Review of van Zelst et al. https://doi.org/10.5194/se-2021-14
Paul Tackley 12.05.2021

This is a well-written and useful article that will serve as a broad introduction to newcomers in numerical modelling. It seems very long for an SE article but as a result fairly complete; there's always a balance between length and level of detail. All that is left for a reviewer to do is point out minor corrections and suggested additions.

As a general point, many points are supported by somewhat random citations (i.e. there are many articles that could be cited, they pick just one or two) while others are supported by exhaustive lists. The somewhat random citations could do with an (e.g.) around them to show that they are not supposed to be complete lists.

l 35-37: Solomatov and Reese (2008) is a good reference to support the importance of grain dynamics, but not magma dynamics. How about (e.g. Keller et a. 2013) to support this.

l 109-110: For this citation an (e.g. …) is certainly needed because there are a huge number of 3D seismic tomography studies dating back to the pioneering works of Dziewonski (e.g. Dziewonski, 1984) - it's really random which ones are chosen here.

l 111-113. Adding some earlier references here might be appropriate, otherwise readers get the impression that these things were developed recently. Adjoint methods: (Liu and Gurnis, 2008; Burstedde et al., 2009). Data assimilation: (Bunge et al., 2003; Ismail-Zadeh et al., 2004; Hier-Majumder et al., 2005).

l 140: Update Ricard "Treatise on Geophysics" reference to the 2nd edition?

l 146: Give a more complete/rigorous definition of a continuum. For example something like what it says in Wikipedia "… certain physical phenomena can be modeled assuming the materials exist as a *continuum, meaning the matter in the body is continuously distributed and fills the entire region of space it occupies*. A continuum is a body that can be continually sub-divided into infinitesimal elements with properties being those of the bulk material."

l 183-185. Curbelo et al. (2019)'s relaxation time analysis doesn't apply to the mantle because they are considering an **ideal gas**, i.e. with the ideal gas equation of state (this is where their equation 4.4 comes from), which does not include elasticity. If you do the same analysis for a visco-elastic material like rock using an appropriate equation of state such as Birch-Murnaghan, then elasticity will be the dominant mechanism to even out pressure variations, i.e. if you apply a pressure perturbation in one place then it will propagate not via viscous relaxation but via elastic (seismic) waves, which are relatively fast. Of course even with seismic waves there is a component of viscous relaxation - this is why there is attenuation - so it would be possible to derive a time-scale for elastic waves to die out due to attenuation. A related point: compression of rock as it descends through the mantle does not occur by viscous relaxation, it occurs by elastic compression.

l 248: This simplification is only correct if density and Cp are constant. (as the authors note later, but it should be stated here as well).

l 264-265: Here their expression of S_2 is the general case of pressure varying in any direction whereas what they write is "…the dominant pressure variation…is the effect of the lithospheric pressure increase with depth". In this case the expression can be simplified: $\nabla p = \rho g \hat{\hat{z}}$ where z=vertical, positive downwards, hence $S_2 = \alpha T \rho g v_z$. This is the expression that is normally used in mantle convection codes.

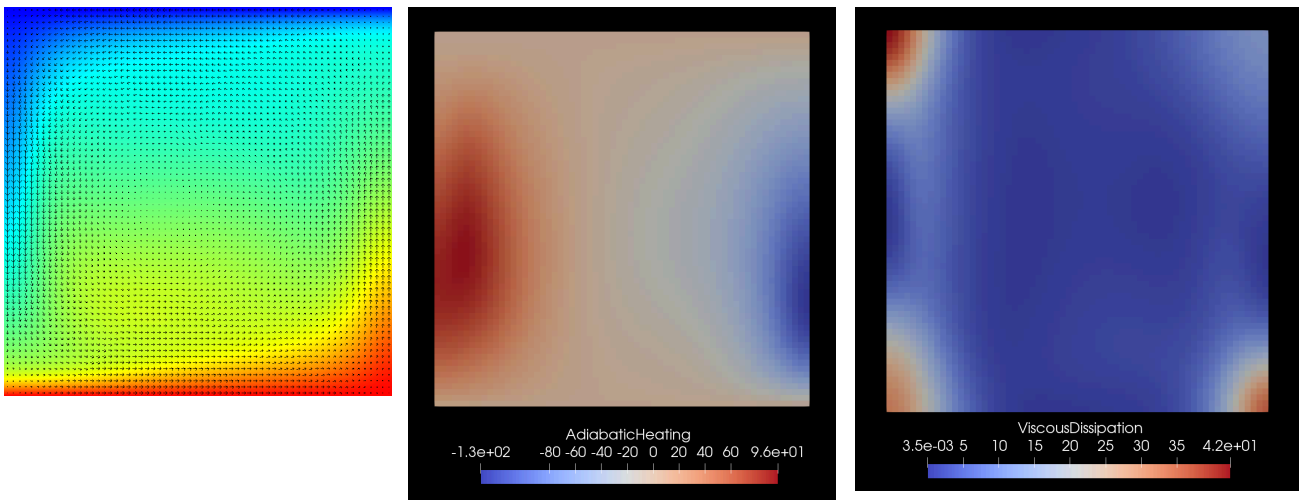
l 289. When considering the various approximations, the order-of-magnitude fractional density error that is expected from ignoring thermally-related density variations is $\alpha \Delta T$, which is O(1%) (α ~10^{-5}, dT~10^3), so errors of this magnitude are considered "par for the course". This is the magnitude of error you get in thermal boundary layers / slabs / plumes with the Boussinesq approximation, and also with the Anelastic Liquid Approximations. It does not "invalidate" either

approximation, it is rather the magnitude of error that is accepted in making the approximation. (an aside: the Anelastic Approximation *is* invalidated by large T deviations from the adiabat, this is why the Anelastic *Liquid* Approximation is preferred for mantle dynamics).

l 294-295. In the convecting mantle, the magnitude of error in ignoring dynamic pressure in the density calculation, hence the resulting energy imbalance, is very small for realistic Earth parameters.
- The magnitude of dynamic pressure can be estimated by how much dynamic topography is generated by mantle flow: ~a few km. Compare this to the mantle depth of ~3000 km - the pressure error (dynamic/lithostatic) is thus in the range 0.1-1%.
- Leng and Zhong (2008) found resulting energy imbalances of up to ~few % because their experiments were at low Rayleigh numbers of 10^4-10^5. Stress and dynamic pressure decrease with increasing Ra as roughly $Ra^{-1/3}$ (assuming that Ra increases because viscosity decreases) so extrapolated to Earth-like Ra of 10^7-10^8, the expected energy imbalance is less than 1%.
- In deciding whether this small error is worth doing something about (i.e. using ALA instead of TALA), one must also consider the accuracy of the numerical pressure solution. For example, a well-documented issue with the finite volume (staggered-grid finite difference) discretisation is artificial pressure overshoots at viscosity jumps (e.g. Deubelbeiss and Kaus, 2008). If large, localised viscosity contrasts exist inside the modelled mantle, this “numerical” dynamic pressure is not something that one would want to use in calculating physical properties - it might result in larger errors than simply ignoring dynamic pressure.

l 308: For the extended Boussineq approximation it is not correct that “adiabatic heating leads to artificial generation of energy in the model”.
(i) Adiabat heating removes energy from the system, it does not generate energy. “Heating” is a bit of a misnomer. This is because “heating” only applies to sinking material: rising material cools. Furthermore, rising material cools at a more rapid rate than sinking material is heated, because the adiabatic gradient is proportional to T, i.e. hot adiabats are steeper than cold adiabats. There is an equal amount of rising material and sinking material (mass conservation). The result of all this is that cooling of rising material exceeds heating of sinking material, so the AH term removes heat from the system. In equilibrium, heat loss due to adiabatic heating is exactly balanced by heat input due to viscous dissipation: the volume-integrals of the two terms are equal and opposite (e.g. Jarvis and McKenzie, 1980).
(ii) I did a quick test EB calculation using StagYY: 1x1 box, $Ra=10^5$ (based on total temperature drop), all properties constant=1, dissipation number=1. This is large Di - it means temperatures increase by a factor of exp(1.0)=2.7 from top to bottom - more than in the Earth. Despite this, integral(adiabat heating) = -integral(viscous dissipation), top heat flux = bottom heat flux, there is no energy imbalance.

```
Viscous dissipation: min =  3.52E-03 ; mean =  3.96E+00 ; max =  4.24E+01
Adiabatic heating  : min = -1.33E+02 ; mean = -3.96E+00 ; max =  9.63E+01
  Top flux and Nu =     4.860    4.860 ; Bot flux and Nu =     4.860    4.860
```



l 340: Use passive tense: “writes” -> “is written”

l 426: “strain rate increases” -> “strain or strain rate increases”

l 459 Either no comma or two commas: “…other variables like chemical composition…” or “other variables, like chemical composition, is…”.

l 519-520. There is some confusion nowadays over the difference between finite difference and finite volume discretizations.
- The staggered-grid (“conservative”) finite difference discretisation used in codes like StagYY, LaMem, I*VIS etc. is an example of a finite volume discretisation, and is normally referred to as such in the broader numerical simulation community and in many papers in our community (e.g. Ogawa et al. 1991; Trompert and Hansen, 1996; Shahnas et al. 2011). So, staggered-grid finite differences = finite volume, but
- it is also possible to have non-staggered grid finite differences that cannot be described as finite volume (e.g. several of the codes in Blankenbach et al. 1989),
- or unstructured-grid finite volume codes that cannot be described as finite difference (e.g. Hüttig and Stemmer, 2008).
- In conclusion, I suggest adding a clarification sentence, for example “We note that the commonly-used staggered grid finite difference discretisation is an example of a finite volume discretisation”.

l 556-557: This explains the “difference” part of “finite difference”; why not also explain the “finite” part, which comes from the mathematical definition of a derivative as being a limit as h, the difference in coordinate, tends to 0: $$\frac{df}{dx} = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

being replaced by a formula in which h is **finite**: $$\frac{df}{dx} \approx \frac{f(a+h) - f(a)}{h}$$

l 584: 10^24 -> 10^{24}.

l 606-607: The most common iterative method used in geodynamic codes is the **multigrid method**. This is what is used in ASPECT, CitCom, StagYY, I3ELVIS, LaMEM, TERRA, etc.

l 620: “top 500” -> “TOP500 list (https://www.top500.org)"

Figure 5: A nice figure, but in the MPI part, bottom right, the 4 processors are all on the same node so actually MPI is not necessary - OpenMP could be used instead. It might be more illustrative to have/use only 1 CPU per node, so that the different nodes are communicating over the network.

l 664-667: The sentences on advection methods need to be rewritten/expanded/clarified.
(i) A distinction should be made between methods designed to treat discontinuities such as a free surface, and methods designed to treat smoothly-varying fields such as temperature.
(ii) The methods they mention (level set, marker chain, volume of fluid) are designed to treat discontinuities. Actually these work well and are widely used. I don’t think there’s anything “notoriously difficult” here. If all you need to track is one discontinuity, then using one of these methods makes more sense than placing particles everywhere in the domain.
(iii) Tracking temperature or other smoothly-varying fields is easier and many methods have been developed over many decades to advect fields while minimising artefacts such as numerical diffusion and dispersion (ripples). These methods work well and are in common use in a variety of fluid dynamics fields including mantle dynamics. For example, finite-element codes ASPECT and Citcom* use the streamline upwind/Petrov-Galerkin (SUPG) method. For finite-volume codes a variety of methods are available such as TVD (Total Variation Diminishing), FCT (Flux-Corrected Transport) and MPDATA (Multidimensional Positive Definite Advection Transport Algorithm), to name a few, and all of these are conservative. Again, if all you need to track is a smoothly-varying

field like temperature, it makes more sense to use one of these methods than to fill the domain with particles.
(iv) For compositional variations that exist everywhere in the domain, or multiple composition fields, it's best to use particles, although you can use one of the methods in (iii) - there isn't a single "compositional fields method" here - actually many possibilities. Some of the field-based methods are tested against particle methods in van Keken et al. (1997) and Tackley and King (2003).

l 670: Other disadvantages of particle-based methods are the introduction of artificial noise and the lack of conservation of advected quantities (when averaged to the grid).

l 670: Particle-based methods are not difficult to parallelise: each process holds the grid cells and particles in its subdomain and then each time step after advection, particles that have crossed to other subdomains are communicated to those subdomains. The only potential issue comes when the subdomains have different volumes, for example as a result of adaptive grid refinement - this can lead to load imbalance (i.e. different #particles in different subdomains).

l 726: Another useful example is Kramer et al. (2020).

l 745: Two more useful community benchmark papers: Travis et al. (1990), which was the US equivalent of Blankenbach et al (1989), and Busse et al. (1994), one of the few 3D benchmark papers.

l 746: Zhong et al. (2008) is not really a community benchmark because it is only testing one code; we don't get an idea of how different codes/methods compare. If you're going to include this why not include Tackley and King (2003) where we at least tested 2 codes (and several methods of treating composition).

l 907-916: Mention that periodic boundary conditions are the natural choice for global simulations.

l 993: Initial conditions do not always determine the model outcome: in long-term simulations or ones obtaining steady-state solutions it can often be that the initial condition is "forgotten". Therefore it would be more accurate to write "initial conditions **can often** determine the model outcome".

l 1026: add Duretz et al. (2011) - for example this is what I am using. It's basically like Kaus et al. (2010) but for the finite volume discretisation.

l 1068 "percentage" -> "fractional"

Section 9: There are a lot of underlined words in this section. I get the impression they are supposed to be hyperlinks, but when I click on them (in Adobe Acrobat Reader) nothing happens. Replace them with normal referencing.

Appendix A: Could list a few more references for numerical modelling, especially ones explaining finite element methods - for example the Zhong et al Treatise chapter or the book by Simpson (2017).

———-

REFERENCES

Bunge, H.-P., C. R. Hagelberg, and B. J. Travis (2003), Mantle circulation models with variational data assimilation: inferring past mantle flow and structure from plate motion histories and seismic tomography, Geophysical Journal International, 152(2), 280-301.

Burstedde, C., O. Ghattas, G. Stadler, T. Tu, and L. C. Wilcox (2009), Parallel scalable adjoint-based adaptive solution of variable-viscosity Stokes flow problems, Computer Methods in Applied Mechanics and Engineering, 198(21-26), 1691-1700.

Busse, F. H., et al. (1994), 3d convection at infinite Prandtl number in Cartesian geometry - a benchmark comparison, Geophys. Astrophys. Fluid Dyn., 75(1), 39-59.

Deubelbeiss, Y., and B. J. P. Kaus (2008), A comparison of Eulerian and Lagrangian numerical techniques for the Stokes equation in the presence of strongly varying viscosity, Phys. Earth Planet. Int., 171(1-4), 92-111.

Duretz, T., D. A. May, T. V. Gerya, and P. J. Tackley (2011), Discretisation errors and free surface stabilization in the finite difference and marker-in-cell method in geodynamic applications: A numerical study, Geochem. Geophys. Geosyst., 12(Q07004), doi:10.1029/2011GC003567

Dziewonski, A. M. (1984), Mapping the Lower Mantle - Determination Of Lateral Heterogeneity In P-Velocity Up to Degree and Order-6, Journal Of Geophysical Research, 89(B7), 5929-.

Huettig, C., and K. Stemmer (2008), The spiral grid: A new approach to discretize the sphere and its application to mantle convection, Geochem. Geophys. Geosyst., 9(2), doi: 10.1029/2007GC001581.

Ismail-Zadeh, A., G. Schubert, I. Tsepelev, and A. Korotkii (2004), Inverse problem of thermal convection: numerical approach and application to mantle plume restoration, Physics of the Earth & Planetary Interiors, 145(1-4), 99-114.

Jarvis, G. T., and D. P. McKenzie (1980), Convection in a compressible fluid with infinite Prandtl number, J. Fluid Mech., 96, 515-583.

Keller, T., D. A. May, and B. J. P. Kaus (2013), Numerical modelling of magma dynamics coupled to tectonic deformation of lithosphere and crust, Geophysical Journal International, 195(3), 1406-1442, doi:10.1093/gji/ggt306.

Kramer, S. C., D. R. Davies, and C. R. Wilson (2020, in review), Analytical solutions for mantle flow in cylindrical and spherical shells, Geoscientific Model Development, doi:10.5194/gmd-2020-194.

Liu, L., and M. Gurnis (2008), Simultaneous inversion of mantle properties and initial conditions using an adjoint of mantle convection, Journal of Geophysical Research: Solid Earth, 113(B8), n/a-n/a, doi:10.1029/2008jb005594.

Ogawa, M., G. Schubert, and A. Zebib (1991), Numerical simulations of 3-dimensional thermal convection in a fluid with strongly temperature-dependent viscosity, J. Fluid Mech., 233, 299-328.

Shahnas, M. H., W. R. Peltier, Z. Wu, and R. Wentzcovitch (2011), The high-pressure electronic spin transition in iron: Potential impacts upon mantle mixing, Journal of Geophysical Research: Solid Earth, 116(B8), B08205, doi:10.1029/2010jb007965.

Simpson, Guy (2017) Practical Finite Element Modeling in Earth Science using Matlab, Wiley, ISBN: 978-1-119-24862-0

Tackley, P. J., and S. D. King (2003), Testing the tracer ratio method for modeling active compositional fields in mantle convection simulations, Geochem. Geophys. Geosyst., 4(4), doi: 10.1029/2001GC000214.

Travis, B. J., C. Anderson, J. Baumgardner, C. W. Gable, B. H. Hager, R. J. O'Connell, P. Olson, A. Raefsky, and G. Schubert (1990), A benchmark comparison of numerical methods for infinite Prandtl number thermal convection in two-dimensional Cartesian geometry, Geophysical & Astrophysical Fluid Dynamics, 55(3-4), 137-160, doi:10.1080/03091929008204111.

Trompert, R. A., and U. Hansen (1996), The application of a finite-volume multigrid method to 3-dimensional flow problems in a highly viscous fluid with a variable viscosity, Geophys. Astrophys. Fluid Dyn., 83(3-4), 261-291.

van Keken, P. E., S. D. King, H. Schmeling, U. R. Christensen, D. Neumeister, and M. P. Doin (1997), A comparison of methods for the modeling of thermochemical convection, J. Geophys. Res., 102(B10), 22477-22495.