A Novel Transformation of the Ice Sheet Stokes Equations and Some of its Properties and Applications

J. Dukowicz

Review by Christian Schoof, University of British Columbia

Overview

The paper builds a unified variational formulation for Stokes flow and the somewhat misnamed Blatter-Pattyn model for ice flow. To get the trivia out of the way, I will call this the Herterich-Blatter-Pattyn model from this point onwards. History may be written by the winners, and Herterich clearly hasn't been one of them, *but* the first instance of the Herterich-Blatter-Pattyn model being formulated was (to my knowledge) the following paper:

- Herterich, K. 1987. On the flow within the transition zone between ice sheet and ice shelf, *in* Dynamics of the West Antarctic Ice Sheet. Proceedings of a workshop held in Utrecht, May 6–8, 1985, pp.!185–202. D. Reidel, Dordrecht,

This predates the commonly used Blatter (1995) and Pattyn (2003) references by eight and sixteen years, respectively. If we're going to fall into the trap of naming things after people, we probably owe it to ourselves to do that accurately. Not that I've been able to convince anyone of this point so far.

Back to the present. It took me a bit of effort in stripping away detail and side notes to understand that the primary advance of the paper is a numerical formulation in which one can dynamically switch between the simpler Herterich-Blatter-Pattyn and the more complete Stokes flow model. The latter applies to flows with arbitrary aspect ratios and abrupt changes in boundary conditions, while the former requires a "shallow" flow.

At face value, such a "switchable" model makes a lot of sense, since ice sheet flow is shallow in most places, but often contains boundary layers (at ice divides, ice stream margins, grounding lines etc) where the Stokes equations must be solved. The present paper uses the variational structure of both models to create a unified formulation, by re-writing the Lagrangian for the Stokes equations so that it takes the form of the Herterich-Blatter-Pattyn Lagrangian with a few extra terms retained. The unified formulation then consists of introducing a flag that activates or deactivates these Stokes correction terms. This is an intriguing idea and deserves to be published. I am not entirely convinced that *The Cryopshere* is the right vehicle as there are a number of technical issues that deserve more thorough scrutiny; *Geoscientific Model Development* would have seemed a more appropriate journal in the EGU stable to use, but really I would have been inclined to go with something like *J. Comp. Phys.*.

As far as I can tell, the modified Stokes Lagrangian does not alter the saddle point strucutre of the Stokes flow problem (in the sense that the solution maximizes the modified Stokes functional with respect to $\tilde{p}$ and minimizes with respect to $\boldsymbol{u}$). Still, the modified Stokes Lagrangian is a fairly ugly object that obscures the natural symmetries of the original Stokes flow Lagrangian and introduces a much larger null space to the elliptic operator. For the usual Stokes flow Lagrangian, that null space consists of rigid body rotations modulo any such motions that are precluded by the boundary conditions (by way of restrictions on the space of admissible functions). In the modified Stokes Lagrangian of the present paper, the vertical velocity w can be changed by adding an arbitrary function of the vertical coordinate z only while leaving the elliptic part of the Lagrangian uchanged; such functions of z are then penalized through the incompressibility constraint.

That ugliness appears to be an unavoidable part of a unified formulation. It does however mean that you would probably not choose to use the modified Lagrangian for the analysis of general Stokes flow problems, except to unify the Stokes and Herterich-Blatter-Pattyn problem computationally. The paper can probably streamlined by focusing on that aspect at the exclusion of some of the more peripheral commentary, and the title of the paper could probably be made more informative.

In terms of pursuing that unification of Stokes and Herterich-Blatter-Pattyn models, there are two significant issues that I can see:

1. While the variational formulation of the Stokes flow problem *is* a saddle point problem, the same is not true of the Herterich-Blatter-Pattyn problem. The latter naturally wants to be solved as an unconstrained minimization problem, with the incompressibilty condition solved a posteriori for the vertical velocity component w. Enforcing incompressibility through a Lagrange multiplier as part of the variational formulation leads to problems that occupy most of the technical material in the paper: unlike the modified Stokes Lagrangian, the elliptic part of the Herterich-Blatter-Pattyn Lagrangian does not contain w at all, which must instead be solved for through the (hyperbolic) incompressibility condition alone — for which the standard function spaces used in Stokes flow solvers are unsuitable. This is perhaps unsurprising, as the standard function spaces used in finite elements are problematic for hyperbolic equations, and discontinuous basis functions (as used in discontinuous Galerkin methods) might be preferable for w in the Herterich-Blatter-Pattyn problem; they might at least alleviate issues such as $\nabla \cdot \boldsymbol{u} = 0$ being an overdetermined problem if we assume that $\boldsymbol{u}$ is represented by P1 basis functions, which work well for the force balance part of the Herterich-Blatter-Pattyn problem.[1]

[1]By contrast, the "P1-E0" spaces advocated here go in a different direction: w is still represented by piecewise linear, and therefore continuous, basis functions, but the incompressibility constraint is weakened through the choice of a "coarser" basis function for the Lagrange multiplier that enforces incompressibility, averaging over two adjacent elements.

The main challenge identified in the paper (which motivates the choice of those P1–E0 soaces, and the others discussed in sections 6,7 and appendix C) is to make sure that the choice of basis functions for the "Blatter-Pattyn system" (that is, the Blatter-Pattyn equations and the incompressibility condition) can also be used for the modified Stokes system. Unfortunately, I think that there are some issues with how this has been addressed in the paper, and I suspect that there is a misunderstanding of what the inf-sup ("LBB") condition really does. More on this below.

2. More practically, I am not convinced that the method developed here will be adopted widely. Consider this: unlike its depth-integrated variants developed by Richard Hindmarsh, the Herterich-Blatter-Pattyn requires a three-dimensional domain to be resolved. This leads to a large number of computational degrees of freedom. In fact, the only advantage relative to a Stokes flow model is that there are fewer variables to be solved for in *the same* computational domain, as you can solve for (u, v) using the elliptic solver, and the compute w and p *a posteriori* if required. As far as I can tell, the Seroussi et al (2012, cited in the manuscript) tiling method makes use of that. The method proposed here, of solving for (u, v, w, p) using the same basis functions for both models without explicitly using the simpler structure of the Herterich-Blatter-Pattyn model, seems to get rid of that last advantage. Which begs the question, why bpther, given that the Stokes model is preferable in terms of the physics it represents. Unfortunately, the last paragraph of the conclusion puts paid to any hope that the paper might address whether the new approach is actually going to be computationally competitive.

To follow up on the first point above, I am concerned that appendix D is misleading. Apologies if this gets a little long below; I am not as much of an expert at this as I'd like to be, so it took me a bit more explanation. My understanding of the inf-sup or Ladyzhenskaya-Babuška-Brezzi condition for Stokes flow problems is the following. Take the Hessian matrix in (55) of the present paper

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{0} \end{pmatrix} \tag{1}$$

and we have to solve $\mathbf{Au} + \mathbf{Bp} = \mathbf{f}$, $\mathbf{B}^{\mathrm{T}}\mathbf{u} = \mathbf{0}$ where $\mathbf{f}$ is the relevant residual of the Stokes equations. $\mathbf{A}$ is positive semi-definite, which makes the solution of this problem (if it exists) equivalent to a saddle point: to find

$$(\mathbf{p}, \mathbf{u}) = \arg\max_{\mathbf{q}} \min_{\mathbf{v}} \left(\frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{v} - \mathbf{v}^{\mathrm{T}}\mathbf{f} + \mathbf{v}\mathbf{B}\mathbf{q} \right). \tag{2}$$

The *purely discrete* inf-sup condition for the existence of a unique saddle point then becomes that i) $\mathbf{A}$ is elliptic[2] on the nullspace of $\mathbf{B}^{\mathrm{T}}$, and that ii) $\mathbf{B}$ is of full rank.

[2]that is, $\mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{v} > c\mathbf{v}^{\mathrm{T}}\mathbf{v}$ for some fixed $c > 0$ and any $\boldsymbol{v}$ that satisfies $\mathbf{B}^{\mathrm{T}}\mathbf{v} = 0$

Appendix D does not address the ellipticity of the matrix $\mathbf{A}$ (which, for the Stokes flow model, is not really an issue), but correctly identifies that $\mathbf{B}$ is of full rank if it can be written as

$$\mathbf{B}^{\mathrm{T}} = \left(\mathbf{M}_{UP}^{\mathrm{T}}, \mathbf{M}_{WP}^{\mathrm{T}}\right) \tag{3}$$

and $\mathbf{M}_{WP}$ is invertible. There is a bit of a sleight of hand here, as the text equates invertibility with $\mathbf{M}_{WP}$ being square (which is obviously a necessary but not a sufficient condition). The text in appendix D actually doesn't invoke saddle point theory directly, but instead eliminates the pressure variable from the elliptic part of the problem using a Schur complement. However, the "proof" in appendix most likely stands as a demonstration of *solvability* for the discretized problem, if we overlook the fact that a square matrix $\mathbf{M}_{WP}$ might still not be invertible.

Solvability is likely to be insufficient, however. As far as I understand (and I would caution that I've only recently revived an interest in mixed finite element methods), *stability* of mixed finite element formulations is not the same as solvability of the discretized problem. Spurious pressure oscillations occur in finite element solutions of the Stokes equations using basis functions that do not satisfy the inf-sup conditions even though (as far as I know) the discretized problem can remain solvable. The pressure oscillations should then be indicative of a lack of convergence under mesh refinement, not of a lack of solvability of the discrete problem.

If we set $n = 1$ for simplicity, then the discretized Stokes flow saddle point problem becomes finding $(\boldsymbol{u}_h, p_h) \in V_h \times Q_h$

$$A(\boldsymbol{u}_h, \boldsymbol{v}_h) + b(v_h, \tilde{p}_h) = \langle \boldsymbol{v}_h, f\rangle \tag{4}$$

$$b(\boldsymbol{u}_h, q_h) = 0. \tag{5}$$

for all $(\boldsymbol{v}_h, q_h) \in V_h \times Q_h$, where V_h and Q_h are finite-dimensional subspaces of $V = \{\boldsymbol{v} \in [H^1(\Omega)]^3 : \boldsymbol{v} \cdot \boldsymbol{n} = 0 \text{ on } \partial\Omega_b\}$ and $Q = L^2(\Omega)$, respectively, endowed with their usual norms. In the weak form of the transformed Stokes problem, $A : V \times V \mapsto \mathbb{R}$ is the elliptic operator

$$A(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega \mu \left(\nabla \boldsymbol{u} : \nabla \boldsymbol{v} + (\nabla \cdot P\boldsymbol{u})(\nabla \cdot P\boldsymbol{v})\right) \mathrm{d}\Omega + \int_{\Omega_b} \beta \boldsymbol{u} \cdot \boldsymbol{v} \, \mathrm{d}\Gamma, \tag{6}$$

$P\boldsymbol{v}$ being the projection of a vector onto the x_1x_2-plane, $P(v_1, v_2, v_3) = (v_1, v_2, 0)$. $b : V \times Q \mapsto \mathbb{R}$ is the incompressibility constraint in weak form

$$b(\mathbf{v}, q) = \int_\Omega q \nabla \cdot \boldsymbol{u} \, \mathrm{d}\Omega \tag{7}$$

My understading is that, in order to ensure convergence of $(\boldsymbol{u}_h, \tilde{p}_h)$ to a weak solution of the Stokes problem, we are looking for a family of subspaces (V_h, Q_h) such that

$$||\boldsymbol{v}_h - \boldsymbol{v}|| + ||q_h - q|| < C_h(||v|| + ||q||) \tag{8}$$

where $C_h \to 0$ as mesh size $h \to 0$, *and* that the inf-sup condition holds in the following "infinite-dimensional" form: 1) A is uniformly elliptic on the kernel of b, $A(\boldsymbol{v}, \boldsymbol{v}) > c\,||\boldsymbol{v}||^2$ for some fixed $c > 0$ and all $\boldsymbol{v} \in V$ such that $b(\boldsymbol{v}, q) = 0$ for all $q \in Q$ and 2) $\inf_{q \in Q_h, q \neq 0} \sup_{\boldsymbol{v} \in V_h, \boldsymbol{v} \neq \boldsymbol{0}} b(\boldsymbol{v}_h, q_h)/(||\boldsymbol{v}_h||\,||q_h||) \geq \beta$ **where β is independent of resolution h.**

If I omitted the last half-sentence and allowed $\beta = \beta_h$ to depend on resolution h rather than enforcing a uniform bound, then condition 2) would simply be equivalent to corresponding block matrix $\boldsymbol{B}^{\mathrm{T}}$ having full rank (that is, not having a non-trivial nullspace.[3] Again, as far as I understand, it is the uniformity of the inf-sup condition under grid refinement that makes establishing *stability* of mixed finite element schemes non-trivial (and more difficult than the dimensional counting argument at the heart of appendix D).

Why is all of this relevant? Key to the proposed scheme is that I *can* find suitable function spaces (V_h, Q_h) that work equally well for the Stokes problem (for which I have constructed the discussion above) and for the Herterich-Blatter-Pattyn problem (for which the invertibility of $\mathbf{M}_{WP}$ is an actual necessary condition for the solvability of the discretized problem). My point is that invertibility of $\mathbf{M}_{WP}$ and stability of the resulting mixed problem for the Stokes equations are probably not the same thing, at least in my understanding of the theory of saddle point problems. This affects pretty much the entire discussion of finite element basis functions in the paper.[4] My point about choice of journals is relevant here: I think this should really be read by reviewers that are expert at mixed finite element problems, rather than glaciological dabblers such as myself.

I hope I have been clear to say that I may be wrong. If I am, the text of the paper should be more explicit about the technical issues that I have raised; it's currently fairly vague in spelling out what the "LBB" (inf-sup) condition actually is, what it does, and whether the basis functions described in sections 6–7 and appendix C satisfy the inf-sup condition. (Also, if you're going to introduce a concept like that, please spell out what the acronym "LBB" stands for before using it, and maybe cite the original places where it comes from.)

My other main points would be the following (partially elaborated under "specific points")

1. The notation in the paper is fairly idiosyncratic, which made it more time-consuming for me to follow various pieces. There is a mix of standard subscript notation, a very much nonstandard "round bracket around subscript means projection onto the horizontal plane" variant to subscript notation, and the use of explicit component

[3] since I would otherwise have $B(\boldsymbol{v}_h, q_h) = 0$ for all $\boldsymbol{v}_h$ and some q_h

[4] There is another issue in play here: the inf-sup condition provides sufficient conditions for a convergence, but not necessary ones. You might also decide that you don't care about convergence of the discretized pressure solution $\tilde{p}_h$, but only of $\boldsymbol{u}_h$, which may conceivably relax the conditions you need to impose. I am not in a position to comment on that, however — you need a real expert reviewer.

notation as in equations (7), (28) and (35) to write scalar quantities that should really be denoted by contractions over subscripts. The paper makes enough assumption about the reader's level of mathematics that I would recommend streamlining this as I don't imagine much of an audience who will be helped rather than hindered by the nonstandard notation. In particular, I would discourage the round bracket notation in favour of a more standard projection operator: if P is the projection onto the horizontal plane, $P(v_1, v_2, v_3) = (v_1, v_2, 0)$, then you can simply replace a_i by $P_i(\boldsymbol{a})$ and retain summation over indices from 1 to 3. I say this even though the author has used it previously: the use of nonstandard notation it may make the paper harder to decipher for numerical analysts, who should really be encouraged to delve deeper into the theory relevant to the paper.
The use of superscripts $^{(s)}$, $^{(b_1)}$ and $^{(b_2)}$ on surface normals is also redundant, as you specify the parts of the domain boundary to which the stated boundary conditions apply. Keep this as simple as possible, because it it certainly isn't simple.

2. The paper is quite ungenerous to the prior literature. I can see the appeal of referencing only your own paper ("DPL" in the present case) as the default reference because you know its content extremely well, but many of the introductory concepts in this manuscript have been developed in other places, which remain uncited. That may put off other practicioners who should read this paper (and who might in fact get some sort of notification, alerting them to your paper if you were to cite them!). It's also unhelpful to any reader who wants to make sense of the field: the may conclude that, really, only DPL is relevant as a prior publication. In particular, I see almost no reference to the extensive numerical analysis literature on Herterich-Blatter-Pattyn and Stokes flow models in glaciology (except my ca. 2010 own effort in that direction, which does get a citation somewhere).

3. The "Improved Blatter-Pattyn or BP+ Approximation" in section 6.2.1. I am unconvinced that it makes sense to include this in the present paper, unles you want to "patent" the idea for eternity. (Who knows, maybe someone will in future refer this approximation by someone else's name, the same fate that befell Herterich?)
The reason why I am unconvinced is that this material further breaks the flow of the paper, without being robust. The supposedly improved approximation remains an *ad hoc*, partial retention of higher order terms in the Stokes flow model, higher order being in sense of the aspect ratio. There is no theoretical justification for doing that, as a partial retention of higher order terms in a model comes with no guarantee of a reduction in model error. In fact, it can make the model error worse. We do see a reduction in model error for the "BP+" model ("HBP+" / "D"?) in the single test that the "improved" model has been subjected to (ISMIP-HOM Test B, figure 9 of the manuscript). However, that is not really a robust demonstration of an "improved" model. It's also unclear whether the supposedly improved model is really competitive relative to solving the full Stokes model in terms of the tradeoff

between accuracy and computational effort. (See my second major point above.)

Specific points

A number of specific points, some of which will replicate elements of what I've written above.

- p3, line 81: "finite element grid" — "grid" might be seen to imply regularity. "finite element basis functions"? I mean, closer scrutiny does indicate that the basis functions used here are restricted to quite regular meshes, so perhaps the terminology is not wrong here, but that required regularity actually needs to be discussed somewhere.

- p3 line 84 "these two elements are so-named because they employ edge-based pressures" — I actually found that confusing. The basis function for pressure is not really defined on edges (since there are plenty of edges in the mesh that don't have a pressure defined on them). They are really P0 basis functions in which two adjacent triangles of the triangulation are assigned the same pressure.
  Also, to make it clear that you are inventing these basis functions, avoid the passive voice here. "I have named these E0 to indicate that pressure is defined on select element edges" or similar. It took me a while to realize that I would look in vain for references to these elements in the literature.

- p 3 line 89 onwards " A conventional ice sheet Stokes model discretized on such a grid is numerically equivalent to an inherently stable positive-definite minimization (i.e., optimization) problem, as demonstrated in Appendix D. This is in contrast to the ubiquitous Stokes finite element practice of needing to use elements that satisfy the "inf-sup" or "LBB" condition for stability (see Elman et al., 2014, and the brief discussion in §4.3.1)." — This honestly confused me. If you're saying that current practice in solving the Stokes equations in ice sheet dynamics ("A conventional ice sheet Stokes model") is to use unconstrained optimization, then please provide a reference. It's very hard to do unconstrained optimization for Stokes flow as you need divergence-free basis functions. If you mean that *your new approach* is equivalent to constructing such a divergence-free basis, then say that instead. However, you really also need to show that the reduction of the single Newton iteration step to a problem for $(u_1, \ldots, u_N, w_1, \ldots, u_N)$ only in appendix D (taking the dimensionality argument for $\mathbf{M}_{WP}$ at face value) really is equivalent to a convergent solution (under mesh refinement) of the unconstrained Stokes optimization problem (that is, the problem in which the incompressibility constraint is directly imposed on admissible $\boldsymbol{u}$). I don't think this is entirely trivial, see also my discussion of the inf-sup criterion above. (Put this another way, for any set of basis functions that allows the problem to be solved, I must be able in principle to eliminate the discretized pressure variable in favour of retaining on the nodal values of velocity, even if I would not choose to do the matrix manipulations involved. By your logic, any such scheme should therefore be robust. Is that true?) I may be mistaken about what you're trying to say here, but it may therefore be worth finding a different way of saying it.

- p 4 line 116. $S = S_{B1} \cup S_{B2}$?

- p 4 line 121–123. See above re: notation.

- eqs (8), (10, (11), (13)–(15). The superscripts on your unit normals seem to be redundant. There is only one (outward-pointing) unit normal to each of these parts of the boundary.

- p6, line 170 "the simplest representation" — I beg to differ. There is a standard way of writing shear stress, as the traction $\Sigma_i = \sigma_{ij} n_j$ minus the normal component of traction $n_i \Sigma_j n_j$, or

$$(\delta_{ij} - n_i n_j)\sigma_{jk} n_k.$$

  That is equation (71) in appendix A, and I think you'd be well advised to just use that (standard) form. (The point about weak solutions is surely that you never need to actually evaluate the shear stress itself from the stress tensor and the normal to the surface; you just need to know the constitutive relation, which here is just

$$(\delta_{ij} - n_i n_j)\sigma_{jk} n_k. = \beta u_i.$$

  I think what the paper says isn't wrong, but manipulations don't seem they will help the unwitting reader, who is just wrapping their head around the basic model formulation. Especially at this point, where you haven't motivated your use of the bracketed indices at all yet, in terms of the mathematics you're doing.

- p 6 line 174, "f_i is a specified frictional sliding force vector." — f_i has the wrong units for a force. It's a traction, but why not just call it an interfacial shear stress? τ_i would probably be a more common symbol.

- p 7 line 180 "...can easily be added" → "...can esaily be added to equation (10)"?

- p 7 "(see DPL, 2010, for a fuller description of the variational principle applied to ice sheet modeling)" — you should refer to
  Chen, Q, M. Gunzburger and M. Perego. 2013. Well-posedness results for a nonlinear Stokes problem arising in glaciology. *SIAM J. Math. Anal.*, 45(5), 2710-2733.
  This is probably the most comprehensive analysis of Stokes flows with sliding in glaciology (and you'll find the relevant functional there, too, naturally).

- p 7 line 200 "...(see Schoof, 2010, in connection with the Blatter-Pattyn model)" — I feel a little unfairly singled out here. There are earlier papers by Coling and Rappaz (MSAN, 1999), Glowinski and Rappaz (2003), Chow et al 92004) Rappaz and Reist (M3AS, 2005) that deserve an equal mention as having contributed to the analysis of the Herterich-Blatter-Pattyn model

- p8 line 208 " As in DPL (2010), arguments enclosed in square brackets, here u_i , P, λ_i ,Λ , indicate those variables that are used in the variation of the functional." — this is an odd way of putting it. A functional is simply a mapping (or function) from some vector space (or perhaps an affine space) into the real numbers. So why not just say that you're enclosing the arguments of the functional (which is a function and therefore takes arguemnts like any other function![5]) in square brackets.

- p 8 line 210 onwards. I don't think this correct, or at least, it seems misleading. In my understanding, there are two ways of imposing Dirichlet conditions: by Lagrange multiplier, or by restricting the space of admissible functions. The latter is not the same as the kind of explicit substitution you're doing here. If I take the functional defined in equation (15) and I *don't* separately impose the constraint $\boldsymbol{u} \cdot \boldsymbol{n} = 0$ at $\S_b$ on the arguments u_i, then I don;t see that taking the first variation of the functional will recover that constraint. If you impose the constraint $\boldsymbol{u}\cdot\boldsymbol{n} = 0$, then you do not need to make the substitution in equation (14) in the friction term of the Stokes Lagrangian
  The substitution here is again quite nonstandard, with little advance warning or real motivation. I think you're really trying to lay the groundwork for the modified Stokes variational principle later (by eliminating w where that's going to be necessary) but here is an awkward place to do it, unless you explain why you're eliminating w.

- p 8 line 218 " Here we use z_b as a shorthand notation for $z_b(x, y)$' — this is surely redundant? Same on line 259.

- page 9 line 232 "...the specified values of velocity are then obtainable a posteriori from (9) or (14)" — this is a strange way of putting it. What is a specified value? One that is prescribed? I don't think that's what you mean. I think you mean that the taking the first variation does not recover the boundary condition $\boldsymbol{u} \cdot \boldsymbol{n} = 0$, see previous point. However, you can't impose that "a posteriori" since you cannot solve the Euler-Lagrange equations for the functional in (15) without using the Dirichlet condition on velocity. The standard way of putting this (I believe) is to restrict the functional to the vector space of suitable smooth velocities (in $[W^{1,1+1/n}(\Omega)]^3$) *that satisfy the Dirichlet condition.*

[5]Unless you insist that a function have to take arguments in $\mathbb{R}^n$ for some finite n.

- page 9 line 254"...as is done in the Blatter-Pattyn approximation (see DPL, 2010)."— is the apropriate reference not one of Herterich (1987), Blatter (1995) or Pattyn (2003(?
- page 10, line 265"The standard Stokes pressure P is some three orders of magnitude larger than the transformed pressure $\tilde{P}$" — asymptotic analysis of the problem will show that, more generally, the "Stokes correction" $\tilde{P}$ scales as $O(\varepsilon)$ when there is signfiicant sliding, where ε is aspect ratio, and of $O(\varepsilon^{1+1/n})$ in the absence of sliding (see the Schoof and Hindmarsh 2010 reference).
- page 12 line 296. It would be a good idea to explain the role of the flag parameters ξ and $\hat{\xi}$ earlier. I spent a page half guessing what they were.
- page 13, equation (32). As above, that is pretty ugly. I can't tell if there is a less confusing form, but might be worth trying.
- p 13 line 353 "The standard (or traditional) Blatter-Pattyn approximation (originally introduced by Blatter, 1995; Pattyn, 2003; later by DPL, 2010; Schoof and Hewitt, 2013)" — as above, Herterich (1987) was the person who introduced this. I don't think the Schoof and Hewitt review paper did much more than describe the theory, as oppposed to contributing to it. If I did anything here, then Schoof and Hindmarsh (2010) provided the first self-consistent asymptotic analysis of the model. If you want to list DPL for its description of the variatoinal formulation, you probably ought to cite the numerical analysis papers I have listed above (Coling and Rappaz, Glowinski and Rappaz, Chow et al, Rappaz and Reist; full citations in my 2010 paper that is in the reference list), who previously dealt with the same variational formulation.
- page 15 line 376 Remark # 1: seems like splitting hairs, especially as you wouldn't bother with this if you wrote down the weak form.
- page 15 line 419 "In summary, the extended Blatter-Pattyn model, (40)-(42), is equivalent to the standard Blatter-Pattyn model, (36), for the horizontal velocities, u,v , except that it also includes two additional equations that determine the pressure P ! and the vertical velocity w , which are usually ignored in the standard Blatter-Pattyn approximation when only the horizontal velocity is of interest. Because of this, we distinguish between the Blatter- Pattyn model that solves for just the two horizontal velocities (i.e., the standard Blatter- Pattyn approximation, (36)), and the Blatter-Pattyn system that solves for all the variables (i.e., the extended Blatter-Pattyn approximation, (40)-(42))." — again, this seems like splitting hairs. Anyone who needs to solve for temperature in an ice sheet (which is a standard part of any ice sheet simulation code) has to solve for the vertical velocity component, so I don't think insisting on a difference between the Blatter-Pattyn model and system is helpful. Solving for $\tilde{P}$ is not particularly relevant as we

know $\tilde{P} = 0$ for the Blatter-Pattyn model. For the sake of not going down rabbit holes, I would omit this.

- p 17 line 437 "The use of the continuity equation to solve for the vertical velocity w is a novel feature of the Blatter-Pattyn approximation since the continuity equation is not normally used for this purpose." — as per my previous comment, I don't think this is true or novel. The vertical velocity component is a pretty important quantity to be able to compute in ice sheet codes.

- p 17 lie 439 onwards. The use of a depth-integrated mass conservation equation is a boilerplate approach for thin film models, so a lengthy discussion without citation seems unnecessary, especially as this doesn't really lead anywhere in developing the novel material in the paper.

- p 17 line 482 "Additional details about the grid and the associated discretization..." — I flagged this above (grid versus mesh), but again: I think it's worth pointing out somewhere that some of the basis functions developed here really do require a regular mesh (grid?) rather than an unstructured triangulation. For the P1-E0 basis functions, I have to be able to match every triangle with precisely one neighbour, not leaving any neighbourless triangles. I don't think that's possible with every fully unstructured mesh. (?)

- page 19 line 504. Italicize the variables u and w.

- page 21 line 526, "...or else they are "saddle point" problems since the Hessian matrix $M(u, w)$ is symmetric but indefinite, with both positive and negative eigenvalues..." — you can be more definitive about this I think: the "saddle point" terminology refers to the fact that there really is a saddle point, where you mimiize the quadratuc form associated with the matrix with respect to the velocity variables and maximize with respect to pressure.

- p 22 line 565 Please define the acronym "LBB" before using it.

- p 23 line 601 Taylor-Hood elements have gone from being an illustration of an element that satisfies the inf-sup condition (line 569) to becoming the standard reference for stable elements. Is it true in general that all elements that satisfy the LBB condition for the Stokes problem will leave M_{WP}^{T} non-invertible on dimensional grounds alone? If so, that is important to point out here.

- Sections 5.1–5.3 You're focusing on velocity solutions here. If there are stability issues with your choice of basis functions, I'd expect these to show up in the pressure solution. Can we see resultss for some of those? Obviously not relevant for Herterich-Blatter-Pattyn, but for the transformed and original Stokes flow problems. Especially pressure along the bed would be useful, but also convergence or lack thereof.

- Figure 7: this is fine, but more convincing would be a grounding line, where you have to worry about a normal stress constraint determining the grounding line location (and that normal stress is in general not cryostatic as assumed by Herterich-Blatter-Pattyn)