Solving REXI terms

Martin Schreiber <M.Schreiber@exeter.ac.uk>, ADD YOURSELF IF YOU ADD THIN

May 31, 2016

This document serves as the basis to understand and discuss various ways how to solve the terms of the REXI approximation given by

$$e^L U_0 \approx \sum_i \beta_i (\alpha_i + L)^{-1} U_0.$$

1 Problem formulation

Exponential integrators provide a form to directly express the solution of a linear operator (non-linear operators are not considered in this work). For a linear PDE given by

$$U_t = L(U)$$

we can write

$$U(t) = e^{Lt} U(0).$$

Furthermore, assuming that the L operator is skew-Hermitian - hence has imaginary Eigenvalues only - we can write this as a Rational approximation of the EXponential Integrator (REXI)

$$e^L U_0 \approx \sum_{n=-N}^{N} Re\left(\beta_n (\alpha_n + L)^{-1} U_0\right),$$

see [2]. This is the already simplified equation where the time step size τ is merged with L. This indeed doesn't make a big difference here.

2 Properties

We will discuss properties and potential misunderstandings in the REXI approximation in this section.

2.1 REXI Coefficient properties

With overbar denoting the complex conjugate, the coefficients in the REXI terms have the following properties:

$$\alpha_{-n} = \bar{\alpha}_n \tag{1}$$

$$\beta_{-n} = \bar{\beta}_n \tag{2}$$

$$Im(\alpha_0) = Im(\beta_0) = 0 \tag{3}$$

2.2 Reduction of REXI terms

Using the REXI coefficient properties we can almost half the terms of the sum to

$$e^L U_0 \approx \sum_{n=0}^{N} Re\left(\gamma_n(\alpha_n + L)^{-1} U_0\right) \tag{4}$$

with

$$\gamma_n := \begin{cases} \beta_0 & for\, n = 0 \\ 2\beta_n & else \end{cases}$$

hence

$$\gamma_{-n} = \bar{\gamma}_n \tag{5}$$

2.3 Reutilization of REXI terms for several coarse time steps

[Based on idea of Mike Ashworth]. Assuming that we're interested in all the solutions at coarse time stamps $T_n := n\Delta T$, is it possible to directly compute them by reutilizing the REXI terms of $T_{n<N}$ for T_N? Reusing REXI terms requires computing REXI terms with the same α_i coefficients. Those alpha coefficients are computed with

$$\alpha_n := h(\mu + i(m + k)).$$

Here, h specifies the sampling accuracy, m is related to the number of REXI terms and k can be assumed constant and is related to the number of poles for the approximation of the Gaussian function. This shows, that REXI terms can be indeed reused! However, we shall also take the β terms into account to see if we can reuse the first partial sum. These β_n coefficients are given by

$$\beta_n^{Re} := h \sum_{k=L_1}^{L_2} Re(b_{n-k}) a_k$$

and

$$b_m = e^{-imh} e^{h^2}.$$

All of these coefficients are constant for given h, but $L_{1/2}$ depend on all a_k and hence the number of total REXI terms. Therefore, we **cannot directly reuse the result of the REXI sum** of the first coarse time step, **but the results of each separate term, the inverse problem** $(\alpha + L)^{-1} U_0$.

Future work: Maybe a reformulation to reuse the previous sum reduction is possible.

2.4 Real values of exponential integrators

Obviously, only real values should be computed by the exponential integrator e^L. There could be the assumption that *REXI also creates only real values with negligible imaginary values*. However, this is not true! It holds that

$$Im\left(\lim_{N\to\infty}\sum_{n=-N}^{N}\left(\beta_n(\alpha_n + L)^{-1}U_0\right)\right) \neq 0.$$

Note, that here we didn't restrict the solution to real values as in (4).

@TERRY: TODO: The way how I derived β and α might be different to your way but I think it's the same. Do you agree in the statement above?

3 Computing inverse of $(\alpha + L)^{-1}$

It was suggested [2] to use a reformulation of this linear operator which is based on an advective shallow-water formulation to compute $\eta(t+\Delta t)$ via a Helmholtz problem and then solve for both velocity components direcctly. However, this reformulation in an ODE-oriented way was only possible with a constant Coriolis term. Here, we will also discuss Matrix formulations.

3.1 Deriving Helmholtz problem for constant f SWE

We can reformulate the SWE (see [4]) into the following formulation

$$((\alpha^2 + f^2) - g\bar{\eta}\Delta)\eta = \frac{f^2 + \alpha^2}{\alpha}\eta_0 - \bar{\eta}\delta_0 - \frac{f\bar{\eta}}{\alpha}\zeta_0.$$

3.1.1 Spectral elements:

For spectral element methods, Gunnar's method was successfully applied to solve this.

3.1.2 Spectral method:

For spectral methods, we used a so-called fast Helmholtz solver to directly solve this very efficiently.

3.2 Deriving Helmholtz problem for f-varying SWE with matrix partitioning

We have to find a matrix-formulation of this reformulation. Following the derivation in [4] we get the system of equations

$$((\alpha^2 + f^2) - g\bar{\eta}\Delta)\eta = \frac{f^2 + \alpha^2}{\alpha}\eta_0 - \bar{\eta}\delta_0 - \frac{f\bar{\eta}}{\alpha}\zeta_0$$

to solve for. Instead of treating every term in the linear operator as being scalar-like, we can a partitioning of the matrix making the linear operator a diagonal matrix L

$$L(U) := \begin{pmatrix} 0 & -\eta_0\partial_x & -\eta_0\partial_y \\ -g\partial_x & 0 & F \\ -g\partial_y & -F & 0 \end{pmatrix} U \tag{6}$$

and all other operators itself also representing a matrix formulation. The term F is then the matrix with varying Coriolis effect

$$F := \begin{bmatrix} cos(\theta_0) & & & & \\ & cos(\theta_1) & & & \\ & & \dots & & \\ & & & cos(\theta_{N-2}) & \\ & & & & cos(\theta_{N-1}) \end{bmatrix}$$

with N the size of the matrix. Then we write the system to solve for as

$$((\alpha^2 + F^2) - g\bar{\eta}\Delta)\eta = \frac{F^2 + \alpha^2}{\alpha}\eta_0 - \bar{\eta}\delta_0 - \frac{F\bar{\eta}}{\alpha}\zeta_0.$$

Now the challenge is to solve for this system of equations with the varying terms in the F matrix. The F^2 terms lead to longitude-constant $\cos^2(\theta)$ terms.

3.2.1 Spectral elements method:

For spectral element methods, Gunnar's method could be applied to solve this.

3.2.2 Spectral methods:

Using spectral space, applying this term could basically mean to shift a solution to a different spectrum. This could allow developing a direct solver for it in spectral space. The real-to-real Fourier transformations results in cos-only eigenfunctions and could be appropriate for this. **[TODO: Just a sketch. Seems to be good to be true, hence probably wrong].**

3.3 Hybridization for SWE

This is related to Colin's idea and is based on writing down the entire formulation in its discretized way, hence before applying solver reformulations as done in the previous section.

Before doing any analytical reformulations it discretizes the equations first (e.g. on a C-grid) and then works on this reformulation.

This focuses on maintaining the conservative properties (e.g. avoiding computational modes) first and then to solve it.

[TODO: Awesome formulation of hybridization on C-grid]

3.4 Iterative solver with complex values

A straight-forward approach is to use an iterative solver which supports complex values. This means that $(\alpha + L)U = U_0$ is solved directly and that's it if we could use already existing solvers.

3.5 Reformulation to real-valued solver

The complex-valued iterative system for $(\alpha + L)U = U_0$ can be reformulated to a real valued system by treating real and imaginary parts separately. This is based on splitting up $U = Re(U) + i\,Im(U) = U^R + i\,U^I$ (see also notes from Terry and Pedro). Similarly, we use $\alpha = Re(\alpha) + i\,Im(\alpha) = \alpha^R + i\alpha^I$. Then the complex system of equations $(\alpha + L)\,U = U_0$ can be written as

$$\left[\begin{array}{c|c} A^R + L & -A^I \\ \hline A^I & A^R + L \end{array}\right] \left[\begin{array}{c} U^R \\ U^I \end{array}\right] = \left[\begin{array}{c} U_0^R \\ 0 \end{array}\right]$$

with A a Matrix with α values on the diagonal. Obviously, the off-diagonal values in partitions given by A^I are a pain in the neck for iterative solvers: they are varying depending on the number of REXI term. We again get a skew Hermitian matrix **[TODO: Check the signs]**.

3.6 Solving real and imaginary parts

We can go one step further and generate a system of equations to solve by eliminating U^I. We first solve the 2nd line for U^I:

$$U^I = -(A^R + L)^{-1} A^I U^R.$$

Putting this in the 1st line

$$(A^R + L)U^R + A^I U^I = U_0^R$$

we get

$$(A^R + L)U^R + A^I (A^R + L)^{-1} A^I U^R = U_0^R.$$

Solving this for U^R, we get

$$\left((A^R + L) + A^I (A^R + L)^{-1} A^I\right) U^R = U_0^R.$$

Inverting stuff is not nice and we multiply both sides from left side with $(A^R + L)$ yielding the following equation:

$$U^R: \ \left((A^R+L)^2 - A^I A^I\right) U^R = (A^R+L) U_0^R$$

We are not finished yet, since we also need the imaginary components of U. The reason for this is that these components, once multiplied with the imaginary component of β, create real values. Solving the 1st line for U^R gives us

$$U^R = (A^R+L)^{-1}\left(A^I U^I + U_0^R\right)$$

Putting this in the 2nd line, we get

$$\begin{aligned} A^I(A^R+L)^{-1}\left(A^I U^I + U_0^R\right) + (A^R+L)U^I &= 0 \\ A^I(A^R+L)^{-1}\left(A^I U^I\right) + (A^R+L)U^I &= -A^I(A^R+L)^{-1}U_0^R \\ A^I A^I U^I + (A^R+L)^2 U^I &= -A^I U_0^R \end{aligned}$$

$$U^I: \ \left(A^I A^I + (A^R+L)^2\right) U^I = -A^I U_0^R$$

Boths things look quite ugly. However, this leads to another important property of PinTing. This allows sovling both contributions independent of each other. Hence, this would give us an additional degree of parallelization.

3.7 Including β and solving for real values only

So far we totally ignored the β coefficient in REXI. This forced us to also care about the imaginary-values solution U^I. If we would be able to put it into the inverse computation, we might be able to compute only the real values. With $(AB)^{-1} = B^{-1}A^{-1}$ we can write

$$\sum_i (\beta_i^{-1})^{-1}(\alpha_i + L)^{-1} U_0 = \sum_i (\alpha_i \beta_i^{-1} + L\beta_i^{-1})^{-1} U_0.$$

We formulate this term to

$$\sum_i (\alpha_i \beta_i^{-1} + L\beta_i^{-1})^{-1} U_0$$

with $\beta_i^{-1}\alpha_i$ again complex valued and $L\beta_i^{-1}$ the linear operator scaled by β_i^{-1}, hence also containing real and complex values. We can now apply the same strategy as before by eliminating U^I and solve for U^R. Let

$$M^R = Re(\alpha_i \beta_i^{-1} + L\beta_i^{-1})$$

and

$$M^I = Im(\alpha_i \beta_i^{-1} + L\beta_i^{-1}).$$

This yields the SoE

$$\left[\begin{array}{c|c} M^R & -M^I \\ \hline M^I & M^R \end{array}\right] \left[\begin{array}{c} U^R \\ U^I \end{array}\right] = \left[\begin{array}{c} U_0^R \\ 0 \end{array}\right]$$

and further

$$U^I = \left(M^R\right)^{-1} \left(-M^I U^R\right).$$

Putting this in 1st line yields

$$M^R U^R + M^I \left(M^R\right)^{-1} M^I U^R = U_0^R$$

$$\left(M^R + M^I \left(M^R\right)^{-1} M^I\right) U^R = U_0^R$$

We can also write

$$\left(I + \left(\left(M^R\right)^{-1} M^I\right)^2\right) U^R = \left(M^R\right)^{-1} U_0^R$$

Now the big question arises what $\left(\left(M^R\right)^{-1} M^I\right)$ is. Seems like computing the time step tendencies $L(U)$ with different constant contributions given by the shifted poles.

4 Interpreting $M^{R/I}$ and $(\alpha + L)^{-1}$ terms

[TODO: Here we assuming that the previous reformulations are really possible and there are probably a lot of bugs in it].

We like to get insight in the meaning of the terms

$$\left(M^R\right)^{-1} M^I$$

Both terms only consist out of a real formulation and should be summarized here as

$$(a + bL)$$

with a a real-valued constant which is e.g. given by $Re(\alpha_i \beta_i^{-1})$ and bL the linear operator L scaled by a real-valued scalar b.

[10 minute brainstorming with John T.]

$(a + bL)U$ can be interpreted as an explicit time stepping method.

$(a + bL)^{-1}U$ can be interpreted as an implicit time stepping method.

The factors a and b can then be interpreted as scaling factors and time step sizes.

5 Final notes

Have fun in reading this. Don't miss out all the errors! ;-)

References

[4] Understanding REXI

[1] Formulations of the shallow-water equations, M. Schreiber, P. Peixoto et al.

[2] High-order time-parallel approximation of evolution operators, T. Haut et al.

[3] Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later, Cleve Moler and Charles Van Loan, SIAM review

[4] Near optimal rational approximations of large data sets, Damle, A., Beylkin, G., Haut, T. S. & Monzon