

An energy-conserving anelastic approximation for strongly stably-stratified fluids

Loren Matilsky

February 25, 2024

1 Introduction

Abstract: When acoustic oscillations are believed to be irrelevant to the dynamics of a fluid, it is useful to employ simplifying approximations to the equations of motion. The two most common of these (which are usually used to treat convection problems) are the Boussinesq approximation (when the background density does not significantly vary across the fluid layer) and the anelastic approximation (when the background density does vary significantly). There are many distinct forms of the anelastic approximation in the literature, and it has often been remarked that they do not properly conserve energy when the fluid is stable to convection. Here we show that the anelastic equations derived by Gough (1969) in fact do conserve energy for arbitrary motions of the fluid, even for strongly stratified background stratification. The key properties of these equations that allow them to conserve energy are (1) the absence of the Lantz-Braginsky-Roberts (LBR) approximation in the momentum equation and (2) the inclusion of a historically neglected term in the internal energy equation. These two properties allow the proper conversion between kinetic and internal energy at the correct order of the formal asymptotic expansion of the equations. We show that the scaling analysis of Gough (1969), which implicitly assumed a single typical value of the background entropy gradient, can be valid even for convective overshoot, where the entropy gradient changes from slightly unstable in the convecting region to stable (sometimes strongly so) in the overshoot region. The requirement for the anelastic equations to be valid for convective overshoot is that the buoyancy frequency be significantly less than the acoustic cutoff frequency.

The anelastic equations consist of an approximation to the continuity and momentum equations, originally derived by assuming small thermal perturbations about a nearly adiabatically stratified hydrostatic reference atmosphere (Batchelor, 1953; Charney & Ogura, 1960). The thermodynamics of the problem thus become “linear,” in the sense that products of thermodynamic variables reduce to linear expressions in the first-order perturbations. The two key consequences of linearized thermodynamics are divergenceless mass flux and the first-order buoyancy force (associated with the first-order perturbed density and pressure) being the primary driver of the flow. Ogura & Phillips (1962) formalized the approximation by expanding the fluid equations in a small parameter ϵ , representing the relative variation

of potential temperature across the fluid layer, and hence the relative magnitude of the thermal perturbations. They recovered the equations of [Batchelor \(1953\)](#) and [Charney & Ogura \(1960\)](#) and showed an assumption about the *time scale* of the motion was necessary, in addition to the assumption of small thermal perturbations. Namely, the dynamical time scale of the buoyantly driven flows must be $O(\epsilon^{-1/2})$ times *larger* than the sound crossing time of the region. Sound waves, which imply rapid temporal variations on the order of the sound crossing time, are thus absent from the anelastic equations, making them ideal for numerical integration, where large time steps are required to capture significant evolution of the system.

In the original asymptotic expansion of [Ogura & Phillips \(1962\)](#), the internal energy equation was replaced by a heat (or entropy) equation for the evolution of potential temperature, *before* non-dimensionalizing the equations. The approach of considering the entropy equation instead of the energy equation before nondimensionalization is repeated in all modern implementations of the anelastic approximation that we are aware of (e.g., [Gilman & Glatzmaier 1981](#); [Lipps & Hemler 1982](#); [Glatzmaier 1984](#); [Lantz 1992](#); [Braginsky & Roberts 1995](#); [Lantz & Fan 1999](#); [Clune et al. 1999](#); [Rogers & Glatzmaier 2005](#); [Brown et al. 2012](#); [Vasil et al. 2013](#); [Wilczyński et al. 2022](#)). The resulting energy equation is also used in all numerical codes we are aware of that utilize the anelastic equations, for example, the `ASH` code ([Brun et al., 2004](#)), the `MagIC` code ([Gastine & Wicht, 2012](#)), the `Rayleigh` code ([Featherstone & Hindman, 2016](#); [Featherstone et al., 2023](#)), the `EULAG` code ([Smolarkiewicz & Prusa, 2004](#)), and the `Dedalus` code ([Burns et al., 2020](#); [Brown et al., 2020](#)).

While nondimensionalizing the entropy equation instead of the internal energy equation may at first appear to be an arbitrary (and harmless) choice, we show in the present work that it leads to an asymptotically inconsistent set of equations that do not conserve energy when the background is stably stratified. [Gough \(1969\)](#), by contrast, took a different approach than [Ogura & Phillips \(1962\)](#) and performed a formal asymptotic expansion in ϵ after nondimensionalizing the internal energy equation. We show that this equation set, which we dub the “Energy-conserving Generalized Gough” (EGG) anelastic equations, conserve energy for arbitrary fluid motions and for all hydrostatic background states (whether stably or unstably stratified).

2 The fully compressible equations

We begin by writing down the unapproximated fully compressible equations of motion for a nonrotating nonmagnetic fluid considered by [Gough \(1969\)](#). These are the continuity equation

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{u}) \quad (1)$$

the momentum equation,

$$\frac{\partial}{\partial t}(\rho \mathbf{u}) = -\nabla \cdot (\rho \mathbf{u} \mathbf{u}) - \nabla P + \rho \mathbf{g} + \nabla \cdot \overleftrightarrow{D}, \quad (2a)$$

$$\text{where} \quad D_{ij} = \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}(\nabla \cdot \mathbf{u})\delta_{ij} \right), \quad (2b)$$

the internal energy equation,

$$\frac{\partial}{\partial t}(\rho U) + \nabla \cdot (\rho U \mathbf{u}) + P \nabla \cdot \mathbf{u} = D_{ij} \frac{\partial u_i}{\partial x_j} + Q - \nabla \cdot \mathbf{F}, \quad (3)$$

and a general equation of state,

$$U = U(P, T). \quad (4)$$

Here, t is the time, the x_i are Cartesian spatial coordinates, ρ is the density, P the pressure, T the temperature, U the internal energy per unit mass, μ the dynamic viscosity, $\mathbf{g} := -\nabla \Phi$ the gravitational acceleration field, Φ the gravitational potential, Q an internal heat source, and \mathbf{F} the combined conductive and radiative heat flux. The gravity \mathbf{g} is assumed to point in the vertical direction $\hat{\mathbf{k}}$ (either the upward Cartesian direction for a plane-parallel fluid layer or the radial direction for spherical shell). Additionally, \mathbf{g} is assumed to depend only on the vertical coordinate q (either the upward Cartesian coordinate x_3 or the radial coordinate r) and to be time-independent (i.e., self-gravity is ignored). The symbol “ \leftrightarrow ” in the viscous stress tensor \overleftrightarrow{D} denotes a second-order tensor, as does the dyadic notation $\mathbf{u}\mathbf{u}$. The subscripts i and j (taking the values 1,2,3) denote vector or tensor components in any of the Cartesian spatial directions. We use the Einstein summation convention and δ_{ij} denotes the Kronecker delta.

These equations are not written in the exact form of [Gough \(1969\)](#) (and use slightly different notation) but are mathematically equivalent. Note that the left-hand side (LHS) of Equation (3) can be written in several other forms which will prove useful:

$$\frac{\partial}{\partial t}(\rho U) + \nabla \cdot (\rho U \mathbf{u}) + P \nabla \cdot \mathbf{u} = \rho \frac{DU}{Dt} - \frac{P}{\rho} \frac{D\rho}{Dt} \quad (5a)$$

$$= \rho \frac{Dh}{Dt} - \frac{DP}{Dt} \quad (5b)$$

$$= \rho T \frac{DS}{Dt}, \quad (5c)$$

where

$$\frac{D}{Dt} := \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \quad (6)$$

is the material (or Lagrangian) derivative,

$$h := U + \frac{P}{\rho} \quad (7)$$

is the specific enthalpy, and

$$S = S(P, T) \quad (8)$$

is the specific entropy.

It will also be helpful to define the following fluid properties associated with the generalized equations of state (4) and (8): the specific heat at constant pressure,

$$C_p = C_p(P, T) := T \left(\frac{\partial S}{\partial T} \right)_P, \quad (9)$$

the squared adiabatic sound speed,

$$c_s^2 = c_s^2(P, T) := \left(\frac{\partial P}{\partial \rho} \right)_S, \quad (10)$$

and the thermal expansion coefficient,

$$\delta = \delta(P, T) := - \left(\frac{\partial \ln \rho}{\partial \ln T} \right)_P. \quad (11)$$

The first law of thermodynamics takes the following forms:

$$TdS = dU - \frac{P}{\rho^2} d\rho \quad (12a)$$

$$= dh - \frac{dP}{\rho} \quad (12b)$$

$$= C_p dT - \frac{\delta}{\rho} dP \quad (12c)$$

$$= \frac{C_p T}{\rho \delta} \left[\frac{dP}{c_s^2} - d\rho \right]. \quad (12d)$$

An equation for the evolution of kinetic energy can be formed from \mathbf{u} dotted into Equation (2),

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho u^2 \right) = -\nabla \cdot \left(\frac{1}{2} \rho u^2 \mathbf{u} \right) + \mathbf{u} \cdot \nabla P - \rho \mathbf{u} \cdot \nabla \Phi + u_i \frac{\partial D_{ij}}{\partial x_j}. \quad (13)$$

Equation (1) multiplied by Φ yields an equation for the evolution of potential energy,

$$\frac{\partial}{\partial t} (\rho \Phi) = -\Phi \nabla \cdot (\rho \mathbf{u}). \quad (14)$$

Adding Equations (3), (13), and (14) yields an equation for the evolution of total energy,

$$\frac{\partial}{\partial t} \left[\rho \left(\frac{1}{2} u^2 + U + \Phi \right) \right] = -\nabla \cdot \left\{ \left[\rho \left(\frac{1}{2} u^2 + U + \Phi \right) + P \right] \mathbf{u} - \mathbf{u} \cdot \overleftrightarrow{D} + \mathbf{F} \right\} + Q. \quad (15)$$

3 The anelastic approximation of Gough (1969)

We will not repeat the full asymptotic expansion in ϵ of Equations (1), (2), (3), and (4) here. Instead, we reiterate the salient assumptions in the case where the horizontally averaged background atmosphere is time-independent and the layer depth is thicker than the typical

pressure scale height¹ The main assumption is that the thermodynamic perturbations from the horizontally averaged background state are small, e.g.,

$$\begin{aligned}\rho &= \bar{\rho}(q) + \rho_1(x_i, t) & \text{with} & \quad \rho_1/\bar{\rho} = O(\epsilon) \ll 1, \\ P &= \bar{P}(q) + P_1(x_i, t) & \text{with} & \quad P_1/\bar{P} = O(\epsilon) \ll 1,\end{aligned}\tag{16}$$

and similarly for T , U , h , C_p , μ , δ , c_s^2 , \mathbf{F} , and Q . Here, the overbars denote horizontal averages and the “1” subscripts denote the perturbations about this average. Note that it is *not* correct to write “ $S = \bar{S}(q) + S_1(x_i, t)$ with $S_1/\bar{S} = O(\epsilon)$.” The fully compressible equations of motion contain only differences in entropy and so no meaningful absolute value of \bar{S} can be defined. Instead, we must write

$$S = \bar{S}(q) + S_1(x_i, t) \quad \text{with} \quad S_1/\bar{C}_p = O(\epsilon) \ll 1.\tag{17}$$

The second assumption is that the coordinate system can be chosen such that there is no mass flux across any horizontal surface, i.e.,

$$\overline{\rho u_i} = 0.\tag{18}$$

In a spherical system, the horizontal average would be a spherically symmetric average and the coordinates would point along the spatially varying curvilinear coordinate directions.

The characteristic length scale of variation of the fluid is assumed to be a typical value H_a for the pressure scale height. The flow is assumed to be buoyantly driven by the $O(\epsilon)$ thermal perturbations, i.e.,

$$|\mathbf{u}| = O(\sqrt{\epsilon g_a H_a}) = O(\sqrt{\epsilon} c_{sa}),\tag{19}$$

where “a” subscripts denote typical atmospheric background-state values. Thus, the squared Mach number of the flow is assumed to be $O(\epsilon)$. The characteristic time scale of variation of the fluid is assumed to be advective, i.e.,

$$\left| \frac{\partial}{\partial t} \right| = O\left(\sqrt{\frac{\epsilon g_a}{H_a}} \right) = O\left(\sqrt{\epsilon} \frac{1}{H_a/c_{sa}} \right).\tag{20}$$

Thus, the characteristic time scale is $O(\epsilon^{-1/2})$ times longer than the time it takes a sound wave to cross a pressure scale height.

Finally, the vertical convective heat flux, which maximally could transport an energy flux of order $\bar{\rho} \bar{T} w \Delta \bar{S}$, where

$$w := \hat{\mathbf{k}} \cdot \mathbf{u}\tag{21}$$

is the vertical velocity and $\Delta \bar{S}$ is the total drop in background entropy across the convecting layer, is assumed to be limited primarily by the thermal diffusion \mathbf{F} . This will be true if the

¹Gough (1969) also considers thin layers, in which the anelastic equations become the Boussinesq equations, and time-dependent (moving) background atmospheres.

conductive heating $-\nabla \cdot \mathbf{F}$ in Equation (3) is at least as large as the viscous and internal heatings. In the case of negligible heatings (high Rayleigh number), one expects

$$\frac{\Delta \bar{S}}{C_{pa}} = O(\epsilon), \quad (22)$$

i.e., the convecting layer should be nearly adiabatically stratified for vigorous convection.

Once all of these scaling assumptions have been made, Equations (1), (2), (3), and (4) are nondimensionalized, each term is expanded in powers of ϵ , terms up to zeroth-order in the continuity equation and first-order in the other equations are retained, and redimensionalization then yields Gough (1969)'s anelastic equations. Note that one consequence of Equation (18) is that the horizontally averaged velocity $\bar{\mathbf{u}}$ is $O(\epsilon)$ *smaller* than the perturbed velocity \mathbf{u}_1 . Hence, only the perturbation velocity \mathbf{u}_1 appears in the equations, and we subsequently drop the subscripts on \mathbf{u} . Thus, by definition,

$$\bar{\mathbf{u}} \equiv 0. \quad (23)$$

Under the Gough (1969)'s anelastic approximation, the continuity equation (1) becomes

$$\nabla \cdot (\bar{\rho} \mathbf{u}) = 0, \quad (24)$$

the momentum equation (2) becomes

$$\frac{\partial}{\partial t}(\bar{\rho} \mathbf{u}) = -\nabla \cdot (\bar{\rho} \mathbf{u} \mathbf{u}) - \nabla P_1 + \rho_1 \mathbf{g} + \nabla \cdot \overleftrightarrow{D} + [-\nabla \bar{P} + \bar{\rho} \mathbf{g}], \quad (25a)$$

$$\text{where now} \quad D_{ij} = \bar{\mu} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}(\nabla \cdot \mathbf{u})\delta_{ij} \right), \quad (25b)$$

the energy equation (3) becomes

$$\begin{aligned} \bar{\rho} \bar{C}_p \frac{\partial T_1}{\partial t} - \bar{\delta} \frac{\partial P_1}{\partial t} &= -\bar{\rho} \mathbf{u} \cdot \left(\nabla h_1 - \frac{1}{\bar{\rho}} \nabla P_1 \right) - \bar{\rho} \bar{T} \mathbf{u} \cdot \nabla \bar{S} \\ &\quad D_{ij} \frac{\partial u_i}{\partial x_j} + Q_1 - \nabla \cdot \mathbf{F}_1 - \rho_1 \mathbf{u} \cdot \mathbf{g} - \bar{T}(\rho_1 \mathbf{u} - \bar{\rho}_1 \bar{\mathbf{u}}) \cdot \nabla \bar{S} \\ &\quad + [\bar{Q} - \nabla \cdot \mathbf{F}], \end{aligned} \quad (26)$$

and the linearized equation of state (4) becomes

$$\bar{T} S_1 = \bar{C}_p T_1 - \frac{\bar{\delta}}{\bar{\rho}} P_1 \quad (27a)$$

$$= U_1 - \frac{\bar{P}}{\bar{\rho}^2} \rho_1 \quad (27b)$$

$$= h_1 - \frac{P_1}{\rho} \quad (27c)$$

$$= \frac{\bar{C}_p \bar{T}}{\bar{\delta} \bar{\rho}} \left[\frac{P_1}{c_s^2} - \rho_1 \right]. \quad (27d)$$

Note that from Equation (27a), the LHS of Equation (26) can be written in terms of the entropy S_1 ,

$$\bar{\rho} \bar{C}_p \frac{\partial T_1}{\partial t} - \bar{\delta} \frac{\partial P_1}{\partial t} = \bar{\rho} \bar{T} \frac{\partial S_1}{\partial t}. \quad (28)$$

Again, these equations are not in the identical form of Gough (1969) but are mathematically equivalent. Note that because the flow is assumed to satisfy Equation (18), Gough (1969) must expand the mass flux (or equivalently, the momentum density) $\mathbf{m} = \rho \mathbf{u}$ in ϵ rather than the velocity \mathbf{u} . Because of Equation (18), we can write

$$\mathbf{m} = \bar{\rho} \mathbf{u} + \rho_1 \mathbf{u} - \bar{\rho}_1 \mathbf{u} + O(\epsilon^2). \quad (29)$$

Each of the two terms $\rho_1 \mathbf{u}$ and $-\bar{\rho}_1 \mathbf{u}$ are $O(\epsilon)$. In most cases, we can thus write $\mathbf{m} \approx \bar{\rho} \mathbf{u}$ to translate from Gough (1969) to the current notation (which uses \mathbf{u} as the field variable), except when multiplying by potentially zeroth order quantities like $\nabla \bar{S}$. We have used Equation (29) to yield the term $-\bar{T}(\rho_1 \mathbf{u} - \bar{\rho}_1 \mathbf{u}) \cdot \nabla \bar{S}$ in Equation (26).

The differentials in Equation (12) can be converted into gradients (e.g., $T \nabla S = \nabla h - \nabla P / \rho$) and the horizontally averaged form of these relations yields

$$\bar{T} \nabla \bar{S} = \bar{C}_p \nabla \bar{T} - \frac{\bar{\delta}}{\bar{\rho}} \nabla \bar{P} + O(\epsilon^2) \quad (30a)$$

$$= \frac{\bar{C}_p \bar{T}}{\bar{\delta} \bar{\rho}} \left[\frac{\nabla \bar{P}}{\bar{c}_s^2} - \nabla \bar{\rho} \right] + O(\epsilon^2). \quad (30b)$$

Note that Gough (1969) uses the superadiabatic temperature gradient

$$\beta := -\frac{\bar{T}}{\bar{C}_p} \hat{\mathbf{k}} \cdot \nabla \bar{S} \quad (31)$$

in place of $\nabla \bar{S}$.

The zeroth order (horizontally averaged) parts of Equations (25) and (26) satisfy

$$-\nabla \bar{P} + \bar{\rho} \mathbf{g} = \nabla(\bar{\rho} \bar{w}^2) \quad (32)$$

and

$$\bar{Q} - \nabla \cdot \bar{\mathbf{F}} = \overline{\bar{\rho} \mathbf{u} \cdot \left(\nabla h_1 - \frac{1}{\bar{\rho}} \nabla P_1 \right)} + \mathbf{g} \cdot \overline{\bar{\rho}_1 \mathbf{u}} - \overline{D_{ij} \frac{\partial u_i}{\partial x_j}} \quad (33)$$

In each of Equations (32) and (33), each term on the right-hand side (RHS) is $O(\epsilon)$ compared to each term on the LHS. In particular, we can approximate

$$\nabla \bar{P} \approx \bar{\rho} \mathbf{g} \quad (34)$$

in terms that are already of first-order in ϵ . Dotting \mathbf{u} into Equation (25a) yields the anelastic kinetic energy equation,

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \bar{\rho} u^2 \right) = -\nabla \cdot \left(\frac{1}{2} \bar{\rho} u^2 \mathbf{u} \right) - \mathbf{u} \cdot \nabla P_1 + \rho_1 \mathbf{u} \cdot \mathbf{g} + u_i \frac{\partial D_{ij}}{\partial x_j} + \mathbf{u} \cdot \nabla(\bar{\rho} \bar{w}^2), \quad (35)$$

where we have used Equation (32) to write the work due to the “turbulent pressure,” $\mathbf{u} \cdot \nabla(\bar{\rho}w^2)$. Using Equations (24) and (27) as needed and then adding Equations (26) and (35) yields Gough (1969)’s total energy equation,

$$\begin{aligned} \frac{\partial}{\partial t} \left[\bar{\rho} \left(\frac{1}{2} u^2 + \bar{T} S_1 \right) \right] = & -\nabla \cdot \left\{ \left[\bar{\rho} \left(\frac{1}{2} u^2 + \bar{T} S_1 \right) + P_1 \right] \mathbf{u} - \mathbf{u} \cdot \overleftrightarrow{D} + \bar{\mathbf{F}} + \mathbf{F}_1 \right\} + \bar{Q} + Q_1 \\ & + \left[-\bar{\rho} \bar{T} \mathbf{u} \cdot \nabla \bar{S} + \mathbf{u} \cdot \nabla(\bar{\rho}w^2) - \bar{T}(\rho_1 \mathbf{u} - \bar{\rho}_1 \bar{\mathbf{u}}) \cdot \nabla \bar{S} \right]. \end{aligned} \quad (36)$$

Because of Equation (23), or more specifically, the condition

$$\bar{w} = 0, \quad (37)$$

each of the rightmost terms in brackets in Equation (36) cannot transport any net energy across the layer. Note that Equation (37) is also a consequence of integrating Equation (24) over volumes bounded by horizontal surfaces, and so the vanishing of all components of $\bar{\mathbf{u}}$ is not strictly necessary for the bracketed terms to conserve energy.

The internal heating terms $\bar{Q} + Q_1$ are assumed “accounted for,” since in modern anelastic codes, they are typically inputs to drive convection. Thus, Equation (36) shows that the total energy integrated over the volume V of the layer,

$$E_{\text{tot}} := \int_V \bar{\rho} \left(\frac{1}{2} u^2 + \bar{T} S_1 \right) dV \quad (38)$$

is conserved if the various fluxes in the divergence on the RHS of Equation (36) vanish on the boundaries. This conservation holds for arbitrary fluid motions that obey Gough (1969)’s anelastic equations (24), (25), (26), and (27) and for arbitrary magnitudes of $|\nabla \bar{S}|$. However, whether the approximation remains *consistent* (i.e., if the thermal perturbations remain small) *does* depend on the magnitude of $|\nabla \bar{S}|$.

Note that from Equation (27b), the perturbed anelastic internal energy is $U_1 = \bar{T} S_1 + (\bar{P}/\bar{\rho}^2)\rho_1$. The latter compressive term $(\bar{P}/\bar{\rho}^2)\rho_1$ is thus eliminated under the anelastic approximation and we instead identify $\bar{T} S_1$ with the internal energy. The disappearance of the compressive term, which was called the “elastic energy” in Eckart & Ferris (1956), is the origin of the term “anelastic,” coined by Jule Charney (see Ogura & Phillips 1962). The potential energy term also disappears, which is a direct consequence of the assumption of zero mean mass flux, Equation (18).

4 The Gough (1969) equations in more familiar form

Modern anelastic codes typically write the equations using the perturbed pressure and entropy (P_1 and S_1) in place the quantities T_1 , P_1 , ρ_1 , and h_1 that appear in Equations (25) and (32). We can convert using Equations (27) and (30) and the approximation (34). In the momentum equation, we find

$$-\nabla P_1 + \rho_1 \mathbf{g} = -\bar{\rho} \nabla \left(\frac{P_1}{\bar{\rho}} \right) - \bar{\delta} \bar{\rho} \left(\frac{S_1}{C_p} \right) \mathbf{g} + \frac{\bar{\delta} \bar{\rho}}{C_p} \left(\frac{P_1}{\bar{\rho}} \right) \nabla \bar{S}. \quad (39)$$

In the energy equation, we find

$$-\bar{\rho}\mathbf{u} \cdot \left(\nabla h_1 - \frac{1}{\bar{\rho}} \nabla P_1 \right) = -\bar{\rho}\bar{T}\mathbf{u} \cdot \nabla S_1 - \bar{\rho}S_1\mathbf{u} \cdot \nabla \bar{T} + \left(\frac{P_1}{\bar{\rho}} \right) \mathbf{u} \cdot \nabla S_1. \quad (40)$$

Using Equations (27), (30), and (34) as needed, we compute, with some effort,

$$-\bar{\rho}S_1\mathbf{u} \cdot \nabla \bar{T} + \left(\frac{P_1}{\bar{\rho}} \right) \mathbf{u} \cdot \nabla S_1 - \rho_1\mathbf{u} \cdot \mathbf{g} = -\bar{\rho}T_1\mathbf{u} \cdot \nabla \bar{S} + O(\epsilon^2). \quad (41)$$

Plugging Equation (39) into Equation (25), and then plugging Equations (28), (40), and (41) into Equation (26), we find

$$\begin{aligned} \frac{\partial}{\partial t}(\bar{\rho}\mathbf{u}) = & -\nabla \cdot (\bar{\rho}\mathbf{u}\mathbf{u}) - \bar{\rho}\nabla \left(\frac{P_1}{\bar{\rho}} \right) - \bar{\delta}\bar{\rho} \left(\frac{S_1}{\bar{C}_p} \right) \mathbf{g} + \underbrace{\frac{\bar{\delta}P_1}{\bar{C}_p} \nabla \bar{S}}_{:=\mathbf{f}_{\text{NLBR}}} + \nabla \cdot \overleftrightarrow{D} \\ & + [-\nabla \bar{P} + \bar{\rho}\mathbf{g}], \end{aligned} \quad (42)$$

and

$$\begin{aligned} \bar{\rho}\bar{T} \frac{\partial S_1}{\partial t} = & -\bar{\rho}\bar{T}\mathbf{u} \cdot \nabla S_1 - \bar{\rho}\bar{T}\mathbf{u} \cdot \nabla \bar{S} - \underbrace{\bar{\rho}T_1\mathbf{u} \cdot \nabla \bar{S}}_{:=Q_{\text{NLBR}}} + D_{ij} \frac{\partial u_i}{\partial x_j} + Q_1 - \nabla \cdot \mathbf{F}_1 \\ & + [\bar{Q} - \nabla \cdot \bar{\mathbf{F}} - \bar{T}(\rho_1\mathbf{u} - \bar{\rho}_1\mathbf{u}) \cdot \nabla \bar{S}], \end{aligned} \quad (43a)$$

$$\text{where} \quad \rho_1 = \frac{P_1}{\bar{c}_s^2} - \frac{\bar{\delta}\bar{\rho}S_1}{\bar{C}_p} \quad (43b)$$

$$\text{and} \quad T_1 = \frac{\bar{T}S_1}{\bar{C}_p} + \frac{\bar{\delta}P_1}{\bar{\rho}\bar{C}_p}. \quad (43c)$$

The final terms in brackets in these equations may be dropped without affecting energy conservation, since they do not transport any net energy. However, dropping them still may still makes the equations an asymptotically inconsistent approximation to the true fully compressible motion, and so we retain them.

The essential terms required for energy conservation when $\nabla \bar{S} \neq 0$ are the “non-LBR” force density

$$\mathbf{f}_{\text{NLBR}} := \frac{\bar{\delta}P_1}{\bar{C}_p} \nabla \bar{S} \quad (44)$$

and the “non-LBR” heating

$$Q_{\text{NLBR}} := -\bar{\rho}T_1\mathbf{u} \cdot \nabla \bar{S}. \quad (45)$$

Both these terms vanish for an adiabatic background state (where $\nabla \bar{S} = 0$), which is expected for a fully (and sufficiently vigorously) convecting layer. Neglecting \mathbf{f}_{NLBR} was first done independently by Lantz (1992) and Braginsky & Roberts (1995) and is referred as the “Lantz-Braginsky-Roberts” (LBR) approximation. The term Q_{NLBR} in the internal energy equation, which was implicitly contained in the equations of Gough (1969), seems to be absent in the other forms of the anelastic equations currently in use, and its neglect seems to have not been explicitly considered.

5 The equivalence of horizontally averaged atmospheres to fixed reference atmospheres

In the formalism of [Gough \(1969\)](#), the horizontally averaged atmosphere, denoted by the overbars, cannot be specified a priori because it depends on the ultimate flow via Equations (32) and (33). Many anelastic numerical codes (e.g., the **Rayleigh**, **EULAG**, and **MagIC** codes, which simulate the anelastic equations in spherical shells) instead treat the background state as a fixed-in-time, spherically symmetric, hydrostatic “reference” state and let the perturbations in S and P about this reference state develop small but nonzero horizontally averaged perturbations. Some codes (e.g., the **ASH** code) alternatively solve for the nonspherical perturbations directly, retaining horizontally averaged terms on the RHS’s of the anelastic equations, similar to the bracketed terms in Equations (25) and (26). As we now show, these two approaches are exactly equivalent to the order of the anelastic equations, provided that the thermal variables do not wander by more than $O(\epsilon)$ away from their preordained reference state values.

Our approach is to define the horizontally averaged profiles as the sum of the preordained reference-state profile (denoted by a tilde) and a horizontally symmetric deviation (denoted by a hat):

$$\bar{P} = \bar{P}(q) = \tilde{P}(q) + \hat{P}(q), \quad (46)$$

$$\bar{\rho} = \bar{\rho}(q) = \tilde{\rho}(q) + \hat{\rho}(q), \quad (47)$$

$$\bar{S} = \bar{S}(q) = \tilde{S}(q) + \hat{S}(q), \quad (48)$$

etc., where we assume (apart from the entropy) that the hatted means are $O(\epsilon)$ compared to the reference-state means. We also assume that the reference state is hydrostatic,

$$\nabla \tilde{P} = \tilde{\rho} \mathbf{g}. \quad (49)$$

We denote the (temporally and horizontally dependent) deviations from the reference state by primes and note that

$$P' := P - \tilde{P} = P_1 + \hat{P}, \quad (50)$$

$$\rho' := \rho - \tilde{\rho} = \rho_1 + \hat{\rho}, \quad (51)$$

$$S' := S - \tilde{S} = S_1 + \hat{S}, \quad (52)$$

etc. The primed quantities (except for the entropy) are always $O(\epsilon)$ compared to the reference-state means.

The linearized equation of state for the primed quantities is exactly analogous to Equation

(27),

$$\tilde{T}S' = \tilde{C}_p T' - \frac{\tilde{\delta}}{\tilde{\rho}} P', \quad (53a)$$

$$= U' - \frac{\tilde{P}}{\tilde{\rho}^2} \rho' \quad (53b)$$

$$= h' - \frac{P'}{\tilde{\rho}} \quad (53c)$$

$$= \frac{\tilde{C}_p \tilde{T}}{\tilde{\delta} \tilde{\rho}} \left[\frac{P'}{\tilde{c}_s^2} - \hat{\rho} \right]. \quad (53d)$$

as is the linearized equation of state for the hatted quantities. We further assume that reference state is defined to satisfy an equation analogous to Equation (30),

$$\tilde{T} \nabla \tilde{S} = \tilde{C}_p \nabla \tilde{T} - \frac{\tilde{\delta}}{\tilde{\rho}} \nabla \tilde{P} \quad (54a)$$

$$= \frac{\tilde{C}_p \tilde{T}}{\tilde{\delta} \tilde{\rho}} \left[\frac{\nabla \tilde{P}}{\tilde{c}_s^2} - \nabla \tilde{\rho} \right]. \quad (54b)$$

To zeroth order in ϵ , Equation (24) becomes simply

$$\nabla \cdot (\tilde{\rho} \mathbf{u}) \equiv 0 \quad (55)$$

Because the RHS of Equations (25) is $O(\epsilon)$ compared to the LHS's we can write (also using Equation (49))

$$\begin{aligned} [-\nabla \bar{P} + \bar{\rho} \mathbf{g}] &= -\nabla \hat{P} + \hat{\rho} \mathbf{g} \\ &= -\tilde{\rho} \nabla \left(\frac{\hat{P}}{\tilde{\rho}} \right) - \tilde{\delta} \tilde{\rho} \left(\frac{\hat{S}}{\tilde{C}_p} \right) \mathbf{g} + \frac{\tilde{\delta} \tilde{\rho}}{\tilde{C}_p} \left(\frac{\hat{P}}{\tilde{\rho}} \right) \nabla \tilde{S}. \end{aligned} \quad (56)$$

Note that the non-LBR force density from Equation (44) is only significant when $\nabla \bar{S}$ is large (in which case $\nabla \bar{S} = \nabla \tilde{S} + O(\epsilon)$), otherwise it is $O(\epsilon^2)$. For all magnitudes of $|\nabla \bar{S}|$, we can thus write

$$\mathbf{f}_{\text{NLBR}} = \frac{\bar{\delta} P_1}{\tilde{C}_p} \nabla \tilde{S} + O(\epsilon^2) \quad (57)$$

Using Equations (56) and (57) and noting that all terms in Equation (42) are of $O(\epsilon)$ (so that we can replace overbars with tildes), we find

$$\frac{\partial}{\partial t} (\tilde{\rho} \mathbf{u}) = -\nabla \cdot (\tilde{\rho} \mathbf{u} \mathbf{u}) - \tilde{\rho} \nabla \left(\frac{P'}{\tilde{\rho}} \right) - \tilde{\delta} \tilde{\rho} \left(\frac{S'}{\tilde{C}_p} \right) \mathbf{g} + \frac{\tilde{\delta} P'}{\tilde{C}_p} \nabla \tilde{S} + \nabla \cdot \overleftrightarrow{D}, \quad (58a)$$

$$\text{where now} \quad D_{ij} = \tilde{\mu} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3} (\nabla \cdot \mathbf{u}) \delta_{ij} \right), \quad (58b)$$

In the energy equation (43), we write

$$\tilde{Q} - \nabla \cdot \tilde{\mathbf{F}} := -\tilde{\rho}\tilde{T}\frac{\partial\hat{S}}{\partial t}. \quad (59)$$

Note that this is simply a *definition* for $\partial\hat{S}/\partial t$. The anelastic equations of Gough (1969) do not specify \hat{S} explicitly, the only requirement being (implicitly through the equation of state) the satisfaction of Equation (32).

Equation (29) can be additionally written

$$\bar{\rho}\mathbf{u} + \rho_1\mathbf{u} - \overline{\rho_1\mathbf{u}} = \tilde{\rho}\mathbf{u} + \rho'\mathbf{u} - \overline{\rho'\mathbf{u}} + O(\epsilon^2). \quad (60)$$

We finally write

$$Q_{\text{NLBR}} = \tilde{\rho}T_1\mathbf{u} \cdot \nabla\tilde{S} + O(\epsilon^2) \quad (61)$$

and

$$-\bar{\rho}\tilde{T}\mathbf{u} \cdot \nabla S_1 - \bar{\rho}\tilde{T}\mathbf{u} \cdot \nabla\tilde{S} = -\tilde{\rho}\tilde{T}\mathbf{u} \cdot \nabla S' - \tilde{T}(\bar{\rho}\mathbf{u}) \cdot \nabla\tilde{S} - \tilde{\rho}\hat{T}\mathbf{u} \cdot \nabla\tilde{S}. \quad (62)$$

Plugging in Equations (59) through (62) into Equation (43) thus yields

$$\begin{aligned} \tilde{\rho}\tilde{T}\frac{\partial S'}{\partial t} = & -\tilde{\rho}\tilde{T}\mathbf{u} \cdot \nabla S' - \tilde{\rho}\tilde{T}\mathbf{u} \cdot \nabla\tilde{S} - \tilde{\rho}T'\mathbf{u} \cdot \nabla\tilde{S} + D_{ij}\frac{\partial u_i}{\partial x_j} + Q' - \nabla \cdot \mathbf{F}' \\ & - \tilde{T}(\rho'\mathbf{u} - \overline{\rho'\mathbf{u}}) \cdot \nabla\tilde{S}, \end{aligned} \quad (63a)$$

$$\text{where} \quad \rho' = \frac{P'}{\tilde{c}_s^2} - \frac{\tilde{\delta}\tilde{\rho}S'}{\tilde{C}_p} \quad (63b)$$

$$\text{and} \quad T' = \frac{\tilde{T}S'}{\tilde{C}_p} + \frac{\tilde{\delta}P'}{\tilde{\rho}\tilde{C}_p}. \quad (63c)$$

Taking the spherical means of Equations (64b) and (64c) and using Equations (49) and Equations (59) then recovers the mean momentum and energy equations (32) and (33). The formulation with fixed reference states (Equations (64a), (64b) and (64c)) is thus seen to be asymptotically equivalent to the formulation with horizontally averaged background states (Equations (24), (42), and (43), combined with Equations (32) and (33)).

6 The Energy-conserving Generalized Gough (EGG) anelastic approximation

To arrive at the final form of the equations, we make one final argument: that the last term in Equation (63a) is negligible. It has zero horizontal mean and therefore cannot affect the net transport of energy. Pointwise, it is always much smaller than the background advection term $-\tilde{\rho}\tilde{T}\mathbf{u} \cdot \nabla\tilde{S}$. It therefore seems safe to neglect.

The final equations, representing what we call the Energy-conserving Generalized Gough (EGG) anelastic approximation, are thus

$$\begin{aligned} \nabla \cdot (\tilde{\rho} \mathbf{u}) &\equiv 0, & (64a) \\ \frac{\partial}{\partial t}(\tilde{\rho} \mathbf{u}) &= -\nabla \cdot (\tilde{\rho} \mathbf{u} \mathbf{u}) - \tilde{\rho} \nabla \left(\frac{P'}{\tilde{\rho}} \right) - \tilde{\delta} \tilde{\rho} \left(\frac{S'}{\tilde{C}_p} \right) \mathbf{g} + \frac{\tilde{\delta} P'}{\tilde{C}_p} \nabla \tilde{S} + \nabla \cdot \overleftrightarrow{D}, & (64b) \\ \text{and } \tilde{\rho} \tilde{T} \frac{\partial S'}{\partial t} &= -\tilde{\rho} \tilde{T} \mathbf{u} \cdot \nabla S' - \tilde{\rho} \tilde{T} \mathbf{u} \cdot \nabla \tilde{S} - \left(\frac{\tilde{T} S'}{\tilde{C}_p} + \frac{\tilde{\delta} P'}{\tilde{\rho} \tilde{C}_p} \right) \tilde{\rho} \mathbf{u} \cdot \nabla \tilde{S} \\ &\quad + D_{ij} \frac{\partial u_i}{\partial x_j} + Q' - \nabla \cdot \mathbf{F}', & (64c) \end{aligned}$$

where \overleftrightarrow{D} is defined in Equation (58b) and it is additionally assumed that the reference state satisfies the hydrostatic condition (49) and the first law of thermodynamics in gradient form, Equation (54).

The EGG kinetic energy equation, derived from \mathbf{u} dotted into Equation (64b), is

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \tilde{\rho} u^2 \right) = -\nabla \cdot \left(\frac{1}{2} \tilde{\rho} u^2 \mathbf{u} \right) - \mathbf{u} \cdot \nabla P' - \tilde{\delta} \tilde{\rho} \left(\frac{S'}{\tilde{C}_p} \right) \mathbf{u} \cdot \mathbf{g} + \frac{\tilde{\delta} P'}{\tilde{C}_p} \mathbf{u} \cdot \nabla \tilde{S} + u_i \frac{\partial D_{ij}}{\partial x_j}, \quad (65)$$

Adding Equations (64c) and (65) yields the EGG total energy equation,

$$\begin{aligned} \frac{\partial}{\partial t} \left[\tilde{\rho} \left(\frac{1}{2} u^2 + \tilde{T} S' \right) \right] &= -\nabla \cdot \left\{ \left[\tilde{\rho} \left(\frac{1}{2} u^2 + \tilde{T} S' \right) + P' \right] \mathbf{u} - \mathbf{u} \cdot \overleftrightarrow{D} + \tilde{\mathbf{F}} + \mathbf{F}' \right\} \\ &\quad + Q' - \tilde{\rho} \tilde{T} \mathbf{u} \cdot \nabla \tilde{S}. \end{aligned} \quad (66)$$

Again using condition (23) (or equivalently, Equation (64a)), the integration of Equation (66) yields conservation of total energy,

$$\tilde{E}_{\text{tot}} := \int_V \tilde{\rho} \left(\frac{1}{2} u^2 + \tilde{T} S' \right) dV = \text{constant}, \quad (67)$$

which mathematically holds for arbitrary fluid motion obeying Equations (64) and for all magnitudes of $|\nabla \tilde{S}|$.

Note that in practice when simulating stiff systems (large $|\nabla \tilde{S}|$) numerically (e.g., Guerrero et al. 2016; Matilsky et al. 2022, 2024), the term

$$\tilde{Q}_{\text{adv}} := -\tilde{\rho} \tilde{T} \mathbf{u} \cdot \nabla \tilde{S} \quad (68)$$

may pointwise be quite large. The degree to which energy is conserved numerically may thus be limited by the precision of the condition

In the streamfunction formulation of ASH and Rayleigh for example (e.g., Clune et al. 1999; Featherstone & Hindman 2016), Equation (37) holds to near machine precision.

To summarize, necessary conditions for anelastic codes implementing background stable layers to conserve energy are (1) including the non-LBR force density

$$\tilde{\mathbf{f}}_{\text{NLBR}} := \frac{\tilde{\delta P'}}{\tilde{C}_p} \mathbf{u} \cdot \nabla \tilde{S} \quad (69)$$

in the momentum equation (i.e., not making the LBR approximation) and (2) including the non-LBR heating term

$$\tilde{Q}_{\text{NLBR}} := - \left(\frac{\tilde{T}S'}{\tilde{C}_p} + \frac{\tilde{\delta P'}}{\tilde{\rho}\tilde{C}_p} \right) \tilde{\rho} \mathbf{u} \cdot \nabla \tilde{S} \quad (70)$$

in the internal energy equation. Sufficient conditions for conserving energy will likely rest in the numerical precision with which the EGG equations (64) are actually solved.

References

- Batchelor, G. K. 1953, Quarterly Journal of the Royal Meteorological Society, 79, 224–235, doi: [10.1002/qj.49707934004](https://doi.org/10.1002/qj.49707934004)
- Braginsky, S. I., & Roberts, P. H. 1995, Geophysical & Astrophysical Fluid Dynamics, 79, 1, doi: [10.1080/03091929508228992](https://doi.org/10.1080/03091929508228992)
- Brown, B. P., Oishi, J. S., Vasil, G. M., Lecoanet, D., & Burns, K. J. 2020, The Astrophysical Journal, 902, L3, doi: [10.3847/2041-8213/abb9a4](https://doi.org/10.3847/2041-8213/abb9a4)
- Brown, B. P., Vasil, G. M., & Zweibel, E. G. 2012, The Astrophysical Journal, 756, 109, doi: [10.1088/0004-637x/756/2/109](https://doi.org/10.1088/0004-637x/756/2/109)
- Brun, A. S., Miesch, M. S., & Toomre, J. 2004, The Astrophysical Journal, 614, 1073, doi: [10.1086/423835](https://doi.org/10.1086/423835)
- Burns, K. J., Vasil, G. M., Oishi, J. S., Lecoanet, D., & Brown, B. P. 2020, Physical Review Research, 2, doi: [10.1103/physrevresearch.2.023068](https://doi.org/10.1103/physrevresearch.2.023068)
- Charney, J. G., & Ogura, Y. 1960, Journal of the Meteorological Society of Japan. Ser. II, 38, 19a–19a, doi: [10.2151/jmsj1923.38.6_19a](https://doi.org/10.2151/jmsj1923.38.6_19a)
- Clune, T., Elliott, J., Miesch, M., Toomre, J., & Glatzmaier, G. 1999, Parallel Comp., 25, 361, doi: [10.1016/s0167-8191\(99\)00009-5](https://doi.org/10.1016/s0167-8191(99)00009-5)
- Eckart, C., & Ferris, H. G. 1956, Reviews of Modern Physics, 28, 48–52, doi: [10.1103/revmodphys.28.48](https://doi.org/10.1103/revmodphys.28.48)
- Featherstone, N. A., & Hindman, B. W. 2016, The Astrophysical Journal, 818, 32, doi: [10.3847/0004-637x/818/1/32](https://doi.org/10.3847/0004-637x/818/1/32)
- Featherstone, N. A., Augustson, K. C., Aurnou, J. M., et al. 2023, The Puzzling Structure of Solar Convection: Window into the Dynamo, arXiv, doi: [10.48550/ARXIV.2305.08823](https://doi.org/10.48550/ARXIV.2305.08823)
- Gastine, T., & Wicht, J. 2012, Icar., 219, 428, doi: [10.1016/j.icarus.2012.03.018](https://doi.org/10.1016/j.icarus.2012.03.018)
- Gilman, P. A., & Glatzmaier, G. A. 1981, The Astrophysical Journal Supplement Series, 45, 335, doi: [10.1086/190714](https://doi.org/10.1086/190714)
- Glatzmaier, G. A. 1984, J. Comp. Phys., 55, 461, doi: [10.1016/0021-9991\(84\)90033-0](https://doi.org/10.1016/0021-9991(84)90033-0)
- Gough, D. O. 1969, J. Atmos. Soc., 26, 448, doi: [10.1175/1520-0469\(1969\)026<0448:taaftc>2.0.co;2](https://doi.org/10.1175/1520-0469(1969)026<0448:taaftc>2.0.co;2)
- Guerrero, G., Smolarkiewicz, P. K., de Gouveia Dal Pino, E. M., Kosovichev, A. G., & Mansour, N. N. 2016, The Astrophysical Journal, 819, 104, doi: [10.3847/0004-637x/819/2/104](https://doi.org/10.3847/0004-637x/819/2/104)
- Lantz, S. R. 1992, PhD thesis, Cornell University, United States

- Lantz, S. R., & Fan, Y. 1999, The Astrophysical Journal Supplement Series, 121, 247, doi: [10.1086/313187](https://doi.org/10.1086/313187)
- Lipps, F. B., & Hemler, R. S. 1982, Journal of the Atmospheric Sciences, 39, 2192–2210, doi: [10.1175/1520-0469\(1982\)039<2192:asaodm>2.0.co;2](https://doi.org/10.1175/1520-0469(1982)039<2192:asaodm>2.0.co;2)
- Matilsky, L. I., Brummell, N. H., Hindman, B. W., & Toomre, J. 2024, The Astrophysical Journal, 962, 189, doi: [10.3847/1538-4357/ad18b2](https://doi.org/10.3847/1538-4357/ad18b2)
- Matilsky, L. I., Hindman, B. W., Featherstone, N. A., Blume, C. C., & Toomre, J. 2022, The Astrophysical Journal Letters, 940, L50, doi: [10.3847/2041-8213/ac93ef](https://doi.org/10.3847/2041-8213/ac93ef)
- Ogura, Y., & Phillips, N. A. 1962, Journal of the Atmospheric Sciences, 19, 173, doi: [10.1175/1520-0469\(1962\)019<0173:saodas>2.0.co;2](https://doi.org/10.1175/1520-0469(1962)019<0173:saodas>2.0.co;2)
- Rogers, T. M., & Glatzmaier, G. A. 2005, The Astrophysical Journal, 620, 432–441, doi: [10.1086/423415](https://doi.org/10.1086/423415)
- Smolarkiewicz, P. K., & Prusa, J. M. 2004, in Turbulent Flow Computation (Kluwer Academic Publishers), 279–312, doi: [10.1007/0-306-48421-8_8](https://doi.org/10.1007/0-306-48421-8_8)
- Vasil, G. M., Lecoanet, D., Brown, B. P., Wood, T. S., & Zweibel, E. G. 2013, The Astrophysical Journal, 773, 169, doi: [10.1088/0004-637x/773/2/169](https://doi.org/10.1088/0004-637x/773/2/169)
- Wilczyński, F., Hughes, D. W., & Kersalé, E. 2022, Journal of Fluid Mechanics, 942, doi: [10.1017/jfm.2022.325](https://doi.org/10.1017/jfm.2022.325)