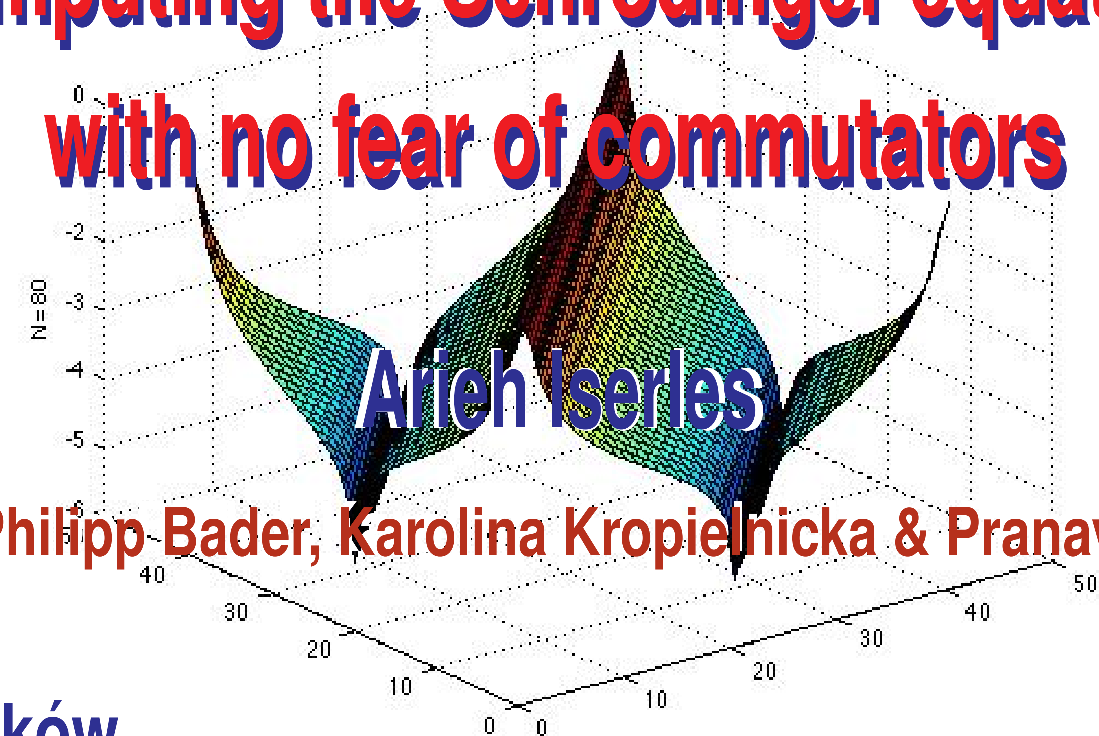
UNIVERSITY OF CAMBRIDGE

CENTRE FOR MATHEMATICAL SCIENCES

DEPARTMENT OF APPLIED MATHEMATICS
& THEORETICAL PHYSICS

Computing the Schrödinger equation with no fear of commutators



Arieh Iserles

(with Philipp Bader, Karolina Kropielnicka & Pranav Singh)

6ECM, Kraków

July 2012

THE PROBLEM

Consider the linear Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial u}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 u}{\partial x^2} - V(x)u, \quad t \geq 0, \qquad -\frac{1}{\sqrt{2m}} \leq x \leq \frac{1}{\sqrt{2m}},$$

where $0 < \hbar \ll 1$, and the interaction potential V is a smooth periodic function, given with smooth periodic initial and boundary conditions.

We can contemplate this equation on a multivariate torus, $\mathrm{i}\hbar\frac{\partial u}{\partial t} = -\frac{\hbar^2}{2m}\Delta u - V(\boldsymbol{x})u$ and our theory scales up to this setting. However, our purpose here is to explore ideas, rather than presenting them in the greatest possible generality.

In physically-interesting coordinates, translating space and time,

$$\frac{\partial u}{\partial t} = \mathrm{i}\varepsilon\frac{\partial^2 u}{\partial x^2} + \mathrm{i}\varepsilon^{-1}V(x)u, \quad x \in [-1, 1],$$

where $0 < \varepsilon \ll 1$ but much larger than the Planck constant $\hbar \approx 1.05 \cdot 10^{-34}$: it is useful to think of the range $10^{-8} \leq \varepsilon \leq 10^{-4}$.

EXPONENTIALS AND SPLITTINGS...

The conventional approach: replace $\dfrac{\partial^2 u}{\partial x^2}$ by a linear combination of function values (or Fourier modes). This results in the linear system

$$\boldsymbol{y}' = (\varepsilon\mathcal{K} + \varepsilon^{-1}\mathcal{D})\boldsymbol{y}, \quad t \geq 0, \qquad \boldsymbol{y}(0) = \boldsymbol{y}_0,$$

whose exact solution is

$$\boldsymbol{y}(t) = \mathrm{e}^{t(\varepsilon\mathcal{K}+\varepsilon^{-1}\mathcal{D})}\boldsymbol{y}_0.$$

It is usual to split the exponential, e.g. with the Strang splitting

$$\mathrm{e}^{\frac{1}{2}t\varepsilon\mathcal{K}}\mathrm{e}^{t\varepsilon^{-1}\mathcal{D}}\mathrm{e}^{\frac{1}{2}t\varepsilon\mathcal{K}} = \mathrm{e}^{t(\varepsilon\mathcal{K}+\varepsilon^{-1}\mathcal{D})} + \mathcal{O}\big(t^3\big)$$

or higher-order splittings. Typically, such high-order splittings are obtained either via the Yošida device or similar techniques, and result in palindromic expressions of the form

$$\mathrm{e}^{\alpha_1 t\varepsilon\mathcal{K}}\mathrm{e}^{\beta_1 t\varepsilon^{-1}\mathcal{D}}\mathrm{e}^{\alpha_2 t\varepsilon\mathcal{K}}\cdots\mathrm{e}^{\beta_r t\varepsilon^{-1}\mathcal{D}}\cdots\mathrm{e}^{\alpha_2 t\varepsilon\mathcal{K}}\mathrm{e}^{\beta_1 t\varepsilon^{-1}\mathcal{D}}\mathrm{e}^{\alpha_1 t\varepsilon\mathcal{K}}.$$

(Palindromy implies even order and helps to preserve unitarity.)

- **Good news**: the splitting separates scales. Exponentials with ε and ε^{-1} are kept apart and the computation of each individual exponential is cheap.
- **Bad news**: we need plenty of exponentials to attain reasonable order. The number of exponentials increases *exponentially* with order and this renders high-order methods very expensive.
- **Ugly news**: the 'scale' of the exponentials doesn't decrease. ideally, we would have liked to combine numerics with asymptotics, namely for the arguments of exponentials to become progressively smaller – while in the present case they are all of just two interlacing orders of magnitude.

The challenge is to develop a methodology which attains all the advantages without any disadvantages: ***A high-order splitting that separates scales and produces an asymptotic expansion in increasing powers of ε.***

ALGEBRA OF OPERATORS

The vector field is a linear combination of two linear operators: ∂_x^2 and (multiplication by) V. Let us consider the free Lie algebra $\mathfrak{F}$ generated by ∂_x^2 and V. Note that

$$\begin{aligned}
[V,\partial_x^2] &= -(\partial_x^2 V) - 2(\partial_x V)\partial_x, \\
[[V,\partial_x^2],\partial_x^2] &= (\partial_x^4 V) + 4(\partial_x^3 V)\partial_x + 4(\partial_x^2 V)\partial_x^2, \\
[[V,\partial_x^2],V] &= -2(\partial_x V)^2, \\
[[[V,\partial_x^2],\partial_x^2],\partial_x^2] &= -(\partial_x^6 V) - 6(\partial_x^5 V)\partial_x - 12(\partial_x^4 V)\partial_x^2 - 8(\partial_x^3 V)\partial_x^3
\end{aligned}$$

and so on. In general, $\mathfrak{F} \subset \mathfrak{G}$, where $\mathfrak{G}$ is the Lie algebra

$$\mathfrak{G} = \left\{ \sum_{k=0}^{n} y_k(x)\partial_x^k \,:\, n \in \mathbb{Z}_+,\ y_0, y_1, \ldots, y_n \text{ smooth \& periodic} \right\}.$$

$$\text{girth}\left(\sum_{k=0}^{n} y_k(x)\partial_x^k \right) = n$$

is the girth of an element in $\mathfrak{G}$.

PROPOSITION *We have* $\mathrm{girth}([X,Y]) = (\mathrm{girth}(X) + \mathrm{girth}(Y) - 1)_+$ *for all* $X, Y \in \mathfrak{G}$.

COROLLARY *Any nested commutator in* $\mathfrak{F}$ *which has at least two more* V*s than* ∂_x^2*s is necessarily zero.*

For example, out of the 41 terms in the Hall basis of all the elements of $\mathfrak{F}$ which can be written with at most 7 'letters',

$$
\begin{gathered}
[[[V,\partial_x^2],V],V], \quad [[[[V,\partial_x^2],V],V],V], \quad [[[[[V,\partial_x^2],\partial_x^2],V],V],V], \\
[[[[[V,\partial_x^2],V],V],V],V], \quad [[[[V,\partial_x^2],V],V],[V,\partial_x^2]], \\
[[[[[[V,\partial_x^2],\partial_x^2],V],V],V],V], \quad [[[[[[V,\partial_x^2],V],V],V],V],V], \\
[[[[[V,\partial_x^2],V],V],V],[V,\partial_x^2]], \quad [[[[V,\partial_x^2],V],V],[[V,\partial_x^2],\partial_x^2]], \\
[[[[V,\partial_x^2],V],V],[[V,\partial_x^2],V]]
\end{gathered}
$$

are all zero.

But this is not the real reason why all this is important! The real reason is that V never walks on its own, it is always multiplied by a large parameter ε^{-1}. The elimination of "ε^{-1}-rich" terms allows us to derive asymptotic splittings!

SYMMETRIC ZASSENHAUS SPLITTING

We commence from the symmetric Baker–Campbell–Hausdorff formula

$$e^{tX} e^{tY} e^{tX} = e^{\mathrm{sBCH}(t;X,Y)},$$

where

$$\begin{aligned}\mathrm{sBCH}(t;X,Y) = {} & t(2X+Y) - t^3\left(\tfrac{1}{24}[[X,Y],X] + \tfrac{1}{12}[Y,[X,Y]]\right)\\ & + t^5\left(\tfrac{7}{360}[X,[X,[X,[X,Y]]]] + \tfrac{1}{360}[Y,[Y,[Y,[X,Y]]]]\right.\\ & - \tfrac{1}{90}[X,[Y,[Y,[X,Y]]]] + \tfrac{1}{45}[Y,[X,[X,[X,Y]]]]\\ & \left. + \tfrac{1}{60}[X,[X,[Y,[X,Y]]]] + \tfrac{1}{30}[Y,[Y,[X,[X,Y]]]]\right)\\ & + \mathcal{O}\left(t^7\right).\end{aligned}$$

It is possible to expand as far as we wish, noting that all the expansion terms live in the free Lie algebra generated by the 'letters' X and Y. Such algebras possess convenient bases that can be constructed algorithmically, e.g. the Hall basis, the Lyndon basis and the Dynkin basis.

The classical Zassenhaus splitting is

$$\mathrm{e}^{t(X+Y)} = \mathrm{e}^{tX}\mathrm{e}^{tY}\mathrm{e}^{t^2 U_2(X,Y)}\mathrm{e}^{t^3 U_3(X,Y)}\mathrm{e}^{t^4 U_4(X,Y)}\cdots,$$

where

$$U_2(X,Y) = -\tfrac{1}{2}[X,Y],$$
$$U_3(X,Y) = \tfrac{1}{3}[Y,[X,Y]] + \tfrac{1}{6}[X,[X,Y]],$$
$$U_4(X,Y) = -\tfrac{1}{24}[[[X,Y],X],X] - \tfrac{1}{8}[[[X,Y],X],Y] - \tfrac{1}{8}[[[X,Y],Y],Y].$$

We are, however, interested in symmetric (i.e., palindromic) splittings. Moreover, powers of t are misleading: we need to reckon with three parameters: ε (small!), the number N of degrees of freedom once we semi-discretize (large!) and the time step Δt (small!). It makes sense to reduce this to a single currency by requiring

$$N \sim \mathcal{O}\left(\varepsilon^{-\rho}\right), \qquad \Delta t \sim \mathcal{O}(\varepsilon^{\sigma})$$

for some $\rho, \sigma > 0$. Note that ∂_x scales like $\mathcal{O}(N) = \mathcal{O}\left(\varepsilon^{-\rho}\right)$.

In the sequel we assume that $\rho = \sigma = \frac{1}{2}$.

SYMMETRIC ZASSENHAUS

We seek a splitting of the form

$$\mathrm{e}^{\mathrm{i}(\Delta t)(\varepsilon\partial_x^2+\varepsilon^{-1}V)} \approx \mathrm{e}^{R_0}\mathrm{e}^{R_1}\cdots\mathrm{e}^{R_s}\mathrm{e}^{T_{s+1}}\mathrm{e}^{R_s}\cdots\mathrm{e}^{R_1}\mathrm{e}^{R_0},$$

where (recalling in line with our 'currency conversion' that $\partial_x \sim \mathcal{O}\left(\varepsilon^{-1/2}\right)$)

$$R_k = R_k(\Delta t, \varepsilon, V) = \mathcal{O}\left(\varepsilon^{k-1/2}\right), \qquad k = 0, 1, \ldots, s,$$
$$T_{s+1} = T_{s+1}(\Delta t, \varepsilon, V) = \mathcal{O}\left(\varepsilon^{s+1/2}\right).$$

Let $\tau = \mathrm{i}\Delta t = \mathcal{O}\left(\varepsilon^{1/2}\right)$. Since $\tau\varepsilon^{-1}\partial_x^2 \sim \mathcal{O}\left(\varepsilon^{1/2}\right)$ and $\tau\varepsilon^{-1}V \sim \mathcal{O}\left(\varepsilon^{-1/2}\right)$, we commence by 'knocking out' the potential:

$$R_0 = \tfrac{1}{2}\tau\varepsilon^{-1}V, \qquad Y_0 = \tau\varepsilon\partial_x^2 + \tau\varepsilon^{-1}V,$$

and rewrite the sBCH formula as

$$\mathrm{e}^{Y_0} = \mathrm{e}^{R_0}\mathrm{e}^{\mathrm{sBCH}(-R_0, Y_0)}\mathrm{e}^{R_0}.$$

In what follows we restrict ourselves to an $\mathcal{O}\left(\varepsilon^{7/2}\right)$ expansion, i.e. $s = 2$. After long calculation:

1 Expanding $\mathsf{sBCH}(-\frac{1}{2}\tau\varepsilon^{-1}V, \tau\varepsilon\partial_x^2 + \tau\varepsilon^{-1}V)$ with the sBCH formula and replacing elements from $\mathfrak{F}$ by terms from $\mathfrak{G}$;

2 Throwing away all zero terms, but also all terms which are $\mathcal{O}(\varepsilon^{\alpha})$ for $\alpha \geq -\frac{7}{2}$;

3 Expressing everything in the Hall basis

we obtain the asymptotic expansion

$$\begin{aligned}\mathsf{sBCH}(X,Y) = {} & \tau\varepsilon H_1 - \tfrac{1}{12}\tau^3\varepsilon H_4 - \tfrac{1}{24}\tau^3\varepsilon^{-1}H_5 - \tfrac{1}{720}\tau^5\varepsilon H_{10} - \tfrac{1}{720}\tau^5\varepsilon^{-1}H_{11} \\ & - \tfrac{1}{180}\tau^5\varepsilon H_{13} - \tfrac{1}{288}\tau^5\varepsilon^{-1}H_{14} + \tfrac{1}{30240}\tau^7\varepsilon^{-1}H_{27} + \tfrac{1}{7560}\tau^7\varepsilon^{-1}H_{32} \\ & - \tfrac{101}{120960}\tau^7\varepsilon^{-1}H_{35} - \tfrac{89}{120960}\tau^7\varepsilon^{-1}H_{40} + \mathcal{O}\left(\varepsilon^{7/2}\right),\end{aligned}$$

where all that survives of the 41 terms in the Hall basis is

$$H_1 = \partial_x^2, \quad H_4 = (\partial_x^4 V) + 4(\partial_x^3 V)\partial_x + 4(\partial_x^2 V)\partial_x^2, \quad H_5 = -2(\partial_x V)^2,$$
$$H_{10} = -\{6(\partial_x^5 V)(\partial_x V) + 12(\partial_x^4 V)(\partial_x^2 V) + 8(\partial_x^3 V)^2\}$$
$$- 24\{(\partial_x^4)(\partial_x V) + (\partial_x^3 V)(\partial_x^2 V)\}\partial_x - 24(\partial_x^3)(\partial_x V)\partial_x^2,$$
$$H_{11} = 8(\partial_x^2 V)(\partial_x V)^2, \quad H_{13} = \{2(\partial_x^5 V)(\partial_x V) - 4(\partial_x^4)(\partial_x^2 V) - 4(\partial_x^3 V)^2\}$$
$$+ 8\{(\partial_x^4 V)(\partial_x V) - (\partial_x^3 V)(\partial_x^2 V)\}\partial_x + 4(\partial_x^3 V)(\partial_x V)\partial_x^2,$$
$$H_{14} = -4(\partial_x^2 V)(\partial_x V)^2, \quad H_{27} = -48(\partial_x^3 V)(\partial_x V)^3,$$
$$H_{32} = 16(\partial_x^3 V)(\partial_x V)^3 + 32(\partial_x^2 V)^2(\partial_x V)^2,$$
$$H_{35} = -8(\partial_x^3 V)(\partial_x V)^3 - 16(\partial_x^2 V)^2(\partial_x V)^2, \quad H_{40} = -32(\partial_x^2 V)^2(\partial_x V)^2.$$

4 We next express the remaining commutators as terms in the larger Lie algebra $\mathfrak{G}$.

But here we have a problem!!! Ultimately, we will replace derivatives by (finite-dimensional) differentiation matrices. Even derivatives yield symmetric matrices and multiplication by i results in skew-Hermitian matrices – ***good!*** But, by the same token, odd derivatives $\Rightarrow$ skew-symmetric matrices $\Rightarrow$ (after multiplication by i) Hermitian matrices – ***bad!*** We must somehow get rid of odd derivatives!!!

5 Replacement of odd derivatives:

The main idea is to replace odd derivatives by linear combinations of even derivatives:

$$y(x)\partial_x = -\tfrac{1}{2}\int_0^x y(\xi)\,\mathrm{d}\xi\,\partial_x^2 - \tfrac{1}{2}y'(x) + \tfrac{1}{2}\partial_x^2\left[\int_0^x y(\xi)\,\mathrm{d}\xi\,\cdot\right],$$

$$y(x)\partial_x^3 = -y'(x)\partial_x^2 - \tfrac{1}{4}\int_0^x y(\xi)\,\mathrm{d}\xi\,\partial_x^4 + \tfrac{1}{4}y'''(x) - \tfrac{1}{2}\partial_x^2[y'(x)\,\cdot\,] + \tfrac{1}{4}\partial_x^4\left[\int_0^x y(\xi)\,\mathrm{d}\xi\,\cdot\right]$$

and so on. The outcome is

$$\begin{aligned}\mathrm{sBCH}(-R_0, Y_0) = {} & \tau\varepsilon\partial_x^2 - \tfrac{1}{12}\tau^3\varepsilon\{-(\partial_x^4 V) + 2\partial_x^2[(\partial_x^2 V)\cdot] + 2(\partial_x^2 V)\partial_x^2] + \tfrac{1}{12}\tau^3\varepsilon^{-1}(\partial_x V)^2 \\ & - \tfrac{1}{720}\tau^5\varepsilon\{6(\partial_x^5 V)(\partial_x V) + 12(\partial_x^4 V)(\partial_x^2 V) + 4(\partial_x^3 V)^2 \\ & \quad - 12\partial_x^2[(\partial_x^3 V)(\partial_x V)\cdot] - 12(\partial_x^3 V)(\partial_x V)\partial_x^2\} - \tfrac{1}{90}\tau^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 \\ & \quad - \tfrac{1}{90}\tau^5\varepsilon\{3(\partial_x^5 V)(\partial_x V) - 2(\partial_x^4 V)(\partial_x^2 V) - 4(\partial_x^3 V)^2 \\ & \quad + 2\partial_x^2[(\partial_x^3 V)(\partial_x V)\cdot] - 4\partial_x^2[(\partial_x^2 V)^2\cdot] + 4(\partial_x^3)(\partial_x V)\partial_x^2 - 8(\partial_x^2)^2\partial_x^2\} \\ & + \tfrac{1}{72}\tau^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 - \tfrac{1}{630}\tau^7\varepsilon^{-1}(\partial_x^3 V)(\partial_x V)^3 \\ & + \tfrac{1}{945}\tau^7\varepsilon^{-1}\{2(\partial_x^3 V)(\partial_x V)^3 + 4(\partial_x^2 V)^2(\partial_x V)^2\} \\ & + \tfrac{101}{15120}\tau^7\varepsilon^{-1}\{(\partial_x^3 V)(\partial_x V)^3 + 2(\partial_x^2 V)^2(\partial_x V)^2\} \\ & + \tfrac{89}{3780}\tau^7\varepsilon^{-1}(\partial_x^2 V)^2(\partial_x V)^2 + \mathcal{O}\left(\varepsilon^{7/2}\right).\end{aligned}$$

6 Arranging in increasing powers of ε:

$$
\begin{aligned}
&\mathrm{sBCH}(-R_0, Y_0)\\
&= \overbrace{\tau\varepsilon\partial_x^2 + \tfrac{1}{12}\tau^2\varepsilon^{-1}(\partial_x V)^2\}}^{\varepsilon^{1/2}}\\
&\quad + \overbrace{\tfrac{1}{360}\tau^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 - \tfrac{1}{6}\tau^3\varepsilon\{\partial_x^2[(\partial_x^2 V)\cdot] + (\partial_x^2 V)\partial_x^2\}}^{\varepsilon^{3/2}}\\
&\quad + \overbrace{\tfrac{1}{12}\tau^3\varepsilon(\partial_x^4 V)}^{\varepsilon^{5/2}}\\
&\quad + \overbrace{\tfrac{1}{180}\tau^5\varepsilon\{-5\partial_x^2[(\partial_x^3 V)(\partial_x V)\cdot] + 4\partial_x^2[(\partial_x^2 V)^2\cdot] - 5(\partial_x^3)(\partial_x V)\partial_x^2 + 4(\partial_x^2 V)^2\partial_x^2\}}^{\varepsilon^{5/2}}\\
&\quad + \overbrace{\tfrac{1}{15120}\tau^7\varepsilon^{-1}\{109(\partial_x^3 V)(\partial_x V)^3 + 622(\partial_x^2 V)^2(\partial_x V)^2\}}^{\varepsilon^{5/2}} + \mathcal{O}\left(\varepsilon^{7/2}\right).
\end{aligned}
$$

Note that the expansion now starts from $\mathcal{O}\left(\varepsilon^{1/2}\right)$: we have got rid of the $\mathcal{O}\left(\varepsilon^{-1/2}\right)$ term!

We now set

$$R_1 = \tfrac{1}{2}\tau\varepsilon\partial_x^2 + \tfrac{1}{24}\tau^3\varepsilon^{-1}(\partial_x V)^2, \qquad Y_1 = \mathrm{sBCH}(-R_0, Y_0).$$

We repeat steps 1–6 and this results in

$$\begin{aligned}\mathrm{sBCH}(-R_1, Y_1) = {} & \overbrace{\tfrac{1}{360}\tau^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 - \tfrac{1}{6}\tau^3\varepsilon\{\partial_x^2[(\partial_x^2 V)\cdot] + (\partial_x^2 V)\partial_x^2\}}^{\varepsilon^{3/2}} \\ & + \overbrace{\tfrac{1}{12}\tau^3\varepsilon(\partial_x^4 V)}^{\varepsilon^{5/2}} \\ & + \overbrace{\tfrac{1}{180}\tau^5\varepsilon\{-\partial_x^2[(\partial_x^3 V)(\partial_x V)\cdot] + 8\partial_x^2[(\partial_x^2 V)^2\cdot] - (\partial_x^3)(\partial_x V)\partial_x^2 + 8(\partial_x^2 V)^2\partial_x^2\}}^{\varepsilon^{5/2}} \\ & + \overbrace{\tfrac{1}{7560}\tau^7\varepsilon^{-1}\{9(\partial_x^3 V)(\partial_x V)^3 + 8(\partial_x^2 V)^2(\partial_x V)^2\}}^{\varepsilon^{5/2}} + \mathcal{O}\left(\varepsilon^{7/2}\right).\end{aligned}$$

Thus, we have disposed of the $\mathcal{O}(\varepsilon^{1/2})$ terms: just one step, knocking out the $\mathcal{O}(\varepsilon^{3/2})$ terms, and we'll be done!

Thus, finally, R_2 is half the leading term in the expansion and

$$T_3 = Y_2 = \mathrm{sBCH}(-R_1, Y_1).$$

To sum up,

$$\begin{aligned}
R_0 &= \tfrac{1}{2}\tau\varepsilon^{-1}V = \mathcal{O}\left(\varepsilon^{-1/2}\right), \\
R_1 &= \tfrac{1}{2}\tau\varepsilon\partial_x^2 + \tfrac{1}{24}\tau^3\varepsilon^{-1}(\partial_x V)^2 = \mathcal{O}\left(\varepsilon^{1/2}\right), \\
R_2 &= -\tfrac{1}{12}\tau^3\varepsilon\{\partial_x^2[(\partial_x^2 V)\,\cdot\,] + (\partial_x^2 V)\partial_x^2\} + \tfrac{1}{120}\tau^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 \\
&= \mathcal{O}\left(\varepsilon^{3/2}\right), \\
T_3 &= \tfrac{1}{12}\tau^3\varepsilon(\partial_x^4 V) + \tfrac{1}{180}\tau^5\varepsilon\{-\partial_x^2[(\partial_x^3 V)(\partial_x V)\,\cdot\,] \\
&\qquad + 8\partial_x^2[(\partial_x^2 V)^2\,\cdot\,] - (\partial_x^3)(\partial_x V)\partial_x^2 + 8(\partial_x^2 V)^2\partial_x^2\} \\
&\quad + \tfrac{1}{7560}\tau^7\varepsilon^{-1}\{9(\partial_x^3 V)(\partial_x V)^3 + 8(\partial_x^2 V)^2(\partial_x V)^2\} = \mathcal{O}\left(\varepsilon^{5/2}\right).
\end{aligned}$$

We have our asymptotic splitting!

Once we discretise using nodal values, e^{R_0} is a diagonal matrix, while matrix/vector products with R_1, R_2 and T_3 can be computed by FFTs as part of a Krylov-subspace-based computation of their exponentials. If we discretise using Fourier expansions, also e^{R_0} is computed with FFT.

Stability

Suppose that each $(2r)$th derivative is replaced by the Hermitian differentiation matrix $\mathcal{K}_{2r}$. Since $\tau = \mathrm{i}\Delta t$ always features with an odd power, $\mathcal{K}_{2r}$, as well as the diagonal matrix that we obtain from discretizing multiplication operators, are multiplied by $\pm\mathrm{i}$ and become skew-Hermitian. For example

$$\begin{aligned} R_0 &\rightsquigarrow \tfrac{1}{2}(\Delta t)\varepsilon^{-1}\mathrm{i}\mathcal{D}_V, \\ R_1 &\rightsquigarrow \tfrac{1}{2}(\Delta t)\varepsilon\mathrm{i}\mathcal{K}_2 - \tfrac{1}{24}(\Delta t)^3\varepsilon^{-1}\mathrm{i}\mathcal{D}_{V_x^2}. \end{aligned}$$

However... Because of replacement of odd derivatives, we have expressions of the form $\mathrm{i}\mathcal{D}\mathcal{H}$ and $\mathrm{i}\mathcal{H}\mathcal{D}$ where both $\mathcal{D}$ and $\mathcal{H}$ are Hermitian – and they are not skew-Hermitian! Is this a problem?

Not at all! Recalling that each term in $\mathfrak{F}$ is multiplied by $\pm\mathrm{i}$, we note that $\mathsf{FLA}(\mathrm{i}\partial_x^2, \mathrm{i}V) \subset \mathfrak{su}(\mathbb{C})$. All our operations, using sBCH but also the replacement of odd derivatives, are Lie-algebra-compliant (linear combinations and commutators), therefore everything we have obtained in the course of our algorithm stays within $\mathfrak{su}(\mathbb{C})$.

As a sanity check, $\mathcal{K}$ symmetric, $\mathcal{D}$ diagonal $\Rightarrow$ $\mathrm{i}(\mathcal{K}\mathcal{D} + \mathcal{D}\mathcal{K})$ skew-Hermitian, hence

$$
\begin{aligned}
R_2 &\leadsto \tfrac{1}{12}(\Delta t)^3 \varepsilon \mathrm{i}(\mathcal{K}_2 \mathcal{D}_{V_x^2} + \mathcal{D}_{V_x^2}\mathcal{K}_2) + \tfrac{1}{120}(\Delta t)^5 \varepsilon^{-1} \mathrm{i}\mathcal{D}_{V_{xx}V_x^2} \\
T_3 &\leadsto \tfrac{1}{12}\tau^3 \varepsilon \mathcal{D}_{\partial_x^4 V} + \tfrac{1}{180}\tau^5 \varepsilon \{8(\mathcal{K}_2 \mathcal{D}_{(\partial_x^2 V)^2} + \mathcal{D}_{(\partial_x^2 V)^2}\mathcal{K}_2) \\
&\qquad - (\mathcal{K}_2 \mathcal{D}_{(\partial_x^3 V)(\partial_x V)} + \mathcal{D}_{(\partial_x^3 V)(\partial_x V)}\mathcal{K}_2)\} \\
&\qquad + \tfrac{1}{7560}\tau^7 \varepsilon^{-1} \{9\mathcal{D}_{(\partial_x^3 V)(\partial_x V)^3} + 8\mathcal{D}_{(\partial_x^2 V)^2(\partial_x V)^2}\}
\end{aligned}
$$

and it is easy to verify that both R_2 and T_3 are skew-Hermitian.

THEOREM *It is true that* $R_k \in \mathfrak{su}(\mathbb{C})$, $k = 0, 1, \ldots, s$, *and* $T_{s+1} \in \mathfrak{su}(\mathbb{C})$.

LEADING WITH THE DERIVATIVE

Instead of $R_0 = \tau\varepsilon^{-1}V$, we start from $R_0 = \tau\varepsilon\partial_x^2$. On the face of it, this makes no sense, because we want to eliminate first the $\mathcal{O}\left(\varepsilon^{-1/2}\right)$ term. However, let's try...

$$R_0 = \tfrac{1}{2}\tau\varepsilon\partial_x^2 = \mathcal{O}\left(\varepsilon^{1/2}\right),$$

$$R_1 = \tfrac{1}{2}\tau\varepsilon^{-1}V = \mathcal{O}\left(\varepsilon^{-1/2}\right),$$

$$R_2 = \tfrac{1}{12}\tau^3\varepsilon^{-1}(\partial_x V)^2 = \mathcal{O}\left(\varepsilon^{1/2}\right),$$

$$R_3 = \tfrac{1}{24}\tau^3\varepsilon\{\partial_x^2[(\partial_x^2 V)\,\cdot\,] + (\partial_x^2 V)\partial_x^2\} + \tfrac{7}{240}\tau^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 = \mathcal{O}\left(\varepsilon^{3/2}\right),$$

$$T_4 = \tfrac{1}{24}\tau^3\varepsilon(\partial_x^4 V) - \tau^5\varepsilon\{\tfrac{1}{30}\partial_x^2[(\partial_x^3 V)(\partial_x V)\,\cdot\,] + \tfrac{1}{30}(\partial_x^3 V)(\partial_x V)\partial_x^2 - \tfrac{1}{60}\partial_x^2[(\partial_x^2 V)^2\,\cdot\,] - \tfrac{1}{60}(\partial_x^2 V)^2\partial_x^2\} - \tau^7\varepsilon^{-1}\{\tfrac{163}{25704}(\partial_x^3 V)(\partial_x V)^3 + \tfrac{211}{12852}(\partial_x^2 V)^2(\partial_x V)^2\} = \mathcal{O}\left(\varepsilon^{5/2}\right).$$

On the face of it, more terms. However... With nodal values e^{R_0} can be computed with a single FFT, R_1 and R_2 are diagonal (the other way around with Fourier expansions)– only $R_3 = \mathcal{O}\left(\varepsilon^{3/2}\right)$ and $T_4 = \mathcal{O}\left(\varepsilon^{5/2}\right)$ require Krylov subspace methods to compute their exponential!

SPACE DISCRETIZATION

In principle, we have three standard options to discretise ∂_x^{2s} to spectral accuracy:

- **Spectral methods:** We project in L_2 the solution on Nth-degree trigonometric polynomials: the unknowns are ***Fourier coefficients***, $\mathcal{K}$ is diagonal and $\mathcal{D}$ a circulant.
- **Spectral collocation:** We interpolate at equally-spaced points by Nth-degree trigonometric polynomials: the unknowns are ***nodal values***, $\mathcal{K}$ is a circulant and $\mathcal{D}$ is diagonal.
- **Pseudospectral method:** We wrap around a finite-difference method an infinite number of times: again, the unknowns are ***nodal values***, $\mathcal{K}$ is a circulant and $\mathcal{D}$ is diagonal.

Except that, having committed already an $\mathcal{O}\left(\varepsilon^{7/2}\right)$ error, we don't need spectral accuracy!

FINITE DIFFERENCES

We discretise in space to the same $\mathcal{O}\left(\varepsilon^{7/2}\right) = \mathcal{O}\left((\Delta x)^7\right)$ (actually, $\mathcal{O}\left((\Delta x)^8\right)$) accuracy using finite differences: $u_m \approx u(m\Delta x, \cdot)$, $m = -N, \ldots, N$, $\Delta x = \frac{1}{N+\frac{1}{2}}$, and

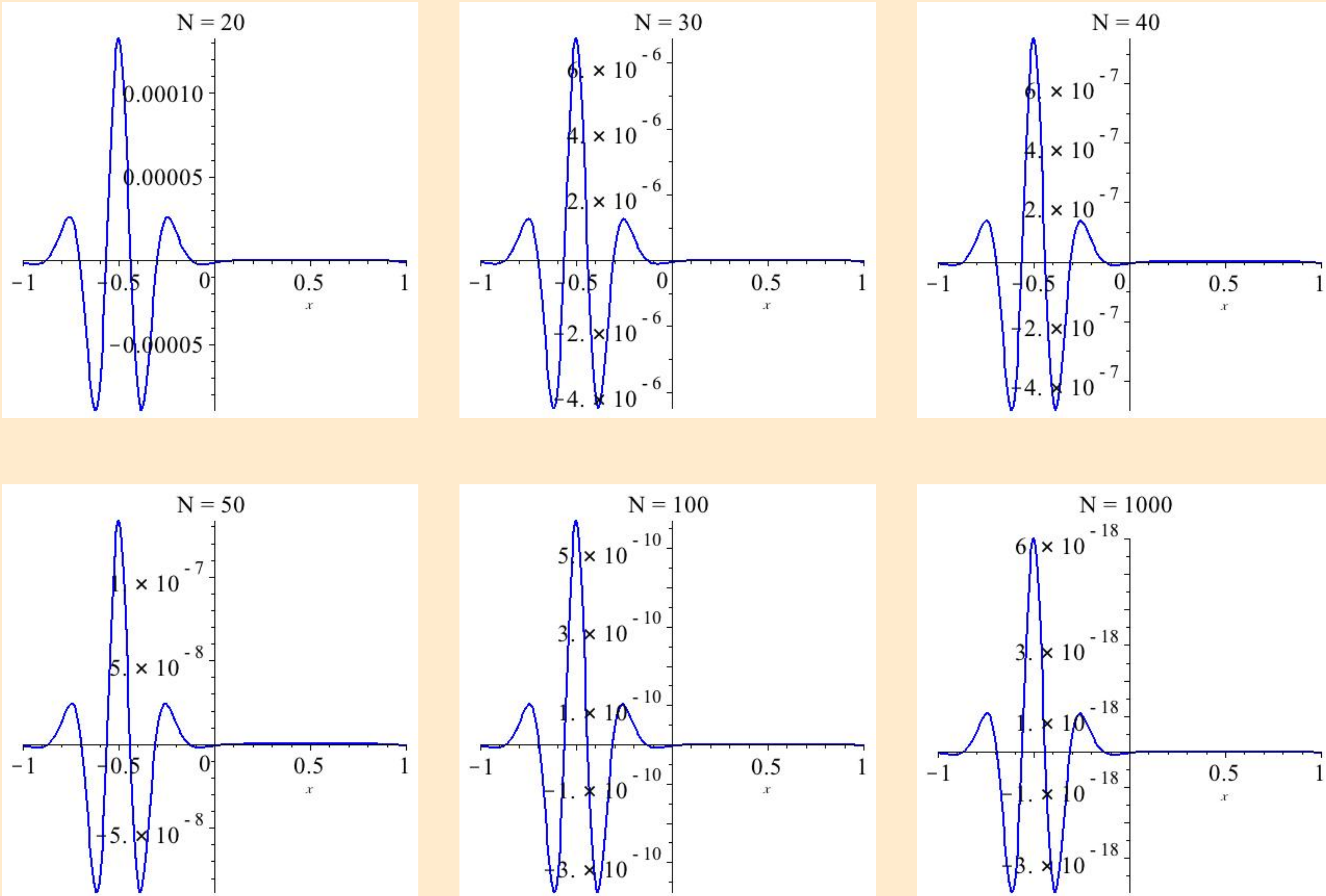
$$u''_m \approx \frac{1}{(\Delta x)^2}\Big(-\tfrac{1}{560}u_{m-4} + \tfrac{8}{315}u_{m-3} - \tfrac{1}{5}u_{m-2} + \tfrac{8}{5}u_{m-1} - \tfrac{205}{72}u_m + \tfrac{8}{5}u_{m+1} - \tfrac{1}{5}u_{m+2} + \tfrac{8}{315}u_{m+3} - \tfrac{1}{560}u_{m+4}\Big).$$

Therefore:

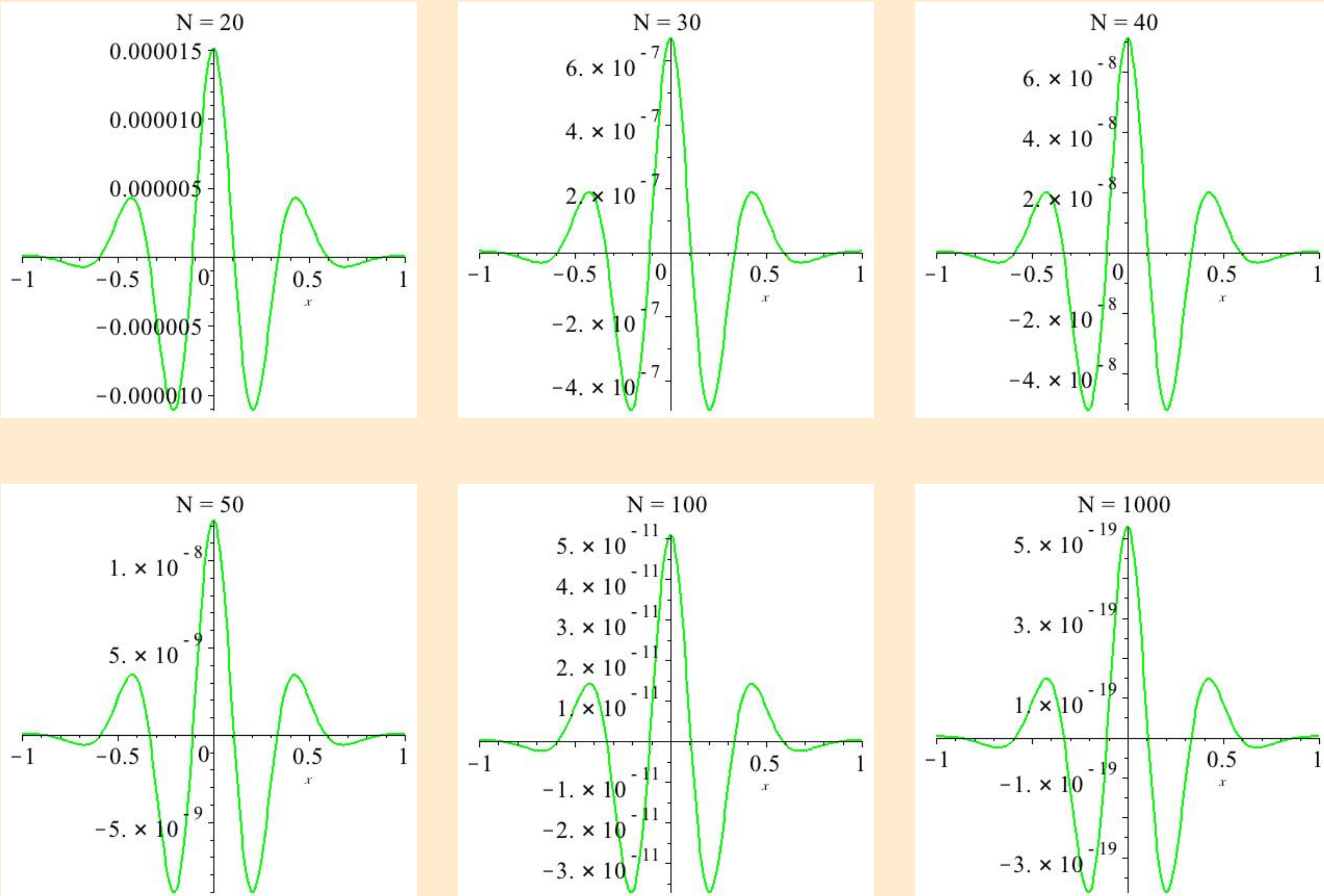
- ! $\mathcal{K}$ is a 9-diagonal circulant, and
- ! $\mathcal{D}$ is a diagonal matrix.

In that case all matrix-vector products are $\mathcal{O}(N)$ operations!

$$u(x) = \frac{1}{2 + \sin \pi x}$$



$$u(x) = \mathrm{e}^{\cos \pi x}$$



So, does the error converge to a multiple of some mystery function?

Yes: in general, for the rth finite difference

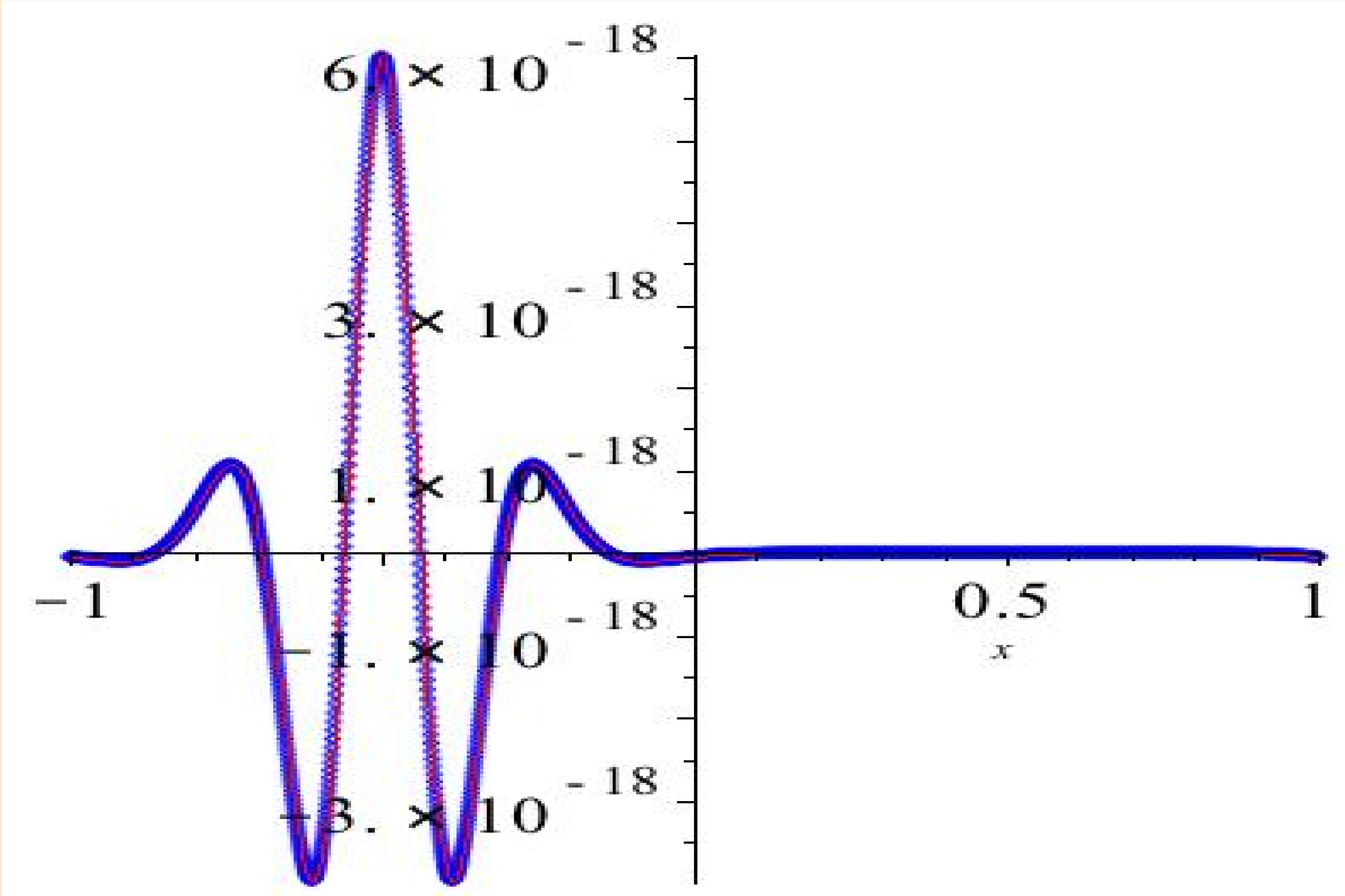
$$u''(x) \approx \frac{1}{(\Delta x)^2} \sum_{m=-r}^{r} a_m u(x + m\Delta x)$$

of order $2r$ we have

$$\sum_{m=-r}^{r} a_m \cos m\theta = -\theta^2 + c_r \theta^{2r+2} + \mathcal{O}\left(\theta^{2r+4}\right),$$

whereby the error is $c_r (\Delta x)^{2r} u^{(2r+2)}(x) + \mathcal{O}\left((\Delta x)^{2r+2}\right)$. In our case $c_4 = -\frac{1}{3150}$, therefore the error is

$$-\tfrac{1}{3150}(\Delta x)^8 u^{(10)}(x) + \mathcal{O}\left((\Delta x)^{10}\right), \qquad \Delta x = \frac{1}{N + \frac{1}{2}}.$$



The error (in blue) and the error estimate (in red) for $u(x) = (2 + \sin \pi x)^{-1}$ and $N (= \mathcal{O}(\varepsilon^{1/2})) = 1000$.

COMPUTING THE EXPONENTIALS

Our method produces a dichotomy:

Arguments of the exponentials are either ***trivial*** or ***small***!

They are trivial if the matrix is diagonal or is a Toeplitz circulant – in both cases computing the exponential is straightforward.

They are small if the argument is $A = \mathcal{O}(\varepsilon^\alpha)$ for some $\alpha > 0$. In that case the mth term in the Krylov basis

$$\{\boldsymbol{v}, A\boldsymbol{v}, A^2\boldsymbol{v}, A^3\boldsymbol{v}\ldots\}$$

is $\mathcal{O}(\varepsilon^{m\alpha})$. Therefore, using estimates of Hochbruck, Lubich & Selhofer, it is possible to prove that we need ridiculously small dimension of a Krylov subspace to compute $\mathrm{e}^{A}\boldsymbol{v}$ to $\mathcal{O}\left(\varepsilon^{7/2}\right)$, say.

Eliminating the derivative first, $R_3 = \mathcal{O}\left(\varepsilon^{3/2}\right)$ and $T_4 = \mathcal{O}\left(\varepsilon^{5/2}\right)$ require dimensions 3 and 2 respectively!

CHALLENGES

- Other pairs (ρ, σ). While the generalisation is straightforward (as a concept: the computation is fiendish), the interesting bit is to explore what are good choices of $\rho, \sigma \in (0, 1)$.
- Time-dependent potentials. Here we need to marry Magnus expansions with symmetric Zassenhaus. Preliminary results indicate that, although the algebra is formidable, this can be done: the underlying free Lie algebra is much more complicated but still embedded in $\mathfrak{G}$ and "girth restrictions" apply.
- Nonlinear Schrödinger. Can we extend our method to the nonlinear Schrödinger equation $\mathrm{i}\hbar u_t = -\frac{\hbar^2}{2m}\nabla u - V(x)u + \lambda|u|^2 u$? Straightforward generalisation does not work but is there a clever way this set of ideas can be extended?
- Other equations. Can this work with the wave equation? With its generalisations, e.g. Klein–Gordon? With Hamiltonian ODEs of the form $H(\boldsymbol{p}, \boldsymbol{q}) = \varepsilon H_1(\boldsymbol{p}, \boldsymbol{q}) + H_2(\boldsymbol{p}, \boldsymbol{q})$, ubiquitous in celestial mechanics?


"That's all Folks!"