EDTRK4 time stepping notes

Navid Constantinou *

May 23, 2016

1 The problem

Say we want to solve numerically the following differential equation satisfied by $u(\boldsymbol{x},t)$:

$$\partial_t u = \mathcal{L}(u) + \mathcal{N}\left(u,t\right) \ , \tag{1}$$

where $\mathcal{L}$ is a linear operator acting on u and $\mathcal{N}$ a nonlinear operator. Consider a discretization so that u is described by the column vector $\boldsymbol{u}$. This correspondence is symbolically written $u(x,t) \leftrightarrow \boldsymbol{u}(t)$ or just $u \leftrightarrow \boldsymbol{u}$. This way (1) becomes:

$$\partial_t \boldsymbol{u} = \boldsymbol{L}\boldsymbol{u} + \boldsymbol{N}\left(\boldsymbol{u},t\right) \ , \tag{2}$$

where $\boldsymbol{L}$ is a matrix acting on $\boldsymbol{u}$ so that $\boldsymbol{L}\boldsymbol{u} \leftrightarrow \mathcal{L}(u)$ and $\boldsymbol{N}$ is a nonlinear operator acting on $\boldsymbol{u}$ and returning a column vector $\boldsymbol{N}\left(\boldsymbol{u},t\right) \leftrightarrow \mathcal{N}(u,t)$.

Consider now the variable $\boldsymbol{v} = \mathrm{e}^{-\boldsymbol{L}t}\boldsymbol{u}$ which satisfies:

$$\partial_t \boldsymbol{v} = \mathrm{e}^{-\boldsymbol{L}t}\boldsymbol{N}\left(\mathrm{e}^{\boldsymbol{L}t}\boldsymbol{v},t\right) \ . \tag{3}$$

The idea is the following: We will time-step (2) using an RK4 time-stepping scheme for solving the nonlinear part, i.e. solving for variable $\boldsymbol{v}$ and then use the (exact) propagator $\mathrm{e}^{\boldsymbol{L}t}$ to solve for the linear part as $\boldsymbol{u} = e^{\boldsymbol{L}t}\boldsymbol{v}$ [1]. If $\boldsymbol{u}_n$ is the value of $\boldsymbol{u}(t)$ at time $t = nh$, where h is the time-step, then $\boldsymbol{u}_{n+1}$ is given as:

$$\boldsymbol{u}_{n+1} = \mathrm{e}^{\boldsymbol{L}h}\boldsymbol{u}_n + f_u\boldsymbol{N}(\boldsymbol{u}_n,t_n) + 2f_{ab}\left[\boldsymbol{N}(\boldsymbol{a}_n,t_n+h/2) + \boldsymbol{N}(\boldsymbol{b}_n,t_n+h/2)\right] + f_c\,\boldsymbol{N}(\boldsymbol{c}_n,t_n+h)\,, \tag{4}$$

where

$$\boldsymbol{a}_n = \mathrm{e}^{\boldsymbol{L}h/2}\boldsymbol{u}_n + \boldsymbol{L}^{-1}\left(\mathrm{e}^{\boldsymbol{L}h/2} - \boldsymbol{I}\right)\boldsymbol{N}(\boldsymbol{u}_n,t_n)\,, \tag{5a}$$

$$\boldsymbol{b}_n = \mathrm{e}^{\boldsymbol{L}h/2}\boldsymbol{u}_n + \boldsymbol{L}^{-1}\left(\mathrm{e}^{\boldsymbol{L}h/2} - \boldsymbol{I}\right)\boldsymbol{N}(\boldsymbol{a}_n,t_n+h/2)\,, \tag{5b}$$

$$\boldsymbol{c}_n = \mathrm{e}^{\boldsymbol{L}h/2}\boldsymbol{a}_n + \boldsymbol{L}^{-1}\left(\mathrm{e}^{\boldsymbol{L}h/2} - \boldsymbol{I}\right)\left[2\boldsymbol{N}(\boldsymbol{b}_n,t_n+h/2) - \boldsymbol{N}(\boldsymbol{u}_n,t_n)\right] \ , \tag{5c}$$

$$f_u = h^{-2}\boldsymbol{L}^{-3}\left\{-4 - \boldsymbol{L}h + \mathrm{e}^{\boldsymbol{L}h}\left[4 - 3\boldsymbol{L}h + (\boldsymbol{L}h)^2\right]\right\} \ , \tag{5d}$$

$$f_{ab} = h^{-2}\boldsymbol{L}^{-3}\left[2 + \boldsymbol{L}h + \mathrm{e}^{\boldsymbol{L}h}\left(-2 + \boldsymbol{L}h\right)\right] \ , \tag{5e}$$

$$f_c = h^{-2}\boldsymbol{L}^{-3}\left[-4 - 3\boldsymbol{L}h - (\boldsymbol{L}h)^2 + \mathrm{e}^{\boldsymbol{L}h}\left(4 - \boldsymbol{L}h\right)\right] \ . \tag{5f}$$

*Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92093–0230, USA.

The last 4 terms on the r.h.s. of (4) are just the RK4 approximation of the term

$$e^{\boldsymbol{L}h} \int_{nh}^{(n+1)h} e^{-\boldsymbol{L}\tau} \boldsymbol{N} \Big(u(t_n + \tau), t_n + \tau \Big) \, \mathrm{d}\tau \, . \tag{6}$$

In the limit $\boldsymbol{L} \to 0$ we should recover the RK4 time-step. Indeed:

$$\boldsymbol{a}_n \to \boldsymbol{u}_n + \frac{h}{2} \boldsymbol{N}(\boldsymbol{u}_n, t_n) \, , \tag{7a}$$

$$\boldsymbol{b}_n \to \boldsymbol{u}_n + \frac{h}{2} \boldsymbol{N}(\boldsymbol{a}_n, t_n + h/2) \, , \tag{7b}$$

$$\boldsymbol{c}_n \to \boldsymbol{u}_n + h \boldsymbol{N}(\boldsymbol{b}_n, t_n + h/2) \, , \tag{7c}$$

while

$$f_u \to h/6 \;, \quad f_{ab} \to h/6 \;, \quad f_c \to h/6 \, , \tag{7d}$$

and after some fiddling around we can see that this gives the RK4 time-step.

However, there is a catch in calculating the coefficients (5). There are a lot of cancelation errors, especially when the eigenvalues of $\boldsymbol{L}$ are close to zero, since in that case we are dividing something vanishingly small over something also vanishingly small. The way over that is to calculate these coefficients by means of complex calculus [2]. For example, any function of $\boldsymbol{L}$ can be evaluated as

$$f(\boldsymbol{L}) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} f(z) \, (z\boldsymbol{I} - \boldsymbol{L})^{-1} \, \mathrm{d}z \, , \tag{8}$$

where $\boldsymbol{I}$ is the identity matrix and Γ is any contour enclosing all eigenvalues of $\boldsymbol{L}$ in the complex plane. This is the generalization of Cauchy's theorem for functions of matrices.

2 An example

Consider the barotropic vorticity equation:

$$\partial_t \zeta + J(\psi, \zeta + \beta y) = -\mu \zeta - (-1)^h \nu_{2h} \triangle^h \zeta \, . \tag{9}$$

with $J(\psi, \zeta) \stackrel{\text{def}}{=} \psi_x \zeta_y - \psi_y \zeta_x$ and $\zeta \stackrel{\text{def}}{=} \triangle \psi$. Consider the Fourier expansion of the vorticity field:

$$\zeta(\boldsymbol{x}, t) = \sum_{\boldsymbol{k}} \tilde{\zeta}_{\boldsymbol{k}}(t) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \;\; , \tag{10}$$

and describe the vorticity field by means of its Fourier coefficients. Written in terms of the Fourier components of $\zeta(\boldsymbol{x}, t) = \sum_{\boldsymbol{k}} \tilde{\zeta}_{\boldsymbol{k}}(t) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$ (9) takes the form:

$$\partial_t \tilde{\zeta}_{\boldsymbol{k}} = \left[-(\mu + \nu_{2h} k^{2h}) + \frac{\mathrm{i}\beta k_x}{k^2} \right] \tilde{\zeta}_{\boldsymbol{k}} - [\widetilde{J(\psi, \zeta)}]_{\boldsymbol{k}} \, , \tag{11}$$

where $[\widetilde{J(\psi, \zeta)}]_{\boldsymbol{k}}$ denotes the Fourier component of $J(\psi, \zeta)$. From (11) we see that in this case the matrix $\boldsymbol{L}$ is diagonal, something that would not be true if we did not chose to describe the vorticity in terms of its Fourier components and instead discretized the vorticity field in physical space.

The fact that $\boldsymbol{L}$ is diagonal simplifies the computations of coefficients (5) immensely since the action of any function $f(\boldsymbol{L})$ on $\boldsymbol{u}$ can be computerd element–by–element. Therefore, all matrix operations in (5) simplify to just operations on the elements of the diagonal of $\boldsymbol{L}$. For example, the matrix integral

$$f(\boldsymbol{L}) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} f(z) \, (z\boldsymbol{I} - \boldsymbol{L})^{-1} \, \mathrm{d}z \,, \tag{12}$$

reduces in a simple integral for each element of the diagonal of $\boldsymbol{L}$, denoted by $L_{\boldsymbol{k}}$,

$$f(L_{\boldsymbol{k}}) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_{\boldsymbol{k}}} \frac{f(z) \, \mathrm{d}z}{z - L_{\boldsymbol{k}}} \,. \tag{13}$$

In the above we are able to chose a different integration contour $\Gamma_{\boldsymbol{k}}$ for each element of $\boldsymbol{L}$. Specifically, we choose as $\Gamma_{\boldsymbol{k}}$ for each integral the unit circle around its integrand pole,

$$\Gamma_{\boldsymbol{k}} = \left\{ L_{\boldsymbol{k}} + \mathrm{e}^{2\pi \mathrm{i} w} \ : \ 0 < w \leq 1 \right\} \,, \tag{14}$$

which implies

$$f\left(L_{\boldsymbol{k}}\right) = \int_0^1 f\left(L_{\boldsymbol{k}} + \mathrm{e}^{2\pi \mathrm{i} w}\right) \, \mathrm{d}w \,. \tag{15}$$

Using M points compute the integral, i.e. $w \in \{1/M, \ldots, (M-1)/M, 1\}$ we get that:

$$f\left(L_{\boldsymbol{k}}\right) \approx \frac{1}{M} \sum_{m=1}^{M} f\left(L_{\boldsymbol{k}} + \mathrm{e}^{2\pi \mathrm{i} m/M}\right) \,. \tag{16}$$

The sum (16) exponentially converges to the actual value of the integral therefore even $M = 32$ or $M = 64$ points are usually enough to give machine precision.

References

[1] S. M. Cox and P. C. Matthews. Exponential time differencing for stiff systems. *J. Comput. Phys.*, 176(2):430–455, 2002.

[2] A.-K. Kassam and L. N. Trefethen. Fourth-order time-stepping for stiff PDEs. *SIAM J. Sci. Comput.*, 26(4):1214–1233, 2005.