Referee report on the thesis of Mgr. Tomáš Gergelits

Patrick E. Farrell

August 28, 2020

The thesis of Mgr. Gergelits is concerned with two closely-related topics: the nonlinear convergence behaviour of the conjugate gradient method in the presence of rounding errors (chapter 2), and the spectrum of second-order elliptic differential operators when the Laplacian is used as a preconditioner (chapters 3 and 4).

The results presented by the author (particularly in chapters 3 and 4) are novel, strikingly beautiful, and elegantly explained. They are also of potentially high importance in the practical solution of problems arising in geophysics, such as in understanding the flow of oil in petroleum reservoirs and of rock in the Earth's mantle. The work has been published in two articles in one of the leading journals in the field. In my opinion, **the thesis clearly demonstrates the author's ability for creative scientific work** and I recommend that he be awarded the doctoral degree of Charles University after (very) minor editions.

Background and novelty

Consider the flow of rock in the Earth's mantle. At the titanic temperatures and pressures in the interior of the Earth, the rock may flow like an (extremely viscous) fluid, and the flow is well-described by the Stokes equations. The main difficulty in solving these equations (after employing a suitable approximation to the pressure Schur complement) is that the viscosity is a strongly nonlinear function of the temperature, and may vary by many orders of magnitude across the domain. This viscosity contrast poses strong difficulties for standard (multigrid) preconditioners employed with the conjugate gradient method, both in the construction of suitable relaxation methods and in the development of coarse grids with good approximation properties.

One possible approach to dealing with this difficulty is to instead solve a simpler problem and use this as a preconditioner for the more challenging one. Following the general operator preconditioning framework expounded by Mardal, Winther, Málek, Strakoš, Kirby and others, a natural suggestion for the simpler problem is to choose a constant viscosity (for which multigrid is effective). The central contribution of this thesis is to precisely and completely characterise the spectrum of the resulting preconditioned operator, and therefore to completely characterise the convergence of the conjugate gradient method in this setting[1].

The results are new and intriguing. The key theorems are of substantial technical depth and combine tools from various areas of mathematics. The closest existing work to this is the cited paper of Nielsen, Tveito and Hackbusch, but Gergelits' work is a very substantial improvement of this result in many different ways: it proves a complete characterisation of the spectrum (rather than a containment); it weakens the regularity requirements on the coefficient; it considers both discretised and continuous problems; and it extends the results to the very important case where the coefficient is tensor-valued. The fact that the results are published in the SIAM Journal on Numerical Analysis is a testament to their novelty.

1 General comments

The thesis is in general very well written. His deep familiarity with and passion for the subject shines through. His work has been influenced very strongly by the superb monograph of Liesen & Strakoš, in particular in its emphasis on the subtle and intricate nonlinear convergence of the conjugate gradient method. Indeed, consulting again my copy of Liesen & Strakoš, I see that Gergelits is acknowledged, indicating that influence has propagated in the other direction also.

One possible criticism that could be offered is that the thesis feels somewhat disjoint between chapter 2 and chapters 3–4. While the fact that the convergence of the conjugate gradient method depends on the entire spectrum motivates understanding the entire spectrum of the preconditioned operator,

[1] The thesis actually considers *scalar*-valued problems, but I expect that the results would extent without difficulty to the *vector*-valued case.

the technical details quoted at length from previous works are not necessary for understanding the subsequent chapters. (A level of exposition similar to that of the SINUM 2019 paper would have been sufficient.)

The technical depth is, however, necessary for the novel research included at the end of chapter 2, on the convergence of the conjugate gradient method with floating point arithmetic being well described by the method in exact arithmetic with a lag arising from the rank deficiency of the Krylov subspace. I suspect that the author's original intention was to write a thesis on this topic, wrote most of chapter 2 with this in mind, then switched emphasis to the results of chapters 3 and 4 when these were subsequently proven. Personally, I would have preferred more detail on the concept of operator preconditioning in the expository chapter, drawn from the book of Málek and Strakoš or other sources. For example, I would like to have seen more emphasis on the fact that the conjugate gradient method for problem (2.1) is *not well-defined* without a preconditioner mapping to and from the appropriate spaces, since this fundamental misunderstanding has been the cause of countless mesh-dependent iteration counts going back decades, and is still not widely understood by practitioners.

The title is too generic; it could apply to many more theses than this one. A more suitable title might be *Preconditioning symmetric second-order differential operators with the Laplacian.* I suggest the author consider changing it (to this or something else).

2 Applications beyond the thesis

Since the solution of problems with large coefficient contrasts is of substantial importance, this thesis bears relevance to many potential applications. However, the details of quite *how* to speed up these calculations with Gergelits' insights are still to be worked out, and in my opinion the thesis would be even stronger if this point had been considered. For example, in the challenging mantle convection problem described above, the results of the thesis will likely reveal that the operator-preconditioned conjugate gradient method will converge poorly, due to the unfavourable spectrum of the preconditioned operator. What can be done about this?

While I have not studied this question myself, I know many others have. One possibility is to precondition the base operator (with $k(x) > 0$) with another operator with a spatially-varying coefficient, but where the coefficient is milder; for example, one could take the square root of the coefficient and use the operator with $\sqrt{k(x)}$ as a preconditioner. It may be the case that multigrid on this preconditioner is effective, and the spectrum of the preconditioned operator is more favourable than employing the constant coefficient 1. If multigrid remains ineffective, the process could be repeated (take the square root again); this will surely yield an operator that multigrid can tackle in a logarithmic number of steps. Perhaps Mgr. Gergelits has better ideas. The theorems in this thesis compare the coefficient $k(x)$ (or $K(x)$) against the constant coefficient 1; an analytical extension that would yield great computational insight would be to develop theorems comparing $k(x)$ with another function $\tilde{k}(x)$ (and analogously for the tensor-valued case). This would allow for the analysis of the scheme outlined above, indicating whether it should be successful or not. Some results in this direction have been achieved in the preprint of Ladecký et al.

3 Suggested editions

- pg. 31: *extense* should be *extends*.
- pg. 33: the author refers several times to the *group* of eigenvalues. I do not believe he means to describe the eigenvalues as a group in the technical sense, which is confusing. I suggest instead the use of the more appropriate word *set*.
- pg. 35–36: the definition of Int should be moved closer to where it is first used (it puzzled me for a while at the bottom of pg. 35 until I decided to read on anyway).
- pg. 36: *derivations* should be *derivatives*.
- pg. 43: (4.2) is missing a λ on the right-hand side.
- pg. 45: More details on the approach of Ladecký et al. should be provided. In particular, the author should compare the advantages and disadvantages of the two approaches.
- pg. 47: (4.17) is missing a λ on the right-hand side.
- pg. 47: Q should be defined as $Q : \Omega \to \mathbb{R}^{2\times 2}$.

- pg. 47: the u in the definition of W below (4.21) is not the same as the u specified as data in the Open problem 4.3.2; a different variable should be used.
- pg. 50 a bracket is missing in the middle equation in the unnumbered trio above (4.34).
- pg. 55 *disretized* should be *discretized*.

4 Recommendation

I strongly recommend that the candidate be awarded the doctoral degree of Charles University.