widely used and well tested libraries that can provide researchers interested in extending it with the support of a large user community. Specifically, we use the DEAL.II library [BHK07, BHK12] for meshes, finite elements and everything discretization related; the TRILINOS library [HBH+05, H+11] for scalable and parallel linear algebra; and P4EST [BWG11] for distributed, adaptive meshes. As a consequence, our code is freed of the mundane tasks of defining finite element shape functions or dealing with the data structures of linear algebra, can focus on the high-level description of what is supposed to happen, and remains relatively compact. The code will also automatically benefit from improvements to the underlying libraries with their much larger development communities. ASPECT is extensively documented to enable other researchers to understand, test, use, and extend it.

Rather than detailing the various techniques upon which ASPECT is built, we refer to the paper by Kronbichler, Heister and Bangerth [KHB12] that gives a detailed description and rationale for the various building blocks.

2.10 Simplifications of the basic equations

There are a number of common variations to equations (1)–(3) that are frequently used in the geosciences. For example, one frequently finds references to the anelastic liquid approximation (ALA), truncated anelastic liquid approximation (TALA), and the Boussinesq approximation (BA). These can all be derived from the basic equations (1)–(3) via various approximation, though there are a number of pitfalls if one wanted to use them in ASPECT. We will discuss these in the following. Since they are typically only provided considering velocity, pressure and temperature, we will in the following omit the dependence on the compositional fields used in previous sections, though this dependence can easily be added back into the equations stated below. A detailed discussion of the approximations introduced below can also be found in [KLv+10].

The three approximations mentioned all start by writing the pressure and temperature as the sum of a (possibly depth dependent) reference value plus a perturbation, i.e., we will write

$$\begin{aligned} p(\mathbf{x},t) &= \bar{p}(z) + p'(\mathbf{x},t), \\ T(\mathbf{x},t) &= \bar{T}(z) + T'(\mathbf{x},t). \end{aligned}$$

Here, barred quantities are reference states and may depend on the depth z (not necessarily the third component of $\mathbf{x}$) whereas primed quantities are the spatially and temporally variable deviations of the temperature and pressure fields from this reference state. In particular, the reference pressure is given by solving the hydrostatic equation,

$$\nabla \bar{p} = \bar{\rho}\mathbf{g},$$

where $\bar{\rho} = \rho(\bar{p}, \bar{T})$ is a *reference density* that depends on depth. $\bar{T}$ is chosen as an adiabatic profile accounting for the fact that the temperature increases as the pressure increases, but also taking into account thermal diffusion effects. It also depends on depth. With these definitions, equations (1)–(2) can equivalently be written as follows:

$$-\nabla\cdot\left[2\eta\left(\varepsilon(\mathbf{u}) - \frac{1}{3}(\nabla\cdot\mathbf{u})\mathbf{1}\right)\right] + \nabla p' = (\rho-\bar{\rho})\mathbf{g} \qquad \text{in } \Omega, \tag{17}$$

$$\nabla\cdot(\rho\mathbf{u}) = 0 \qquad \text{in } \Omega. \tag{18}$$

Likewise, when omitting entropic effects, the temperature equation reads as

$$\begin{aligned} &\rho C_p\left(\frac{\partial T}{\partial t} + \mathbf{u}\cdot\nabla T\right) - \nabla\cdot k\nabla T \\ &\qquad = \rho H + 2\eta\left(\varepsilon(\mathbf{u}) - \frac{1}{3}(\nabla\cdot\mathbf{u})\mathbf{1}\right) : \left(\varepsilon(\mathbf{u}) - \frac{1}{3}(\nabla\cdot\mathbf{u})\mathbf{1}\right) + \alpha T\left(\mathbf{u}\cdot\nabla p\right) \quad \text{in } \Omega. \end{aligned} \tag{19}$$

Scott: I don' see this dis cussed i [KLv+10], bu one needs t make $\bar{T}$ als take into ac count diffusio so that it dis appears fron the tempera ture equatio in the ALA.

Starting from these equations, the approximations discussed in the next few subsections make use of the fact that for the flows for which these approximations are valid, the perturbations p', T' are much smaller than typical values of the reference quantities $\bar{p}, \bar{T}$. The equations for these approximations are almost always given in terms of non-dimensionalalized quantities. We will for now stick with the dimensional form because it expresses in a clearer way the approximations that are made. The non-dimensionalization can then be done on each of the forms below separately.

2.10.1 The anelastic liquid approximation (ALA)

The *Anelastic Liquid Approximation (ALA)* is based on two assumptions. First, that the density variations $\rho(p,T) - \bar{\rho}$ are small and in particular can be accurately described by a Taylor expansion:

$$\rho(p,T) \approx \bar{\rho} + \frac{\partial \rho(\bar{p},\bar{T})}{\partial p} p' + \frac{\partial \rho(\bar{p},\bar{T})}{\partial T} T'.$$

Here, $\frac{\partial \rho(\bar{p},\bar{T})}{\partial T}$ is related to the thermal expansion coefficient, and $\frac{\partial \rho(\bar{p},\bar{T})}{\partial p}$ to the compressibility.

The second assumotion is that the variation of the density from the reference density can be neglected in the mass balance and temperature equations. This yields the following system of equations for the velocity and pressure equations, using the definition of the adiabatic hydrostatic pressure field $\bar{p}$:

$$-\nabla \cdot \left[2\eta \left(\varepsilon(\mathbf{u}) - \frac{1}{3}(\nabla \cdot \mathbf{u})\mathbf{1} \right) \right] + \nabla p' = \left(\frac{\partial \rho(\bar{p},\bar{T})}{\partial p} p' + \frac{\partial \rho(\bar{p},\bar{T})}{\partial T} T' \right) \mathbf{g} \qquad \text{in } \Omega, \tag{20}$$

$$\nabla \cdot (\bar{\rho}\mathbf{u}) = 0 \qquad \text{in } \Omega. \tag{21}$$

Timo: need to figure out how to solve (21) need $\nabla\rho$!
Scott: This does not match (18) in [KLv+10]

For the temperature equation we arrive at the following:

$$\bar{\rho} C_p \left(\frac{\partial T'}{\partial t} + \mathbf{u} \cdot \nabla T' \right) - \nabla \cdot k \nabla T' = \rho H + 2\eta \left(\varepsilon(\mathbf{u}) - \frac{1}{3}(\nabla \cdot \mathbf{u})\mathbf{1} \right) : \left(\varepsilon(\mathbf{u}) - \frac{1}{3}(\nabla \cdot \mathbf{u})\mathbf{1} \right) + \alpha T' (\mathbf{u} \cdot \nabla p) \quad \text{in } \Omega. \tag{22}$$

Timo: this is not quite correct, you need reference density and simplified adiabatic heating to get:

$$\bar{\rho} C_p \left(\frac{\partial T'}{\partial t} + \mathbf{u} \cdot \nabla T' \right) - \nabla \cdot k \nabla T' = \bar{\rho} H + 2\eta \left(\varepsilon(\mathbf{u}) - \frac{1}{3}(\nabla \cdot \mathbf{u})\mathbf{1} \right) : \left(\varepsilon(\mathbf{u}) - \frac{1}{3}(\nabla \cdot \mathbf{u})\mathbf{1} \right) + \alpha \bar{\rho} T' (\mathbf{u} \cdot \mathbf{g}) \quad \text{in } \Omega. \tag{23}$$

2.10.2 The truncated anelastic liquid approximation (TALA)

The *Truncated Anelastic Liquid Approximation (TALA)* further simplifies the ALA by assuming that the variation of the density due to pressure variations is small, i.e., that

$$\rho(p,T) \approx \bar{\rho} + \frac{\partial \rho(\bar{p},\bar{T})}{\partial T} T'.$$

This does not mean that the density is not pressure dependent – it will, for example, continue to be depth dependent because the hydrostatic pressure grows with depth. It simply means that the deviations from the reference pressure are assumed to be so small that they do not matter in describing the density. Because the pressure variation is induced by the flow field (the static component pressure is already taken care of by

the hydrostatic pressure), this assumption in essence means that we assume the flow to be very slow, even beyond the earlier assumption that we can neglect inertial terms when deriving (1)–(2).

This further assumption then transforms (20)–(21) into the following equations:

$$-\nabla\cdot\left[2\eta\left(\varepsilon(\mathbf{u})-\frac{1}{3}(\nabla\cdot\mathbf{u})\mathbf{1}\right)\right]+\nabla p' = \frac{\partial\rho(\bar{p},\bar{T})}{\partial T}T'\mathbf{g} \qquad \text{in } \Omega, \tag{24}$$

$$\nabla\cdot(\bar{\rho}\mathbf{u}) = 0 \qquad \text{in } \Omega. \tag{25}$$

2.10.3 The extended Boussinesq approximation (EBA) and Boussinesq approximation (BA)

If we further assume that the reference density is constant, $\bar{\rho}(\bar{p},\bar{T})=\rho_0$, – in other words, the density does not only not depend on the pressure variations p' as assumed in the TALA, but also does not depend on the much larger hydrostatic pressure $\bar{p}$ nor on the reference temperature $\bar{T}$ – then we can further simplify the mass conservation equations of the TALA to $\nabla\cdot\mathbf{u}=0$. We then obtain the following set of equations that also uses the incompressibility in the definition of the viscous strain:

$$-\nabla\cdot[2\eta\varepsilon(\mathbf{u})]+\nabla p' = \frac{\partial\rho(\bar{p},\bar{T})}{\partial T}T'\mathbf{g} \qquad \text{in } \Omega, \tag{26}$$

$$\nabla\cdot\mathbf{u} = 0 \qquad \text{in } \Omega. \tag{27}$$

These equations are identical in form for the *Boussinesq approximation (BA)* which only differs from the EBA in neglecting compressibility effects in the temperature equation (shear heating and simplified adiabatic heating).

2.10.4 The Boussinesq approximation: Incompressibility

The original Boussinesq approximation assumes that the density can be considered constant in all occurrences in the equations with the exception of the buoyancy term on the right hand side of (1). The primary result of this assumption is that the continuity equation (2) will now read

$$\nabla\cdot\mathbf{u}=0.$$

This makes the equations *much* simpler to solve: First, because the divergence operation in this equation is the transpose of the gradient of the pressure in the momentum equation (1), making the system of these two equations symmetric. And secondly, because the two equations are now linear in pressure and velocity (assuming that the viscosity η and the density ρ are considered fixed). In addition, one can drop all terms involving $\nabla\cdot\mathbf{u}$ from the left hand side of the momentum equation (1) as well as from the shear heating term on the right hand side of (3); while dropping these terms does not affect the solution of the equations, it makes assembly of linear systems faster. In addition, in the incompressible case, one needs to neglect the adiabatic heating term $\frac{\partial\rho}{\partial T}T\mathbf{u}\cdot\mathbf{g}$ on the right hand side of (3).

From a physical perspective, the assumption that the density is constant in the continuity equation but variable in the momentum equation is of course inconsistent. However, it is justified if the variation is small since the momentum equation can be rewritten to read

$$-\nabla\cdot 2\eta\varepsilon(\mathbf{u})+\nabla p_d = (\rho-\rho_0)\mathbf{g},$$

where p_d is the *dynamic* pressure and ρ_0 is the constant reference density. This makes it clear that the true driver of motion is in fact the *deviation* of the density from its background value, however small this value is: the resulting velocities are simply proportional to the density variation, not to the absolute magnitude of the density.

As such, the Boussinesq approximation can be justified. On the other hand, given the real pressures and temperatures at the bottom of the earth mantle, it is arguable whether the density can be considered to be almost constant. Most realistic models predict that the density of mantle rocks increases from somewhere

around 3300 at the surface to over 5000 kilogram per cubic meters at the core mantle boundary, due to the increasing lithostatic pressure. While this appears to be a large variability, if the density changes slowly with depth, this is not in itself an indication that the Boussinesq approximation will be wrong. To this end, consider that the continuity equation can be rewritten as $\frac{1}{\rho}\nabla\cdot(\rho\mathbf{u})=0$, which we can multiply out to obtain

$$\nabla\cdot\mathbf{u}+\frac{1}{\rho}\mathbf{u}\cdot\nabla\rho=0.$$

The question whether the Boussinesq approximation is valid is then whether the second term (the one omitted in the Boussinesq model) is small compared to the first. To this end, consider that the velocity can change completely over length scales of maybe 10 km, so that $\nabla\cdot\mathbf{u}\approx\|u\|/10\text{km}$. On the other hand, given a smooth dependence of density on pressure, the length scale for variation of the density is the entire earth mantle, i.e., $\frac{1}{\rho}\mathbf{u}\cdot\nabla\rho\approx\|u\|0.5/3000\text{km}$ (given a variation between minimal and maximal density of 0.5 times the density itself). In other words, for a smooth variation, the contribution of the compressibility to the continuity equation is very small. This may be different, however, for models in which the density changes rather abruptly, for example due to phase changes at mantle discontinuities.

In summary, models that use the approximation of incompressibility solve the following set of equations instead of (1)–(3):

$$-\nabla\cdot[2\eta\varepsilon(\mathbf{u})]+\nabla p=\rho\mathbf{g} \qquad \text{in } \Omega, \tag{28}$$

$$\nabla\cdot\mathbf{u}=0 \qquad \text{in } \Omega, \tag{29}$$

$$\rho C_p\left(\frac{\partial T}{\partial t}+\mathbf{u}\cdot\nabla T\right)-\nabla\cdot k\nabla T=\rho H+2\eta\varepsilon(\mathbf{u}):\varepsilon(\mathbf{u}) \qquad \text{in } \Omega, \tag{30}$$

where the coefficients $\eta,\rho,\mathbf{g},C_p$ may possible depend on the solution variables.

Note: As we will see in Section 7, it is easy to add new material models to ASPECT. Each model can decide whether it wants to use the Boussinesq approximation or not. The description of the models in Section 5.66 also gives an answer which of the models already implemented uses the approximation or considers the material sufficiently compressible to go with the fully compressible continuity equation.

2.10.5 Almost linear models

A further simplification can be obtained if one assumes that all coefficients with the exception of the density do not depend on the solution variables but are, in fact, constant. In such models, one typically assumes that the density satisfies a relationship of the form $\rho=\rho(T)=\rho_0(1-\beta(T-T_0))$ with a small thermal expansion coefficient β and a reference density ρ_0 that is attained at temperature T_0. Since the thermal expansion is considered small, this naturally leads to the following variant of the Boussinesq model discussed above:

$$-\nabla\cdot[2\eta\varepsilon(\mathbf{u})]+\nabla p=\rho_0(1-\beta(T-T_0))\mathbf{g} \qquad \text{in } \Omega,$$

$$\nabla\cdot\mathbf{u}=0 \qquad \text{in } \Omega,$$

$$\rho C_p\left(\frac{\partial T}{\partial t}+\mathbf{u}\cdot\nabla T\right)-\nabla\cdot k\nabla T=\rho H+2\eta\varepsilon(\mathbf{u}):\varepsilon(\mathbf{u}) \qquad \text{in } \Omega,$$

If the gravitational acceleration $\mathbf{g}$ results from a gravity potential φ via $\mathbf{g} = -\nabla\varphi$, then one can rewrite the equations above in the following, commonly used form:[5]

$$-\nabla \cdot [2\eta\varepsilon(\mathbf{u})] + \nabla p_d = -\beta\rho_0 T\mathbf{g} \qquad \text{in } \Omega, \tag{31}$$

$$\nabla \cdot \mathbf{u} = 0 \qquad \text{in } \Omega, \tag{32}$$

$$\rho C_p \left(\frac{\partial T}{\partial t} + \mathbf{u}\cdot\nabla T\right) - \nabla \cdot k\nabla T = \rho H + 2\eta\varepsilon(\mathbf{u}) : \varepsilon(\mathbf{u}) \qquad \text{in } \Omega, \tag{33}$$

where $p_d = p + \rho_0(1+\beta T_0)\varphi$ is the dynamic pressure, as opposed to the total pressure $p = p_d + p_s$ that also includes the hydrostatic pressure $p_s = -\rho_0(1+\beta T_0)\varphi$. Note that the right hand side forcing term in (31) is now only the deviation of the gravitational force from the force that would act if the material were at temperature T_0.

Under the assumption that all other coefficients are constant, one then arrives at equations in which the only nonlinear terms are the advection term, $\mathbf{u}\cdot\nabla T$, and the shear friction, $2\eta\varepsilon(\mathbf{u}) : \varepsilon(\mathbf{u})$, in the temperature equation (33). This facilitates the use of a particular class of time stepping scheme in which one does not solve the whole set of equations at once, iterating out nonlinearities as necessary, but instead in each time step solves first the Stokes system with the previous time step's temperature, and then uses the so-computed velocity to solve the temperature equation. These kind of time stepping schemes are often referred to as *operator splitting* methods. A particular operator splitting method, used in earlier ASPECT versions, solves first the Stokes equations and then uses a semi-explicit time stepping method for the temperature equation where diffusion is handled implicitly and advection explicitly; this algorithm is often called *IMPES* (it originated in the porous media flow community, where the acronym stands for *Im*plicit *P*ressure, *E*xplicit *S*aturation) and is explained in more detail in [KHB12]. However, since then the algorithm in ASPECT has been rewritten to use an implicit time stepping algorithm also for the temperature equation because this allows to use larger time steps.

2.10.6 Compressible models

In the compressible case, the conservation of mass equation in equation (2) becomes $\nabla\cdot(\rho\mathbf{u}) = 0$ instead of $\nabla\cdot\mathbf{u} = 0$, which is nonlinear and no longer symmetric to the ∇p term in the force balance equation (1), making solving and preconditioning the resulting linear and nonlinear systems difficult. To make this work in ASPECT, we consequently reformulate this equation. Dividing by ρ and applying the product rule of differentiation gives

$$\frac{1}{\rho}\nabla\cdot(\rho\mathbf{u}) = \nabla\cdot\mathbf{u} + \frac{1}{\rho}\nabla\rho\cdot\mathbf{u}.$$

We will now make two basic assumptions: First, the variation of the density $\rho(p,T,\mathbf{x},\mathfrak{c})$ is dominated by the dependence on the (total) pressure; in other words, $\nabla\rho \approx \frac{\partial\rho}{\partial p}\nabla p$. This assumption is primarily justified by the fact that, in the Earth mantle, the density increases by at least 50% between Earth crust and the core-mantle boundary due to larger pressure there. Secondly, we assume that the pressure is dominated by the static pressure, which can be written as $\nabla p \approx \nabla p_s \approx \rho\mathbf{g}$. This is essentially motivated by the slowness of the movement in the Earth or, alternatively, based on the fact that the viscosity is so large. This finally allows us to write

$$\frac{1}{\rho}\nabla\rho\cdot\mathbf{u} \approx \frac{1}{\rho}\frac{\partial\rho}{\partial p}\nabla p\cdot\mathbf{u} \approx \frac{1}{\rho}\frac{\partial\rho}{\partial p}\nabla p_s\cdot\mathbf{u} \approx \frac{1}{\rho}\frac{\partial\rho}{\partial p}\rho\mathbf{g}\cdot\mathbf{u}$$

so we get

$$\nabla\cdot\mathbf{u} = -\frac{1}{\rho}\frac{\partial\rho}{\partial p}\rho\mathbf{g}\cdot\mathbf{u} \tag{34}$$

where $\frac{1}{\rho}\frac{\partial\rho}{\partial p}$ is often referred to as the compressibility.

[5] Note, however, that ASPECT does not solve the equations in the form given in (31)–(33). Rather, it takes the original form with the real density, not the variation of the density. That said, you can use the formulation (31)–(33) by implementing a material model (see Section 7.3.1) in which the density in fact has the form $\rho(T) = \beta\rho_0 T$ even though this is not physical.

In the implementation used in ASPECT, this equation replaces (2). It has the advantage that it retains the symmetry of the Stokes equations if we can treat the right hand side of (34) as known. We do so by evaluating ρ and $\mathbf{u}$ using the solution from the last time step (or values extrapolated from previous time steps).

2.11 Approximations

1. Define adiabat as constant for BA models.
   2. given the density:

$$\rho(depth, p, T, \dots) = A(depth) + B(p - p_{ref}) + C(T - T_{ref}) + \dots$$

with an adiabat p_{ref} and T_{ref} as a function of depth, introduce a density approximation $\hat{\rho} = \rho(depth, p_{ref}, T_{ref})$ by default computed using adiabatic temperature and pressure (or one of them adiabatic) but can be overwritten.

Formulation:

- ALA: $\hat{\rho} = \rho(z)$, $T_{ref} = T(z)$, adiabatic heating (simplified), shear heating; mass density = implicit adiabatic, buoyancy = full, temperature = adiabatic
- "TALA": like ALA, except buoyancy approximation = adiabatic pressure
- "EBA": $\rho = 1$, $T_{ref} = 0$, still with adiabatic heating (simplified), shear heating; mass density = incompressible, temperature = adiabatic, buoyancy = adiabatic pressure;

  Assert: material.compressible=false
- "BA": EBA, but turn off shear heating and adiabatic heating.
- "current": full/full/full
- "better?"

2.11.1 Stress tensor approximation

The stress tensor in the conservation of momentum (diffusion) is either

$$\nabla \cdot \tau = \nabla \cdot \left[2\eta \left(\varepsilon(\mathbf{u}) - \frac{1}{3}(\nabla \cdot \mathbf{u})\mathbf{1} \right) \right]$$

or

$$\nabla \cdot \tau = \nabla \cdot [2\eta\varepsilon(\mathbf{u})]$$

Currently controlled by material_model.is_compressible.

2.11.2 Mass density approximation

simplify $\nabla \cdot (\rho u) = 0$ in one of the following ways:

- "ask material model": use "full" if is_compressible and "incompressible" otherwise
- "fuller?":

$$\nabla \cdot \mathbf{u} = -\frac{1}{\rho}\frac{\partial \rho}{\partial p}\rho\mathbf{g} \cdot \mathbf{u} - \alpha\rho\mathbf{g} \cdot \mathbf{u}$$

- "full":

$$\nabla \cdot \mathbf{u} = -\frac{1}{\rho}\frac{\partial \rho}{\partial p}\rho\mathbf{g} \cdot \mathbf{u}$$

- "adiabatic":

$$\nabla \cdot \mathbf{u} = -\frac{1}{\rho}\frac{\partial \rho}{\partial p}\hat{\rho}\mathbf{g} \cdot \mathbf{u}$$

- "implicit adiabatic":

$$\nabla \cdot \mathbf{u} + \frac{1}{\rho}\frac{\partial \rho}{\partial p}\hat{\rho}\mathbf{g} \cdot \mathbf{u} = 0$$

- "incompressible":

$$\nabla \cdot \mathbf{u} = 0$$

question: what if ρ is a function of depth and we could give the exact $\nabla\rho$?
"full" is probably inconsistent because it doesn't use the reference density

2.11.3 Buoyancy density approximation

we write the RHS of conservation of momentum ρg as

- "full": ρg
- "adiabatic pressure": $\rho(depth, p_{ref}, T)g$

2.11.4 Temperature density approximation

temperature equation:

$$\begin{aligned}\rho C_p \left(\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T\right) - \nabla \cdot k\nabla T = \rho H \\ + 2\eta\left(\varepsilon(\mathbf{u}) - \frac{1}{3}(\nabla \cdot \mathbf{u})\mathbf{1}\right) : \left(\varepsilon(\mathbf{u}) - \frac{1}{3}(\nabla \cdot \mathbf{u})\mathbf{1}\right) \\ + \alpha T\left(\mathbf{u} \cdot \nabla p\right)\end{aligned}$$

- "full": use full density
- "adiabatic": use density approximation in all terms

Note: simplified adiabatic heating allows different way of handling gravity

2.12 Questions

- how do people solve $\nabla \cdot (\rho u) = 0$ with $\rho = \rho(z)$?
- Should we allow very simple setups (pick all constants as 1)? Then we would need things like prefactors for shear heating. Otherwise picking constants like viscosity are somewhat more complicated...
- The current material.is_compressbile setting is needed for BA/EBA to get the simplified diffusion operator. Should this be a separate setting?
- Should we model $\rho(z)$ using a linear adiabatic pressure and setting compressibilty to $\nabla\rho$ or should we introduce the depth derivative explicitly?
- Can we improve "full" further and include additional terms?
- for WB: eqn (22) needs reference density and simplified adiabatic heating