Enhanced Rotational Mixing in the Radiative Zones of Massive Stars

Adam S. Jermyn,[1]* Christopher A. Tout[1] and Shashikumar M. Chitre[1,2]

[1] *Institute of Astronomy, University of Cambridge, Madingley Rd, Cambridge CB3 0HA, UK*
[2] *Centre for Basic Sciences, University of Mumbai, India*

ABSTRACT

Convection in the cores of massive stars becomes anisotropic when they rotate. This anisotropy leads to a misalignment of the thermal gradient and the thermal flux, which in turn results in baroclinicity and circulation currents in the upper radiative zone. We show that this induces a much stronger meridional flow in the radiative zone than previously thought. This drives significantly enhanced mixing, though this mixing does not necessarily reach the surface. The extra mixing takes on a similar form to convective overshooting, and is relatively insensitive to the rotation rate above a threshold, and may help explain the large overshoot distances inferred from observations. This has significant consequences for the evolution of these stars by enhancing core-envelope mixing.

Key words: Stars - evolution; Stars - interiors; Stars - rotation

1 INTRODUCTION

The problem of meridional mixing in the radiative zones of rotating stars has been studied for nearly a century (Edding-ton 1929; Maeder & Zahn 1998). The understanding that has emerged is based on the famed von Zeipel theorem (Zeipel 1924), which states that in a radiative zone there is no way to satisfy the equations of hydrostatic and thermal equilib-rium without allowing the pressure and temperature to vary separately. Hence in rotating stars thermodynamic quanti-ties are not solely determined by the pressure; they exhibit baroclinicity. This baroclinicity drives a circulation current that mixes material in the meridional plane.

The scale of this current is proportional to the heat flux, to the reciprocal of the entropy gradient and to $\Omega^2 R/g$, where Ω is the angular velocity and g and R are the typi-cal gravity and cylindrical radial coordinate in the radiative zone. The interpretation of this effect, either as due to tran-sient damping or a genuinely driven flow, has led to some controversy (Busse 1981) but it is generally agreed that the mixing itself occurs (Osaki 1982).

Underlying the standard analysis of this mixing is the assumption that baroclinicity is generated locally by the re-quirements of thermal and hydrostatic equilibrium. A key result of this work is that this fails in stars with central convection zones[1]. In such stars rotation distorts convective

motions, which set a baroclinic boundary condition at the base of the radiative zone. This boundary condition is of-ten more significant than the locally-generated baroclinicity and so results in enhanced circulation. This scenario has not been adequately studied and constitutes a novel driver for meridional flows in the radiative zones of massive rotating stars.

A related scenario that has been studied extensively is that of circulation driven by mechanical pumping in the convection zone. For instance Garaud & Brummell (2008) and Garaud & Acevedo Arreguin (2009) examined the trans-mission of mechanical forcing across the tachocline and Ga-raud & Bodenheimer (2010) extended this to stars with mul-tiple convection zones, where there may be interesting inter-actions between the zones mediated by the intervening ra-diative zone. A key insight of these studies which also holds for our scenario is that the circulation cannot be transmit-ted and propagated without significant mechanical stresses in the radiative zone. Such stresses have also been consid-ered as drivers of circulation (Zahn 1992; Maeder & Zahn 1998) and so must be taken into account in any analysis of these phenomena.

The possibility of enhanced mixing in these systems is particularly relevant in light of the recent discovery of grav-itational waves from merging binary black holes (Abbott et al. 2016). This has led to increased interest in forma-tion mechanisms for massive stellar cores. Many such mech-anisms are quite sensitive to the magnitude of rotational mixing, particularly in massive stars (Marchant, Pablo et al.

---

* E-mail: adamjermyn@gmail.com
[1] These are those with masses above 1.2M$_\odot$

2016), and so understanding this phenomenon in more depth is extremely important.

In Section 2 we review the derivation of the meridional velocity by Eddington (1929), Sweet (1950) and Zeipel (1924). Briefly, the centrifugal effect distorts isobars, causing a flux anisotropy which violates thermal equilibrium (Fig. 1, top panel). Thermal equilibrium is then restored by inducing baroclinicity (Fig. 1, middle panel) and hence introducing a circulation current (Fig. 1, bottom panel).

In Section 3 we argue that turbulent anisotropy in the convection zone makes the zone baroclinic. This arises because the Coriolis effect couples different components of the velocity[2] and so produces heat transport at an angle to the entropy gradient, as shown in Fig. 2. As in the case of radiative zones this results in a circulation current. Notably this is significantly enhanced relative to what would occur in a radiative zone because the radial entropy gradient is much smaller in convection zones.

A further consequence of turbulent anisotropy is that the heat flux in the zone is highly aspherical[3]. This is important in massive stars with convective cores because this flux perturbation produces baroclinicity at the base of the radiative zone. In Section 5 we argue that this baroclinic boundary condition serves to drive additional mixing in the radiative zone. This mixing is often orders of magnitude greater than what the Eddington–Sweet calculation alone provides. We then comment on the prospects for similar mechanisms in stars of lower mass.

In Section 6 we argue that baroclinic mixing mechanism is not fundamentally altered by chemical composition gradients. In particular we show that such gradients play the same role in this mechanism as they do in the Eddington–Sweet circulation and hence only produce corrections of order unity.

Finally in Section 7 we argue that these circulation currents are actually the source of the extra mixing normally attributed to convective overshooting. We then discuss the physical interpretation of the circulation in light of historical difficulties with the Eddington–Sweet circulation.

2   ORIGINS OF MERIDIONAL FLOW

The meridional velocity arises from the impossibility of thermal and hydrostatic equilibrium when the pressure gradient is misaligned with the thermal flux. In the case considered by Eddington (1929) and von Zeipel (Zeipel 1924) this arises because of a bending of the isobars owing to $O(\Omega^2)$ centrifugal effects. We now review this derivation.

To begin note that, in the absence of differential rotation, there exists an effective potential satisfying

$$\nabla^2 \Phi = 4\pi G \rho - 2\Omega^2, \tag{1}$$

where ρ is the density and G is Newton's gravitational constant. With hydrostatic equilibrium we also have

$$\nabla P = -\rho \nabla \Phi \tag{2}$$

[2] We also estimate the effect of centrifugally–induced baroclinicity and find it to be quite a bit smaller.
[3] This is related to the baroclinicity because the convective flux depends strongly on the entropy gradient.
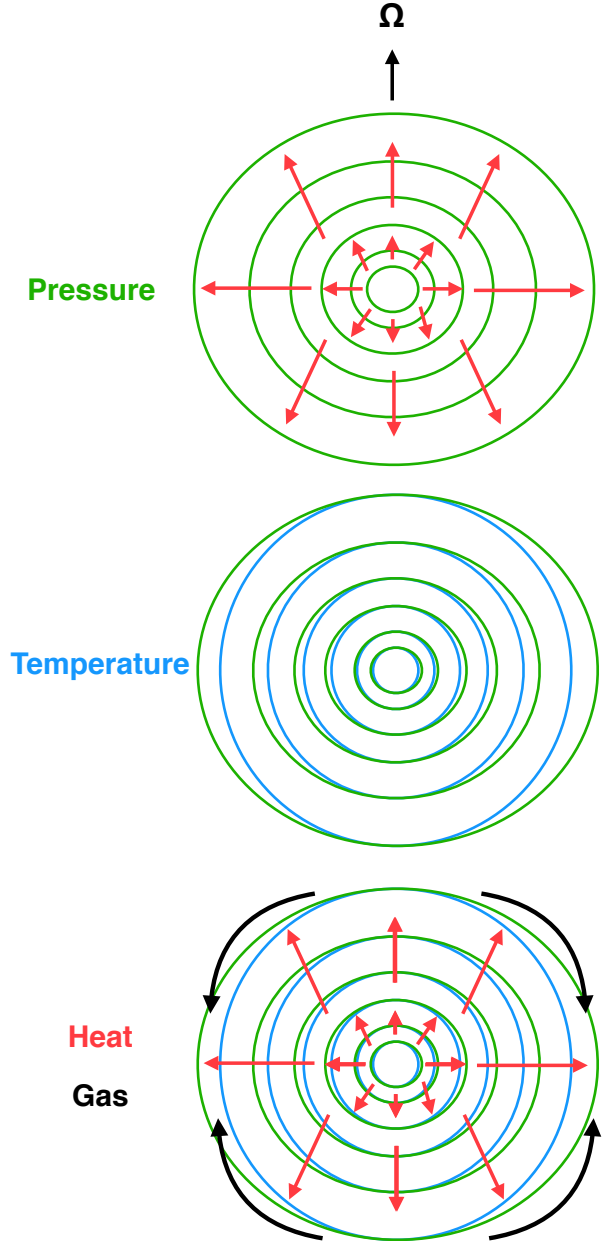

Figure 1. Centrifugal effects distort isobars. Variation in the magnitude of the pressure gradient along isobars leads to a flux anisotropy (top panel). This pushes the system out of thermal equilibrium, and a new equilibrium is established with baroclinicity (mismatched entropy and pressure surfaces, middle panel) and circulation currents (bottom panel). The length of the arrows is not meaningful. Likewise the geometry of the gas circulation is only meant schematically.

This means that

$$\nabla \times \left(\frac{1}{\rho} \nabla P \right) = -\nabla \times \nabla \Phi = 0 \tag{3}$$

and so

$$\nabla \rho \times \nabla P = 0, \tag{4}$$

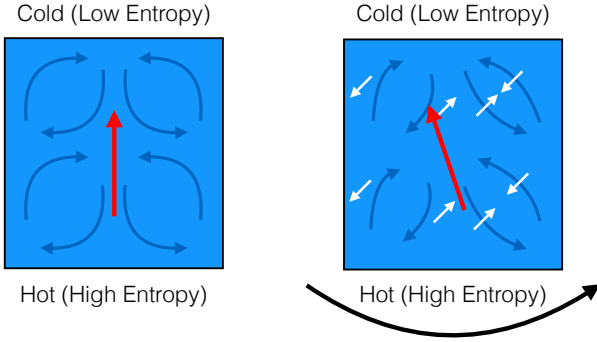hence isobars and isochors coincide. From equation (2) it

Figure 2. Non-rotating convection (left, blue arrows) transports heat along the entropy gradient. Rotating convection (right, blue arrows) is distorted by the Coriolis force (white arrows) and so transports heat at an angle relative to the entropy gradient.

follows that these coincide with surfaces of constant Φ and so P and ρ may be written solely as functions of Φ.

The energy balance in stars is given by

$$c_p \rho T \frac{\partial s}{\partial t} = \frac{\partial F_i}{\partial x_i} + \rho \varepsilon - \rho c_p T u_i \frac{\partial s}{\partial x_i}, \tag{5}$$

where c_p is the specific heat capacity at constant pressure, $\boldsymbol{F}$ is the diffusive flux, ε is the specific rate of energy generation from nuclear processes, s is the dimensionless entropy, ρ is the density, T is the temperature, u_i is the velocity and summation is implicit over repeated indices. In the case of radiation we may write the diffusive flux as

$$F_i = -\rho c_p \chi \frac{\partial T}{\partial x_i}, \tag{6}$$

where χ is the thermal diffusivity. In steady state $\partial_t s$ vanishes and with no nuclear burning we may set $\varepsilon = 0$ so that equation (5) becomes

$$\frac{\partial F_i}{\partial x_i} - \rho c_p T u_i \frac{\partial s}{\partial x_i} = 0. \tag{7}$$

Now suppose that there is no meridional flow, such that $u = 0$. We have then

$$\frac{\partial F_i}{\partial x_i} = 0. \tag{8}$$

So from equation (6)

$$\frac{\partial}{\partial x_i} \left(\rho c_p \chi \frac{\partial T}{\partial x_i} \right) = 0. \tag{9}$$

All these quantities are thermodynamic variables which may be written as functions of P and ρ using the equation of state. Because P and ρ are functions only of Φ equation (9) may be written as

$$\frac{\partial}{\partial x_i} \left(q(\Phi) \frac{\partial \Phi}{\partial x_i} \right) = 0, \tag{10}$$

where q is a function only of potential. Expanding the divergence yields

$$\frac{dq}{d\Phi} |\nabla \Phi|^2 + q(\Phi) \nabla^2 \Phi = 0 \tag{11}$$

and hence

$$|\nabla \Phi|^2 = -\frac{q(\Phi) \nabla^2 \Phi}{dq/d\Phi}. \tag{12}$$

The right-hand side may be evaluated with equation (1) to find a function only of Φ. The left-hand side, by contrast, may be written as

$$|\nabla \Phi|^2 = \left| \boldsymbol{g} - \Omega^2 R e_r \right|^2, \tag{13}$$

where R is the distance from the rotation axis, e_r is the cylindrical radial unit vector and $\boldsymbol{g}$ is the acceleration due to gravity. This generally varies along isobars because the direction of $\boldsymbol{g}$ changes while that of the centrifugal acceleration does not, and the magnitude of the centrifugal acceleration varies while, to leading order, that of gravity does not. As a result equation (12) cannot be satisfied simultaneously with equation (4). This is the original Eddington–Sweet argument for radiative zones (Zeipel 1924).

From this it follows that there must at a minimum be either a meridional flow or baroclinicity or both. The original Eddington–Sweet argument does not address the possibility of baroclinicity as an alternative and so just concludes that a meridional flow arises with scale set by equation (7). In Appendix A we address this point and demonstrate that in fact both a meridional flow and differential rotation generally arise, with this equation setting the scale of the flow.

3 ANISOTROPIC CONVECTION

We now turn to the convection zone with the aim of showing that convective anisotropy results in baroclinicity. When viewed on large length-scales convective turbulence acts to diffuse heat[4]. The effective diffusivity tensor is of the form

$$\mathsf{D}_{ij} \approx \langle \delta x_i \delta u_j \rangle. \tag{14}$$

Physically this just means that material located at $\delta \boldsymbol{x}$ relative to the centre of an eddy is transported along with velocity $\delta \boldsymbol{u}$.

In the presence of an entropy gradient, this diffusivity gives rise to a heat flux

$$F_i = -c_p \rho T \mathsf{D}_{ij} \frac{\partial s}{\partial x_j}. \tag{15}$$

Crucially, the convective viscosity is not isotropic (Lesaffre et al. 2013; Gough 1978; Unno 1957). This is because the Coriolis acceleration of a fluid parcel owing to its velocity $\boldsymbol{u}$ is

$$\boldsymbol{a}_c = 2\boldsymbol{u} \times \boldsymbol{\Omega}, \tag{16}$$

where $\boldsymbol{\Omega}$ is the angular velocity of rotation. Because this expression contains a cross product, the Coriolis effect generically leads to a coupling between different components of the velocity, and hence between different components of the position and velocity of the eddy. In rotating systems this means that the heat flux is not aligned with the entropy gradient (Kichatinov & Rudiger 1993; Lesaffre et al. 2013), and it is this effect that drives baroclinicity.

---

[4] More formally, the convective diffusivity arises from applying the Renormalization Group to the Navier–Stokes equation. Upon integrating out short-range degrees of freedom we obtain the effective couplings among longer-range degrees of freedom (See e.g. Yakhot & Orszag 1986). The motion is random and so leads to diffusion rather than advection.
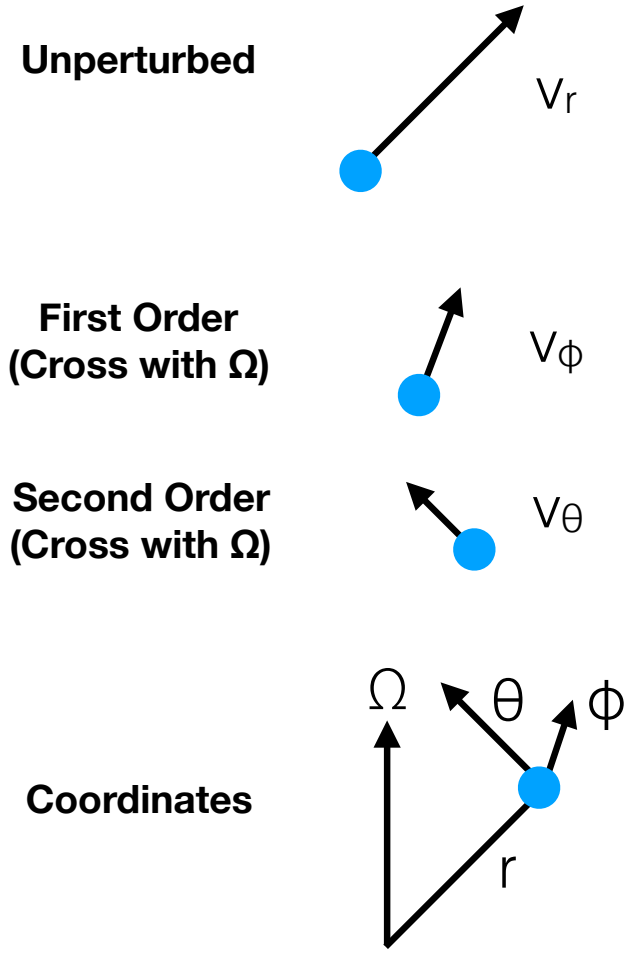
Unperturbed

v_r

**First Order
(Cross with Ω)**

v_ϕ

**Second Order
(Cross with Ω)**

v_θ

Ω θ ϕ

Coordinates

r

Figure 3. The radial motion v_r of a convective eddy is shown unperturbed. The Coriolis force acting on this creates an azimuthal component u_θ at first order. Acting on this the Coriolis force creates a latitudinal component u_θ at second order. So two applications of the Coriolis effect are required to generate motion along $\boldsymbol{e}_\theta$ from motion along $\boldsymbol{e}_r$.

In spherical coordinates the diffusivity tensor may be written as

$$\mathbf{D} = D_0 \mathbf{D}' = D_0 \begin{pmatrix} 1 & \epsilon_{\theta r} & \epsilon_{\phi r} \\ \epsilon_{r\theta} & c_1 & \epsilon_{\phi\theta} \\ \epsilon_{r\phi} & \epsilon_{\theta\phi} & c_2 \end{pmatrix}, \tag{17}$$

where D_0 is a scalar function (Gough 2012). When $\Omega = 0$ this tensor is diagonal because of spherical symmetry. The terms on the diagonal do not vanish when $\Omega = 0$, so we let c_1 and c_2 be constants of order unity. Perturbing away from this limit we argue that all components involving the azimuthal direction are linear in Ω because a single application of the Coriolis effect suffices to correlate meridional motion with azimuthal motion. By contrast $\epsilon_{r\theta}$ and $\epsilon_{\theta r}$ are quadratic in Ω at leading order because it takes two applications of this effect to correlate one component of meridional motion with another, as shown in Fig. 3. This form is in agreement with the works of Kitchatinov (2013) as well as those of Jermyn et al. (2018) and Lesaffre et al. (2013), each of which have

been verified with data from 3D simulations of rotating convection.

To see how this drives baroclinicity we first calculate the corrections that the convective anisotropy introduces in the equation of thermal equilibrium. For this we assume that there is no meridional flow in the convection zone. We then introduce a meridional flow in the convection zone and demonstrate that it does not change the overall magnitude of the baroclinicity.

Inserting equations (17) and (15) into equation (8) we find

$$(\nabla + \nabla \ln(\rho T D_0)) \cdot (\mathbf{D}' \cdot \nabla s) = 0. \tag{18}$$

Taking the system to be axisymmetric, we may write the function D_0 as

$$D_0 = D_{0,0}(P,T) + \Omega^2 D_{0,1}(P,T,\theta) + \mathcal{O}(\Omega^3). \tag{19}$$

The dependence of $D_{0,1}$ on θ and the independence of $D_{0,0}$ on the same reflects the symmetry of the problem: only contributions coupling to the rotation may depend on latitude because in the absence of rotation latitude is an arbitrary coordinate. In addition, we write

$$\nabla s = \boldsymbol{e}_r s_r + \boldsymbol{e}_\theta s_\theta, \tag{20}$$

where s_θ is $\mathcal{O}(\Omega^2)$, and

$$\nabla \ln(\rho T D_0) = \boldsymbol{e}_r t_r + \boldsymbol{e}_\theta t_\theta, \tag{21}$$

where t_θ is $\mathcal{O}(\Omega^2)$. In both cases we have used axisymmetry, so that $\partial/\partial_\phi \equiv 0$. Inserting these into equation (18) gives

$$0 = (\nabla + t_r \boldsymbol{e}_r + t_\theta \boldsymbol{e}_\theta) \cdot [\boldsymbol{e}_r (s_r + \epsilon_{\theta r} s_\theta) + \boldsymbol{e}_\theta (s_r \epsilon_{r\theta} + s_\theta c_1)]. \tag{22}$$

Expanding equation (22) then gives us

$$0 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 (s_r + \epsilon_{\theta r} s_\theta) \right) + \frac{1}{r \sin\theta} \frac{\partial}{\partial \theta} (\sin\theta (s_r \epsilon_{r\theta} + s_\theta c_1))$$
$$+ t_r (s_r + \epsilon_{\theta r} s_\theta) + t_\theta (s_r \epsilon_{r\theta} + s_\theta c_1). \tag{23}$$

We now drop terms which we know to be of order Ω^3 and higher so that

$$0 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 s_r \right) + \frac{1}{r \sin\theta} \frac{\partial}{\partial \theta} (\sin\theta (s_r \epsilon_{r\theta} + s_\theta c_1)) + t_r s_r. \tag{24}$$

As before terms which couple to latitude must be at least second order in Ω. This includes derivatives in θ of terms which themselves are non-zero for $\Omega = 0$. This allows us to neglect the term $\epsilon_{r\theta} \partial_\theta s_r$ but does not allow any such simplification for the remaining terms because those which are $\mathcal{O}(\Omega^2)$ do not necessarily become of higher order when differentiated with respect to latitude. Thus we have

$$0 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 s_r \right) + \frac{1}{r} \left(s_r \frac{\partial \epsilon_{r\theta}}{\partial \theta} + c_1 \frac{\partial s_\theta}{\partial \theta} \right)$$
$$+ \frac{\cot\theta}{r} (s_r \epsilon_{r\theta} + s_\theta c_1) + t_r s_r. \tag{25}$$

Reorganising terms we write

$$0 = \frac{\partial \epsilon_{r\theta}}{\partial \theta} + \frac{c_1}{s_r} \frac{\partial s_\theta}{\partial \theta} + \cot\theta \left(\epsilon_{r\theta} + \frac{s_\theta c_1}{s_r} \right) + 2 + r t_r + \frac{\partial \ln s_r}{\partial \ln r}. \tag{26}$$

Now $\epsilon_{r\theta}$, $\partial_\theta \epsilon_{r\theta}$, s_θ and $\partial_\theta s_\theta$ vanish when $\Omega = 0$ because

these terms can only be non-zero when spherical symmetry is broken. So

$$0 = \frac{\partial \epsilon_{r\theta}}{\partial \theta} + \frac{c_1}{s_r}\frac{\partial s_\theta}{\partial \theta} + \cot\theta\left(\epsilon_{r\theta} + \frac{s_\theta c_1}{s_r}\right) + \Delta\left(rt_r + 2 + \frac{\partial \ln s_r}{\partial \ln r}\right),$$
(27)

where Δ refers to the difference between the expression evaluated in the rotating and non-rotating cases. Hence the term in parentheses is at least of order Ω^2.

Because the fourth term in equation (27) does not contain an explicit dependence on θ it may well vanish in a barotropic setting if the other terms all vanish as well. However if the other terms do not vanish then either their explicit dependence on θ must cancel among them, in which case it must be that s_θ is non-zero, or else they must introduce a θ dependence into the fourth term. In either case the system becomes baroclinic, and by the same argument as in the Eddington–Sweet case and Appendix A this comes along with a circulation current of the order of the implied baroclinicity.

We have argued that $\epsilon_{r\theta}$ is $O(\Omega^2)$. The only relevant[5] time-scale other than the rotation rate is the Brünt–Väisälä frequency $|N|$, so up to a dimensionless factor of order unity we expect

$$\epsilon_{r\theta} \approx \frac{\Omega^2}{|N|^2}.$$
(28)

Because $\epsilon_{r\theta}$ vanishes on the equator by symmetry we expect its variation over latitudes to be of order unity. Hence

$$\frac{\partial \epsilon_{r\theta}}{\partial \theta} \approx \frac{\Omega^2}{|N|^2},$$
(29)

which is in good agreement with the results of Kichatinov & Rudiger (1993). This may also be seen by noting that these quantities are locally determined and are perturbed to second order, so the magnitude of each is just its characteristic scale (unity in the case of $\epsilon_{r\theta}$, D_0 in the case of ΔD_0) multiplied by the dimensionless parameter $\Omega^2/|N|^2$.

The quantities t_r, s_θ and s_r and their derivatives, by contrast, are primarily determined non-locally by the perturbing term $\epsilon_{r\theta}$ throughout the atmosphere. To find them we must integrate outward from the centre of the star, where they all vanish by symmetry. We approximate these integrals by the integral of the magnitude of the perturbation from the centre to the point of interest. The relevant radial length scale is the pressure scale height

$$h \equiv -\frac{\partial \ln P}{\partial r},$$
(30)

so radial derivatives of these quantities are characterised by

$$\left|\frac{d}{dr}\Delta(rt_r)\right| \approx \frac{d}{dr}\left|\frac{s_\theta}{s_r}\right| \approx \left|\frac{d}{dr}\Delta\frac{\partial \ln s_r}{\partial \ln r}\right| \approx \frac{1}{h}\left(\frac{\Omega}{|N|}\right)^2.$$
(31)

Integrating up from the core where $P = P_c$ then gives

$$\Delta(rt_r) \approx \frac{s_\theta}{s_r} \approx \Delta\frac{\partial \ln s_r}{\partial \ln r} \approx \alpha,$$
(32)

where

$$\alpha \equiv \int_{\ln P}^{\ln P_c} \frac{\Omega(\mathcal{P})^2}{|N|(\mathcal{P})^2}\,d\ln\mathcal{P}$$
(33)

and $\mathcal{P}$ is the pressure. We use this as a proxy for the radial coordinate because even in highly baroclinic regions the pressure gradient is predominantly radial. Note that α is related[6] to the small angle λ between ∇p and $\nabla\rho$ by

$$\lambda \approx h|s_r\alpha|.$$
(34)

In Appendix C we demonstrate that the dominant contribution to this is that owing to turbulent anisotropy. Hence we may neglect centrifugal effects.

4 CIRCULATION IN THE CONVECTION ZONE

We now examine the consequences for circulation in the convection zone. We know that it is not possible for rotating radiative zones to avoid circulation currents because they cannot satisfy the condition of thermal equilibrium without becoming baroclinic, and, once baroclinic, a flow is generically driven with magnitude set by the baroclinicity. The same is true for convection zones and so there is a meridional flow. Incorporating such a flow[7] and recalling the assumption of axisymmetry, this equation becomes[8]

$$\frac{r}{D_0}u_r = \frac{\partial \epsilon_{r\theta}}{\partial \theta} + \frac{c_1}{s_r}\frac{\partial s_\theta}{\partial \theta} + \cot\theta\left(\epsilon_{r\theta} + \frac{s_\theta c_1}{s_r}\right)$$
$$+ \Delta\left(rt_r + 2 + \frac{\partial \ln s_r}{\partial \ln r}\right),$$
(35)

where u_r is the radial component of the flow velocity. From this we see that the magnitude of the flow is set by the typical magnitude of the remaining terms, so we expect

$$u_r \approx \frac{D_0}{r}\alpha',$$
(36)

where α' describes the total perturbation to equation (27). This is related to α by adding the local perturbations given by equations (28) and (29). Hence

$$\alpha' \equiv \frac{\Omega^2}{|N|^2} + \alpha.$$
(37)

This is similar to the result of Roxburgh (1991), except that the denominator of the perturbing parameter here is correctly identified as $|N|^2$ rather than g/r.

We now determine the relative scales of the radial and latitudinal flows u_r and u_θ. In steady state conservation of mass gives $\nabla \cdot (\rho\boldsymbol{u}) = 0$, so

$$\frac{r}{\sin\theta}\frac{\partial(u_\theta\sin\theta)}{\partial\theta} + ru_\theta\frac{\partial \ln\rho}{\partial\theta} = -\frac{\partial(r^2 u_r)}{\partial r} - u_r r^2\frac{\partial \ln\rho}{\partial r}.$$
(38)

---

[5] The dynamical frequency-scale $\sqrt{g/h}$ is also present but is not relevant because the perturbation acts upon convective eddies which move with frequency-scale $|N|$ and does so with coupling proportional to that scale.

[6] See Appendix B.

[7] This is done by starting the derivation with equation (7) rather than equation (8) and propagating the effects through to equation (27).

[8] Note that s_θ is $O(\Omega^2)$ and u_θ is at least $O(\Omega)$ so we have dropped their product.

We may drop the term containing $\partial_\theta \ln \rho$ because it is higher-order in Ω than $\partial_r \ln \rho$. So

$$\frac{r}{\sin\theta}\frac{\partial(u_\theta \sin\theta)}{\partial\theta} = -\frac{\partial(r^2 u_r)}{\partial r} - u_r r^2 \frac{\partial \ln\rho}{\partial r}, \quad (39)$$

We can evaluate the derivatives approximately by noting that D_0, and hence u_r, have characteristic scale h as does ρ. This is because all of these owe their spatial dependence primarily to the variation of thermodynamic quantities, which ultimately all vary with the same spatial scale as P. This produces the scaling relation

$$u_\theta \approx u_r \frac{r}{h}, \quad (40)$$

and thence

$$u_\theta \approx \frac{D_0}{h}\alpha'. \quad (41)$$

From simple mixing length theory of convection

$$D_0 \approx \frac{1}{3}hu_c, \quad (42)$$

where u_c is the convective velocity (Gough 1978) Inserting this into equations (36) and (41) we find

$$u_r \approx \frac{1}{3}u_c\frac{h}{r}\alpha' \quad (43)$$

and

$$u_\theta \approx \frac{1}{3}u_c\alpha'. \quad (44)$$

As argued in Appendix A, the actual baroclinicity is reduced by virtue of our including the meridional circulation. In particular, the bulk of the violation of thermal equilibrium implied by equation (15) is resolved by circulation and not by baroclinicity. So in what follows when we discuss baroclinicity we are talking about that which would exist in the absence of a meridional circulation[9]. This is the quantity which sets the magnitude of the circulation.

The flow given by equations (43) and (44) generally acts to damp the flux anisotropy (the no-circulation baroclinicity) because it is driven by this effect. In Appendix D we show that this modifies equation (33) so that

$$\alpha = P^y \int_P^{P_c} \frac{\Omega(\mathcal{P})^2}{|N|(\mathcal{P})^2}\frac{\mathrm{d}\mathcal{P}}{\mathcal{P}^{1+y}}. \quad (45)$$

Note that we work with α rather than directly with the baroclinicity because it is α which evolves according to equation (31). The effect of the damping is then just to change how we weight the average of the perturbation rather than to change the fundamental scalings.

As a final consideration, equations (28) and (29) do not hold for arbitrarily large Ω. In particular, convective anisotropy is known to saturate for Ω large relative to $|N|$ (Lesaffre et al. 2013; Kitchatinov 2013). So we use the prescription

$$\epsilon = \epsilon_{\max}\min\left(1, \frac{\Omega^2}{|N|^2}\right), \quad (46)$$

where $\epsilon_{\max} \approx 0.2$ (Jermyn et al. 2018), and hence find that

$$\alpha \approx P^y \int_P^{P_c} \epsilon_{\max}\min\left(1, \frac{\Omega(\mathcal{P})^2}{|N|(\mathcal{P})^2}\right)\frac{\mathrm{d}\mathcal{P}}{\mathcal{P}^{1+y}}. \quad (47)$$

Furthermore, in the limit of rapid rotation it is no longer the case that there need be a circulation of order the implied baroclinicity, because the baroclinicity alone suffices to restore thermal equilibrium[10]. This does not mean that the flux imbalance is less than equation (47) implies but it does imply that y may be smaller in this limit and that equations (44) and (43) may not hold in rapidly rotating convection zones.

As a test of this model, Fig. 4 shows the radial and latitudinal velocities, (43) and (44), in the solar convection zone compared with the circulation velocity inferred from helioseismic observations (Rajaguru & Antia 2015). The solar model envelope, computed by G. Houdek and D. O. Gough, is the model used by Chaplin et al. (2005) in their study of the power spectral density of solar p-modes. We fitted the parameter y so as to minimise the root mean square error in the logarithm of the velocity components. This produced $y = 0.2$, in agreement with the general magnitude we expect. The variation in the radial component of the flow is well modelled by equation (43) except near the surface, where the sharp density gradient means that inertial effects become increasingly important relative to thermal considerations. It is therefore worth noting that equation (43) is really a lower bound on the radial velocity set by the condition of thermal equilibrium: greater velocities are, of course, permitted. The overall magnitude and trend in the θ component of the flow is reasonably captured by equation (44), but the details are not. In particular, the dips in the observed u_θ are due to cell boundaries in the meridional flow structure, and these geometric features are not captured by our simplified analysis. Furthermore we generally predict velocities which are larger than what is observed in deeper regions. This could be a result of geometric or magnetic effects near the tachocline or else could indicate that we ought to have used a more precise prescription in equation (46).

A further test is provided by the simulations of Brun & Palacios (2009), who find that in three-dimensional simulations of slowly rotating ($\Omega < |N|$) convecting giant stars of order 5 to 10 per-cent of the kinetic energy resides in the meridional circulation, such that the circulation velocity is of order 30 per-cent of the convective velocity. This is in good agreement with equations (43) and (44) when $\Omega > |N|$.

5   EFFECTS IN MASSIVE STARS

In massive stars it is typical to have a convective core and a radiative envelope. If the convection is anisotropic then the core is baroclinic and a meridional flow is present. Here we argue that this produces baroclinicity in the radiative zone, which ultimately drives a circulation current there[11]. We then examine the decay of this circulation into the radiative zone and examine how these effects scale in stellar models.

---

[9] That is, we are using it as a proxy for the anisotropy of the flux distribution.

[10] See Appendix A.

[11] This follows the standard Eddington–Sweet argument once we establish the magnitude of the baroclinicity.
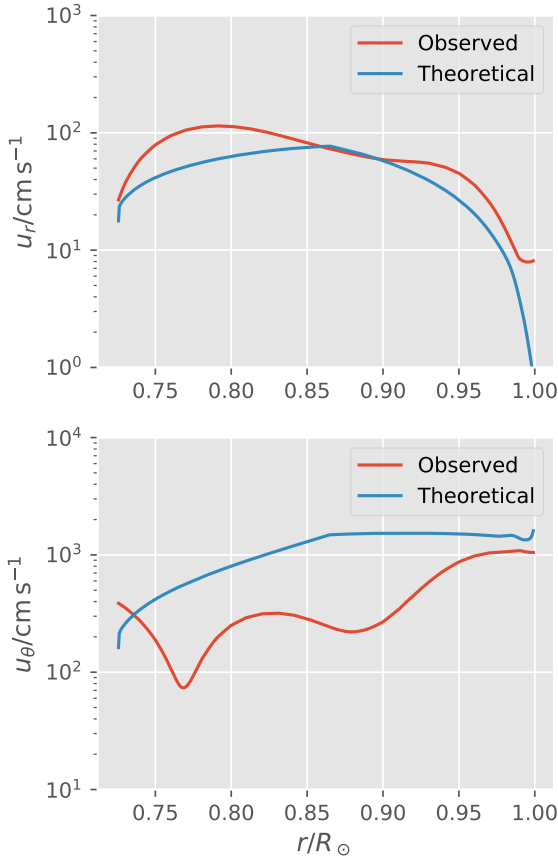
Figure 4. The root mean square meridional velocity taken over latitudinal slices of the data is shown (red) for radial (top) and angular (bottom) components ([Rajaguru & Antia 2015]). In blue are the predicted velocities with $\Omega \approx 2.5 \times 10^{-6}$Hz.

5.1 Transmitting Baroclinicity

To analyse the transmission of baroclinicity we begin by noting that in radiative zones

$$hs_r = h\frac{\partial s}{\partial r} \tag{48}$$

$$= h\left[\frac{\partial \ln P}{\partial r} - \gamma\frac{\partial \ln \rho}{\partial r}\right] \tag{49}$$

$$= h\left[(1+\gamma)\frac{\partial \ln P}{\partial r} + \gamma\frac{\partial \ln T}{\partial r}\right] \tag{50}$$

$$= -(1+\gamma) + \gamma\frac{d\ln T}{d\ln P}, \tag{51}$$

where we have taken the mean molecular weight μ to be constant and made use of equation ([B10]) as well as the ideal gas law $P = \rho k_B T/\mu$. In radiative zones $d\ln T/d\ln P$ is generally small compared to the adiabatic gradient, so $hs_r \approx -(1+\gamma)$ is of order unity. Thus according to equation ([B12])

$$\lambda \approx \alpha. \tag{52}$$

At the tachocline, the boundary of the convection zone, we expect

$$\alpha \approx P_{\text{tachocline}}^y \int_{P_{\text{tachocline}}}^{P_c} \epsilon_{\max} \min\left(1, \frac{\Omega(\mathcal{P})^2}{|N|(\mathcal{P})^2}\right) \frac{d\mathcal{P}}{\mathcal{P}^{1+y}}, \tag{53}$$

because α accumulates perturbations to the entropy gradient[12]. In this way effects which would be small in the radiative zone, which is difficult to perturb because $|N|$ tends to be quite large there, are enhanced because they accumulate in the convection zone, where $|N|$ is small, and are then transmitted to the radiative zone at the tachocline.

To more precisely examine this transmission we write the heat flux as

$$\mathbf{F} = \mathbf{e}_r F_0(r) + A(r)\mathbf{e}_r \cos\theta + B(r)\mathbf{e}_\theta \sin\theta, \tag{54}$$

where

$$F_0 = \frac{L}{4\pi r^2} \tag{55}$$

is the unperturbed flux of the star, L its luminosity and A and B are of order αF_0 at the tachocline. There is no $\cos\theta$ term along $\mathbf{e}_\theta$ because $\mathbf{e}_\theta \cdot \mathbf{F}$ must vanish for $\theta = 0$ and $\theta = \pi$, and we neglect higher-order harmonics for simplicity. Within the radiative zone and in the absence of a meridional flow, the thermal flux obeys equation ([6]), which may be written as

$$\mathbf{F} = -k\nabla T, \tag{56}$$

where the thermal conductivity

$$k \equiv \rho c_p \chi \tag{57}$$

depends solely on the temperature, pressure and opacity. Hence

$$\nabla T = -\frac{1}{k}\left[\mathbf{e}_r F_0(r) + A(r)\mathbf{e}_r \cos\theta + B(r)\mathbf{e}_\theta \sin\theta\right]. \tag{58}$$

It follows that when α is small

$$\frac{\partial_\theta T}{r\partial_r T} \approx \frac{B\sin\theta}{F_0}, \tag{59}$$

which is of order α. Thus the heat flux transmits baroclinicity from the convection zone into the radiative zone to an extent given by equation ([52]).

Note that the *mechanical* transmission of currents across the tachocline has been studied in great detail by [Garaud & Bodenheimer] ([2010]) and in the context of that work these stars are in a limit of efficient transmission, being neither mechanically nor thermally constrained[13]. Hence it is possible that mechanical pumping plays a significant role as well.

5.2 Decay Profile

Within the radiative zone the flux perturbation generally decays. We now aim to determine the scale over which this occurs. In equilibrium and in the absence of any meridional flow or heat generation, the flux obeys the conservation law

$$\nabla \cdot \mathbf{F} = 0 \tag{60}$$

and, by equation ([56]),

$$\nabla \times \mathbf{F} = -k\nabla \times \nabla T - \nabla k \times \nabla T = \nabla \ln k \times \mathbf{F}. \tag{61}$$

---

[12] It is important to note that it is α, not α', which matters in radiative zones. This is because the additional factor of $\Omega^2/|N|^2$ which arises through convective anisotropy vanishes in radiative zones.

[13] See Appendix A for further details.

From

$$\nabla \cdot \boldsymbol{F} = 0 \tag{62}$$

and equation (54) we obtain

$$\frac{\partial A}{\partial r} + \frac{2}{r}(A + B) = 0 \tag{63}$$

and from equation (61)

$$\frac{\partial B}{\partial r} + \frac{1}{r}(B - A) = \frac{\partial \ln k}{\partial r} B - \frac{2}{\pi} \int_0^\pi \frac{\partial \ln k}{\partial \theta} \left[F_0 + \frac{1}{r} A(r) \cos \theta \right] \sin \theta d\theta \tag{64}$$

$$= \frac{\partial \ln k}{\partial r} B - \frac{2F_0}{\pi r} \int_0^\pi \frac{\partial \ln k}{\partial \theta} \sin \theta d\theta. \tag{65}$$

Because we have assumed a chemically homogeneous star we can relate k in the limit of small α to the derivatives of k in temperature and pressure. That is,

$$\frac{\partial \ln k}{\partial \theta} \approx \frac{\partial \ln k}{\partial \ln T} \frac{\partial \ln T}{\partial \theta} + \frac{\partial \ln k}{\partial \ln P} \frac{\partial \ln P}{\partial \theta}, \tag{66}$$

where the thermodynamic derivatives of k are taken with respect to the unperturbed state and so are independent of θ. The pressure only acquires a dependence on θ through the centrifugal force, so

$$\frac{\partial \ln P}{\partial \theta} \approx \frac{\Omega^2 R}{g} \approx \frac{\Omega^2 r}{g} \sin \theta. \tag{67}$$

The temperature depends on θ through equation (58) so

$$\frac{\partial \ln T}{\partial \theta} = -\frac{rB(r)}{kT} \sin \theta. \tag{68}$$

Inserting equations (66), (67) and (68) into equation (73) we find

$$\frac{\partial B}{\partial r} + \frac{1}{r}(B - A) = \frac{\partial \ln k}{\partial r} B - \frac{F_0}{r}\left(-\frac{rB}{kT} \frac{\partial \ln k}{\partial \ln T} + \frac{\Omega^2 r}{g} \frac{\partial \ln k}{\partial \ln P} \right). \tag{69}$$

From equation (58) we find

$$\frac{F_0}{kT} \approx |\nabla \ln T| \approx -\frac{d \ln P}{dr} \frac{d \ln T}{d \ln P} = \frac{1}{h} \frac{d \ln T}{d \ln P}. \tag{70}$$

So

$$\frac{\partial B}{\partial r} + \frac{1}{r}(B - A) = \left(\frac{\partial \ln k}{\partial r} - \frac{1}{h} \frac{d \ln T}{d \ln P} \frac{\partial \ln k}{\partial \ln T} \right) B - F_0 \frac{\Omega^2 r}{g} \frac{\partial \ln k}{\partial \ln P}. \tag{71}$$

The final term in this equation is important near the surface, where it is necessary to reproduce the usual Eddington–Sweet circulation. However deeper in the star it may be neglected because there $B/F_0 \approx \alpha \gg \Omega^2 r/g$ and logarithmic derivatives of k with respect to each of pressure and temperature are of order unity. Hence we find

$$\frac{\partial B}{\partial r} + \frac{1}{r}(B - A) = \left(\frac{\partial \ln k}{\partial r} - \frac{1}{h} \frac{d \ln T}{d \ln P} \frac{\partial \ln k}{\partial \ln T} \right) B, \tag{72}$$

which may be written as

$$\frac{\partial B}{\partial r} + \frac{1}{r}(B - A) \approx -\frac{b}{h} B, \tag{73}$$

where

$$b \equiv \frac{d \ln k}{d \ln P} + \frac{d \ln T}{d \ln P} \frac{\partial \ln k}{\partial \ln T}. \tag{74}$$

Combining equations (63) and (73) we find

$$\frac{\partial}{\partial r} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} -\frac{2}{r} & -\frac{2}{r} \\ +\frac{1}{r} & -\frac{1}{r} - \frac{b}{h} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}. \tag{75}$$

Treating b and r as constants, the eigenvalues of this system are

$$\lambda_\pm = -\frac{1}{r} \left[\frac{3}{2} + \frac{br}{2h} \pm \frac{1}{2} \sqrt{\frac{b^2 r^2}{h^2} - 2b\frac{r}{h} - 7} \right]. \tag{76}$$

In and around the cores of stars r is typically smaller than h, and b is of order unity, so for simplicity we neglect terms involving b and find that the slowest-decaying mode is

$$\lambda_- \approx -\frac{3}{2r} - i\frac{\sqrt{7}}{2r}, \tag{77}$$

so

$$\Re(\lambda) \approx -\frac{3}{2r}. \tag{78}$$

Taking only this mode, because it is the one that persists for the largest distance, we find that

$$\frac{d \ln A}{d \ln r} = \frac{d \ln B}{d \ln r} = \Re(\lambda). \tag{79}$$

Hence

$$\frac{A(r)}{A(r_c)} \approx \frac{B(r)}{B(r_c)} \approx \left(\frac{P}{P_{\text{tachocline}}} \right)^\beta, \tag{80}$$

where r_c is the radius of the convective core and

$$\beta \equiv -h\Re(\lambda_-) \approx \frac{3h}{2r}. \tag{81}$$

Thus the baroclinicity is of order

$$\alpha \approx \alpha(r_c) \left(\frac{P}{P_{\text{tachocline}}} \right)^\beta. \tag{82}$$

This generally prevents solutions which do not involve $O(\alpha)$ meridional flows for the same reason as outlined in section 2.

It is worth noting that equation (77) indicates that the modes have an oscillatory component with characteristic scale $2r/7$. This suggests that the cell structure of the flow varies on that radial scale, though the level of detail in our calculations is not sufficient to determine more detail of how this manifests in the circulation pattern.

The standard Eddington–Sweet derivation for baroclinic radiative zones finds that the radial circulation velocity is of order that required to balance the flux anisotropy. That is,

$$\rho c_p T u_r s_r \approx \nabla \cdot F \tag{83}$$

(Eddington 1929). Noting that hs_r is of order unity we find

$$u_r \approx \frac{h\nabla \cdot F}{\rho c_p T}. \tag{84}$$

The length-scale associated with the flux anisotropy is r, because that is the latitudinal scale, and the flux scale of the anisotropy is set by αF, so

$$u_r \approx \alpha \frac{hF}{r \rho c_p T}. \tag{85}$$

Using equation (40) we find

$$u \approx \alpha \frac{F}{\rho c_p T}. \tag{86}$$

In an efficient convection zone the luminosity equals the convective luminosity. This is well approximated by the power flux of a fluid moving with the convection speed, so

$$L \approx 4\pi r^2 \rho u_c^3,$$ (87)

where u_c is the convection speed, the heat flux outside of the core may be related to that at the boundary of the core by

$$F \approx \rho_c u_{c,core}^3 \left(\frac{r_c}{r}\right)^2.$$ (88)

and so

$$\frac{u_r}{u_{c,core}} \approx \left(\frac{\rho_c}{\rho}\right)\left(\frac{u_{c,core}}{c_s}\right)^2 \left(\frac{r_c}{r}\right)^2 \left(\frac{P}{P_{tachocline}}\right)^\beta \alpha(r_c),$$ (89)

where c_s is the sound speed and

$$\alpha(r_c) = P_{tachocline}^y \int_{P_{tachocline}}^{P_c} \epsilon_{max} \min\left(1, \frac{\Omega(\mathcal{P})^2}{|N|(\mathcal{P})^2}\right) \frac{d\mathcal{P}}{\mathcal{P}^{1+y}}.$$ (90)

It is worth noting that the ratio $P/P_{tachocline}$ scales exponentially in the radial coordinate while the flow speed scales as a power law in this ratio. The latter scaling persists even after accounting for the additional heat flux transported by the meridional flow. To see this note that the rate at which α changes with r is proportional to the flux transported latitudinally by the meridional flow because this flow damps α. This flux scales as α because the flow transports material radially at a rate which scales as α and the radial gradient is largely independent of α. This means that $d\alpha/dr \approx -(h/r^2)\alpha$, which produces a power-law in P, which merely modifies the exponent β. The modification is approximately

$$\beta \to \beta + \frac{h}{r},$$ (91)

applied to equation (81), so that

$$\beta \approx \frac{5h}{2r}.$$ (92)

As one final simplification, it is worth noting that when the exponent β is unity equation (86) may be written as

$$u_r \approx \left(\frac{F}{P_c}\right)\alpha_{core}.$$ (93)

This form has the disadvantage of not generalising should the exponent differ from one and of lacking the clear dimensionless ratios of equation (89), but has the advantage that the origin of the flow is clear: rotationally-driven ansiotropy forces a meridional flow to carry a portion of the flux, with the speed set by the energy density P.

5.3   Scaling in Stellar Models

To understand these effects quantitatively we define

$$\tau \equiv \frac{h}{u_r}$$ (94)

as the analogue of the local Eddington–Sweet time. Noting that h is approximately the distance $z = R - r$ to the surface of the star we see that τ is an estimate of the time needed to bring material to the surface. Fig. 5 shows this time as a function of mass coordinate for three different stellar models made with the Cambridge STARS code (Eggleton 1971; Pols et al. 1995). These calculations were made with a surface rotation of $u_{rot} = 3 \times 10^6 \, \mathrm{cm \, s^{-1}}$ and we assume that the
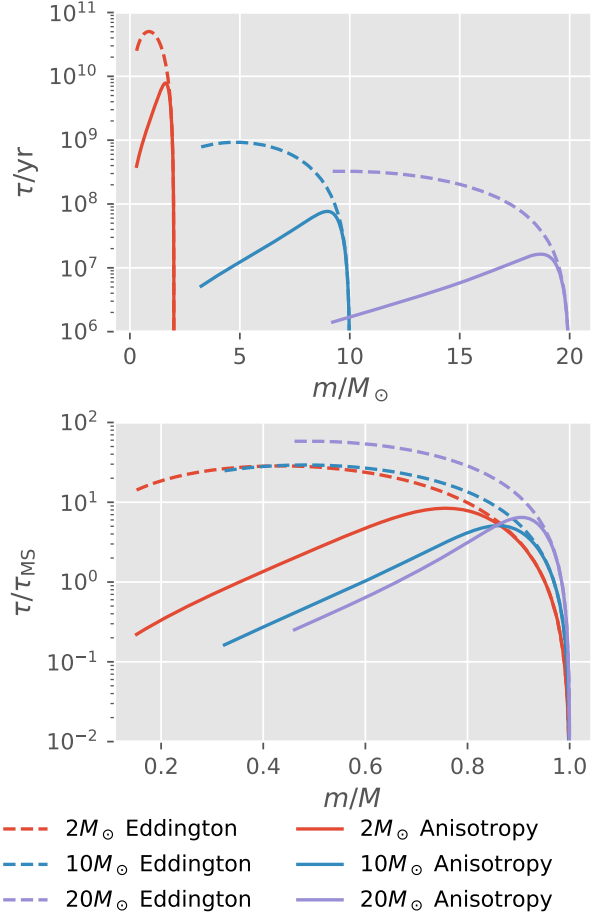
Figure 5. The analogue of the Eddington–Sweet time $\tau = z/u_r$ is shown as a function of mass m for three stellar models on the main sequence with masses of $2M_\odot$, $10M_\odot$ and $20M_\odot$. For each model results from both the Eddington–Sweet mechanism and the mechanism presented in this paper are shown. These calculations were made with $\beta = 3/2$ and rigid rotation with a surface velocity of $u_{rot} = 3 \times 10^6 \, \mathrm{cm \, s^{-1}}$. The bottom panel shows the same timescales τ normalised to the main-sequence lifetime τ_{MS}, with the mass normalised to the total mass of the star M.

mixing rate cannot fall below that of the Eddington circulation. For simplicity we took $\beta = 3/2$ though in practice this parameter varies with r.

A few implications are clear from this figure. First, rotating massive stars mix near the core much more rapidly than would normally be expected without accounting for convective anisotropy, which increases the effective core mass and acts very much like convective overshooting (Prather & Demarque 1974), a topic we explore in more detail in Section 7. This enhancement arises because $|N|$ in the core is generally much smaller than $\sqrt{g/R}$, and so the convective boundary condition drives baroclinicity much more than the simple Eddington–Sweet mechanism. Secondly, this effect damps strongly as we approach the surface and is eventually overtaken by the Eddington circulation, so there would not necessarily be strong observable chemical signatures of the enhanced mixing except by virtue of

making material more readily available to other mixing processes.

To understand how effectively this mechanism transports material through the star we now examine the travel time from the tachocline to a point at radius r, given by

$$\tau' \equiv \int_{r_c}^{r} \frac{dr}{u_r}. \tag{95}$$

This is shown in Figure 6. In each of the stellar models in Fig. 5 τ' is less than the lifetime of the star almost everywhere and so, at least at this rotation rate, a significant amount of material ought to reach the core. By contrast, the time

$$\tau'' \equiv -\int_{r_{\mathrm{surface}}}^{r} \frac{dr}{u_r}. \tag{96}$$

to mix to the surface exceeds the main-sequence lifetime in most of the star. Hence at least at this rotation rate the amount of material mixed from the core to the surface from this region ought to be minimal. For more rapidly rotating stars this effect is stronger, principally because the Eddington–Sweet circulation is stronger, and considerably more material may reach the surface from the core.

A similar story may be seen in Fig. 7, which shows the radial velocity of the circulation as a function of position in the star. The results are for the same stellar models as in Fig. 5 and, once more, the Eddington curves only contain the Eddington–Sweet circulation while the Anisotropy curves contain both effects. Because the anisotropy-induced circulation declines towards the surface it is clear from this where the Eddington–Sweet mechanism becomes dominant. In the top panel, which was made with a surface velocity of $u_{\mathrm{rot}} = 30\,\mathrm{km\,s^{-1}}$, this point is typically around a mass fraction of 0.9. By contrast in the bottom panel, which was made with a surface velocity of $u_{\mathrm{rot}} = 300\,\mathrm{km\,s^{-1}}$, the transition occurs between 0.5 and 0.7. This is because the anisotropy-driven circulation saturates at much lower rotation rates and so does not scale with rotation in this regime while the Eddington–Sweet mechanism scales quadratically with rotation rate.

It is worth noting that a similar effect is possible in stars of lower mass, where the anisotropy in the outer convection zone induces baroclinicity in the radiative core. This effect is more strongly damped because the characteristic damping scale is h, which is much shorter at an outer radiative–convective boundary than at a core convective–radiative boundary. Equivalently, the damping goes as $P_{\mathrm{tachocline}}/P$ in this case, and $P_{\mathrm{tachocline}}$ may be quite small relative to the core pressure. In addition the velocity damps towards the core because the temperature increases, making it easier to dissipate the flux accumulation. Finally, the anisotropy for a given rotation is smaller because the relevant $|N|$ is greater. Still, in the Sun $h \approx R_\odot/10$ near the tachocline, so this likely causes some mixing between the convection zone and the material a few tenths of a solar radius below it.

6 CHEMICAL COMPOSITION GRADIENTS

In the convection zones of massive stars chemical composition gradients are generally wiped out by turbulent mixing. So there are significant composition gradients only in the radiative zones of these bodies.
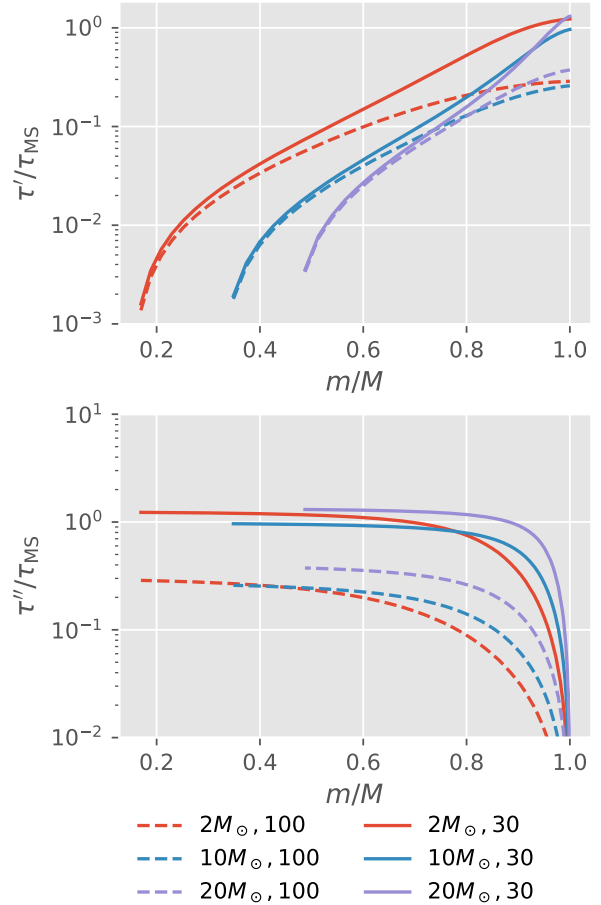

Figure 6. The time τ' required for material to flow to the core (top) and the time τ'' required for material to flow to the surface (bottom) are shown as functions of mass fraction m/M for three stellar models on the main sequence with masses of $2M_\odot$, $10M_\odot$ and $20M_\odot$. Both times are normalised to the main-sequence lifetime τ_{MS}. For each model the velocity u_r is assumed to be the sum of that owing to the Eddington–Sweet mechanism and that owing to the mechanism presented in this paper. These calculations were made with $\beta = 3/2$ and rigid rotation with a surface velocities of $u_{\mathrm{rot}} = 3 \times 10^6\,\mathrm{cm\,s^{-1}}$ (solid) and $u_{\mathrm{rot}} = 10^7\,\mathrm{cm\,s^{-1}}$ (dashed). The former correspond respectively to 4×10^{-3}, 2×10^{-3} and 10^{-3} of the surface breakup rate for the three stars, while the latter correspond to 4×10^{-2}, 2×10^{-2} and 10^{-2} of the breakup rate respectively.

Because the circulation in radiative zones is driven by the same criterion of thermal equilibrium as the Eddington–Sweet circulation such gradients act on it in the same manner as they do that mechanism. The only place where there might be a difference in how composition gradients impact the mechanism we have introduced is in the transmission of baroclinicity from the convection zone to the radiative zone. However this transmission is achieved by the core emitting an aspherical heat flux and this is not at all changed by the chemistry in the vicinity of the core. Hence the effect of composition ought to be the same in both mechanisms. This has been examined in detail by Meynet & Maeder (2000), who provide a prescription for correcting the circulation velocity
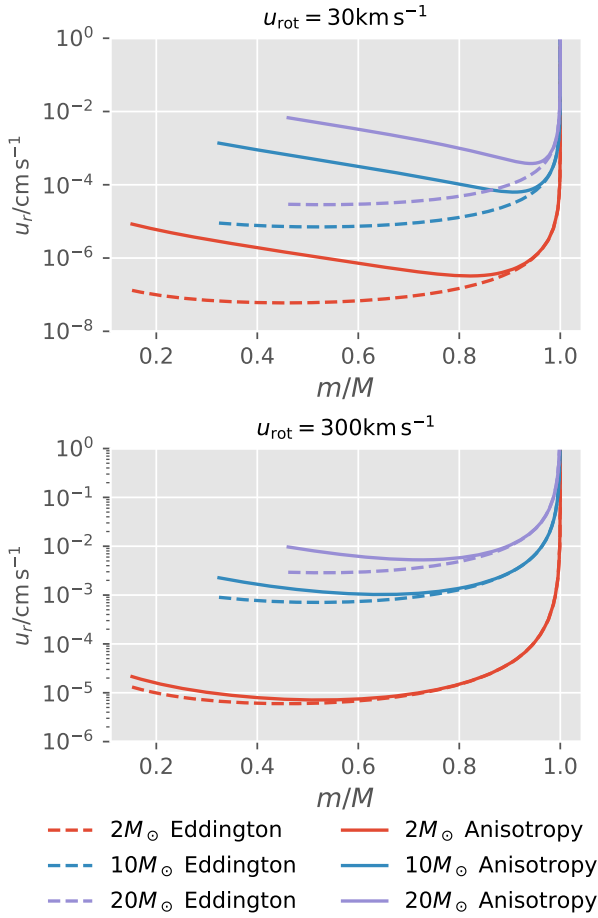
Figure 7. The radial velocity u_r is shown as a function of mass fraction m/M for three stellar models on the main sequence with masses of $2M_\odot$, $10M_\odot$ and $20M_\odot$. For each model results from both the Eddington–Sweet mechanism and the mechanism presented in this paper are shown. These calculations were made with $\beta = 3/2$ and for simplicity assume rigid rotation. The top panel has surface velocity of $u_{\rm rot} = 3 \times 10^6 \, \rm cm \, s^{-1}$, corresponding respectively to 4×10^{-3}, 2×10^{-3} and 10^{-3} of the surface breakup velocity for the three stars. The bottom panel has $u_{\rm rot} = 3 \times 10^7 \, \rm cm \, s^{-1}$, corresponding respectively to 0.4, 0.2 and 0.1 of the breakup velocity.

and find that such corrections are typically of order unity. This is therefore what we expect.

7 CONVECTIVE OVERSHOOTING

The fact that these circulations drive extra mixing in radiative zones close to convective cores suggests that this may be the source of extra mixing which is commonly, though probably erroneously, called convective overshooting or penetration (Prather & Demarque 1974). Such extra mixing naturally prolongs the main-sequence lifetime as required to match a number of evolved binary systems (Schröder et al. 1997; Farmer et al. 2015; Eggleton & Yakut 2016) and so is of significant interest for the purposes of stellar modelling.

Convective overshooting is often incorporated in stellar models by modelling a region beyond the convection zone as

isentropic and well-mixed. This region has width

$$l_{\rm ov} = \alpha_{\rm ov} h, \tag{97}$$

where $\alpha_{\rm ov}$ is a dimensionless parameter[14] (Schröder et al. 1997). A long-standing problem with this prescription is that the overshoot distance $\alpha_{\rm ov} h$ required to match observations is often much larger than what calculations of the stiffness of the radiative-convective boundary suggest (Saslaw & Schwarzschild 1965). What appears more likely is that the physical overshooting is small as predicted by stiffness calculations but that there is additional mixing caused by the baroclinic mechanism developed in the previous section.

The most immediate comparison with observations that can be made to test this hypothesis is to calculate $\alpha_{\rm ov}$ from our model. To do so we calculate $l_{\rm ov}$ such that in the lifetime of the star the material within this distance of the tachocline is well-mixed with that at the tachocline. This is just the statement that

$$\frac{l_{\rm ov}}{u_r} \approx \tau_{\rm MS}, \tag{98}$$

or

$$l_{\rm ov} \approx u_r \tau_{\rm MS}, \tag{99}$$

where $\tau_{\rm MS}$ is the main-sequence lifetime of the star. This is not quite right, however, because vertical chemical mixing is less efficient than the circulation velocity alone suggests. As suggested by Maeder & Zahn (1998) the effective mixing is reduced by a factor of

$$w \equiv \frac{r u_r}{30 D_{\rm h}}, \tag{100}$$

where $D_{\rm h}$ is the horizontal diffusivity. From Appendix A we have

$$D_{\rm h} \approx \nu \approx \frac{T}{\rho S}, \tag{101}$$

where ν is the turbulent viscosity, S is the shear owing to differential rotation and T is the magnitude of the turbulent stress. With equation (A23) we find

$$D_{\rm h} \approx \frac{u h \Omega}{S}. \tag{102}$$

Hence

$$w \approx \frac{r u_r}{30 u h} \left(\frac{S}{\Omega} \right). \tag{103}$$

From equation (40) we find

$$w \approx \frac{1}{30} \left(\frac{S}{\Omega} \right). \tag{104}$$

In Appendix A we found $S \gtrsim \Omega$. For simplicity, and as a reflection of the uncertainties involved in determining S, we use

$$w \approx \frac{1}{30}. \tag{105}$$

Hence equation (99) becomes

$$l_{\rm ov} \approx \frac{1}{30} u \tau_{\rm MS}. \tag{106}$$

---

[14] Confusingly, this is also called $f_{\rm ov}$ by some authors.

Inserting equation (89) into equation (106) we find

$$l_{\rm ov} \approx \frac{u_{\rm c,core}}{30} \tau_{\rm MS} \left(\frac{\rho_{\rm c}}{\rho}\right) \left(\frac{u_{\rm c,core}}{c_{\rm s}}\right)^2 \left(\frac{r_{\rm c}}{r}\right)^2 \left(\frac{P}{P_{\rm tachocline}}\right)^\beta \alpha_{\rm core}. \tag{107}$$

In massive stars with even mild rotation, $\Omega > |N|$ everywhere in the core so

$$\alpha_{\rm core} = P_{\rm tachocline} \int_{P_{\rm tachocline}}^{P_{\rm c}} \min\left(\epsilon_{\rm max}, \frac{\Omega(\mathcal{P})^2}{|N|(\mathcal{P})^2}\right) \frac{d\mathcal{P}}{\mathcal{P}^2} \tag{108}$$

$$= \epsilon_{\rm max} P_{\rm tachocline} \int_{P_{\rm tachocline}}^{P_{\rm c}} \frac{d\mathcal{P}}{\mathcal{P}^2} \tag{109}$$

$$\approx \epsilon_{\rm max}, \tag{110}$$

because $P_{\rm tachocline} \ll P_{\rm c}$. Hence

$$l_{\rm ov} \approx \frac{\epsilon_{\rm max}}{30} u_{\rm c,core} \tau_{\rm MS} \left(\frac{\rho_{\rm c}}{\rho}\right) \left(\frac{u_{\rm c,core}}{c_{\rm s}}\right)^2 \left(\frac{r_{\rm c}}{r}\right)^2 \left(\frac{P}{P_{\rm tachocline}}\right)^\beta \tag{111}$$

$$\approx \frac{\epsilon_{\rm max}}{30} u_{\rm c,core} \tau_{\rm MS} \left(\frac{u_{\rm c,core}}{c_{\rm s,core}}\right)^2 \left(\frac{r_{\rm c}}{r}\right)^2 \left(\frac{P}{P_{\rm tachocline}}\right)^{\beta-1}, \tag{112}$$

where we have used the fact that $P \propto \rho c_{\rm s}^2$. Noting that near the tachocline

$$r \approx r_{\rm c} \tag{113}$$

and

$$P \approx P_{\rm tachocline} e^{-\delta r/h}, \tag{114}$$

where δr is the distance to the tachocline, we find that

$$l_{\rm ov} \approx \frac{\epsilon_{\rm max}}{30} u_{\rm c,core} \tau_{\rm MS} \left(\frac{u_{\rm c,core}}{c_{\rm s,core}}\right)^2 e^{-l_{\rm ov}(\beta-1)/h}. \tag{115}$$

Inserting equation (97) we find

$$\alpha_{\rm ov} \approx \frac{\epsilon_{\rm max}}{30} \frac{u_{\rm c,core}}{h} \tau_{\rm MS} \left(\frac{u_{\rm c,core}}{c_{\rm s,core}}\right)^2 e^{-\alpha_{\rm ov}(\beta-1)} \tag{116}$$

and making use of equation (92)

$$\alpha_{\rm ov} \approx \frac{\epsilon_{\rm max}}{30} \frac{u_{\rm c,core}}{h} \tau_{\rm MS} \left(\frac{u_{\rm c,core}}{c_{\rm s,core}}\right)^2 e^{-\alpha_{\rm ov}(5h/2r-1)}. \tag{117}$$

This may be solved in terms of the Lambert W function or else numerically. Fig. 8 shows $\alpha_{\rm ov}$ as a function of stellar mass for a fine grid of stellar models made with the Cambridge STARS code (Eggleton 1971; Pols et al. 1995). We have taken $\epsilon_{\rm max}$ to be 0.2 as suggested by various turbulence closure schemes (Kitchatinov 2013; Jermyn et al. 2018). The stellar lifetime was computed with the main-sequence fit of Eggleton et al. (1989). The variation of $\alpha_{\rm ov}$ with mass suggests that it rises rapidly in the window between 1 and $2\,\mathrm{M}_\odot$ and then asymptotically, shortly thereafter, $\alpha_{\rm ov} \approx 0.25$. The variation is primarily driven by variation in $h/r_{\rm c}$ because the convective Mach number varies little with mass. This behaviour is in very good agreement with the findings of Ribas et al. (2000), Moravveji et al. (2016) and Stancliffe et al. (2015), while yielding a slight overestimate relative to those of Claret & Torres (2017), who find an asymptote of $\alpha_{\rm ov} \approx 0.2$.

Note that even very slowly rotating stars have $\Omega > |N|$, as $|N| \lesssim 10^{-8}\mathrm{s}^{-1}$ in core convection zones. For instance KIC 10526294, which rotates with a period of 188d and hence has $\Omega \approx 4 \times 10^{-7}\mathrm{s}^{-1}$, is a sufficiently rapid rotator that we
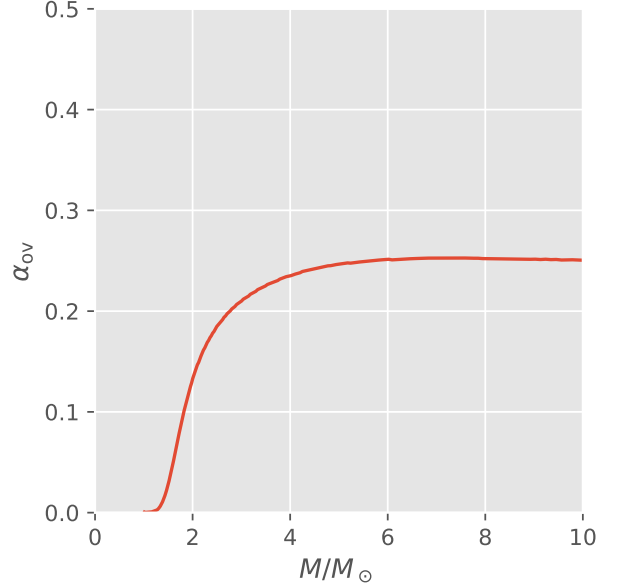

Figure 8. The convective overshoot parameter $\alpha_{\rm ov}$ is shown as a function of stellar mass, computed with equation (117) and $\epsilon_{\rm max} \approx 0.2$ as various turbulence closure schemes suggest (Kitchatinov 2013; Jermyn et al. 2018).

expect it to follow equation (117) and indeed this is approximately what is seen (Moravveji et al. 2015). For stars with rotation periods in excess of thousands of days it might be possible to observe the quadratic dependence of circulation Ω that we predict, though such objects are quite rare.

An alternate way to parametrise convective overshooting is as an additional diffusivity in the radiative zone of the form

$$D_{\rm overshoot} = D_{\rm convective} e^{-2\delta r/(f_{\rm ov}h)} \tag{118}$$

(Herwig et al. 1997; Paxton et al. 2011), where $f_{\rm ov}$ is the exponential overshoot parameter, δr is the distance to the radiative-convective boundary and

$$D_{\rm convective} \equiv u_{\rm c,core} h. \tag{119}$$

There are two ways to compare this prescription with our calculations. First, we may note from a practical perspective that the $f_{\rm ov}$ and $\alpha_{\rm ov}$ prescriptions agree with one another, and so the fact that our calculated $\alpha_{\rm ov}$ are consistent with those inferred from observations implies consistency with the $f_{\rm ov}$ prescription. Secondly, we can calculate an effective $f_{\rm ov}$ from our model. Because we predict different spatial behaviour from that of equation (118) there is no unique way to do this and different prescriptions yield different results. Indeed to self-consistently map from mixing via circulation currents to an exponential convective overshoot model requires a detailed comparison of stellar models with each prescription which we leave for future work. Nevertheless one straightforward way to perform this mapping is to insist that one scale-height from the convective-radiative boundary, the effective diffusivity owing to the meridional circulation

$$D_{\rm merid} \approx uh \tag{120}$$

equals the overshoot diffusivity. Following the same proce-

dure as before to calculate u and incorporating the correction to chemical mixing we find

$$D_{\mathrm{merid}} \approx \frac{\epsilon_{\max}}{30} u_{\mathrm{c,core}} h \left(\frac{u_{\mathrm{c,core}}}{c_{\mathrm{s,core}}}\right)^2 \left(\frac{r_{\mathrm{c}}}{r}\right)^2 \left(\frac{P}{P_{\mathrm{tachocline}}}\right)^{\beta-1}. \quad (121)$$

Letting $D_{\mathrm{merid}} = D_{\mathrm{overshoot}}$ and inserting equation (119) we obtain

$$e^{-2\delta r/(f_{\mathrm{ov}} h)} \approx \frac{\epsilon_{\max}}{30} \left(\frac{u_{\mathrm{c,core}}}{c_{\mathrm{s,core}}}\right)^2 \left(\frac{r_{\mathrm{c}}}{r}\right)^2 \left(\frac{P}{P_{\mathrm{tachocline}}}\right)^{\beta-1}. \quad (122)$$

Noting that near the tachocline $r \approx r_{\mathrm{c}}$ and

$$P \approx P_{\mathrm{tachocline}} e^{-\delta r/h}, \quad (123)$$

we find

$$e^{-2\delta r/(f_{\mathrm{ov}} h)} \approx \frac{\epsilon_{\max}}{30} e^{-(\beta-1)\delta r/h} \left(\frac{u_{\mathrm{c,core}}}{c_{\mathrm{s,core}}}\right)^2. \quad (124)$$

Hence

$$\left[\frac{2}{f_{\mathrm{ov}}} - (\beta-1)\right] \frac{\delta r}{h} = \ln\frac{30}{\epsilon_{\max}} - 2\ln\frac{u_{\mathrm{c,core}}}{c_{\mathrm{s,core}}}. \quad (125)$$

Letting $\delta r = h$ we find

$$\frac{2}{f_{\mathrm{ov}}} - (\beta-1) = \ln\frac{30}{\epsilon_{\max}} - 2\ln\frac{u_{\mathrm{c,core}}}{c_{\mathrm{s,core}}}. \quad (126)$$

Once more inserting equation (92) and taking $r \approx r_{\mathrm{c}}$ we find

$$\frac{2}{f_{\mathrm{ov}}} - \left(\frac{5h}{2r_{\mathrm{c}}} - 1\right) = \ln\frac{30}{\epsilon_{\max}} - 2\ln\frac{u_{\mathrm{c,core}}}{c_{\mathrm{s,core}}}. \quad (127)$$

Finally solving for f_{ov} yields

$$f_{\mathrm{ov}} = \frac{2}{\ln\frac{30}{\epsilon_{\max}} - 2\ln\frac{u_{\mathrm{c,core}}}{c_{\mathrm{s,core}}} + \left(\frac{5h}{2r_{\mathrm{c}}} - 1\right)}. \quad (128)$$

Fig. 9 shows f_{ov} as a function of stellar mass with the same stellar models and parameters as were used for Fig. 8. The results qualitatively capture the inferred dependence on stellar mass, but we overestimate f_{ov} by a factor of approximately 2 (Stancliffe et al. 2015; Moravveji et al. 2016). This is not surprising given the ad-hoc nature of our matching between equations (118) and (120) but it is encouraging that the overall magnitude is approximately correct and that the dependence on mass reproduces what is observed.

As a check of these calculations, Fig. 10 shows the spatial dependence of both the circulation diffusivity and the convective overshoot diffusivity for the three stellar models considered for Fig. 5. The overshoot diffusivity was calculated with the exponential parametrisation and $f_{\mathrm{ov}} = 0.09$, which is representative from Fig. 9. Note that, in each case, the crossover occurs roughly one scale-height from the edge of the convective core, in keeping with our matching to the exponential parametrisation. This both confirms the results of equation (127) and Fig. 9 and demonstrates the nature of the matching procedure we have used to infer f_{ov}. It also highlights that the penetrative convective overshoot, if present, is by far the dominant process near the core, and it is only further from the core that the meridional circulation becomes more important. However for the purposes of material mixing even the much slower meridional circulation is sufficient to chemically homogenise the system on evolutionary time-scales and so either could be responsible for the inferred chemical mixing near the tachocline.
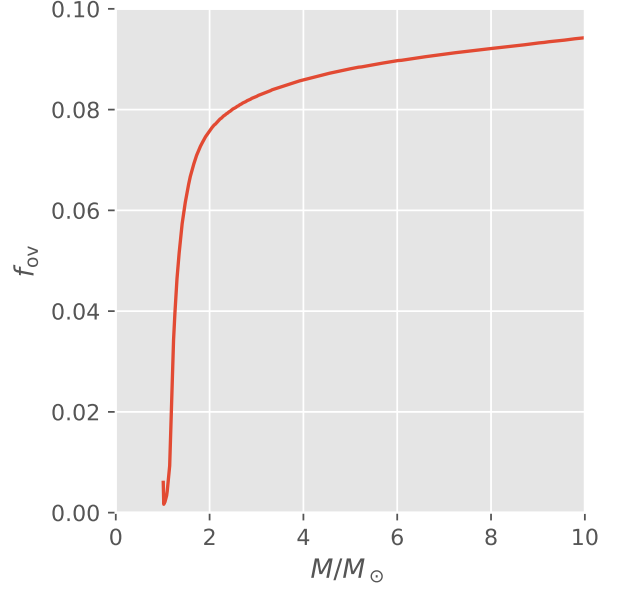
Figure 9. The convective overshoot parameter f_{ov} is shown as a function of stellar mass, computed with equation (127). The input stellar models were the same as those in Figure 8.

8 IMPLEMENTATION IN STELLAR MODELS

While we have not yet implemented these enhanced circulation currents in stellar models, there are a variety of ways in which this could be done. As a rough approximation one could set the convective overshoot parameters according to equations (117) and (127). This is only an approximation because the circulation currents exhibit a different spatial dependence from that of convective overshooting, but it is straightforward to do.

A more complete accounting of this mechanism would require tracking the parameter α' in convection zones as well as α in radiative zones. This may be done either by directly computing these parameters according to the prescriptions in section 3 or by expanding the basic equations of stellar structure to incorporate the first several spherical harmonics. Of particular importance are the harmonics in the heat flux $\boldsymbol{F}$ as well as those in density. In addition a turbulence closure model is needed to calculate $\mathbf{D}$ in the convection zone, because this mechanism ultimately relies on convective anisotropy.

9 INTERPRETATION

There are two further matters which are worth emphasising. First, none of what we argue here is in contradiction of the Eddington–Sweet circulation calculations: we are simply extending them to account for a baroclinic boundary condition. Secondly, the origin of the velocity field we posit lies in turbulence and so the objection that an inviscid fluid cannot support such circulation currents (see e.g. Busse 1981) is simply not applicable. As we have shown in Appendix A this objection is not even relevant in radiative zones. Thus
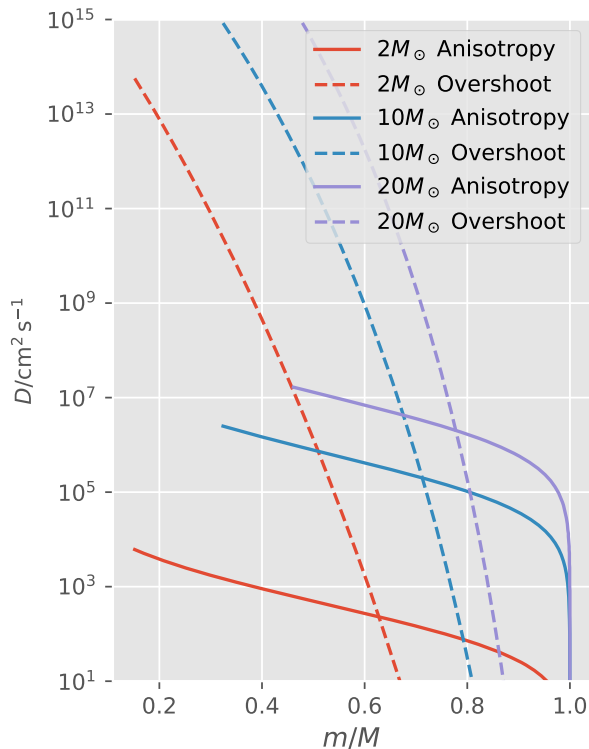
Figure 10. The effective circulation chemical diffusivity $D = hu_r/30$ and convective overshoot diffusivity are shown as functions of mass fraction m/M for three stellar models on the main sequence with masses of $2M_\odot$, $10M_\odot$ and $20M_\odot$. The circulation diffusivity was calculated with $\beta = 3/2$ and only includes that owing to convective anisotropy and not the Eddington–Sweet mechanism. The overshoot diffusivity was modelled with an exponential with $f_{ov} = 0.09$. These calculations assume rigid rotation and a surface velocity of $u_{rot} = 30\,\mathrm{km\,s^{-1}}$.

we arrive at the conclusion that such circulations are physical.

However there may be magnetic barriers to mixing. There is strong differential rotation near the solar tachocline so it seems likely that there are also strong toroidal magnetic fields present in this region. These fields may work to inhibit transverse motion. At the same time though they serve to increase the turbulent anisotropy, which may partially act to counter this inhibiting effect. The net effect of magnetic fields in determining the solar meridional flow cannot be too significant because the predicted meridional velocities match observations even in the tachocline. This suggests that our non-magnetic scaling analysis may suffice, though we cannot say for certain that magnetic fields are generally irrelevant.

10   CONCLUSIONS

We have demonstrated that the meridional flow rate in the radiative zones of massive rotating stars in many cases is set by anisotropy in the central convection zone. This enhances mixing in massive stars as well as core-envelope angular momentum coupling in all stars with a radiative zone overlying a convection zone. Importantly, the additional mixing

is approximately independent of the rotation rate above a relatively low threshold, and so ought to be distinguishable from other sources of mixing.

These effects may aid the formation of massive binary black hole pairs (Marchant, Pablo et al. 2016) through enhanced mixing at long periods. More generally, such mixing would lead to stronger core-envelope mixing thus leading to massive stars growing larger burnt-out cores and living longer owing to mixing of fuel into the core. This mechanism mimics convective overshooting, and we have argued that it suffices to explain the anomalous overshoot distances which have been inferred from observations. In addition, such baroclinicity-driven mixing is likely to occur in stars with outer convection zones and inner radiative zones, though with more limited effect. This represents a fundamentally new phenomenon in stellar mixing and we expect it to have wide-reaching consequences for stellar evolution.

REFERENCES

Abbott B. P., et al., 2016, Phys. Rev. Lett., 116, 061102
Balbus S. A., Latter H., Weiss N., 2012, MNRAS, 420, 2457
Brun A. S., Palacios A., 2009, The Astrophysical Journal, 702, 1078
Busse F. H., 1981, Geophysical and Astrophysical Fluid Dynamics, 17, 215
Caleo A., Balbus S. A., 2016, Monthly Notices of the Royal Astronomical Society, 457, 1711
Caleo A., Balbus S. A., Potter W. J., 2015, Monthly Notices of the Royal Astronomical Society, 448, 2077
Canuto V. M., Hartke G. J., 1986, Astronomy and Astrophysics, 168, 89
Chaplin W. J., Houdek G., Elsworth Y., Gough D. O., Isaak G. R., New R., 2005, MNRAS, 360, 859
Claret A., Torres G., 2017, The Astrophysical Journal, 849, 18
Eddington A. S., 1929, MNRAS, 90, 54
Eggenberger P., Maeder A., Meynet G., 2005, A&A, 440, L9
Eggleton P. P., 1971, MNRAS, 151, 351
Eggleton P. P., Yakut K., 2016, preprint, (arXiv:1611.05041)
Eggleton P. P., Fitchett M. J., Tout C. A., 1989, ApJ, 347, 998
Farmer R., Fields C. E., Timmes F. X., 2015, ApJ, 807, 184
Galperin B., Sukoriansky S., Anderson P. S., 2007, Atmospheric Science Letters, 8, 65
Garaud P., 2002, Monthly Notices of the Royal Astronomical Society, 335, 707
Garaud P., Acevedo Arreguin L., 2009, ApJ, 704, 1
Garaud P., Bodenheimer P., 2010, ApJ, 719, 313
Garaud P., Brummell N. H., 2008, ApJ, 674, 498
Gough D. O., 1978, in Belvedere G., Paternò L., eds, Proceedings

of the Workshop on Solar Rotation. University of Catania, p. 337

Gough D. O., 2012, ISRN Astronomy and Astrophysics, 2012, 987275

Herwig F., Bloecker T., Schoenberner D., El Eid M., 1997, A&A, 324, L81

Jeans. J. H., 1926, Monthly Notices of the Royal Astronomical Society, 86, 328

Jermyn A. S., Lesaffre P., Tout C. A., Chitre S. M., 2018, MNRAS

Kichatinov L. L., Rudiger G., 1993, A&A, 276, 96

Kitchatinov L. L., 2013, in Kosovichev A. G., de Gouveia Dal Pino E., Yan Y., eds, IAU Symposium Vol. 294, Solar and Astrophysical Dynamos and Magnetic Activity. pp 399–410 (arXiv:1210.7041), doi:10.1017/S1743921313002834

Kulenthirarajah L., Garaud P., 2018, preprint, (arXiv:1803.11530)

Lesaffre P., Chitre S. M., Potter A. T., Tout C. A., 2013, MNRAS, 431, 2200

Maeder A., 1997, A&A, 321, 134

Maeder A., Zahn J.-P., 1998, A&A, 334, 1000

Marchant, Pablo Langer, Norbert Podsiadlowski, Philipp Tauris, Thomas M. Moriya, Takashi J. 2016, A&A, 588, A50

Menou K., Balbus S. A., Spruit H. C., 2004, The Astrophysical Journal, 607, 564

Meynet G., Maeder A., 2000, A&A, 361, 101

Miesch M. S., Toomre J., 2009, Annual Review of Fluid Mechanics, 41, 317

Moravveji E., Aerts C., Pápics P. I., Triana S. A., Vandoren B., 2015, A&A, 580, A27

Moravveji E., Townsend R. H. D., Aerts C., Mathis S., 2016, ApJ, 823, 130

Osaki Y., 1982, PASJ, 34, 257

Paxton B., Bildsten L., Dotter A., Herwig F., Lesaffre P., Timmes F., 2011, ApJS, 192, 3

Pols O. R., Tout C. A., Eggleton P. P., Han Z., 1995, MNRAS, 274, 964

Prat V., Guilet J., Viallet M., Müller E., 2016, A&A, 592, A59

Prather M. J., Demarque P., 1974, ApJ, 193, 109

Rajaguru S. P., Antia H. M., 2015, ApJ, 813, 114

Ribas I., Jordi C., Giménez Á., 2000, MNRAS, 318, L55

Roxburgh I. W., 1991, Differential Rotation of Fully Convective Pre-Main Sequence Stars. Springer Netherlands, Dordrecht, pp 83–88, doi:10.1007/978-94-011-3580-1_7, https://doi.org/10.1007/978-94-011-3580-1_7

Saslaw W. C., Schwarzschild M., 1965, ApJ, 142, 1468

Schröder K.-P., Pols O. R., Eggleton P. P., 1997, MNRAS, 285, 696

Spruit H. C., 2002, A&A, 381, 923

Spruit H. C., Knobloch E., 1984, A&A, 132, 89

Stancliffe R. J., Fossati L., Passy J.-C., Schneider F. R. N., 2015, A&A, 575, A117

Sweet P. A., 1950, MNRAS, 110, 548

Unno W., 1957, ApJ, 126, 259

Yakhot V., Orszag S. A., 1986, Journal of Scientific Computing, 1, 3

Zahn J.-P., 1992, A&A, 265, 115

Zahn J.-P., 1993, ”Space Sci. Rev.”, 66, 285

Zeipel H. v., 1924, Monthly Notices of the Royal Astronomical Society, 84, 665

APPENDIX A: ROLE OF DIFFERENTIAL ROTATION

The derivation of the Eddington–Sweet circulation relies on it being impossible to satisfy equation (9) when the system is perfectly barotropic. One might think from the structure

of the argument that this is just a mathematical difficulty rather than a physical one. After all, valid solutions could exist if one introduces an arbitrarily small differential rotation, allowing the cross-product in equation (4) to be non-zero.

While this is mathematically a valid concern, it runs up against physical difficulties because what matters is how rapidly the diffusivity and other thermodynamic variables may vary along an isobar. Suppose that equation (9) fails to be satisfied by a dimensionless amount of order ϵ. Differential rotation may arise and introduce baroclinicity, or misalignment of isobars and isotherms. We define the angle of misalignment to be λ [15]. If $\lambda \ll 1$, logarithmic derivatives of the diffusivity along isobars must be at least of order ϵ/λ in order to satisfy equation (9). These derivatives ought to be of order unity so we must have $\lambda \approx \epsilon$ or $\lambda > \epsilon$.

To break the degeneracy between differential rotation and meridional circulation we invoke the non-magnetised vorticity equation, which states that

$$\frac{\partial \boldsymbol{\omega}}{\partial t} = \boldsymbol{\omega} \cdot \nabla \boldsymbol{v} - \boldsymbol{\omega} \nabla \cdot \boldsymbol{v} - \boldsymbol{v} \cdot \nabla \boldsymbol{\omega} + \frac{1}{\rho^2} \nabla \rho \times \nabla p + \nabla \times \left(\frac{1}{\rho} \nabla \cdot \mathbf{T} \right),$$
(A1)

where $\boldsymbol{\omega}$ is the vorticity, $\boldsymbol{v}$ is the total velocity including rotation, ρ is the density, p is the pressure and $\mathbf{T}$ is the total fluid stress excluding magnetic effects. We now let $\boldsymbol{u}$ be the meridional velocity and assume axisymmetry, such that

$$\boldsymbol{v} = \boldsymbol{u}(R, z) + \boldsymbol{e}_\phi \Omega(R, z) R,$$
(A2)

where z is the cylindrical vertical coordinate and R is the cylindrical radial coordinate. Then

$$\boldsymbol{\omega} \cdot \nabla \boldsymbol{v} = \boldsymbol{\omega} \cdot \nabla \boldsymbol{u} + \boldsymbol{\omega} \cdot \nabla(\Omega R \boldsymbol{e}_\phi)$$
(A3)

$$= \boldsymbol{\omega} \cdot \nabla \boldsymbol{u} - \boldsymbol{e}_R \Omega \omega_\phi + \boldsymbol{e}_\phi \boldsymbol{\omega} \cdot \nabla(\Omega R),$$
(A4)

$$\boldsymbol{v} \cdot \nabla \boldsymbol{\omega} = \boldsymbol{u} \cdot \nabla \boldsymbol{\omega} + \Omega R \boldsymbol{e}_\phi \cdot \nabla \boldsymbol{\omega}$$
(A5)

$$= \boldsymbol{u} \cdot \nabla \boldsymbol{\omega} - \Omega \omega_\phi \boldsymbol{e}_R + \Omega \omega_R \boldsymbol{e}_\phi,$$
(A6)

and

$$\boldsymbol{\omega} \nabla \cdot \boldsymbol{v} = \boldsymbol{\omega} \nabla \cdot \boldsymbol{u} + \boldsymbol{\omega} \nabla \cdot (\boldsymbol{e}_\phi \Omega R)$$
(A7)

$$= \boldsymbol{\omega} \nabla \cdot \boldsymbol{u}.$$
(A8)

Putting this together we find

$$\boldsymbol{\omega} \cdot \nabla \boldsymbol{v} - \boldsymbol{\omega} \nabla \cdot \boldsymbol{v} - \boldsymbol{v} \cdot \nabla \boldsymbol{\omega}$$
(A9)

$$= \boldsymbol{\omega} \cdot \nabla \boldsymbol{u} - \boldsymbol{u} \cdot \nabla \boldsymbol{\omega} - \boldsymbol{\omega} \nabla \cdot \boldsymbol{u}$$
(A10)

$$+ \boldsymbol{e}_\phi \boldsymbol{\omega} \cdot \nabla(\Omega R) - \Omega \boldsymbol{e}_\phi \omega_R.$$
(A11)

With this equation (A1) becomes

$$\frac{\partial \boldsymbol{\omega}}{\partial t} = \boldsymbol{\omega} \cdot \nabla \boldsymbol{u} - \boldsymbol{\omega} \nabla \cdot \boldsymbol{u} - \boldsymbol{u} \cdot \nabla \boldsymbol{\omega}$$

$$+ \boldsymbol{e}_\phi \boldsymbol{\omega} \cdot \nabla(\Omega R) - \Omega \boldsymbol{e}_\phi \omega_R$$

$$+ \frac{1}{\rho^2} \nabla \rho \times \nabla p + \nabla \times \left(\frac{1}{\rho} \nabla \cdot \mathbf{T} \right).$$
(A12)

Suppose that differential rotation which satisfies equation (9) arises with no meridional flow. So $\boldsymbol{u} = 0$ and

$$\boldsymbol{\omega} = \boldsymbol{e}_z \frac{1}{R} \frac{\partial(\Omega R^2)}{\partial R} - \boldsymbol{e}_R R \frac{\partial \Omega}{\partial z}.$$
(A13)

---

[15] See Appendix B.

This rotation must then also satisfy equation (A12). In steady state this reduces to

$$0 = \boldsymbol{e}_\phi \boldsymbol{e}_z \cdot \nabla(\Omega^2 R) + \frac{1}{\rho^2}\nabla\rho \times \nabla p + \nabla \times \left(\frac{\nabla \cdot \mathbf{T}}{\rho}\right). \quad (A14)$$

We refer to the first term as advective because it is related to the advection of angular momentum. The second is the thermal wind term (Balbus et al. 2012) and the third reflects the turbulent stress.

Because equation (A14) is a vector equation it contains three scalar equations,

$$0 = \boldsymbol{e}_z \cdot \nabla(\Omega^2 R) + \frac{1}{r\rho^2}\left(\frac{\partial\rho}{\partial r}\frac{\partial P}{\partial\theta} - \frac{\partial P}{\partial r}\frac{\partial\rho}{\partial\theta}\right) + \boldsymbol{e}_\phi \cdot \nabla \times \left(\frac{\nabla \cdot \mathbf{T}}{\rho}\right), \quad (A15)$$

$$0 = \boldsymbol{e}_R \cdot \nabla \times \left(\frac{\nabla \cdot \mathbf{T}}{\rho}\right) \quad (A16)$$

and

$$0 = \boldsymbol{e}_z \cdot \nabla \times \left(\frac{\nabla \cdot \mathbf{T}}{\rho}\right), \quad (A17)$$

where r is the spherical radial coordinate. Given P and ρ satisfying equation (9) it is possible to solve for Ω with the first of these equations because $\boldsymbol{\omega}$ is just a function of Ω. The remaining two equations must then be satisfied by that solution. Because they depend on the turbulent flux there is no reason to expect them to be satisfied by a rotation profile arrived at independently. As a result the vorticity equation is unlikely to be satisfied without a meridional flow.

From this we conclude that satisfying both the vorticity equation and thermal equilibrium requires a meridional flow. Now suppose that differential rotation permits baroclinicity of order λ. If $\lambda \ll \epsilon$ then thermal equilibrium relies on the meridional flow. If $\lambda \gg \epsilon$ then thermal equilibrium could be satisfied without a meridional flow. Because the flow appears in the vorticity equation at first order multiplied by the rotation, and the differential rotation also appears at first order multiplied by the rotation, the flow must be at least as large as the differential rotation, possibly larger if turbulent stresses dominate. In this case then the flow is at least of order λ and hence also suffices to satisfy thermal equilibrium. This means that, no matter what, there is a meridional flow which is at least of order ϵ.

In the remainder of this appendix we show that if thermal equilibrium is satisfied without a meridional flow the differential rotation and/or baroclinicity which results *mechanically* drives a meridional flow. We furthermore show that this flow exceeds in magnitude that which would have been required to satisfy thermal equilibrium, and hence conclude that a flow of this order ought to generally arise.

A1 Radiative Zones

We first consider radiative zones. To begin note that because the system is axisymmetric the thermal wind term has magnitude

$$\frac{1}{\rho^2}\left|\boldsymbol{e}_\phi \cdot \nabla p \times \nabla \rho\right| = \frac{1}{\rho^2}\left|\nabla p \times \nabla\rho\right| \quad (A18)$$

$$= \frac{\lambda}{\rho^2\gamma}\left|\nabla P\right|\left|\nabla\rho\right|, \quad (A19)$$

where we have used the definition λ as the small angle between the density and pressure gradients. Taking both derivatives to produce factors of order h^{-1} we find

$$\frac{1}{\rho^2}\left|\nabla p \times \nabla\rho\right| \approx \frac{\lambda P}{h^2\rho^2\gamma}. \quad (A20)$$

Using $P = \rho g h$ and neglecting γ because it is of order unity yields

$$\left|\boldsymbol{e}_\phi \cdot (\nabla p \times \nabla\rho)\right| \approx \lambda\frac{g}{h}. \quad (A21)$$

There are now two cases to consider. First suppose that in radiative zones the stress term in equation (A15) balances the thermal wind term. Then

$$\left|\boldsymbol{e}_\phi \cdot \nabla \times \left(\frac{\nabla \cdot \mathbf{T}}{\rho}\right)\right| \approx \frac{\mathsf{T}}{\rho h^2} \approx \frac{g}{h}\lambda, \quad (A22)$$

where we have approximated derivatives by factors of h^{-1} and where $\mathsf{T} = |\mathbf{T}| = \max(\mathsf{T}_{ij})$. Using equation (A1) we may now calculate the circulation implied by the vorticity equation. Because the meridional circulation enters the meridional components of the vorticity equation in the form of gradients multiplied by Ω, we have

$$u \approx \frac{1}{h\rho\Omega}\mathsf{T} \quad (A23)$$

$$\approx \lambda\frac{g}{\Omega}. \quad (A24)$$

In essence we are balancing terms like $\boldsymbol{\omega} \cdot \nabla\boldsymbol{u}$ and $\boldsymbol{\omega}\nabla \cdot \boldsymbol{u}$ with the stress induced by the differential rotation and using $|\boldsymbol{\omega}| \approx \Omega$. The rate at which heat is deposited per unit volume by this flow is of order

$$Q = \rho c_p T u|\nabla s| \approx \rho c_p T\frac{u}{h}, \quad (A25)$$

because, in radiative zones, the dimensionless entropy gradient is on the order of h^{-1}. Noting that c_p is of order k_{B} and so $P \approx \rho c_p T$ for an ideal gas, we find

$$Q \approx \lambda P\frac{g}{h\Omega}. \quad (A26)$$

The flux anisotropy is

$$\delta F \approx \epsilon F. \quad (A27)$$

This deposits heat at a rate

$$\delta Q = \nabla \cdot \delta\boldsymbol{F} \approx \epsilon\frac{F}{h}, \quad (A28)$$

where we have used the fact that the characteristic scale of the flux anisotropy is the pressure scale height. So

$$\frac{Q}{\delta Q} \approx \frac{\lambda}{\epsilon}\left(\frac{gP}{\Omega F}\right) > \frac{gP}{\Omega F} = \frac{P\sqrt{gh}}{F}\left(\frac{g}{\Omega^2 h}\right)^{1/2} \gg \frac{P}{F}\sqrt{gh}, \quad (A29)$$

when $\lambda \geq \epsilon$ and $\Omega^2 \ll g/h$. Noting that

$$\sqrt{gh} = \sqrt{\frac{P}{\rho}} = c_s\sqrt{\gamma}, \quad (A30)$$

where c_s is the adiabatic sound speed, and that the heat flux in a radiative zone is bounded above[16] by $c_s P$, we find that $Q > \delta Q$. The heat transported in this way is therefore greater

---

[16] To see this note that $F = 16\sigma|\nabla T^4|/3\kappa\rho$. If the medium is optically thick over a scale height then $\kappa\rho h > 1$, so $F < 16\sigma h|\nabla T^4| =$

than the initial flux disturbance. As a result this flow does more to relieve the violation of thermal equilibrium than the differential rotation and so, in practice, a much milder differential rotation develops, with the circulation on the order of the baroclinicity.

A similar outcome can be seen in the case where the advective term balances the thermal wind term. Recall that the advective term is

$$\boldsymbol{e}_z \cdot \nabla(\Omega^2 R) \approx \Omega^2 |R\nabla \ln \Omega|, \tag{A31}$$

where we have assumed that the differential rotation is not strongly preferentially along $\boldsymbol{e}_R$. In this case

$$|R\nabla \ln \Omega| \approx \lambda \frac{g}{h\Omega^2}. \tag{A32}$$

Noting that $\lambda \geq \Omega^2 r/g$ we find[17]

$$|R\nabla \ln \Omega| \geq \frac{r}{h} \approx 1, \tag{A33}$$

where we have approximated r/h as unity because this is the case in the bulk of the star. Now note that the turbulent viscosity is of order (Maeder 1997)

$$\nu \approx \xi \chi \frac{S^2}{N^2}, \tag{A34}$$

where ξ is a dimensionless constant of order unity, S is the shear, $N^2 \approx g/h$ is the square of the Brünt–Väisälä frequency and

$$\chi = \frac{k}{\rho c_p} \tag{A35}$$

is the thermal diffusivity with k the thermal conductivity. We are able to use this expansion because the differential rotation which balances the thermal wind term is vertical and so not subject to centrifugal stabilisation. In order for this prescription to apply however we require that

$$\frac{N^2}{c'S^2}\left(\frac{\nu_{\mathrm{micro}}}{\chi}\right) < 1, \tag{A36}$$

where ν_{micro} is the microscopic viscosity and c' is a constant which has been variously estimated as lying between 10^{-2} and 1 Kulenthirarajah & Garaud (2018). We shall later verify that this holds in most cases of interest. Even when it does not a comparable diffusivity may be provided by turbulence in the non-stratified direction Zahn (1993). Following equation (36) of Menou et al. (2004) we find that so long as the colatitude θ is such that $\sin^2 \theta$ is small compared with the reciprocal of the left-hand side of equation (A36) this effect is active. Furthermore it is active if the specific angular momentum decreases outward. Hence even when equation (A36) is not precisely satisfied equation (A34) likely still applies in much of the star.

$4acT^4h|\nabla \ln T|$. The onset of convection occurs when $|\nabla \ln T| > (2/5)|\nabla \ln P| = 2/5h$, so $F < 8acT^4/5$. Noting that radiation pressure is given by $P_{\mathrm{rad}} = aT^4/3$, $F/P < (24/5)c(P_{\mathrm{rad}}/P)$. The sound speed is $c_{\mathrm{s}} = \sqrt{\gamma P/\rho} = c\sqrt{\gamma P/\rho c^2}$, so $F/Pc_{\mathrm{s}} \lesssim (P_{\mathrm{rad}}/P)\sqrt{P/\rho c^2}$. The first factor is at most unity by definition and the second is much less than unity (for non-relativistic gases) or of order unity (for relativistic gases), so the bound holds.

[17] We show in Appendix C that λ receives a contribution from the centrifugal term of order $\Omega^2 r/g$, but there may be additional contributions from other effects such as flux anisotropies. For the purposes of this argument though it suffices to have a lower bound.

Taking $\xi = 1$ and noting that

$$\mathsf{T} \approx \rho\nu S \tag{A37}$$

we find

$$\rho^{-1}\mathsf{T} \approx \frac{kh}{\rho c_p g}S^3. \tag{A38}$$

In a system with differential rotation

$$S \approx |R\nabla\Omega| \tag{A39}$$

so

$$\rho^{-1}\mathsf{T} \approx \frac{kh}{\rho c_p g}\Omega^3|R\nabla \ln \Omega|^3. \tag{A40}$$

Using equations (A32) and the first equality of (A24) we may now calculate the circulation implied by the vorticity equation. So we have

$$u \approx \frac{1}{h\rho\Omega}\mathsf{T} \approx \lambda^3 \frac{k}{\rho c_p g}\frac{g^3}{h^3\Omega^4}. \tag{A41}$$

Noting that $F = |k\nabla T|$ and $|\nabla T| \approx T/h$ we find

$$u \approx \lambda^3 \frac{F}{\rho c_p T}\frac{g^2}{h^2\Omega^4}, \tag{A42}$$

Noting again that $\lambda \gtrsim \Omega^2 r/g$,

$$u \gtrsim \lambda \frac{F}{\rho c_p T}\frac{g}{h\Omega^2}\left(\frac{r}{h}\right)^2. \tag{A43}$$

Because the star is spinning below breakup, $\sqrt{g/h}$ is greater than Ω. We also have $r > h$ so, when $\lambda \geq \epsilon$,

$$u > \lambda \frac{F}{\rho c_p T} \geq \epsilon \frac{F}{\rho c_p T}. \tag{A44}$$

The heat deposited by this flow is again (A25) of order

$$Q = \rho c_p Tu|\nabla s| \approx \rho c_p T\frac{u}{h}, \tag{A45}$$

because, in radiative zones, the dimensionless entropy gradient is on the order of h^{-1}. With equation (A44) we find

$$Q > \epsilon\frac{F}{h} \approx \delta Q. \tag{A46}$$

Once more the heat transported by the circulation is greater than the initial flux disturbance, so we expect a much milder differential rotation to develop, with the circulation on the order of the baroclinicity.

We now determine the conditions under which inequality (A36) gives rise to a vertical instability. Using equation (A39) this may be written as

$$\frac{N^2}{|R\nabla\Omega|^2}\left(\frac{\nu_{\mathrm{micro}}}{\chi}\right) < c'. \tag{A47}$$

Near the cores of stars, which is where we shall ultimately be interested in this expression, the viscosity is predominantly radiative[18], so that Jeans. (1926)

$$\nu_{\mathrm{micro}} \approx \frac{\chi c_p T}{4c^2}, \tag{A48}$$

---

[18] Higher up in the atmosphere the particle viscosity dominates, but there h/r is small and equation (A33) indicates that the shear is large, so we do not expect that limit to violate equation (A36).

where $c \approx 3 \times 10^{10} \mathrm{cm\,s^{-1}}$ is the speed of light. Hence the condition is that

$$\frac{N^2}{|R\nabla\Omega|^2} \left(\frac{c_\mathrm{p}T}{4c^2}\right) < c'. \tag{A49}$$

We now need a better estimate of the differential rotation than that provided by equation (A33). Near the cores of these stars, where equation (A36) is hardest to satisfy because N^2 and T are largest, we argue in Section 3 that $\epsilon \approx \Omega^2/|N|_\mathrm{core}^2$, with an upper bound of order unity, where $|N|_\mathrm{core}^2$ is the average Brünt–Väisälä frequency in the convection zone. Hence if $\lambda \geq \epsilon$ then with equation (A32) we find

$$\frac{\min(|N|_\mathrm{core}/\Omega, 1)^2 N^2 h}{g} \left(\frac{c_\mathrm{p}T}{4c^2}\right) < c'. \tag{A50}$$

Noting that $g/h \approx N^2$ we find

$$\min\left(1, \frac{|N|_\mathrm{core}^2}{\Omega^2}\right)\left(\frac{c_\mathrm{p}T}{4c^2}\right) < c'. \tag{A51}$$

Noting that $c_P T \approx c_\mathrm{s}^2$, we obtain

$$\min\left(1, \frac{|N|_\mathrm{core}^2}{\Omega^2}\right)\left(\frac{c_\mathrm{s}^2}{4c^2}\right) < c', \tag{A52}$$

or

$$\min\left(1, \frac{|N|_\mathrm{core}}{\Omega}\right)\left(\frac{c_\mathrm{s}}{c}\right) < 2c', \tag{A53}$$

In the cores of massive stars such as those we consider in Section 5, $c_\mathrm{s} \approx 3^7 \mathrm{cm^2\,s^{-2}}$ and we take $R \approx 0.2 R_\odot \approx 10^{10}\mathrm{cm}$ so

$$\min\left(1, \frac{|N|_\mathrm{core}}{\Omega}\right) < 2 \times 10^3 c'. \tag{A54}$$

Using $2 \times 10^3 c' > 1$ this reduces to

$$\Omega > \frac{|N|_\mathrm{core}}{2 \times 10^3 c'}. \tag{A55}$$

In the stars of interest $|N|_\mathrm{core} \approx 10^{-8}\mathrm{s^{-1}}$, so even with the pessimistic estimate $c' \approx 10^{-2}$ the cutoff frequency is around $10^{-9}\mathrm{s^{-1}}$, which accommodates the vast majority of stars. Away from the core we show in Section 5 that the flux anisotropy drops off so the shear falls to that given by equation (A33). As this occurs h falls, equation (A33) becomes a tighter bound, NT falls and the criterion becomes looser, so we do not expect this to represent a significant limitation, though there could be stars for which equation (A36) fails at some intermediate point. Such objects are potentially of interest, though even when the criterion is not precisely satisfied horizontal turbulence of comparable magnitude may still be active in a region in the vicinity of the poles. This means that at a minimum turbulence is present in many stars, even those which rotate quite slowly, though the latitudinal extent of the circulation may be restricted in the slowest-rotating stars.

A similar and in many ways more detailed analysis was made by Caleo et al. (2015), but they neglected turbulent stresses, which we find to be crucial to the angular momentum balance. Caleo & Balbus (2016) studied linear modes in the solar radiative zone and found them to be asymptotically stable, though many such modes grow significantly on the short term. This is consistent with the findings of Galperin et al. (2007) who show that such modes

which are not asymptotically unstable nevertheless generate turbulence for arbitrarily small shears, scaling in the manner predicted by Canuto & Hartke (1986) and Zahn (1993). This has been verified in simulations (Prat et al. 2016) and is what we assume here.

Likewise, it is useful to note that this argument may be recast in the language of Garaud & Bodenheimer (2010) by writing their parameter

$$\sigma \equiv \frac{N}{\Omega}\sqrt{\frac{\nu}{K}} = \frac{S}{\Omega}\sqrt{\xi} \approx \frac{S}{\Omega}. \tag{A56}$$

It follows that equation (A59) places an upper bound on σ of order $S/\Omega \approx |R\nabla\ln\Omega|$. So long as this is at least of order unity the flow is not mechanically constrained (Garaud 2002), and we have argued that this is the case. The scale over which mechanically forced perturbations decay is of order R/σ (Garaud & Bodenheimer 2010), so should σ be large this would be quite small. We shall argue below that σ is of order unity but may in principle grow large, in which case this may set constraints on the extent of the circulation. However because the flow in this case is locally forced by turbulent stresses it is less clear if this scale applies. In particular, because the stresses are driven by differential rotation and baroclinicity, which in turn result from perturbations to the heat flux, the scale on which perturbations mechanically damp is much less important than that on which the heat flux becomes spherically symmetric. We find this to be of order r in Section 5 and which suggests that σ does not actually set the decay length for this circulation.

Finally, with knowledge of the nature of turbulence in these systems we can, in principle, infer the extent of the differential rotation which arises. In practice our knowledge of these instabilities is incomplete, and we have entirely neglected magnetic fields, so the result of such an inference may not be trustworthy. Nevertheless it is interesting to consider and we do so now. Because baroclinic instabilities are unlikely to be active in these systems (Spruit & Knobloch 1984) it seems more likely that the azimuthal vorticity balance is between the advective terms balance the baroclinic one and that the resulting differential rotation generates shear turbulence. In this case the baroclinicity is reduced by the meridional circulation such that the shears drive a circulation which is just able to ensure thermal equilibrium. Setting $Q = \delta Q$, we see from equation (A42) that

$$\lambda^3 \frac{g^2}{h^2\Omega^4} \approx \epsilon, \tag{A57}$$

so that all subsequent inequalities become equalities. Hence the baroclinicity is of order

$$\lambda \approx \left(\frac{\epsilon h^2\Omega^4}{g^2}\right)^{1/3}, \tag{A58}$$

which is quite small even for ϵ of order unity. With this we also find

$$\frac{S}{\Omega} \approx |R\nabla\ln\Omega| \approx \left(\frac{\epsilon g}{h\Omega^2}\right)^{1/3}, \tag{A59}$$

so as $\Omega \to 0$ the differential rotation becomes small but the relative differential rotation becomes large. This behaviour cuts off when $\Omega \approx |N|$ in the convection zone, at which point

$\epsilon \approx \epsilon_{\text{max}}$ and

$$|R\nabla \ln \Omega| \approx \left(\frac{g\epsilon_{\text{max}}}{h|N|_{\text{c,core}}^2}\right)^{1/3} \approx \left(\epsilon_{\text{max}}\frac{c_s^2}{u_{c,\text{core}}^2}\right)^{1/3}. \qquad (A60)$$

For the stars of interest c_s in the core is of order $3\times10^7\text{cm s}^{-1}$ and $u_{c,\text{core}} \approx 10^3\text{cm s}^{-1}$. Taking $\epsilon_{\text{max}} \approx 0.2$ and we find the limiting relative differential rotation is $|R\nabla \ln \Omega| \lesssim 300$, corresponding to relatively mild absolute shears of order $3\times10^{-6}\text{s}^{-1}$. This represents an upper bound on the *r*elative shear in that any additional stresses or instabilities would serve to reduce the requisite shear. For the more typical case of a star with $h \approx R/2 \approx R_\odot$ rotating at 30km s^{-1} equation (A59) gives $|R\nabla \ln \Omega| \approx 3$.

In practice we do not expect the differential rotation to be so extreme. Other factors, notably the generation of a magnetic field by dynamo action, ought to become increasingly important as the rotation rate decreases and serve to further bound the differential rotation (Spruit 2002; Eggenberger et al. 2005). It is also possible that when the differential rotation becomes large it serves to in some fashion suppress the convective flux anisotropy by providing mechanical feedback to the convection zone. Without a robust formalism for evaluating these possibilities, however, we leave them for the future.

A2 Convection Zones

A similar argument holds in convection zones with two modifications. First, without any differential rotation or circulation currents, the meridional components of the vorticity equation are imbalanced due to the Λ-effect (Kitchatinov 2013). We neglect this effect because it does not generically balance the differential rotation which emerges from equation (A15) because that has a physically distinct origin and hence, in either case, there is a conflict between the meridional and azimuthal components of the vorticity equation which must be resolved by a meridional circulation. When, as we assume, there is no geometric tuning between the thermal wind term and the Λ-effect the magnitude of this conflict is set by the greater of the two effects, and hence must be at least that implied by equation (A15). Thus we may reasonably find a lower bound on the circulation velocity implied by the requirements of thermal equilibrium without considering the Λ-effect.

The second modification is to account for the fact that the flux in convection zones is modulated by the entropy gradient rather than by the temperature gradient. The entropy gradient is reduced relative to the bare thermodynamic gradients by a factor of $|N|^2h/g$, and so the baroclinicity required to restore thermal equilibrium is reduced by the same factor. Thus $\lambda \geq \epsilon|N|^2h/g$.

With these modifications in hand we may begin. Once more there are two cases. First suppose that $\Omega < |N|$. Proceeding as before we note that in convecting regions the differential rotation perturbs the stresses by an amount of order $\nu|R\nabla\Omega|$, where $\nu \approx h^2|N|$ (Gough 1978) is the turbulent

viscosity. Thus

$$\rho^{-1}\mathsf{T} \approx \nu|\nabla v| \qquad (A61)$$
$$\approx \nu|R\nabla\Omega| \qquad (A62)$$
$$\approx \nu\Omega|R\nabla \ln \Omega| \qquad (A63)$$
$$\approx h^2\Omega|N||R\nabla \ln \Omega|. \qquad (A64)$$

Because $|N| > \Omega$ this term, rather than the advective one, balances the thermal wind term in equation (A15). Thus

$$\rho^{-1}\mathsf{T} \approx \lambda hg. \qquad (A65)$$

From equation (A24) we have

$$u \approx \frac{1}{h\rho\Omega}\mathsf{T} \qquad (A66)$$
$$\approx \lambda\frac{g}{\Omega} \qquad (A67)$$
$$\approx g\frac{|N|}{\Omega|N|}\lambda \qquad (A68)$$
$$\approx h|N|\frac{g/h}{\Omega|N|}\lambda, \qquad (A69)$$

where v_c is the convection speed. Because $g/h \gg \Omega^2$ and $|N| > \Omega$ we have

$$u \approx v_c\frac{g/h}{\Omega|N|}\lambda \qquad (A70)$$
$$\geq \epsilon u_c\frac{|N|}{\Omega} \qquad (A71)$$
$$\geq \epsilon u_c. \qquad (A72)$$

Because v_c sets the scale of heat transport it follows that u more than suffices to transport the flux anisotropy ϵ.

Next suppose that $\Omega > |N|$. In this case the thermal wind term $\lambda g/h$ is of order $|N|^2$, because the flux anisotropy cannot be larger than of order unity (Jermyn et al. 2018). If we balance this with the advective term we find

$$|R\nabla \ln \Omega| \approx \frac{|N|^2}{\Omega^2}, \qquad (A73)$$

while if we balance it with the turbulent stress we find with equation (A64)

$$|R\nabla \ln \Omega| \approx \frac{|N|}{\Omega}. \qquad (A74)$$

The former is smaller because the advective term is larger and so it is the advective term which accomplishes this balance. The stress is therefore

$$\rho^{-1}\mathsf{T} \approx h^2|N|\Omega|R\nabla \ln \Omega| \qquad (A75)$$
$$\approx h^2|N|^2. \qquad (A76)$$

So, as when $\Omega < |N|$,

$$u \approx \frac{1}{h\rho\Omega}\mathsf{T} \qquad (A77)$$
$$\approx h|N|\frac{|N|}{\Omega} \qquad (A78)$$
$$\approx u_c\frac{|N|}{\Omega}. \qquad (A79)$$

In this case we find instead that the vorticity equation can be balanced by a circulation which does not balance the flux anisotropy. Thus in the case of rapid rotation baroclinicity may play a more significant role than the circulation in balancing the heat equation.

It follows then that in all but the case of rapidly rotating convection the dominant effect in restoring thermal equilibrium is the circulation current and not the introduction of baroclinicity and differential rotation. The restoring circulation therefore takes on the characteristic scale of the violation of thermal equilibrium, in this case scaling like the velocity associated with heat transport and ϵ.

It is instructive to compare this conclusion with the results of three-dimensional simulations of rotating convection zones. These simulations typically find a circulation current somewhat larger than that required to maintain thermal equilibrium yet of the same order (Miesch & Toomre 2009). This has been attributed to the requirements of the vorticity equation in the presence of anisotropic turbulent stresses[19] because, though such effects scale in the same way as the flux ansiotropy, they are generally found to be several times larger. However it is also possible that the flux anisotropy is ultimately the driver of these circulations. Nothing in our arguments precludes a larger velocity than the minimum which thermal imbalance requires, so an alternate interpretation of the simulations is that the requirement of thermal equilibrium pumps energy into the circulation and the flow geometry then arranges itself to match this input to the dissipation. The final circulation rate is not constrained except insofar as it must be at least the minimum required to restore thermal equilibrium[20].

APPENDIX B: BAROCLINICITY

In this Appendix we consider the contribution to the baroclinicity owing to both convective anisotropy and centrifugal effects. First though we define the angle λ between the pressure and entropy gradients as

$$\lambda \equiv \sin^{-1}\left(\frac{|\nabla P \times \nabla \rho|}{|\nabla P||\nabla \rho|}\right). \tag{B1}$$

Noting that in a nearly-adiabatic region

$$|\nabla \ln \rho| \approx \gamma^{-1}|\nabla \ln P|, \tag{B2}$$

where γ is the adiabatic index, we find that for small angles

$$\lambda \approx \frac{|\nabla \ln P \times \nabla \ln \rho|}{|\nabla \ln P||\nabla \ln \rho|} \tag{B3}$$

$$\approx \gamma \frac{|\nabla \ln P \times \nabla \ln \rho|}{|\nabla \ln P|^2}. \tag{B4}$$

Using the fact that ∇P and $\nabla \rho$ both lie in the meridional plane we find

$$\lambda \approx \gamma \frac{|e_\phi \cdot (\nabla \ln P \times \nabla \ln \rho)|}{|\nabla \ln P|^2} \tag{B5}$$

$$\approx \gamma \left| e_\phi \cdot ((h\nabla \ln P) \times (h\nabla \ln \rho)) \right|, \tag{B6}$$

where

$$h \equiv |\nabla \ln P|^{-1} \tag{B7}$$

is the pressure scale height. If we approximate the pressure gradient as radial then

$$\lambda \approx \left| e_\phi \cdot (-e_r \times (h\nabla \ln \rho)) \right| \tag{B8}$$

$$\approx \frac{h}{r}|\partial_\theta \ln \rho|. \tag{B9}$$

For an ideal gas we may write[21]

$$s = \ln \frac{P}{\rho^\gamma}, \tag{B10}$$

so

$$\lambda \approx \gamma h|s_\theta|. \tag{B11}$$

This parametrises the baroclinicity. Because it is of order unity we neglect the factor of γ and use equation (32) to find

$$\lambda \approx h|s_r \alpha|. \tag{B12}$$

This is our leading order estimate of the baroclinicity in the convection zone.

APPENDIX C: CENTRIFUGAL ACCELERATION

The contribution to baroclinicity from the centrifugal acceleration is quite a bit smaller than that from convective anisotropy. The pressure gradient distorts by an amount of order $\rho\Delta(\nabla\Phi) \approx \rho\Omega^2 R$. If the density were not to adjust as well this would influence λ as

$$\Delta\lambda \approx h\Delta(\nabla \ln p) \approx \frac{h}{p}\rho\Omega^2 R = \frac{\Omega^2 R}{g}. \tag{C1}$$

However in reality the density gradient adjusts too and it is instead the entropy gradient which is perturbed by the requirement of thermal equilibrium Eddington (1929). The corresponding perturbation to the baroclinicity is therefore multiplied by a factor of hs_r, so

$$\Delta\lambda \approx \frac{\Omega^2 R}{g}hs_r, \tag{C2}$$

which is much less than $hs_r\alpha$ because $N^2 \ll g/R$. Thus the total baroclinicity is well approximated by equation (B12).

APPENDIX D: DAMPING

The term that appears in equation (35) associated with the meridional flow is

$$Q_{\text{flow}} \equiv \frac{r}{D_0 s_r}(u_r s_r + u_\theta s_\theta) \tag{D1}$$

$$= \frac{r}{D_0}\left(u_r + u_\theta \frac{s_\theta}{s_r}\right). \tag{D2}$$

Inserting equations (36) and (41) we find

$$Q_{\text{flow}} \approx \alpha' + \alpha'\frac{r}{h}\frac{s_\theta}{s_r}, \tag{D3}$$

---

[19] These include the Λ-effect.
[20] Otherwise the energy input would rise as a thermal imbalance developed.

[21] In massive stars the effects of radiation pressure may matter, but these just serve to alter γ and do not fundamentally change our analysis, so we omit them for simplicity.

and with equation (32) we arrive at

$$Q_{\text{flow}} \approx \alpha' + \alpha'\alpha\frac{r}{h}. \tag{D4}$$

Because both α and α' are quadratic in the rotation we drop the second term and find

$$Q_{\text{flow}} \approx \alpha'. \tag{D5}$$

Because the terms in equation (35) are those which contribute to α and α', and because this must act to damp the baroclinicity[22], we find

$$\frac{\mathrm{d}\alpha}{\mathrm{d}\ln P} = Q_{\text{flow}} - \frac{\mathrm{d}}{\mathrm{d}\ln P}\left[\int_{\ln P}^{\ln P_{\mathrm{c}}} \frac{\Omega(\mathcal{P})^2}{|N|(\mathcal{P})^2}\mathrm{d}\ln\mathcal{P}\right] \tag{D6}$$

$$= \alpha' - \frac{\Omega^2}{|N|^2}. \tag{D7}$$

Because the term proportional to α' is responsible for damping α', the coefficient may matter, so we write

$$\frac{\mathrm{d}\alpha}{\mathrm{d}\ln P} = y'\alpha' - \frac{\Omega^2}{|N|^2}, \tag{D8}$$

where y' is a dimensionless factor of order unity determined by the geometry of the circulation. With equation (37) this becomes

$$\frac{\mathrm{d}\alpha}{\mathrm{d}\ln P} = y\alpha - \frac{\Omega^2}{|N|^2}, \tag{D9}$$

where y is another dimensionless factor of order unity. Applying the boundary condition $\alpha(P_{\mathrm{c}}) = 0$, which must be true because all latitudinal derivatives vanish at the origin, we find

$$\alpha = P^y \int_P^{P_{\mathrm{c}}} \frac{\Omega(\mathcal{P})^2}{|N|(\mathcal{P})^2}\frac{\mathrm{d}\mathcal{P}}{\mathcal{P}^{1+y}} \tag{D10}$$

as desired.

One might worry that the turbulent diffusivity found in Appendix A could also contribute significantly to the damping, but this is not the case. The turbulent thermal diffusivity is of order the viscosity given by equation (A34), so

$$\chi_{\text{turb}} \approx \nu \tag{D11}$$

Using equation (A23) we find the circulation to be of order

$$u \approx \frac{\mathsf{T}}{h\rho\Omega} \approx \nu\frac{S}{h\Omega}, \tag{D12}$$

producing effective thermal diffusivity

$$\chi_{\text{circulation}} \approx u_r h \approx \nu\frac{S}{\Omega}\left(\frac{h}{r}\right). \tag{D13}$$

Near the cores of stars $h \approx r$ and we have argued that $S \gtrsim \Omega$, so this is larger than the turbulent thermal diffusivity, so the latter may safely be neglected.

This paper has been typeset from a TeX/LaTeX file prepared by the author.

---

[22] This is because the work which powers the flow comes from the baroclinicity, or equivalently the flux anisotropy, and so thermodynamically the only option is for the flow to damp it.