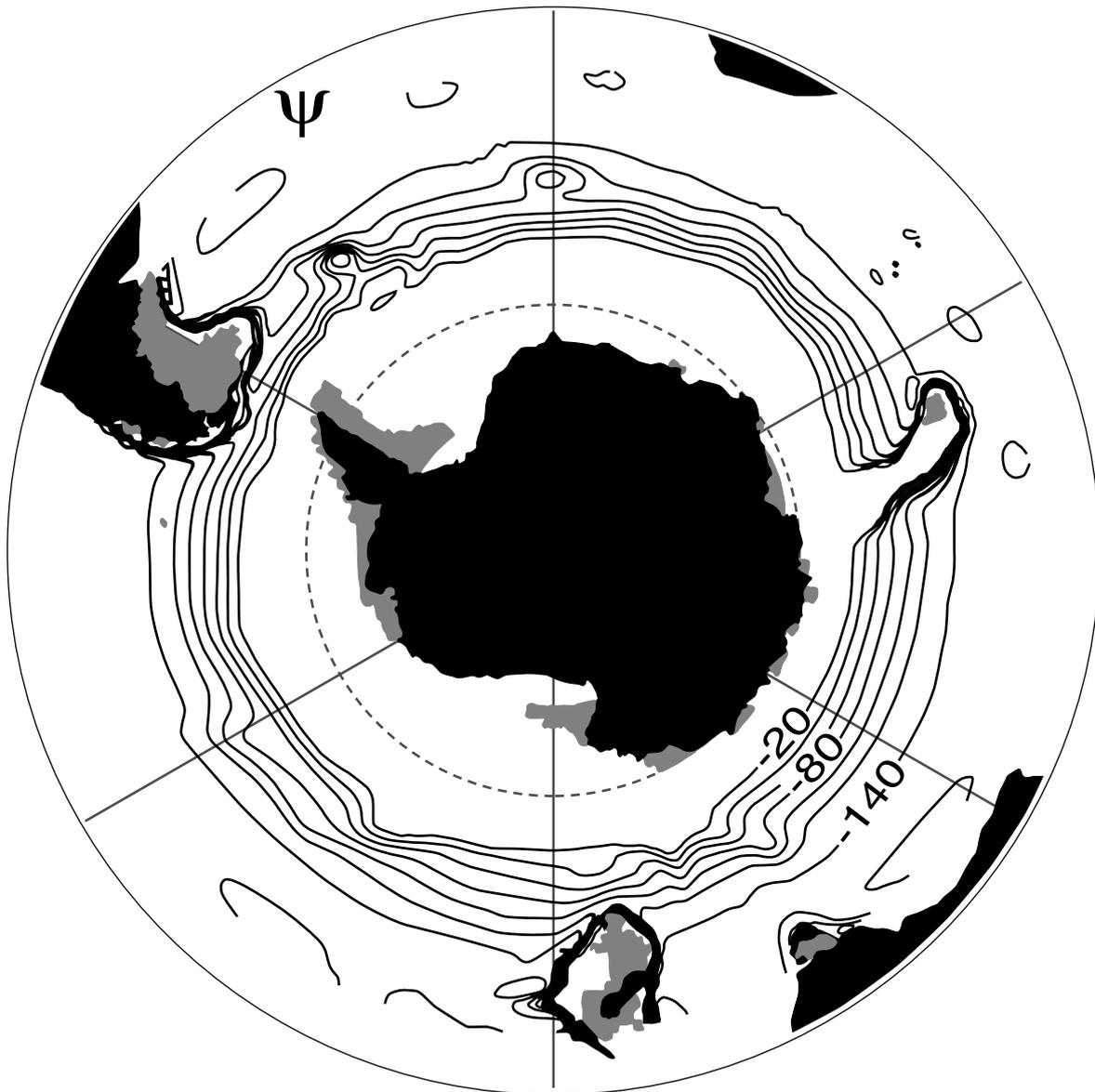
Fig. 4.6.11   Mass streamfunction in Sverdrups from the analytical model of Marshall (1995b), in which fluid parcels negotiate a variable bottom topography while conserving density and potential vorticity. The inviscid, adiabatic circulation follows characteristics which lie between the *f/H* contours found in a homogeneous ocean and the *f* contours found in a strongly stratified ocean.

4.6.4.1 Subantarctic Mode Water and Antarctic Intermediate Water

Subantarctic Mode Water (SAMW) is formed by deep winter convection on the equatorward side of the ACC (Hanawa and Talley, Chapter 5.4; McCartney, 1977). The deep convection imprints the SAMW with its characteristic properties: a vertically well-mixed (hence low potential vorticity) layer that is rich in oxygen (**Fig. 4.6.3**). These tracers allow the SAMWs to be tracked from their formation regions to the southern hemisphere

subtropical gyres, where they renew the waters of the lower thermocline (McCartney, 1982).

The Antarctic Intermediate Water (AAIW) is identified by a salinity minimum layer that descends near the Subantarctic Front (**Fig. 4.6.3b**; Hanawa and Talley, Chapter 5.4). Early authors traced the salinity minimum core layer poleward to where it outcropped in the cold, fresh Antarctic Surface Water, and inferred that circumpolar sinking along this layer renewed the AAIW tongue. The apparent spreading was thought to be driven

by either wind, buoyancy loss, or mixing along isopycnals, depending on the author. McCartney (1982) presented an alternative view, that AAIW is made up of the densest SAMW produced in the Pacific, the end-product of gradual cooling, freshening and loss of buoyancy by air–sea fluxes along the long SAMW circulation path across the Indian and Pacific Oceans. England *et al.* (1993) showed a similar mechanism operating in a GCM. From the southeast Pacific, AAIW spreads north and west in the Pacific and through Drake Passage. Further modification by air–sea fluxes and mixing in the southwest Atlantic produces a cooler and fresher variety of AAIW in the southwest Atlantic (Molinelli, 1981; Piola and Georgi, 1982; Piola and Gordon, 1989; Talley, 1996).

High-quality hydrographic and tracer sections collected during WOCE have focused attention on the distribution and circulation of SAMW and AAIW. The SAMW becomes progressively cooler, fresher and denser across the Indian and Pacific basins (McCartney, 1982). Within each basin, the lighter SAMW varieties are injected further west, and are restricted to the southwest corner of the subtropical gyre. Dense varieties entering the subtropical gyres on their eastern sides travel around the gyres and extend to lower latitudes. While the circulation of the denser AAIW shares some similarities with that of the overlying SAMW, there are important differences as well. In contrast to the near-circumpolar formation of SAMW, the formation of 'new', well-ventilated AAIW is limited to the southeast Pacific and southwest Atlantic, as indicated by the gradual decrease in AAIW oxygen and potential vorticity across the Atlantic and Indian basins (Gordon and Molinelli, 1982; Talley, 1996). For example, at the WOCE SR3 section south of Tasmania (the eastern limit of the Atlantic–Indian variety of AAIW), the oxygen and chlorofluorocarbon (CFC) saturations of AAIW are only 65% and 10–20%, respectively; the 'ventilation age' of AAIW based on CFC-11 is 22–26 years (Rintoul and Bullister, 1999). AAIW enters the lower thermocline of the Atlantic and Pacific subtropical gyres mainly in the southeast quadrant; in contrast, oxygen-rich AAIW enters the Indian basin preferentially in the southwest (Fine, 1993; Toole and Warren, 1993). In the western boundary currents of the subtropical gyres, modified 'older' AAIW returns to the south, resulting in multiple varieties of AAIW with similar densities

at some longitudes (e.g. 40°E, Read and Pollard, 1993; 140°E, Rintoul and Bullister, 1999), and a shift in AAIW properties between basins (Piola and Georgi, 1982).

Few estimates have been made of the formation rate of SAMW and AAIW. Even less attention has been paid to the rest of the circulation loop associated with export of SAMW/AAIW to lower latitudes: what is the fate of SAMW/AAIW exported from the Southern Ocean? What water masses are modified to supply the SAMW/AAIW, and where and how does this occur?

Sloyan and Rintoul (2000a,b) have used a box inverse model, which explicitly includes air–sea buoyancy fluxes and the water mass transformations they drive (e.g. Walin, 1982; Speer and Tziperman, 1992), to examine the formation and circulation of SAMW and AAIW. Comparing the transport in density layers at each Southern Ocean chokepoint shows that 18 Sv more SAMW/AAIW leaves the Indian Ocean sector south of Australia than enters the basin south of Africa (**Fig. 4.6.12**). This production rate represents the net effect of air–sea buoyancy fluxes, diapycnal mixing and meridional exchange with the subtropical gyre. The net production in the Indian sector is balanced by net consumption in the other basins. SAMW is formed in the Indian sector both by cooling and freshening of subtropical water carried south in the Agulhas Current and its extension, and by water flowing northward across the ACC (and gaining buoyancy) in the Ekman layer. SAMW is carried from the Indian to the Pacific by the ACC, where it is modified by mixing and air–sea exchange, and ultimately returns to the Indian basin via the Indonesian passages (Gordon, Chapter 4.7). The SAMW therefore participates in an Indian–Pacific 'throughflow gyre' in which warm throughflow water is converted to cool SAMW in the Indian Ocean, and SAMW entering the Pacific basin is converted back to water warm enough to supply the throughflow.

In the zonal integral, 34 Sv of Upper Circumpolar Deep Water (UCDW) is upwelled south of the ACC and converted to IW densities by air–sea buoyancy flux (**Fig. 4.6.12**; see also the discusion in Section 4.6.5 and **Fig. 4.6.14**). This transformation by air–sea fluxes of heat and fresh water is largely compensated by diapycnal mixing south of 40°S. However, the compensation is not immediate. While the zonally integrated export of SAMW/
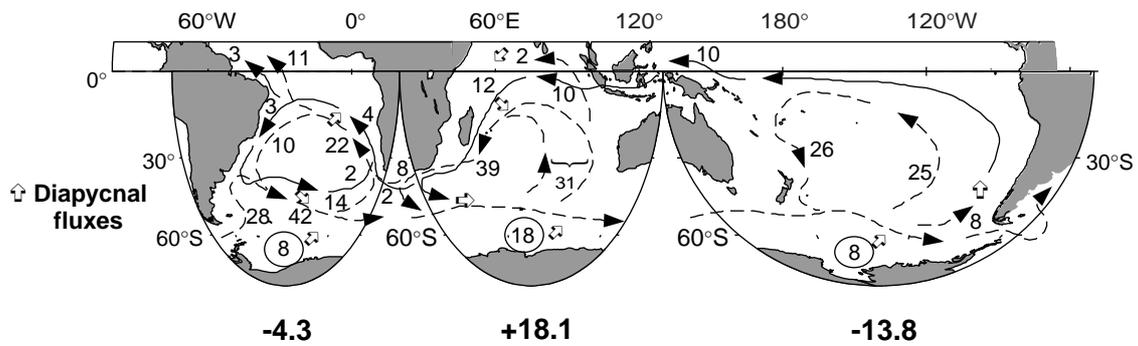
Fig. 4.6.12   A summary of the circulation and formation of SAMW and AAIW, from the inverse model of Sloyan and Rintoul (2000b). Numbers give volume fluxes in Sv of thermocline water (solid line, neutral density $< 26.0 \, kg \, m^{-3}$) and intermediate water (SAMW/AAIW, dashed line, neutral density $26.0 < \gamma_n < 27.4 \, kg \, m^3$). Open arrows represent diapycnal fluxes driven by air–sea exchange and interior mixing. Circled numbers in each of the Southern Ocean sectors represent conversion of Upper Circumpolar Deep Water to SAMW and AAIW. Numbers in boxes are the net convergence ($+$ve) or divergence ($-$ve) of SAMW/AAIW in each sector of the Southern Ocean due to meridional and diapycnal fluxes; mass is conserved by a compensating divergence in zonal transport of the ACC.

AAIW from the Southern Ocean is small (< 5 Sv), the gross exchange is large (about ± 80 Sv). The inflow of 'new' SAMW/AAIW to the subtropical gyres is roughly balanced by an outflow of 'old' SAMW/AAIW, whose properties have been modified by mixing during their transit of the gyres. Part of the returning SAMW/AAIW is then converted to denser UCDW by diapycnal mixing.

Because the SAMW and AAIW are renewed by air–sea interaction on decadal time scales, they provide a good place to look for evidence of changes in forcing over such time scales. Several recent studies have identified decadal changes in the SAMW and AAIW layers in the south Indian and Pacific oceans (e.g. Bindoff and Church, 1992; Johnson and Orsi, 1997; Bindoff and McDougall, 2000). Wong *et al*. (1999) have shown that these patterns are generally coherent throughout the Pacific. These changes are consistent with warming and freshening at the surface outcrops of these layers, as predicted to occur in coupled climate models forced by increasing greenhouse gas concentrations (Bindoff and McDougall, 1994).

4.6.4.2 Circumpolar Deep Water

Two prominent core layers underlie the salinity minimum of the Antarctic Intermediate Water throughout the Southern Ocean. An oxygen minimum layer (**Fig. 4.6.3c**) is used to define the Upper Circumpolar Deep Water (UCDW) (Callahan, 1972). At slightly greater depth (and density) lies the salinity maximum of the Lower Circumpolar

Deep Water (LCDW) (**Fig. 4.6.3b**). The LCDW layer is supplied by saline North Atlantic Deep Water (NADW) exported from the Atlantic. The addition of 'new' NADW in the South Atlantic produces an oxygen and salinity maximum between the UCDW and LCDW entering the basin through Drake Passage (Reid *et al*., 1977). The high-salinity signature of the NADW/LCDW can be traced to low latitudes of the abyssal Indian and Pacific basins (Reid and Lynn, 1971). LCDW enters the low-latitude basins primarily in a series of deep western boundary currents (see Hogg, Chapter 4.5). Within the subtropical basins, the LCDW is slowly modified by mixing with surrounding fresher water, and the oxygen concentrations are reduced by biological consumption. The return of the slightly less dense, low-oxygen water from the Indian and Pacific basins supplies the oxygen minimum of the UCDW (Callahan, 1972). Both types of CDW spread poleward and upward across the Southern Ocean, ultimately outcropping at the sea surface south of the ACC. As described in the previous section and in more detail in Section 4.6.5, the outcropping of CDW and the resulting water mass transformation by air–sea fluxes provides the main connection between the lower and upper limbs of the global overturning circulation.

4.6.4.3 Antarctic Bottom Water

The WOCE decade has seen substantial progress in our understanding of where, how and at what rate AABW is formed in the Southern Ocean. In the

Weddell Sea several major programmes have contributed to this advance (see Fahrbach *et al.*, 1998, for a review). The new results are broadly consistent with the picture of circulation and water mass formation in the Weddell developed by earlier investigators. Relatively warm Circumpolar Deep Water (sometimes called Warm Deep Water in the Weddell) enters the basin from the east in the Weddell gyre and is converted to Weddell Sea Deep and Bottom Water through ice–ocean–atmosphere interactions along the southern and western margin. Two mixing scenarios have been proposed for the formation of Weddell Sea Bottom Water: (1) mixing of Winter surface water, Warm Deep Water, and Western Shelf Water, whose salinity has been enriched by brine rejected during sea ice formation (Foster and Carmack, 1976); and (2) mixing of Ice Shelf Water (formed by cooling and freshening of Western Shelf Water beneath the vast ice shelves in the southern Weddell Sea) with Weddell Sea Deep Water and Warm Deep Water (Carmack and Foster, 1975; Foldvik *et al.*, 1985).

Fahrbach *et al.* (1994) maintained a line of current meter moorings and repeat hydrographic sections across the central Weddell Sea between 1989 and 1993. They estimate a Weddell gyre transport of 30 Sv, almost all of which is carried in narrow currents along the continental slope. Recent model results of Beckmann *et al.* (1999) indicate that Fahrbach's section does not cut through the centre of the Weddell gyre, which has a maximum transport exceeding 50 Sv, in agreement with observations along the Greenwich meridian (Schröder and Fahrbach, 1999). By considering inflow and outflow in density layers across this section, Fahrbach *et al.* (1991) infer a net conversion of 3–4 Sv of Winter and Warm Deep Water to Weddell Sea Deep and Bottom Water. Direct measurements of the outflow of bottom water (potential temperature $< -0.7°C$) from moorings in the western boundary current show a mean of 1.7 Sv, and a range from 0.8 to 3.9 Sv; additional bottom water export occurs offshore of the narrow boundary current (Fahrbach *et al.*, 1994). Dense shelf water also escapes from the Weddell Sea to the Scotia Sea through gaps in the island chain separating the basins (Whitworth *et al.*, 1994).

The 700-km-long drift of Ice Station Weddell in 1992 further refined our understanding of ice–ocean–atmosphere interactions in the western Weddell Sea (Gordon *et al.*, 1993; Gordon, 1998). Taken together, the results of recent programmes identify a number of distinct AABW sources along the southern and western rim of the Weddell Sea, each with a characteristic temperature, salinity and stable isotope signature. The relative importance of the two mixing scenarios described above varies along the rim of the Weddell. Entrainment of warmer and saltier deep water found over the slope largely determines the ultimate properties of deep and bottom water leaving the Weddell Sea. Gordon (1998) estimates the formation of Weddell Sea Bottom Water (WSBW) (potential temperature $< -0.7°C$) during the period of the Ice Station drift to be 4.0–4.8 Sv. Mensch *et al.* (1998) estimate a similar formation rate (about 5 Sv of deep and bottom water) from tracer measurements obtained during the Ice Station.

Satellite observations of a large polynya in the Weddell Sea in the late 1970s, and the large changes in temperature and salinity of Weddell Sea Deep Water that resulted (Gordon, 1982), first sparked interest in the variability of Weddell waters. A number of recent studies have documented variations of deep and bottom water properties. Gordon (1998) concludes that the average salinity of the WSBW formed during the Ice Station is too low to provide the end-member required to account for the deep water of the Weddell gyre, and suggests that the bottom water forming at the present time contains more Ice Shelf Water. Nöst and Österhus (1998) describe the impact of several large grounded icebergs north of the Filchner depression, which have caused the cessation of high-salinity shelf water formation there and led to a cooling and freshening in the depression. Further afield, Coles *et al.* (1996) and Hogg and Zenk (1997) have observed cooling and freshening (on isopycnals) of AABW spreading north in the South Atlantic, which they attribute to changes in open ocean convective events in the Weddell Sea.

Recent studies have also identified or confirmed a number of important sources of AABW outside the Weddell Sea. Rintoul (1998) shows that the Adélie coast (140–150°E) is likely to be a more significant source of bottom water than previously appreciated. The Adélie Land Bottom Water is evident in Fig. 4.6.3 as a thin layer of cold, fresh, dense, high-oxygen water found over the continental slope and rise of Antarctica. He argues on the

basis of Worthington's (1981) volumetric census, and T–S properties at the sills bounding the Australian–Antarctic Basin, that the Adélie Land Bottom Water accounts for up to 25% of the global volume of AABW. The concentration of CFC-11 in plumes of Adélie Land Bottom Water is high ($>2.9\,\mathrm{pmol\,kg^{-1}}$ in 1991; Rintoul and Bullister, 1999), and the CFC saturation (35%) is similar to that in plumes observed in the southwest Weddell Sea. High CFC concentrations observed over the continental slope at 30°E in the eastern Weddell Sea (Mantisi *et al.*, 1991) indicate that a source exists along the Enderby coast or in Prydz Bay as well, as suggested earlier by Jacobs and Georgi (1977).

Recent studies have also identified regions of the Antarctic coastline where AABW is *not* formed. For example, Fahrbach *et al.* (1994) show that AABW is not formed in the eastern Weddell Sea, primarily because of the inability to confine shelf water (and increase its salinity) over the narrow shelf there. In the Amundsen and Bellingshausen Seas, nearly undiluted warm Circumpolar Deep Water intrudes onto the shelf and melts large amounts of continental ice, preventing formation of dense saline water (Hellmer *et al.*, 1998).

Foster (1995) confirmed the observation of Carmack and Killworth (1978) that the Wilkes Land coast east of 150°E forms water that is not quite dense enough to sink to the seafloor. Sinking of Antarctic shelf waters to deep, rather than abyssal, layers is potentially of significant importance to ventilation of the deep ocean at lower latitudes, since these lighter layers are less confined by bathymetry and are free to spread northward. The formation of 'not quite bottom' water around Antarctica has so far received little attention, and the transport is unknown. Weppernig *et al.* (1996) conclude from the distribution of stable isotopes that a large fraction ($>80\%$) of the dense Ice Shelf Water leaving the continental shelf in the Weddell Sea supplies the WSDW at intermediate depth, rather than sinking to the bottom to supply WSBW.

A powerful demonstration of the utility of transient tracer measurements is the use of CFC inventories to estimate the formation rate of AABW (Orsi *et al.*, 1999). They find that about 8 Sv (4.9 Sv in the Atlantic and 3.2 Sv in the Indian–Pacific) of new bottom water (a 50:50 mix of dense shelf water and entrained deep water) must sink across the 2500 m isobath to explain the observed CFC-11 inventory in the AABW layer (water with neutral density greater than $28.27\,\mathrm{kg\,m^{-3}}$). Much of the 3.2 Sv formed in the Indian–Pacific sectors is likely produced by sources in the Australian Antarctic Basin (e.g. Adélie Land), since the average CFC-11 concentration there is 2.5 times as large as that of the Atlantic and Pacific basins (**Fig. 4.6.13**).

The CFC-11 inventory of Orsi *et al.* (1999) provides a strong integral constraint on the formation of AABW. Broecker *et al.* (1997) have noted that PO_4^* ($PO_4^* = PO_4 + O_2/175 - 1.95\,\mathrm{\mu mol\,kg^{-1}}$, where 175 is the average molar Redfield ratio of O_2 consumption to remineralization in the deep sea and 1.95 is an arbitrary constant.) also provides a constraint on the relative production of dense ventilated water in the Southern Ocean and the North Atlantic: roughly equal contributions from the two source regions are required to explain the PO_4^* value observed in the deep Indian and Pacific Oceans. The PO_4^* budget thus requires sinking of about 15 Sv of ventilated shelf water around Antarctica, a factor of three to four higher than implied by the CFC-11 inventory or by transport measurements near the sources. Part of this discrepancy may be explained by exchange of ventilated shelf water with deep water lighter than neutral density of $28.27\,\mathrm{kg\,m^{-3}}$. However, given that the CFC content of the lighter deep water is so much lower than that of the dense abyssal layer it is difficult to see how this process can account for the discrepancy. Another possibility is that AABW production in recent decades is substantially smaller than the average during the last millennium (Broecker *et al.*, 1997). The conflict between estimates of ventilated deep water production based on CFC and PO_4^* remains unresolved.

WOCE sections, moored arrays, and other recent measurements have helped to map out the system of deep western boundary currents and through-passage flows that carry AABW and LCDW to lower latitudes (e.g. Mantyla and Reid, 1995; Rhein *et al.*, 1998; Whitworth *et al.*, 1999; Zenk *et al.*, 1999; see also Hogg, Chapter 4.5).

4.6.5 The Southern Ocean and the global overturning circulations

The global overturning circulation is often taken to be synonymous with the circulation loop formed by sinking and export of NADW from the
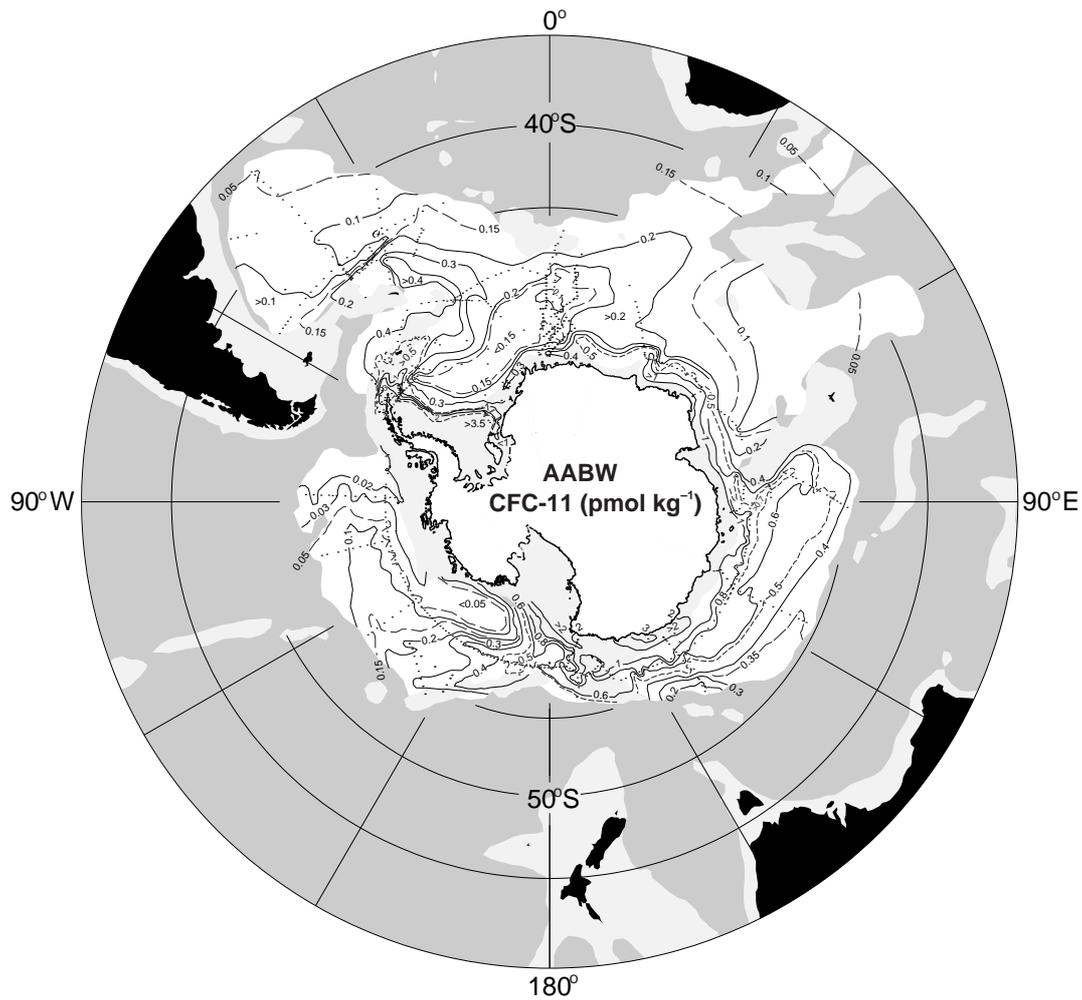
Fig. 4.6.13 Inventory of CFC-11 in the AABW layer (between neutral density 28.27 kg m^3 and the seafloor), from Orsi *et al.* (1999). The layer-mean CFC-11 concentrations in the Atlantic (Weddell–Enderby), Indian (Australian Antarctic), and Pacific basins are 0.17, 0.47 and 0.18 pmol kg^{-1}, respectively.

North Atlantic, upwelling elsewhere in the ocean, and a return flow of lighter water. The ACC carries NADW from the Atlantic to the other ocean basins and so forms an important link in this circulation path. But the significance of the Southern Ocean to the global overturning goes beyond the passive role of redistributing NADW. Upwelling and buoyancy forcing in the Southern Ocean drive the water mass transformations required to close the NADW cell, as well as several additional overturning circulations, both shallow and deep. The nature of the overturning circulation is also intimately related to the dynamics of the ACC, as described above.

To balance the sinking of NADW in the North Atlantic, somewhere deep water must be converted to lighter water. Given that no zones of concentrated upwelling from the abyss to the thermocline have been observed, the traditional assumption is that the upwelling is more or less broadly distributed in the ocean interior. This assumption, together with conservation of potential vorticity, has profound implications for the deep flow, as illustrated in the simple and elegant theory of the abyssal circulation of Stommel and Arons (1960).

Direct measurements of mixing in the ocean pycnocline, however, typically find diffusivities an order of magnitude too low to support the upwelling required to balance the sources of deep water (Ledwell *et al.*, 1993; Toole *et al.*, 1994). Radiocarbon distributions also argue against widespread upwelling of deep water through the pycnocline (e.g. Toggweiler and Samuels, 1993). On the other hand, recent measurements suggest

diapycnal mixing throughout much of the water column is enhanced over rough topography (Polzin *et al.*, 1997; Ledwell *et al.*, 2000). Diapycnal mixing may also be enhanced near the boundaries of the ocean (Wunsch, 1970; Armi, 1978). To date there are few direct measurements of diapycnal mixing, and so it is not yet clear if (non-uniform) interior mixing will prove to be sufficient to balance the sinking at high latitude.

An alternative view is that the required conversion of dense to light water occurs primarily in the Southern Ocean, where deep isopycnals outcrop and are exposed to air–sea fluxes of heat and fresh water (Döös and Coward, 1997; Toggweiler and Samuels, 1998). Here we reconsider the meridional overturning in the Southern Ocean, this time from the perspective of its connection to the rest of the world ocean.

As explained in the introduction, an active overturning circulation in the Southern Ocean was inferred by early investigators on the basis of the prominent core layers (**Figs 4.6.2** and **4.6.3**) which extend across the ACC (e.g. Sverdrup, 1933). Deep water spreads south and upward across the ACC, and is balanced by sinking and northward flow of both lighter intermediate water and denser bottom water. This picture is broadly consistent with the requirements of closure of the NADW overturning – import of NADW to the Southern Ocean is balanced by export of IW and BW, as observed in the South Atlantic (Rintoul, 1991; Saunders and King, 1995) – and suggests that the water mass conversions taking place in the Southern Ocean are a key element of the overall cell. The 'intermediate' cell (conversion of DW to IW) is consistent with northward Ekman transport driven by the strong westerly winds: divergent Ekman transport drives upwelling of deep water, which in steady state must be converted to lighter water by buoyancy input from the atmosphere as it is driven north across mean density contours (**Fig. 4.6.1**; Toole, 1981; Speer *et al.*, 2000).

However, in the ensuing 60 years, few attempts have been made to quantify the flow paths in Sverdrup's diagram. One recent exception is Schmitz's (1995, 1996a,b) attempts to synthesize global view from a large number of published estimates for individual branches of the overturning circulation. Summing his estimates for the three individual basins at about 40°S, the overturning in the Southern Ocean consists of 53 Sv of deep water

(NADW/UCDW) flowing south, balanced by 48 Sv of bottom water (AABW/LCDW) and 5 Sv of intermediate water (SAMW/AAIW) flowing north. Schmitz's summary of 'best guess' values from the literature suggests that in the zonal integral the deep cell (DW to BW) is much stronger than the intermediate cell (DW to IW) of Sverdrup's diagram.

Schmitz's circulation scheme is derived from a number of individual estimates, which may not be internally consistent. His results are, however, very similar to estimates from a recent Southern Ocean inverse model, which provides an internally consistent solution that explicitly includes air–sea buoyancy forcing and diapycnal mixing (Sloyan and Rintoul, 2000c). In the zonal integral across roughly 30°S they find 52 Sv of deep water flowing south, balanced by 46 Sv of lower deep and bottom water and 6 Sv of intermediate water flowing north (**Fig. 4.6.14**; see caption for layer definitions).

These results suggest the deep overturning cell, in which dense AABW/LCDW exported to the Indian and Pacific is balanced by import of slightly less dense DW, dominates the overall Southern Ocean overturning. Note that the contribution of NADW to the zonally integrated overturning is small by comparison: for example, the 10 Sv poleward flow of lower NADW (neutral density between 28.0 and 28.2) is more than compensated by strong equatorward flow in this density class in the Indian and Pacific. The intermediate cell (conversion of DW to IW) is weak in the zonal integral at 40°S (although the gross exchanges in this density class are large, as described in Section 4.6.4.1).

The observations of flow entering and leaving the Southern Ocean across 30–40°S imply significant poleward transport in density layers shallower than the Drake Passage sill (neutral density of $27.4 - 28.0 \, \text{kg m}^{-3}$). In addition, the net northward transport of light water is much smaller than the Ekman transport, suggesting that much of the Ekman transport returns poleward at similar density south of 40°S. Substantial poleward transport in density layers not blocked by topography implies divergence of the interfacial form stress, or equivalently, the eddy buoyancy flux (Section 4.6.3). Speer *et al.* (2000) show that there are strong gradients in isopycnal thickness (or potential vorticity) across the ACC in the UCDW layer; eddy mixing will therefore tend to smooth out the gradient, resulting in a volume flux to the south. Meridional gradients of isopycnal thickness are
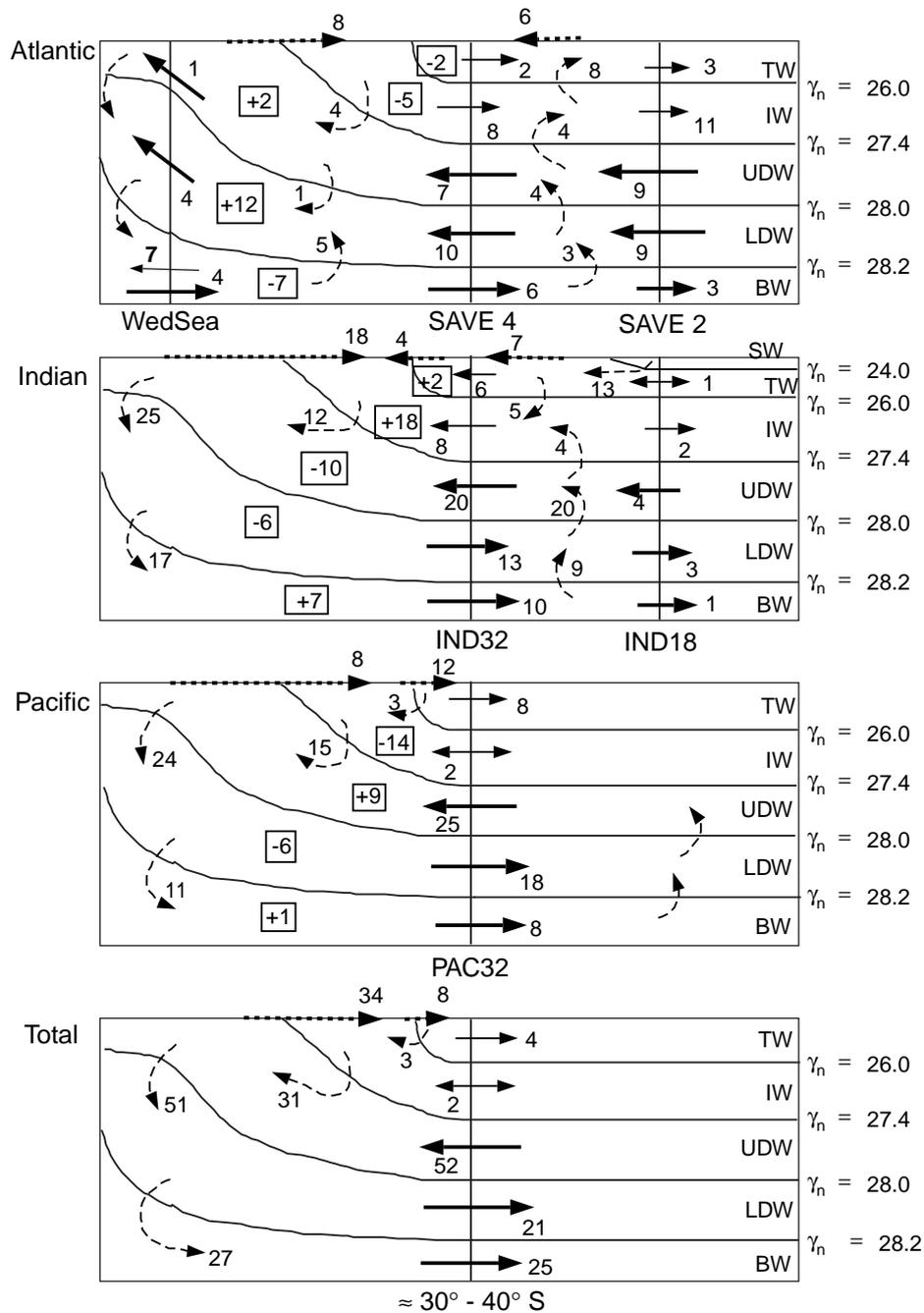
Fig. 4.6.14 A schematic five-layer view of the overturning circulation (units of $10^6\,\mathrm{m}^3\,\mathrm{s}^{-1}$ in each southern hemisphere basin, and the zonal sum, from the inverse model of Sloyan and Rintoul (2000c). SW, Surface Water; TW, Thermocline Water; IW, Intermediate Water; UDW, Upper Deep Water; LDW, Lower Deep Water; and BW, Bottom Water. The neutral surfaces used to define each layer are shown. Air–sea flux-driven diapycnal fluxes are shown by bold, dashed arrows at the sea surface. Net diapycnal fluxes due to interior mixing are indicated by thin dashed arrows; their location along the isopycnal is schematic only. Numbers in boxes represent the net convergence ($+$ve) or divergence ($-$ve) of a particular layer in that sector of the Southern Ocean. Two-headed arrows in intermediate water highlight that the net flux is the difference between nearly balancing northward and southward fluxes. Labels beneath each plot indicate the hydrographic sections used in the inverse model (e.g. PAC32 is a section at about 32°S in the Pacific).

small in the LCDW layer, which lies below the topography where mean geostrophic meridional flow is possible. In the terminology of Section 4.6.3.2, these observations suggest the real ocean is closer to case I.

Because the overturning circulation in numerical models depends on how the model is forced (e.g. restoring or surface flux boundary conditions), how eddies are parameterized, and whether the model is in steady state, they do not provide conclusive evidence for or against deep versus shallow compensation of the Ekman transport. Nevertheless, a number of recent simulations, both coarse and fine resolution, support the idea that a significant fraction of the Ekman transport is balanced by a return flow shallower than the topography (case I). For example, Hirst and McDougall (1998) show that a coarse-resolution level model with the Gent and McWilliams (1990) parameterization of eddy-induced advection is consistent with case I: in density coordinates, there is very little meridional overturning associated with the Ekman transport between 40°S and 60°S. While the degree of cancellation between the eddy-induced advection and the Ekman transport (nearly complete in their model) depends on the value chosen for the diffusivity, the model improvements that result when such a parameterization is used support the notion that eddy-driven transport likely plays an important role in the overturning circulation of the Southern Ocean. (The simulation is also improved by the decrease in horizontal mixing permitted when the eddy parameterization is used.) Killworth and Nanneh's (1994) analysis of the zonal momentum budget in isopycnal layers in FRAM supports this conclusion: at the latitudes of Drake Passage, almost all (at the southern side) to about half (at the northern side) of the northward transport of light water returns south at densities that do not intersect topography.

In summary, observations suggest significant poleward flow in layers above topography, which must be driven by divergence of the eddy (standing and/or transient) interfacial form stress, in addition to a poleward flow in density layers blocked by topography. A variety of numerical simulations also suggest that the presence of eddies permits a southward flux at densities above topography, which compensates a large fraction of the northward Ekman flux. In QG simulations this compensation must be complete, since no diabatic transport is permitted. In PE models that either resolve or adequately parameterize the effect of eddies, there is still a large degree of compensation. The time and zonal mean flow at constant depth, in which the Ekman mass transport returns at depths blocked by topography, is thus decoupled from the time and zonal mean flow at constant density, in which most of the Ekman transport returns at densities above topography.

4.6.6 Conclusions

Substantial progress has been made in understanding the circulation of the Southern Ocean during the 'WOCE decade'. This progress has relied on advances in observations, theory and modelling. Observations collected during WOCE represent a significant achievement, given the challenges posed by the remote and often hostile nature of the Southern Ocean. Highlights include a circumpolar survey of high-quality hydrographic, tracer and ADCP data (including some of the first repeat sections obtained in the region), a number of mooring arrays, float and drifter deployments, and satellite measurements of sea surface height and temperature. Analytical models have provided insight into the mechanisms responsible for setting the transport of the ACC. The last decade has also seen rapid development of numerical models of the Southern Ocean, including the first GCMs to incorporate stratification, realistic topography, and to resolve (or 'permit') eddies. The ability of coarse-resolution models to simulate the Southern Ocean has also improved significantly, in part due to more effective parameterizations of the effect of eddies. Fine-resolution models have achieved sufficient realism that we can use them to estimate the magnitude of individual terms in the momentum or vorticity budgets, to identify sites of strong topographic influence, and to describe qualitatively the circumpolar structure of the complex, filamented ACC, although it is still not clear what resolution is necessary for a truly realistic simulation.

The momentum, vorticity and buoyancy budgets – and as a consequence, the zonal and meridional circulations – are intimately linked. The dynamics of the ACC differ in character from those of strong currents in other ocean basins that are zonally blocked. The now established fact that

the zonal wind stress is balanced by bottom form stress means that the Sverdrup balance is upset by interactions with topography: meridional flows driven by the wind stress curl are returned in flows balanced by bottom pressure torques rather than in viscous boundary layers. In addition, for a significant part of the northward Ekman flux to return in density layers that do not intersect topography at some latitude (as models and observations indicate), a strong, deep-reaching ACC is needed to maintain the interfacial form stress divergence required for dynamical balance of the return flow. Without strong top-to-bottom flow, the ACC could not achieve a dynamical state that is almost non-viscous, supercritical with respect to Rossby wave propagation and meridionally constrained in narrow fronts.

It has long been recognized that eddy fluxes play an important part in the dynamics and thermodynamics of the Southern Ocean. In the last decade the central role of eddies has become even more clear. Eddies carry heat poleward and momentum downward across density surfaces; the momentum transfer helps establish the correlation between pressure and topography that provides the bottom form stress to balance the wind; and a substantial fraction of the meridional overturning circulation is dynamically balanced by the divergence of interfacial form stress. In this sense, the dynamical analogy between the ACC system and the mid-latitude troposphere is even more complete than previously appreciated.

The high-quality WOCE hydrographic and tracer sections have provided new insights into the formation and circulation of Southern Ocean water masses. Buoyancy exchange with the atmosphere drives substantial water mass transformations, converting both light water to dense water (e.g. over the extensions of the subtropical western boundary currents, and near the Antarctic margin) and dense to light (over much of the Southern Ocean, where northward Ekman transport combines with heat and freshwater input to convert deep water to intermediate water). Tracers have been used to refine estimates of the rate of Antarctic Bottom Water formation, to identify the main circulation pathways, and to quantify the 'ventilation age' of Southern Ocean water masses.

Many of these ideas have come together to provide a new appreciation of the significance of the Southern Ocean in the global climate system. The ACC provides the interbasin connection required for a global thermohaline circulation to exist. Perhaps a more important link between the Southern Ocean and the global overturning circulation is the transformation of water masses driven by air–sea forcing and diapycnal mixing. Various lines of evidence suggest that diapycnal mixing rates in the main thermocline are too slow to support the traditional view that sinking of deep water is balanced by upwelling uniformly distributed over the ocean. Much of the conversion of cold to warm water that is required to close the NADW overturning circulation appears to take place in the Southern Ocean, where deep water outcrops and is exposed to air–sea buoyancy forcing. The NADW cell, however, is overwhelmed in the zonal integral by an even stronger deep overturning involving conversion of upper deep water to denser lower deep and bottom water in the Southern Ocean, and conversion of lower to upper deep water in the deep basins of the Indian and Pacific Oceans.

Analyis of WOCE data from the Southern Ocean is at an early stage. We anticipate further progress will be made as the full suite of observations collected during WOCE (e.g. hydrography and tracers, floats, altimetry, moorings) are synthesized. Advances in theory and modelling of the ACC also continue at a rapid rate. Among the important open questions to be addressed by these analyses and future observational programmes are: What is the absolute transport of the ACC, and how and why does it vary in time? How sensitive are the water mass conversions taking place in the Southern Ocean, and the overturning circulations of which they are part, to changes in atmospheric forcing? What are the relative contributions of 'deep' and 'bottom' waters produced along the Antarctic margin to the ventilation of the deep sea? How representative are the WOCE-era measurements? Can we detect and interpret changes between measurements made during WOCE and historical or future observations?

Many, but not all, of these questions will be answered as analysis of the WOCE data set continues. However, while a major step forward, the WOCE observations are still sparse in space and time. In a region as remote as the Southern Ocean, there will always be a strong reliance on remote sensing and autonomous instruments. For example, the development of profiling floats now provides

the opportunity to sample the variability of the Southern Ocean on broad spatial scales for the first time. The floats complement, but do not replace, the high-density repeat sampling along fixed cruise tracks required for transport estimates. Direct, coherent and sustained *in-situ* measurements of absolute velocity on large spatial scales, as required to improve our understanding of the barotropic flow, remain beyond our present technological capacity. However, the combination of a highly accurate and well-resolved geoid from planned satellite gravity missions and satellite altimetery will provide an unprecedented opportunity to resolve the absolute flow of individual jets in the ACC, and and make possible quantitative studies of eddy–mean flow interactions in the Southern Ocean. The Southern Ocean also remains a great challenge for ocean models, which suffer their greatest difficulties where strong currents and eddies interact with bottom topography, a process central to the dynamics of the ACC. While significant progress been made in the

theory, observation and modelling of the Southern Ocean, we still do not have a realistic theoretical picture of the dynamical processes that regu-late the response of the ACC to wind stress and thermohaline forcing. Only with such a picture will it be possible to place the WOCE observations into their proper context and determine their limitations. The challenge is to use and extend the WOCE data and models to help build that picture.