Comments on the manuscript 'Automating Finite Element Methods for Geodynamics via Firedrake' by Davies, Kramer, Ghelichkhan and Gibson

General Comments

I have read the paper with great interest and find it well-written and also well-structured. The approach of the authors to take the reader from the most simple isoviscous 2D Cartesian box problem through a sequence of more and more challenging applications up to a full 3D simulation of mantle convection is nice. That these are all (mostly) standard community benchmarks makes them easily relatable. The embedded code pieces are well explained and easy to follow.

Overall I think the paper demonstrates that there is a great potential in the approach taken by Firedrake for the Geodynamics community. And it is interesting to see how that approach of automated finite element analysis has developed and been extended since the 2013 Vynnytska paper on FEniCS and mantle convection simulation.

I fully recommend publishing the manuscript once the details mentioned here and by the two other reviewers have been taken care of.

Specific Comments

- Listings 4, 5, 6: w.r.t. the `ExtrudedMesh()` function it might be helpful to the reader to mention that the default value for the `layer_height` parameter is `1 / layers`, which explains why e.g. $r_{\max}$ does not factor into the call in your script.

- line 571: 'Specification of the matrix type matfree (line 16) for the combined system ensures that we do not explicitly assemble its associated sparse matrix, instead computing the matrix-vector multiplications required by the Krylov iterations as they arise.'

  I am not sure that I fully understand that part. Does this imply that the full matrix (27) of the Stokes system is not assembled, but that the matvecs needed in the Schur complement approach are computed involving e.g. assembled G? Or is it a fully matrix-free implementation. Besides the need to assemble K for the AMG, that is. Especially if it means the latter case, that would be a feat to showcase more clearly and add a reference how that is handled in Firedrake.

- Listing 5, line 17: 'maximal polynomial degree of gradient of velocity' Not sure why that is 2? For a bi-quadratic Q2 element that should hold w.r.t. the gradient on the reference element, but wouldn't that change when you map to the actual element using a quadratic polynomial (iso-parametric), since the Jacobian of that would be of degree 1? Or is that unimportant for the Nitsche coefficient, or included in the fudge factor sort of?

- I have one question concerning the 3D examples with time-varying viscosity. Especially to the discussion in Section 6. As viscosity changes, is the matrix K reassembled in each time-step and does that then imply that the AMG preconditioner is re-computed? Or do you keep the initial AMG hierarchy? In the latter case can you give any idea on how large a change in viscosity needs to be before that would lead to a deterioration in the convergence?

- Section 7: As this is the sort of simulation a lot of geodynamicists would like to run, I'd appreciated more technical details here. Like what was the mesh resolution and on how many cores was it run. What was the time-to-solution? Was it the same architecture as in Section 6? I know that runtimes are hard to extrapolate from one machine to another on, but that would at least give some impression. Maybe in a future publication one could compare Firedrake and maybe Aspect? Just to demonstrate that its competitive, which I'd expect.

- line 296+297: Eqns (22) and (23) IMHO the '$= 0$' is incorrect and (23) should read '$b_j = -F_{\text{energy}}$'?

- line 342: 'indefinite matrices'? Should that not read 'singular matrices', as far as I know, indefinite implies positive and negative (and potentially also zero) eigenvalues? So the additional nullspace would not affect the indefiniteness of the matrix.

- Figure 1: Just for my understanding, the different benchmark results reported in Blankenbach et al. for Case 1a/1c were very close together, only 1-2% apart, so is that why there is only one line visible? As opposed to e.g. Figure 3, where one has a visible range for reported results?

- Figure 10: Given the mentioned 'change in gradient' it would be nice to have another datapoint inbetween 1536 and 12288 cores. But that's probable not possible given the 4096 elements per core and fixed aspect ratio requirement? Is it possible to present these results in a maybe more appealing fashion?
- I concur with Cedric Thieulot in that the treatment of the continuity equation is relatively uncommon. The standard approach being to keep the derivative with the velocity and shift it over from pressure to the velocity test-function in the momentum equation, not only, but also for symmetry. In a Taylor-Hood approach the divergence of a function from the velocity space would be inside the pressure space. Can you comment on your choice? Is your approach maybe motivated by the fact that you then do not get a derivative of $\bar{\rho}$ in the compressible formulation?

Suggestions

- line 213: Wouldn't 'Ritz-Galerkin' be more appropriate here?
- line 484: Maybe insert `\left ... \right` for scaling of parentheses?
- line 318+319: Should the derivative not be w.r.t. the expansion coefficient 'u_j' (non-bold u)? '$\mathbf{u}_j$' (bold u) seems to indicate the j-th component of the vector field $\mathbf{u}$?
- line 701: Maybe write 'Panel (a)'?
- Personally I found the 3 pages discussion plus the conclusion a little bit lengthy as some aspects are repeatedly mentioned. Maybe that could be shortened and the extensibility of Firedrake to aspects that are not directly FEM be give more space? Like how could one combine it with e.g. a Eulerian-Lagrangian type approach for the energy equation or particle tracing for chemical species advection, ... Sec. 2 mentions 'handwritten C-kernels' and the discussion 'Firedrake's simple API', but that remains a little vague.

Technical Corrections

- line 192: 'constituative' → 'constitutive'
- line 269: 'F_{Stokes}' seems to be only introduced later in (24)
- line 346: 'all' → 'any' ?
- line 382: 'unit vector (k)' → 'unit vector ($\hat{\mathbf{k}}$)' ?
- line 423: 'andsurface' → 'and surface'
- line 497: 'of of' → 'of'
- line 723: 'trilinear' → remove
- line 915: 'the final term' → to me that looks like the 'last but one term'?
- line 1056: King 2009, was probably available online 2009, but the volume it is in is from 2010
- line 1121: Maybe you could fix the capitalisation for Nitsche. That would be 'Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind.' p 9-15, despite its age it even has a DOI: 10.1007/BF02995904
- line 1132: 'ACT' → 'ACM' and its pages 1-27 (24 was the article number)
- line 1145: 'Discussions' → remove
- Figure 10: 'multgird' → 'multigrid'