

# High accuracy computational methods for the semiclassical Schrödinger equation



Pranav Singh

King's College

DAMTP, Centre for Mathematical Sciences

University of Cambridge

A thesis submitted for the degree of

*Doctor of Philosophy*

June 2016



# Abstract

The computation of Schrödinger equations in the semiclassical regime presents several enduring challenges due to the presence of the small semiclassical parameter. Standard approaches for solving these equations commence with spatial discretisation followed by exponentiation of the discretised Hamiltonian via exponential splittings.

In this thesis we follow an alternative strategy—we develop a new technique, called the symmetric Zassenhaus splitting procedure, which involves directly splitting the exponential of the undiscretised Hamiltonian. This technique allows us to design methods that are highly efficient in the semiclassical regime. Our analysis takes place in the Lie algebra generated by multiplicative operators and polynomials of the differential operator.

This Lie algebra is completely characterised by Jordan polynomials in the differential operator, which constitute naturally symmetrised differential operators. Combined with the  $\mathbb{Z}_2$ -graded structure of this Lie algebra, the symmetry results in skew-Hermiticity of the exponents for Zassenhaus-style splittings, resulting in unitary evolution and numerical stability.

The properties of commutator simplification and height reduction in these Lie algebras result in a highly effective form of *asymptotic splitting*: exponential splittings where consecutive terms are scaled by increasing powers of the small semiclassical parameter. This leads to high accuracy methods whose costs grow quadratically with higher orders of accuracy.

Time-dependent potentials are tackled by developing commutator-free Magnus expansions in our Lie algebra, which are subsequently split using the Zassenhaus algorithm. We present two approaches for developing arbitrarily high-order Magnus–Zassenhaus schemes—one where the integrals are discretised using Gauss–Legendre quadrature at the outset and another where integrals are preserved throughout.

These schemes feature high accuracy, allow large time steps, and the quadratic growth of their costs is found to be superior to traditional approaches such as Magnus–Lanczos methods and Yoshida splittings based on traditional Magnus expansions that feature nested commutators of matrices.

An analysis of these operatorial splittings and expansions is carried out by characterising the highly oscillatory behaviour of the solution.

---

**Keywords:** Semiclassical Schrödinger equations, time-dependent potentials, exponential splittings, Zassenhaus splitting, Magnus expansions, Lanczos iterations, Magnus–Zassenhaus schemes, commutator free, high-order methods, asymptotic analysis, Lie algebras, Jordan polynomials, symmetrised differential operators, spectral collocation.

Dedicated to my beloved grandfather, Mr Vijai Bahadur Singh (1924–2017).



## Acknowledgements

It has been real a privilege to be a student of Arie Iserles and I would like to express my gratitude to him for being a great mentor throughout my time at Cambridge. It would be pointless to attempt listing the ways in which he has helped me, for there are way too many, but at the very least I would like to say that my mathematical toolbox would have been much emptier without his guidance and life much duller without his enthusiastic and humorous nature. I will dearly miss our coffee breaks!

One of the better parts of my stay at DAMTP was when Karolina Kropielnicka was also around—her enthusiasm and curiosity is infectious, and working with her was truly a joy. I thoroughly enjoyed collaborating with her on the research projects that this thesis is largely based around. Above everything, though, Karolina has always been a great friend.

I cannot thank Caroline Lasser enough for all the ways in which she has helped shape not only my understanding of the topics that this thesis revolves around, but also the broader research ideas that I am likely to pursue over the coming years. Caroline has always been supportive, encouraging, and extremely kind, and for this I owe her a lot.

I am indebted to Hans Munthe-Kaas and Antonella Zanna for inviting me to Bergen to visit their research group and for the many fruitful discussions that proved crucial for developing my algebraic ideas. However, my gratitude doesn't just extend to this visit—in the first few months of my PhD at Cambridge, when I was just learning the ropes, I had the privilege of sharing an office with Hans and Anders Hansen, and learning a great deal of mathematics from them.

Kurusch Ebrahimi-Fard has also been a great mentor, helping me not only with algebraic ideas and techniques, but also guiding and supporting me through the academic world.

I am also grateful to Marjorie Batchelor and Dimitar Grantcharov for much guidance on algebraic notions and for many helpful discussions.

I would like to thank S. Arun-Kumar, my Master's thesis supervisor, who suggested some very helpful computational techniques that proved crucial in

---

making an efficient solver for deriving the complicated algebraic expressions appearing in the Zassenhaus splittings and forming conjectures regarding the algebra of the symmetrised differential operators.

I am grateful to S. Arun-Kumar, Subhashish Bannerjee, Nilanjana Datta, Arie Iserles, Narayanan Kurur, Sanjiva Prasad and Alexei Shadrin for beautiful introductions to the three subjects at the intersection of which lies this present work—numerical mathematics, algebra and quantum mechanics.

I am grateful for the generous financial support provided by King's College of the University of Cambridge in the form of the King's College Studentship, which supported my entire doctoral research at Cambridge, as well as travel grants for SciCADE 2013 and TNADE 2013, the financial support of Hans Munthe-Kaas and the University of Bergen for sponsoring my visit to Norway, the financial support of MFO Oberwolfach which allowed me to attend two Oberwolfach meetings, and ICMAT Madrid for financially supporting me for the workshop TNADE 2013.

Without the help of my friend and office mate Evangelos Papoutsellis I would have probably never got started writing this thesis. I am also grateful to him, as well as my dear friend Gil Ramos, for guiding me through the academic machinery at Cambridge. In addition to the numerous coffee break discussions, I also had the pleasure of the company of Helge Dietert, Gil Ramos and Marcus Webb on various travels around the world.

I will be eternally grateful to my grandfather Mr V. B. Singh who taught me mathematics and physics when I was young, and whose unshakeable belief in me has been a great source of strength. Nor would I have been able to do without the unconditional support of my parents Dr. Anil Kumar Singh and Dr. Sudha Singh during my doctoral studies.

Above all, I would like to thank my wife, Winnie Singh, for her love, joyful company and the incredible support in my research throughout my time at Cambridge.



# Contents

<b>List of Notation</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Zassenhaus splittings . . . . .	2
1.2 Algebraic structures . . . . .	4
1.3 Magnus–Zassenhaus schemes . . . . .	5
1.4 Outline of the thesis . . . . .	6
<b>2 Numerical tools</b>	<b>9</b>
2.1 Spatial discretisation strategies . . . . .	9
2.1.1 Norms and inner products . . . . .	13
2.1.2 Higher order finite difference methods . . . . .	15
2.1.3 Spectral collocation and pseudospectral methods . . . . .	18
2.1.4 Spectral methods . . . . .	20
2.1.5 Comparisons . . . . .	22
2.1.6 Our discretisation choices . . . . .	24
2.2 Approximation of matrix exponentials . . . . .	26
2.2.1 Padé methods . . . . .	27
2.2.2 Krylov subspace methods . . . . .	28
2.2.3 Trotter splitting and the BCH formula . . . . .	31
2.2.4 Strang splitting and the symmetric BCH . . . . .	34
2.2.5 Yoshida splittings . . . . .	36
2.2.6 Lie groups, Lie algebras and the exponential map . . . . .	37
2.3 The Magnus expansion . . . . .	39
2.3.1 Magnus expansion . . . . .	41
2.3.2 Graphical representation of the Magnus expansion . . . . .	43
2.3.3 Power truncation . . . . .	45
2.3.4 Time symmetry and gain of power . . . . .	47
2.3.5 Discretisation of integrals . . . . .	47

<b>3</b>	<b>Schrödinger equations and their numerical solutions</b>	<b>53</b>
3.1	The time-dependent Schrödinger equation . . . . .	53
3.2	The semiclassical parameter . . . . .	55
3.2.1	Rapid oscillations . . . . .	57
3.3	Traditional methods for the semiclassical TDSE . . . . .	58
3.3.1	Discretisation . . . . .	59
3.3.2	Direct exponentiation . . . . .	60
3.3.3	Lanczos iterations . . . . .	61
3.3.4	Splitting methods . . . . .	62
3.4	Time-dependent potentials . . . . .	64
3.4.1	Magnus expansions . . . . .	65
<b>4</b>	<b>Commutator-free asymptotic splittings</b>	<b>71</b>
4.1	Zassenhaus splitting . . . . .	73
4.1.1	Symmetrisation of Zassenhaus . . . . .	74
4.1.2	Zassenhaus under semiclassical scaling . . . . .	76
4.2	Splitting the undiscretised Hamiltonian . . . . .	77
4.2.1	Solving commutators . . . . .	78
4.2.2	The Lie algebra $\mathfrak{G}$ and reduction of height . . . . .	80
4.2.3	The first stage of Zassenhaus splitting . . . . .	83
4.3	Loss of stability . . . . .	84
4.3.1	Restoring skew-Hermiticity . . . . .	86
4.3.2	Some structural observations . . . . .	88
<b>5</b>	<b>A Lie algebra of symmetrised differential operators</b>	<b>91</b>
5.1	Notations . . . . .	93
5.2	Associative algebra of Jordan polynomials . . . . .	96
5.2.1	Explicit form of the coefficients . . . . .	100
5.2.2	The generating function for the coefficients . . . . .	104
5.3	Lie algebra of Jordan polynomials . . . . .	110
5.3.1	Characterisation of the Lie algebra . . . . .	115
5.4	Tables of coefficients . . . . .	117
5.5	A finite-dimensional example . . . . .	118
<b>6</b>	<b>Symmetric Zassenhaus splittings</b>	<b>121</b>
6.1	Working with symmetrised differential operators . . . . .	122
6.1.1	The most general case of interest . . . . .	122
6.1.2	A more specific case . . . . .	123
6.1.3	Schrödinger equation in the language of $\mathfrak{G}$ . . . . .	124
6.1.4	Consequences for exponential splitting schemes . . . . .	125

6.1.5	Relation to other algebras . . . . .	128
6.2	The first Zassenhaus splitting . . . . .	129
6.3	Zassenhaus splitting of the second kind . . . . .	132
6.4	Computation of exponentials . . . . .	133
6.4.1	Evaluating matrix–vector products . . . . .	134
6.4.2	Estimating Lanczos iterations . . . . .	135
6.5	Splittings for different scaling laws . . . . .	137
6.5.1	Number of Lanczos iterations . . . . .	137
6.6	Termination of the Zassenhaus algorithm . . . . .	138
6.7	Cost of Zassenhaus splittings . . . . .	138
6.8	Scaling choices and global cost concerns . . . . .	142
6.9	Stability . . . . .	143
6.10	Numerical experiments . . . . .	144
6.10.1	Finding a reference solution . . . . .	146
6.10.2	Comparison with Yoshida splittings . . . . .	149
<b>7</b>	<b>Time-dependent potentials and Magnus expansions</b>	<b>151</b>
7.1	Truncation by powers of the semiclassical parameter . . . . .	153
7.2	Approach 1: Discretised integrals . . . . .	155
7.2.1	The self-adjoint basis of Munthe-Kaas & Owren . . . . .	156
7.2.2	Computations in the Lie algebra $\mathfrak{H}$ . . . . .	157
7.2.3	Commutator-free Magnus expansion . . . . .	159
7.3	Approach 2: Undiscretised integrals . . . . .	160
7.3.1	Computing with integrals intact . . . . .	161
7.3.2	Simplifying integrals over Magnus polytopes . . . . .	162
7.3.3	Time symmetry and gain of order . . . . .	164
7.3.4	A simplifying notation . . . . .	165
7.4	Exponentiation of Magnus expansions . . . . .	166
<b>8</b>	<b>Magnus–Zassenhaus schemes</b>	<b>171</b>
8.1	Approach 1: Discretised integrals . . . . .	172
8.2	Approach 2: Undiscretised integrals . . . . .	176
8.2.1	Evaluation of integrals . . . . .	177
8.3	Numerical results . . . . .	179
8.3.1	Finding a reference solution . . . . .	181
8.4	High order schemes with undiscretised integrals . . . . .	182
8.4.1	Revisiting time symmetry . . . . .	183
8.4.2	Identifying odd and even components . . . . .	184
8.4.3	Time-symmetrised rewriting . . . . .	186
8.4.4	Zassenhaus splitting after time-symmetrisation . . . . .	187

8.4.5	Numerical results . . . . .	189
8.5	Costs of Magnus–Zassenhaus schemes . . . . .	192
<b>9</b>	<b>Formal error analysis</b>	<b>195</b>
9.1	Highly oscillatory solutions in the semiclassical regime . . . . .	196
9.1.1	The first derivative in the semiclassical limit . . . . .	197
9.1.2	Time-dependent potentials . . . . .	198
9.1.3	Higher derivatives in the semiclassical limit . . . . .	200
9.1.4	Size of symmetrised differential operators . . . . .	204
9.1.5	Size of commutators . . . . .	205
9.2	Oscillatory solutions for truncated Magnus expansions . . . . .	206
9.2.1	Conservation of the perturbed Hamiltonian . . . . .	207
9.2.2	The first non-trivial Magnus expansion . . . . .	208
9.2.3	The second non-trivial Magnus expansion . . . . .	210
9.2.4	Higher order Magnus expansions . . . . .	212
9.3	Error Analysis for Magnus Expansions . . . . .	213
9.3.1	Proof outline . . . . .	214
9.3.2	Error bounds for the truncated Magnus expansion . . . . .	216
9.3.3	Error in discretisation . . . . .	221
9.3.4	A summary of the error bounds . . . . .	223
9.4	Error Analysis for Zassenhaus Splittings . . . . .	224
9.4.1	Error in the first step of the symmetric Zassenhaus algorithm . . . . .	225
9.4.2	Error in the second step of the symmetric Zassenhaus algorithm . . . . .	228
9.4.3	Local error in a single step of $\mathcal{Z}_{1,1}^{[2]}$ . . . . .	229
9.4.4	Error bounds for the first non trivial Zassenhaus splitting . . . . .	230
<b>10</b>	<b>Conclusions and future work</b>	<b>233</b>
10.1	Summary of the thesis . . . . .	233
10.2	Zassenhaus splittings in practice . . . . .	234
10.3	Related equations of quantum mechanics . . . . .	235
10.4	Extension of the algebraic framework . . . . .	237
10.4.1	Commuting elements of the Lie idealiser . . . . .	237
10.4.2	Non-commuting elements of the Lie idealiser . . . . .	238
10.4.3	Wigner equation . . . . .	239
10.4.4	Matrix valued potentials . . . . .	240
	<b>Bibliography</b>	<b>241</b>

# List of Notation

Against each entry is the page at which the notation is introduced.

## Differential operators:

$\partial_x, \partial_x^k$	-	The partial differential operators $\frac{\partial}{\partial x}$ and $\frac{\partial^k}{\partial x^k}$ , respectively.
$\partial_x^k u$	-	The $k$ th partial derivative of $u$ with respect to $x$ , $\frac{\partial^k u}{\partial x^k}$ .
$\Delta u$	-	The Laplacian of $u$ i.e., $\Delta u = \sum_{i=1}^d \frac{\partial^2 u}{\partial x_i^2}$ .
$\partial_t$	-	The partial differential operator $\frac{\partial}{\partial t}$ .
$u', u''$	-	The first and second derivatives of $u$ with respect to its only parameter, respectively.
$u^{(k)}$	-	The $k$ th derivative of $u$ with respect to its only parameter.
$H$	53	The Hamiltonian operator.
$\partial_x^k[f \cdot]$	86	The differential operator $\partial_x^k$ composed with the multiplicative operator $f$ , also written as $\partial_x^k \circ f$ .

## Discretisation parameters:

$\varepsilon$	56	The semiclassical parameter, $\varepsilon \ll 1$ .
$M$	11	The number of spatial grid points. Usually, $M = 2N + 1$ is odd and scales as $M = \mathcal{O}(\varepsilon^{-1})$ for semiclassical Schrödinger equation.
$\Delta x$	9	The resolution of the spatial grid. Typically, $\Delta x = 2/M$ .
$h$	26	The time step. Usually, $h = \mathcal{O}(\varepsilon^\sigma)$ , $\sigma > 0$ , for the semiclassical Schrödinger equation.
$\sigma$	76	The parameter governing scaling of the time step with $\varepsilon$ for the semiclassical Schrödinger equation.

## Values on the grid:

$u_n$	9	The value of $u$ at the $n$ th grid point.
$u'_n$	9	The value of $\partial_x u$ at the $n$ th grid point.
$t_n$	26	$n$ th time step. Usually, $t_{n+1} = t_n + h$ .
$u^n$	40	The time-stepped numerical solution approximating the exact solution $u(t_n)$ at time $t_n$ .

### Discretisation of functions:

$\mathbf{u}$	10	The vector of values of $u$ at grid points.
$\mathbf{u}^n$	26	The time-stepped solution approximating the discretised solution at time $t_n$ , $\mathbf{u}(t_n)$ .
$\mathbf{u}'(t)$	60	The vector of values of $\partial_t u$ at grid points.
$\mathbf{u}'$	10	The vector of values of $\partial_t u$ at grid points (except in Section 2.1 where this is the vector of values of $\partial_x u$ at grid points).
$\mathbf{u}''$	16	The vector of values of $\partial_x^2 u$ at grid points (only used in Section 2.1).
$\hat{\mathbf{u}}$	19	The Fourier transform of the vector $\mathbf{u}$ , $\hat{\mathbf{u}} = \mathcal{F}\mathbf{u}$ .

### Discretisation of operators:

$\mathcal{K}$	25	The degree one differentiation matrix (using spectral collocation, unless specified otherwise).
$\mathcal{K}_k$	11	The degree $k$ differentiation matrix discretising $\partial_x^k$ .
$\mathcal{L} \rightsquigarrow L$	11	The discretisation of the linear operator $\mathcal{L}$ by the matrix $L$ , for instance $\partial_x \rightsquigarrow \mathcal{K}$ .
$\Theta \rightsquigarrow \Omega$	49	The temporal discretisation of integrals in $\Theta$ using quadrature methods to an integral-free expression $\Omega$ .
$\mathcal{F}$	19	The Fourier transform matrix.
$\mathcal{F}^{-1}$	19	The inverse Fourier transform matrix, $\mathcal{F}^{-1} = \mathcal{F}^*$ .
$\mathcal{D}_f$	17	The diagonal matrix featuring values of $f$ at the grid points.
$\mathcal{D}_{a_n}$	17	The diagonal matrix whose diagonal entries are given by the sequence $a_{-N} \dots a_0 \dots a_N$ .

### Norms and inner products:

$\ \mathbf{u}\ _p, \ \mathbf{u}\ _{\ell^p}$	14	The $\ell^p$ norm on the grid $\ \mathbf{u}\ _p = (\sum_n  u_n ^p \Delta x)^{\frac{1}{p}}$ for $1 \leq p < \infty$ .
$\ \mathbf{u}\ _\infty, \ \mathbf{u}\ _{\ell^\infty}$	14	The $\ell^\infty$ norm $\ \mathbf{u}\ _\infty = \max_n  u_n $ .
$\ u\ _p, \ u\ _{L^p(X, \mathbb{C})}$	14	The $L^p$ norm $\ u\ _p = (\int_X  u(x) ^p dx)^{\frac{1}{p}}$ for $1 \leq p < \infty$ .
$\ u\ _\infty, \ u\ _{L^\infty(X, \mathbb{C})}$	14	The $L^\infty$ norm $\ u\ _\infty = \operatorname{ess\,sup}_{u \in X}  u $ .
$\langle u, v \rangle$	13	The $L^2$ inner product, $\langle u, v \rangle = \int \bar{u}(x)v(x) dx$ .
$\langle \mathbf{u}, \mathbf{v} \rangle$	13	The $\ell^2$ inner product, $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^* \mathbf{v} \Delta x$ .
$\mathcal{A}^T$	14	The transpose of the matrix $\mathcal{A}$ .
$\mathcal{A}^*$	13	The adjoint of $\mathcal{A}$ in the context of the inner product $\langle \cdot, \cdot \rangle$ . For vectors and matrices, this is equivalent to $\overline{\mathcal{A}}^T$ , the conjugate transpose.

$\ \mathcal{A}\ _p$	15	The $\ell^p$ norm of matrix $\mathcal{A}$ , induced by the norm $\ \cdot\ _p$ on $\ell^p$ spaces, or the $L^p$ norm of the operator $\mathcal{A}$ induced by the $L^p$ norm of functions.
---------------------	----	---

$\rho(\mathcal{A})$	15	The spectral radius of $\mathcal{A}$ .
---------------------	----	--

### Function spaces:

$\mathcal{G}, \mathcal{H}$	-	Function spaces.
----------------------------	---	------------------

$\mathcal{H}$	-	A Hilbert space.
---------------	---	------------------

$\mathcal{P}_M$	15	The space of degree $M$ polynomials in $x$ .
-----------------	----	--

$\mathcal{T}_M$	18	The space of degree $M$ trigonometric polynomials in $x$ .
-----------------	----	--

$C_p^\infty(X, \mathbb{R}), C_p^\infty(X, \mathbb{C})$	9	The space of periodic smooth real-valued and complex-valued functions on $X$ , respectively.
--	---	--

$L^2(X, \mathbb{R}), L^2(X, \mathbb{C})$	53	The space of real and complex-valued Lebesgue-measurable square integrable functions over $X$ , respectively, such that $\ u\ _2 < \infty$ .
--	----	--

### Splitting methods:

BCH	32	The Baker–Campbell Hausdorff formula.
-----	----	---------------------------------------

sBCH	34	The symmetric Baker–Campbell Hausdorff formula.
------	----	---

$\mathcal{S}_{2n}$	36	The order $2n$ Yoshida splitting.
--------------------	----	-----------------------------------

### Lie algebra and Lie groups:

$\mathfrak{g}, \mathfrak{f}, \mathfrak{h}$	38	Lie algebras.
--	----	---------------

$G$	124	A Lie group.
-----	-----	--------------

$\varphi^{[\mathcal{A}]}(u, t)$	38	The flow of the differential equation
---------------------------------	----	---------------------------------------

$$u(t)' = \mathcal{A}(u(t))$$

describing the evolution of  $u$  from 0 to  $t$ ,

$$u(t) = \varphi^{[\mathcal{A}]}(u(0), t), \quad \forall t \geq 0.$$

exp	39	The exponential map from a Lie algebra to its Lie group $\exp : \mathfrak{g} \rightarrow G$ , which coincides with matrix exponential for matrix groups.
-----	----	--

$U(m)$	38	The Lie group of complex $m \times m$ unitary matrices.
--------	----	---

$\mathfrak{u}(n)$	39	The Lie algebra of complex $m \times m$ skew-Hermitian matrices.
-------------------	----	--

$M(m, \mathbb{R}), M(m, \mathbb{C})$	37	The matrix Lie group of $m \times m$ real-valued and complex-valued matrices.
--------------------------------------	----	---

$I_m$	118	The $m \times m$ identity matrix.
-------	-----	-----------------------------------

$O_m$	118	The $m \times m$ zero matrix.
-------	-----	-------------------------------

## Magnus expansions:

$\varphi_{t_2, t_1}^{[\mathcal{A}]}$	41	The flow of the differential equation
		$u(t)' = \mathcal{A}(t)u(t)$
		describing the evolution of $u$ from $t_1$ to $t_2$ ,
		$u(t_2) = \varphi_{t_2, t_1}^{[\mathcal{A}]} u(t_1), \quad \forall t_2 \geq t_1.$
$\Theta(t_2, t_1)$	41	The Magnus expansion for evolution from $t_1$ to $t_2$ .
$\Theta(h)$	42	The Magnus expansion for evolution from 0 to $h$ .
$\Theta^{[k]}(h)$	43	The components of Magnus expansion with $k$ integrals.
$\Theta_p(h)$	46	The Magnus expansion truncated by powers of $h$ ,

$$\Theta_p(h) = \Theta(h) + \mathcal{O}(h^{p+1}).$$

## Graphical representation of Magnus expansions:

$\hat{\mathbb{T}}_k$	44	The auxiliary set of trees required for defining the trees in the Magnus expansion, indexed by number of integrals.
$\mathbb{T}_k$	44	The set of trees appearing in the Magnus expansion, indexed by number of integrals.
$\tau$	44	A tree used for graphical representation of the Magnus expansion.
$\mathcal{C}_\tau(h)$	44	The term corresponding to the tree $\tau$ .
$\mathcal{C}_\tau \rightsquigarrow \tau$	44	The tree corresponding to the term $\mathcal{C}_\tau$ .
$\alpha(\tau)$	45	The coefficient for the tree $\tau$ in the Magnus expansion.
$\mathbb{F}_k$	45	The set of trees indexed by powers of $h$ .
$\mathbb{E}_k$	154	The set of trees indexed by powers of $\varepsilon$ .

## Algebra:

$\mathcal{A}$	-	An associative algebra.
$\mathcal{C}$	-	A commutative algebra.
$f \circ g$	-	The composition of the functions $f$ and $g$ .
$A \circ B$	-	The composition of the operators $A$ and $B$ .
$a \cdot b$	-	An associative product.
$\mathcal{P}(d)$	115	The ring of polynomials in $d$ with constant coefficients.
$a \bullet b$	93	The Jordan product of $a$ and $b$ , usually the anti-commutator $\frac{1}{2}(a \cdot b + b \cdot a)$ in the context of an associative product such as ‘ $\circ$ ’.



$[a, b]$	31	The Lie bracket of $a$ and $b$ , usually the commutator $a \cdot b - b \cdot a$ in the context of an associative product such as ‘ $\circ$ ’.
$\text{ad}_a$	42	The adjoint map $\text{ad}_a(b) = [a, b]$ .
$\mathcal{I}$	95	A Lie ideal.
$\text{LA}(S)$	78	The Lie algebra generated by the set $S$ .

### Jordan polynomials:

$\langle a \rangle_k^d$	95	The degree $k$ Jordan monomial in $d$ with coefficient $a$ , $\langle a \rangle_k^d = a \bullet d^k = \frac{1}{2}(a \cdot d^k + d^k \cdot a).$
$\mathfrak{F}_k^d$	96	The linear space of Jordan monomials of degree $k$ in $d$ .
$\mathfrak{G}^d$	96	The Lie algebra of Jordan polynomials in $d$ .
$\pi_{n,i}^{k,l}$	97	The coefficients appearing in the simplification of products in the associative algebra $(\mathfrak{G}^d, \cdot, +)$ .
$\lambda_{n,i}^{k,l}$	111	The coefficients appearing in the simplification of commutators in the Lie algebra $(\mathfrak{G}^d, [\cdot, \cdot], +)$ .

### Combinatorics:

$\delta_{ij}, \delta_{i,j}$	-	The Kronecker delta function, $\delta_{i,i} = 1$ and $\delta_{i,j} = 0$ for $i \neq j$ .
$\binom{n}{k}$	-	The binomial coefficient.
$\mathcal{B}_n(x)$	-	The $n$ th Bernoulli polynomial.
$B_n$	-	The $n$ th Bernoulli number.
$\gamma$	-	The Euler constant $\gamma \approx 0.577$ .

### Lie algebras of differential operators:

$\mathfrak{S}$	80	The Lie algebra of linear differential operators.
$\mathcal{M}_f$	123	The left multiplication operator, $\mathcal{M}_f(g) = fg$ .
$\langle f \rangle_k$	124	The symmetrised linear differential operator $\langle f \rangle_k = \mathcal{M}_f \bullet \partial_x^k = \frac{1}{2}(\mathcal{M}_f \circ \partial_x^k + \partial_x^k \circ \mathcal{M}_f).$
$\mathfrak{F}_k$	123	The linear space of Jordan monomials of degree $k$ in $\partial_x$ with function coefficients.
$\mathfrak{G}$	124	The Lie algebra of Jordan polynomials in $\partial_x$ with function coefficients.
$\mathfrak{e}$	125	The Lie triple system of Jordan monomials of even degree in $\partial_x$ with function coefficients.
$\mathfrak{o}$	125	The Lie algebra of Jordan monomials of odd degree in $\partial_x$ with function coefficients.

$\mathfrak{H}$	125	The Lie algebra of skew-Hermitian Jordan polynomials in $\partial_x$ with function coefficients.
$\text{ht}(\mathcal{L})$	81	The <i>height</i> or degree of the linear differential operator $\mathcal{L}$ .

### Zassenhaus splittings:

$W^{[k]}, \mathcal{W}^{[k]}$	131	Exponents in the Zassenhaus splitting in operator form.
$\mathcal{Z}_{n,\sigma}^{[i]}$	131	Symmetric commutator-free Zassenhaus splitting of kind $i = 1, 2$ with accuracy $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$ in operator form.
$\tilde{W}^{[k]}, \tilde{\mathcal{W}}^{[k]}$	133	Exponents in the Zassenhaus splitting following spatial discretisation.
$\tilde{\mathcal{Z}}_{n,\sigma}^{[i]}$	133	Symmetric commutator-free Zassenhaus splitting of kind $i = 1, 2$ with accuracy $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$ following spatial discretisation.
$\mathcal{Z}_{n,\sigma}^{\Theta[\star]}$	174	Magnus–Zassenhaus scheme with accuracy $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$ starting from integral-free ( $\star = M$ ) or integral-preserving ( $\star = I$ ) Magnus expansion.

### Magnus expansions for semiclassical Schrödinger equations:

$\tilde{\Theta}_p^\varepsilon(h)$	154	The Magnus expansion truncated by powers of $\varepsilon$ .
$\tilde{\Theta}_p^{\varepsilon[M]}(h), \tilde{\Theta}_p^{\varepsilon[I]}(h)$	159	The integral-free and integral-preserving Magnus expansions, respectively, truncated by powers of $\varepsilon$ .
$\tilde{\mathcal{B}}_n(h, \zeta)$	165	The $n$ th rescaled Bernoulli polynomial.
$\mu_{j,k}(h)$	165	The integral over the line

$$\mu_{j,k}(h) = \int_0^h \tilde{\mathcal{B}}_j^k(h, \zeta) V(\zeta) d\zeta,$$

where  $V$  is the potential function.

$\Lambda[f]_{a,b}(h)$	165
-----------------------	-----

The integral over the triangle

$$\Lambda[f]_{a,b}(h) = \int_0^h \int_0^\zeta f(h, \zeta, \xi) \partial_x^a V(\zeta) \partial_x^b V(\xi) d\xi d\zeta,$$

where  $V$  is the potential function.

$F^e, F^o$	184
------------	-----

The even and odd components, respectively, of  $F$ ,

$$F^e(\zeta) = \frac{1}{2} (F(\zeta) + F(-\zeta)), \quad F^o(\zeta) = \frac{1}{2} (F(\zeta) - F(-\zeta)).$$

$U$	184
-----	-----

The potential function with origin re-shifted to the middle of the interval,  $U(\zeta) = V(t + \frac{h}{2} + \zeta)$ .

$\mu_{\star,j,k}(h)$  185 The integral over the line,

$$\mu_{\star,j,k}(h) = \int_0^h \tilde{B}_j^k(h, \zeta) U^\star \left( \zeta - \frac{h}{2} \right) d\zeta, \quad \star \in \{e, o\}.$$

$\Lambda[f]_{\star,a,b}(h)$  185 The integral over the triangle

$$\int_0^h \int_0^\zeta f(h, \zeta, \xi) \left[ \partial_x^a U \left( \zeta - \frac{h}{2} \right) \partial_x^b U \left( \xi - \frac{h}{2} \right) \right]^\star d\xi d\zeta,$$

where  $\star \in \{e, o\}$ .

### Miscellaneous:

expbump 24 The bump function

$$\text{expbump}(x) = \begin{cases} \exp \left( -\frac{1}{1-x^2} \right) & |x| < 1, \\ 0 & |x| \geq 1. \end{cases}$$

expbump<sub>p</sub> 24 The periodic bump function

$$\text{expbump}_p(x) = \text{expbump}(1 - \sin(\pi(x + 1/2))).$$

### Basics:

$\mathbb{N}$	-	The set of natural numbers.
$\mathbb{Z}$	-	The set of integers.
$\mathbb{Z}^+$	-	The set of non-negative integers.
$\mathbb{Z}_n$	-	The set of integers modulo $n$ , $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ .
$\mathbb{Q}$	-	The set of rational numbers.
$\mathbb{R}$	-	The set of real numbers.
$\mathbb{R}^+$	-	The set of non-negative real numbers.
$\mathbb{C}$	-	The set of complex numbers.
$\mathbb{R}^M$	-	The set of real valued vectors of length $M$ .
$\mathbb{C}^M$	-	The set of complex valued vectors of length $M$ .
$\text{Re}(c)$	-	The real part of the complex number $c$ .
$\text{Im}(c)$	-	The imaginary part of the complex number $c$ .
$ c $	-	The absolute value of the complex number $c$ .
$i$	-	The imaginary unit.
$\bar{c}$	-	The complex conjugate of the complex number $c$ .
$\lceil x \rceil$	-	The ceiling of $x$ —the smallest integer greater than or equal to $x$ .
$\oplus, \bigoplus$	-	The direct sum.

$\mathcal{O}()$	-	The big O notation.
id	-	The identity operator.

# Chapter 1

## Introduction

Equations of quantum mechanics are notoriously difficult to solve. Under most circumstances it is not possible to solve them with pen and paper, and even computational methods encounter several enduring challenges. Depending on the equation of quantum mechanics under consideration, the development of effective numerical methods might require overcoming considerable technical hurdles presented by high-dimensionality, non-linearity, unboundedness, stochasticity and time-dependence of potentials, for instance.

The focus of this thesis is the one-dimensional, linear, semiclassical Schrödinger equations. Some results of this thesis, in particular the analysis of size of commutators and the analysis of error, can be extended to the case of Schrödinger equations in multiple dimensions and non-semiclassical regime, while others, such as the analysis of cost, will not translate directly.

The aim of this thesis will be to lay down some algorithmic procedures, algebraic observations and analysis approaches that can be used for developing and analysing a series of high order methods, some of which may be found to be inferior to existing methods while others are clearly superior. Thus, the techniques developed here can be understood to be as much, if not more, of a central part of the results as the concrete numerical methods presented. The extension of these techniques to more general cases, however, is beyond the scope of this thesis.

Even when we consider a single dimension and time-independent potentials, efficient computational solution of these equations is far from simple due to the presence of the small semiclassical parameter,  $\varepsilon$ , which plays a role similar to the reduced Planck's constant,  $\hbar \approx 1.054 \times 10^{-34} \text{ J} \cdot \text{s}$ , and which is a source of numerous difficulties. In principle, it is possible to scale  $\varepsilon = \hbar = 1$  by working in atomic units, but this restricts us to length, time and mass scales that are exceedingly small ( $10^{-11}\text{m}$ ,  $10^{-17}\text{s}$  and  $10^{-30}\text{kg}$ , respectively). When simulations for even moderately larger scales are required,  $\varepsilon$  can take very small values, and the range  $10^{-8} \leq \varepsilon \leq 10^{-2}$  is not unrealistic.

Small values of  $\varepsilon$  result in rapid oscillations of frequency  $\mathcal{O}(1/\varepsilon)$  in both space and

time. Resolving these oscillations typically requires fine grids and small time steps of size  $\mathcal{O}(\varepsilon)$ , making standard methods like Strang splitting prohibitively expensive<sup>1</sup> and inaccurate. Higher-order methods such as *Yoshida splittings* allow larger time steps while maintaining reasonable accuracy. However, their costs grow exponentially as  $\mathcal{O}(3^n)$  once we consider higher orders of accuracy. In this thesis we demonstrate an approach for overcoming this exponential blow-up by resorting to some algebraic techniques.

## 1.1 Zassenhaus splittings

Instead of working in the language of matrices, following the discretisation of the Hamiltonian, our narrative develops directly in the Lie algebra generated by the infinite dimensional operators  $V$  and  $\partial_x^2$ , which are the constituents of the Hamiltonian. The resulting methods, called the *symmetric Zassenhaus splittings*, can be designed to arbitrarily high-order accuracies with costs that grow quadratically, not exponentially.

In Chapters 4 and 6, we will develop asymptotic exponential splittings of the form

$$\mathcal{Z}_{n,\sigma} = e^{\frac{1}{2}W^{[0]}} e^{\frac{1}{2}W^{[1]}} \dots e^{\frac{1}{2}W^{[n]}} e^{\mathcal{W}^{[n+1]}} e^{\frac{1}{2}W^{[n]}} \dots e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[0]}},$$

where consecutive exponents are scaled by increasing powers of  $\varepsilon$ ,

$$\begin{aligned} W^{[0]} &= \mathcal{O}(\varepsilon^{\sigma-1}), \\ W^{[k]} &= \mathcal{O}(\varepsilon^{(2k-1)\sigma-1}), \quad k = 1, \dots, n, \\ \mathcal{W}^{[n+1]} &= \mathcal{O}(\varepsilon^{(2n+1)\sigma-1}), \end{aligned}$$

and where the size of the exponents and the accuracy are analysed in the *currency* of the inherent semiclassical parameter  $\varepsilon$  after tying the choice of spatial grid resolution  $\Delta x$  and the time step  $h$  to  $\varepsilon$  using the power laws,

$$\Delta x = \mathcal{O}(\varepsilon), \quad h = \mathcal{O}(\varepsilon^\sigma), \quad \sigma > 0.$$

$\mathcal{Z}_{n,\sigma}$  is called a symmetric Zassenhaus splitting and features an  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  accuracy.

At this stage the splitting is in an operatorial form— $W^{[k]}$ s are still infinite dimensional and unbounded operators. The  $\mathcal{O}(\varepsilon^{(2k-1)\sigma-1})$  size of  $W^{[k]}$ , thus, is merely an indication of how large this term will be upon discretisation (typically via spectral collocation) where

---

<sup>1</sup> (Jin, Markowich & Sparber 2011) note that when we are only concerned with observables, the time step in the Strang splitting has no restrictions on account of  $\varepsilon$ . However, as will become evident in Figure 6.8, higher order methods such as Yoshida and Zassenhaus splitting do outperform the Strang splitting. This is because they are higher order in time. Thus, even when time steps are not constrained by  $\varepsilon$ , higher order methods do remain desirable.

$\partial_x$  is replaced by the skew-symmetric differentiation matrix  $\mathcal{K}$ , whose spectral radius is

$$\rho(\mathcal{K}) = \mathcal{O}((\Delta x)^{-1}) = \mathcal{O}(\varepsilon^{-1}).$$

Although  $\partial_x$  is an unbounded operator, we abuse notation and write

$$\mathcal{K} = \mathcal{O}(\varepsilon^{-1}), \quad \partial_x = \mathcal{O}(\varepsilon^{-1}),$$

to indicate that it is eventually replaced by a matrix of size<sup>2</sup>  $\mathcal{O}(\varepsilon^{-1})$ . In Chapter 9 we will see how this notation can be made rigorous without discretisation considerations.

After discretisation, the outermost exponent  $W^{[0]}$  in the Zassenhaus splitting is exponentiated by using two FFTs since it typically happens to be circulant matrix, while  $W^{[1]}$  is typically a diagonal matrix and exponentiated directly. The remaining exponents do not possess a structure that lends them to direct exponentiation. However, they are very small,  $W^{[k]} = \mathcal{O}(\varepsilon^{(2k-1)\sigma-1})$ , and we find that very few Lanczos iterations are required to exponentiate them to the required accuracy of  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$ .

Much like the comparable order  $2n+2$  Yoshida splitting, which will be shown to feature an  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  accuracy under these scaling choices, the cost of  $\mathcal{Z}_{n,\sigma}$  is dominated by the  $\mathcal{O}(M \log M)$  cost of FFTs where  $M = \mathcal{O}((\Delta x)^{-1}) = \mathcal{O}(\varepsilon^{-1})$  is the number of grid points in spatial discretisation. However, unlike the  $2 \times 3^n$  FFTs required in the Yoshida splitting, the number of FFTs in the Zassenhaus splittings grow quadratically in  $n$ .

While the complexity results are theoretically very satisfactory, the appearance of a larger constant in the growth of costs for Zassenhaus splittings means that Yoshida splittings remain more effective for moderately high orders.

**Note:** It is natural to question why the reference method (as far as cost is concerned) is taken to be the Yoshida splitting, since it is far from the most efficient high-order method. The choice is motivated by the fact that (i) on the one hand, the quadratic growth of cost for Zassenhaus splittings will not be matched by more optimised Yoshida type splittings in terms of complexity and (ii) on the other hand, as we will note in Section 6.7, for moderately high order, even the unoptimised Yoshida splitting is more inexpensive than the Zassenhaus splittings due to the constants in the costs.

Nevertheless, as we see in its application to splitting the Magnus expansions, the Zassenhaus splitting proves to be highly flexible and efficient, significantly improving upon the Yoshida splitting. Here the choice of Yoshida splitting as a reference

---

<sup>2</sup>In Section 9.3.3, we will challenge the scaling of  $\Delta x$  with  $\varepsilon$ , arguing for a slightly finer spatial resolution. However, due to considerations that become clear only in Chapter 9,  $\partial_x$  will still scale as  $\mathcal{O}(\varepsilon^{-1})$  and the analysis of size and error essentially remains unaffected, therefore. The only consequence is that the cost of each FFT works out to be a bit higher. For the sake of simplicity, we will continue to assume  $\Delta x = \mathcal{O}(\varepsilon)$  till Chapter 9 since it makes analysing sizes and errors easier.

makes sense for a different reason—the absence of any optimised high order splitting that can be directly applied to the Magnus expansion.

## 1.2 Algebraic structures

The remarkable features of Zassenhaus methods are due to the structural properties of the Lie algebra  $\mathfrak{G}$  of symmetrised differential operators discussed in Chapter 5. These symmetrised operators are of the form

$$\langle f \rangle_k = \frac{1}{2}(f \circ \partial_x^k + \partial_x^k \circ f),$$

and discretise directly to the form

$$\langle f \rangle_k \rightsquigarrow \frac{1}{2}(\mathcal{D}_f \mathcal{K}^k + \mathcal{K}^k \mathcal{D}_f),$$

where  $\mathcal{D}_f$  is the diagonal matrix with values of  $f$  along the diagonal and  $\mathcal{K}^k$  is the  $k$ th differentiation matrix. Assuming that  $f$  is real valued, it can be seen that this form is skew-symmetric for odd  $k$  and symmetric for even  $k$ .

We will find that once the Schrödinger equation is re-written in the language of this algebra, the nested commutators that arise in the development of symmetric Zassenhaus splittings can be solved in  $\mathfrak{G}$  by using the simplification rule,

$$[\langle f \rangle_k, \langle g \rangle_l] = \sum_{n=0}^{\frac{k+l-1}{2}} \sum_{i=0}^{2n+1} \lambda_{n,i}^{k,l} \langle (\partial_x^i f)(\partial_x^{2n+1-i} g) \rangle_{k+l-2n-1},$$

where the coefficients  $\lambda$  are known.

A crucial property of this algebra is the height reduction that is evident in the simplification rule. Since  $\langle f \rangle_k$  is discretised as  $\frac{1}{2}(\mathcal{D}_f \mathcal{K}^k + \mathcal{K}^k \mathcal{D}_f)$ , and  $\rho(\mathcal{K}) = \mathcal{O}(\varepsilon^{-1})$  we use the shorthand

$$\langle f \rangle_k = \mathcal{O}(\varepsilon^{-k}).$$

In the simplification rule, we find that the commutator of the  $\mathcal{O}(\varepsilon^{-k})$  and  $\mathcal{O}(\varepsilon^{-l})$  terms,  $\langle f \rangle_k$  and  $\langle g \rangle_l$ , is simplified to a linear combination of terms where the largest term is of the form  $\langle \rangle_{k+l-1}$ , which is of size  $\mathcal{O}(\varepsilon^{-k-l+1})$ . This is a tighter estimate than one can obtain by using simple commutator bounds such as

$$\|[X, Y]\| \leq 2 \|X\| \|Y\|.$$

This property of height reduction lends an extra power of  $\varepsilon$  to each commutator, making the asymptotic splittings possible.

Moreover, we find that the Lie algebra  $\mathfrak{G}$  possesses a  $\mathbb{Z}_2$ -graded structure which, when



combined with the symmetrised structure of its elements ensures that the exponents appearing in Zassenhaus-style splittings are always skew-Hermitian. This proves crucial for unitary evolution and the numerical stability of our methods.

### 1.3 Magnus–Zassenhaus schemes

In the case of semiclassical Schrödinger equations featuring time-dependent potentials, direct exponentiation of the Hamiltonian, whether discretised or undiscretised, is not an option. In principle one must resort to the Magnus expansion, whose exponential gives us the formal solution.

The Magnus expansion is an infinite series of nested integrals of nested commutators of the Hamiltonian. Thus, unlike the case of time-independent potentials, where the commutators are encountered only when we develop Zassenhaus splittings and do not appear when we resort to Yoshida splittings, a few nested commutators are inevitable here.

In standard practice, one truncates the Magnus expansion, discretises the integrals via quadrature rules and discretises the Hamiltonian via spectral collocation. This results in Magnus expansions that feature nested commutators of matrices. These expansions are then exponentiated via Lanczos iterations, leading to Magnus–Lanczos schemes, or via Yoshida splittings. We find that both these approaches become very expensive in the case of the semiclassical Schrödinger equation.

Instead, we once again resort to working with the undiscretised Hamiltonian in the language of the Lie algebra  $\mathfrak{G}$ , whereby we benefit from commutator simplification and height reduction. Consequently we arrive at Magnus expansions in Chapter 7 that are free of nested commutators and which feature terms that are asymptotically smaller in powers of  $\varepsilon$ . This allows us to use Magnus expansions with relatively larger time steps.

These Magnus expansions, when combined with traditional approaches for computing the exponential, remain expensive, even though they are less expensive than the Magnus expansions featuring nested matrix commutators. Directly exponentiating these via Lanczos iterations is highly expensive due to the large number of iterations resulting from the large spectral radius of the exponent. Splitting the Magnus expansion with Yoshida splittings, followed by Lanczos iterations, on the other hand, allows us to separate large but structurally favourable exponents from the small but structurally unfavourable exponents and, therefore, seems like a reasonable approach. However, even this requires  $(8n^2 + 8n + 4) \times 3^n$  FFTs for a method with  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  accuracy.

Fortunately, the Zassenhaus splitting algorithm proves to be flexible enough to be combined with the Magnus expansion, resulting in the highly efficient Magnus–Zassenhaus schemes presented in Chapter 8. These benefit from all the favourable properties of the Zassenhaus splittings—they feature the same number of exponents, which are of the same

size and which posses nearly the same structure. Consequently, these are very efficient methods whose implementation follows along the same lines and whose costs are only marginally higher but still quadratic in  $n$  for  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  accuracy. Even for moderately high orders, these turn out to be more inexpensive than the alternatives. Moreover, working in the Lie algebra of symmetrised differential operators once again guarantees unitary evolution and numerical stability of the schemes.

We present two versions of the Magnus–Zassenhaus schemes—the splittings  $\mathcal{Z}_{n,\sigma}^{\Theta[I]}$  that preserve integrals in the Magnus expansion and throughout the subsequent application of the Zassenhaus splitting algorithm, and the splittings  $\mathcal{Z}_{n,\sigma}^{\Theta[M]}$  that commence from the integral-free Magnus expansion of (Munthe–Kaas & Owren 1999), where the integrals have been approximated by Gauss–Legendre quadrature at the outset. The former requires more involved mathematics, particularly as we go towards higher orders, but has the advantage of flexibility—the integrals appearing in this splitting could be computed exactly for some potentials which are known analytically or could be approximated using a variety of quadrature rules.

## 1.4 Outline of the thesis

In Chapter 2, we summarise some well known numerical methods which can be skipped by a reader familiar with the subject, referring to it for the finer details of our methods and derivations as needed. Section 2.1 gives an introduction of the various spatial discretisation options available to us, while the discretisation choices that we will use throughout this work are described in Section 2.1.6. Section 2.2 describes some methods for approximating the matrix exponential, including the standard exponential splitting methods. An introduction to the Magnus expansions, with an emphasis on some properties that prove useful in our applications, are provided in Section 2.3. Of particular note are the practical numerical methods based on these expansions, which are discussed in Section 2.3.5.

In Chapter 3, we introduce the semiclassical Schrödinger equations with time-dependent and time-independent potentials. In Section 3.2, we note that the semiclassical parameter induces spatio-temporal oscillations with wavelength of  $\mathcal{O}(\varepsilon)$ . Some standard high-order methods such as Yoshida splittings and the Magnus–Lanczos schemes are also discussed briefly in this chapter.

We commence upon the development of our symmetric Zassenhaus splittings in Chapter 4 by resorting to a recursive application of the sBCH formula. These methods are discussed in Section 4.1.1 after making the choice of spatial discretisation in accordance with the oscillatory nature of the solution and allowing the time step to scale with powers of the semiclassical parameter  $\varepsilon$  as  $h = \mathcal{O}(\varepsilon^\sigma)$ ,  $\sigma > 0$ . In order to arrive at commutator-free splittings, however, we have to resort to working with undiscretised operators.

In Section 4.2, we see how commutators occurring in these expansions can be systemat-

ically solved, leading to the first observation of height reduction. However, in Section 4.3 we find that these methods suffer from a major flaw—the exponents arrived at by following commutator simplifications lose skew-Hermiticity upon discretisation via standard methods. A remedy is explored in Section 4.3.1.

The full resolution of this problem, however, is only achieved after the theoretical results of Chapter 5 have been established. This chapter marks a break from the narrative of the thesis, taking place in a highly abstract and algebraic setting. We develop the theory of the Lie algebra of the Jordan polynomials that appear as the symmetrised differential operators in the context of the Schrödinger equation. We study the properties of commutator simplification, height reduction, and  $\mathbb{Z}_2$ -grading in this highly abstract setting. These structural properties, when translated to the Schrödinger equation, are responsible for the remarkable properties of the Zassenhaus splittings.

Having established the necessary algebraic tools, the symmetric Zassenhaus splittings are developed in Chapter 6. The implementation of these splittings requires computation of various exponentials, the details of which are discussed in Section 6.4. We analyse the local (per time step) costs of these splittings as we approach higher orders in Section 6.7 and the global costs for various choices of  $\sigma$  in Section 6.8. Since the Magnus–Zassenhaus schemes of Chapter 8 are structurally very similar, most of the details of implementation as well as the cost analysis remain similar.

In Chapter 7, we develop commutator-free Magnus expansions by working in the Lie algebra of symmetrised differential operators. In Section 7.2, we develop integral-free and commutator-free Magnus expansions  $\tilde{\Theta}_p^{\varepsilon[M]}$ . Integral-preserving Magnus expansions  $\tilde{\Theta}_p^{\varepsilon[I]}$  are developed in Section 7.3. In Section 7.4 we find that the cost of these Magnus expansions, when combined with standard methods for exponentiation, remains high.

In order to exponentiate the Magnus expansions  $\tilde{\Theta}_p^{\varepsilon[M]}$  and  $\tilde{\Theta}_p^{\varepsilon[I]}$  efficiently, we split their exponentials using the Zassenhaus splitting algorithm in Chapter 8, arriving at the Magnus–Zassenhaus schemes  $\mathcal{Z}_{n,\sigma}^{\Theta[M]}$  and  $\mathcal{Z}_{n,\sigma}^{\Theta[I]}$  in Section 8.1 and Section 8.2, respectively.

Integral-preserving Magnus expansions suffer from one drawback—they feature terms of size  $\mathcal{O}(\varepsilon^{4\sigma-1})$ ,  $\mathcal{O}(\varepsilon^{6\sigma-1})$ ,  $\mathcal{O}(\varepsilon^{8\sigma-1})$ ,  $\dots$  that don't feature in the integral-free versions  $\mathcal{Z}_{n,\sigma}^{\Theta[M]}$ . Consequently, the exponentiation of these becomes more expensive. In Section 8.4, we develop a method for discarding such terms by exploiting the time symmetry of the truncated Magnus expansions. We analyse the costs of the resulting Magnus–Zassenhaus schemes in Section 8.5 and find that these methods are more cost effective than standard methods for high, as well as low, orders of accuracy.

Working directly with undiscretised operators is crucial for the results presented in this thesis. However, shorthand notations such as  $\partial_x = \mathcal{O}(\varepsilon^{-1})$  inherently seem to assume discretisation. Moreover the accuracy of the Magnus expansion and the various splittings has, in general, only been studied for the case of matrices, when solving ODEs. When solving PDEs like the Schrödinger equation, we need a greater degree of rigour. In Chap-

ter 9 we attempt to present such an analysis for the operatorial splittings and expansions presented in this thesis. In Section 9.1, we proceed by exploiting the property of energy preservation in order to show the emergence of highly oscillatory behaviour in the semi-classical regime, which allows us to make the notation  $\partial_x = \mathcal{O}(\varepsilon^{-1})$  and the analysis of errors more rigorous. Following the approach of (Hochbruck & Lubich 2003), we are then able to derive very favourable error bounds for the Magnus expansion in Section 9.3. This analysis is then used in Section 9.4 for the analysis of symmetric Zassenhaus splittings by expressing the sBCH series in terms of the Magnus expansion.

In Chapter 10, we summarise the results of the thesis. We briefly outline some future work and research directions that seem appealing in light of our results.

# Chapter 2

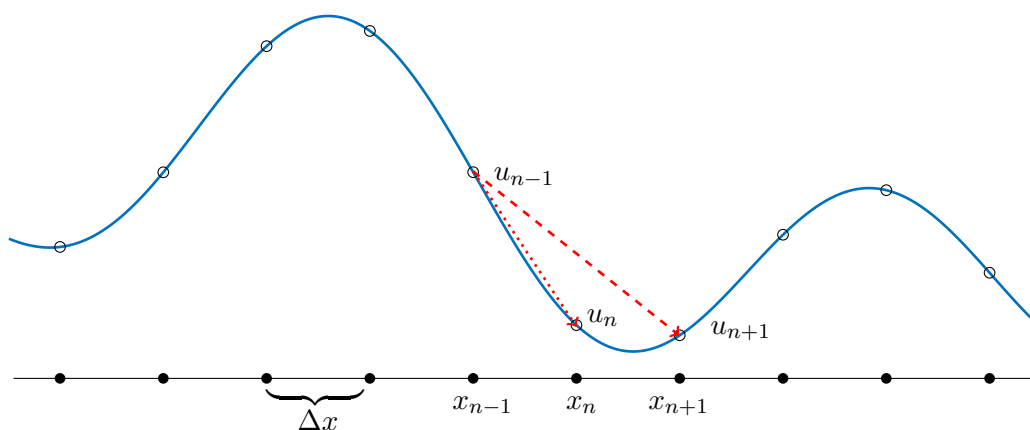
## Numerical tools

### 2.1 Spatial discretisation strategies

Consider the discretisation of a smooth periodic function  $u \in C_p^\infty[-1, 1]$  with period 2 over a uniform grid. Let  $u_n = u(x_n)$  be the value of  $u$  at the  $n$ th grid point,

$$x_n = n\Delta x, \quad n = -N, \dots, 0, \dots, N,$$

where  $\Delta x = 2/(2N + 1)$ . The first derivative of  $u$  at  $x_n$ ,  $u'_n = u'(x_n)$ , can be approximated



**Figure 2.1:** Discretisation on an equispaced grid

as

$$u'_n \approx \frac{1}{\Delta x}(u_n - u_{n-1}), \quad n = -N, \dots, 0, \dots, N, \quad (2.1)$$

indicated by the dotted red arrow in Figure 2.1, or through the (more symmetric) central differences,

$$u'_n \approx \frac{1}{2\Delta x}(u_{n+1} - u_{n-1}), \quad n = -N, \dots, 0, \dots, N, \quad (2.2)$$

indicated by the dashed red arrow in Figure 2.1.

**Note:** In this section, since  $u$  is a function of  $x$  alone, we use  $u'$  to denote the first derivative of  $u$  with respect to  $x$ ,

$$u' = \partial_x u = \frac{du}{dx}.$$

When we study numerical methods for the Schrödinger equation, we encounter dependence on both time  $t$  and position  $x$ . In such cases we will reserve  $u'$  for the derivative with respect to  $t$  and use  $\partial_x u$  for the derivative with respect to  $x$ .

The methods (2.1, 2.2) approximate the derivative using differences of the nodal values at neighbouring nodes and are called *finite difference methods*. The approximation of the second derivative using central finite difference method,

$$u_n'' \approx \frac{1}{(\Delta x)^2} (u_{n+1} - 2u_n + u_{n-1}), \quad n = -N, \dots, 0, \dots, N, \quad (2.3)$$

is also well known.

Depending on the boundary conditions, matters could become more complicated for  $n = -N$  and  $n = N$  where the values of  $u_{n-1}$  or  $u_{n+1}$  may not be defined. However we will always work with periodic boundaries at  $\pm 1$  for the purpose of spatial discretisation. In this case

$$u_{-N-1} = u(-1 + (\Delta x)/2 - \Delta x) = u(1 - (\Delta x)/2) = u_N$$

and

$$u_{N+1} = u(1 - (\Delta x)/2 + \Delta x) = u(-1 + (\Delta x)/2) = u_{-N},$$

since  $u_N = u(2N/(2N+1)) = u(1 - (\Delta x)/2)$  and  $u_{-N} = u(-2N/(2N+1)) = u(-1 + (\Delta x)/2)$ .

Let  $\mathbf{u}$  and  $\mathbf{u}'$  be the vector of values of  $u$  and the exact derivative  $u' = \partial_x u$  on the equispaced grid, respectively,

$$\mathbf{u} = \begin{pmatrix} u_{-N} \\ \vdots \\ u_0 \\ \vdots \\ u_N \end{pmatrix}, \quad \mathbf{u}' = \begin{pmatrix} u'_{-N} \\ \vdots \\ u'_0 \\ \vdots \\ u'_N \end{pmatrix}.$$

(2.1) can be written as

$$\begin{pmatrix} u'_{-N} \\ \vdots \\ u'_0 \\ \vdots \\ u'_N \end{pmatrix} \approx \frac{1}{\Delta x} \begin{pmatrix} 1 & 0 & \dots & -1 \\ -1 & 1 & \dots & 0 \\ 0 & \vdots & \ddots & 0 \\ 0 & \vdots & \ddots & 0 \\ 0 & \dots & -1 & 1 \end{pmatrix} \begin{pmatrix} u_{-N} \\ \vdots \\ u_0 \\ \vdots \\ u_N \end{pmatrix},$$

or

$$\mathbf{u}' \approx \mathcal{K}_1 \mathbf{u},$$

where

$$\mathcal{K}_1 = \frac{1}{\Delta x} \begin{pmatrix} 1 & 0 & \dots & -1 \\ -1 & 1 & \dots & 0 \\ 0 & \vdots & \ddots & 0 \\ 0 & \dots & -1 & 1 \end{pmatrix}$$

is a differentiation matrix that approximates the differential operator  $\partial_x$ ,

$$\mathcal{K}_1 \approx \partial_x = \frac{d}{dx}.$$

We say that  $\mathcal{K}_1$  is a discretisation of  $\partial_x$  and write  $\partial_x \rightsquigarrow \mathcal{K}_1$  to indicate this. Of course, the discretisation is not unique. Let  $\tilde{\mathcal{K}}_1$  be the differentiation matrix corresponding to (2.2), then  $\partial_x \rightsquigarrow \tilde{\mathcal{K}}_1$  is an alternative discretisation of  $\partial_x$ . Similarly the differentiation matrix

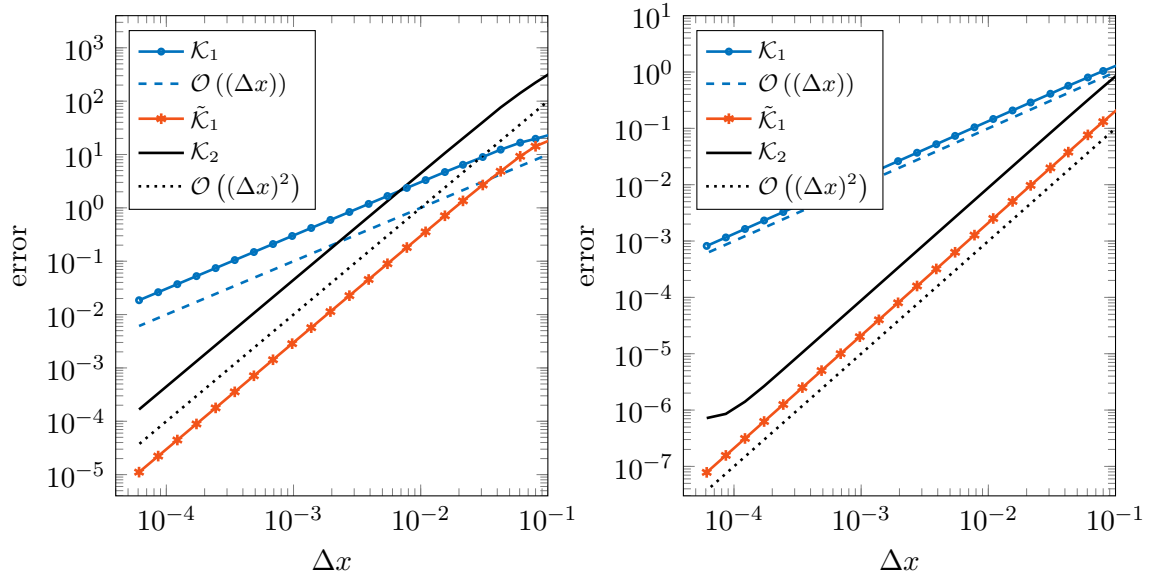
$$\mathcal{K}_2 = \frac{1}{(\Delta x)^2} \begin{pmatrix} -2 & 1 & 0 & \dots & 0 & 1 \\ 1 & -2 & 1 & \dots & 0 & 0 \\ 0 & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \dots & 0 & 1 & -2 \end{pmatrix}$$

approximates  $\partial_x^2$  using the scheme (2.3).

**Note:**

$$\mathcal{K}_2 \neq (\mathcal{K}_1)^2, \quad \mathcal{K}_2 \neq (\tilde{\mathcal{K}}_1)^2.$$

$\mathcal{K}_1$ ,  $\tilde{\mathcal{K}}_1$  and  $\mathcal{K}_2$  are sparse circulant matrices of size  $M \times M$ , where  $M = 2N + 1$ . We can approximate the first or second derivative of  $u$  at the grid points by evaluating the matrix–vector product  $\mathcal{K}_1 \mathbf{u}$  or  $\mathcal{K}_2 \mathbf{u}$ , respectively. These matrices are sparse, and the evaluation of  $\mathcal{K}_1 \mathbf{u}$ , for instance, merely requires the evaluation of (2.1) for  $n = -N, \dots, N$  at a total cost of  $\mathcal{O}(M)$  operations.



**Figure 2.2:**  $\ell^\infty$  error in the approximation of  $\partial_x u$  by  $\mathcal{K}_1 u$  and  $\tilde{\mathcal{K}}_1 u$  (which correspond to the methods (2.1) and (2.2), respectively), and the approximation of  $\partial_x^2 u$  by  $\mathcal{K}_2 u$  (which corresponds to the method (2.3)) where we take  $u(x) = \sin(2\pi(x-1/7))^4 + \sin(5\pi(x-1/3))^2$  (left) and  $u(x) = \exp(\sin(\pi x))$  (right). The method (2.1) is seen to be a first order approximation, while the central difference methods (2.2) and (2.3) are second order.

As evident in Figure 2.2, (2.1) is an order one method,

$$u'_n = \frac{1}{\Delta x}(u_n - u_{n-1}) + \mathcal{O}(\Delta x),$$

while (2.2) and (2.3) are second order methods,

$$\begin{aligned} u'_n &= \frac{1}{2\Delta x}(u_{n+1} - u_{n-1}) + \mathcal{O}((\Delta x)^2) \\ u''_n &= \frac{1}{(\Delta x)^2}(u_{n+1} - 2u_n + u_{n-1}) + \mathcal{O}((\Delta x)^2). \end{aligned}$$

The method (2.1) is characterised by the stencil

$$\frac{1}{\Delta x} \quad \begin{array}{c} \text{---} \bigcirc \text{---} \bigcirc \end{array} \quad .$$

The stencils for (2.2) and (2.3) are

$$\frac{1}{\Delta x} \quad \begin{array}{c} \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \end{array}$$



and

$$\frac{1}{(\Delta x)^2} \quad \begin{array}{c} \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \\ \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \end{array} \quad ,$$

respectively. Since the stencils of (2.2) and (2.3) span three grid points, they are said to have a *bandwidth* of three, while (2.1) has a bandwidth of two. We refer the reader to (Iserles 2008) for detailed analysis of these methods.

### 2.1.1 Norms and inner products

Note that the differentiation matrices  $\tilde{\mathcal{K}}_1$  and  $\mathcal{K}_2$  are skew-symmetric and symmetric, respectively. This preserves an important characteristic of  $\partial_x$  and  $\partial_x^2$ , which are skew-adjoint and self-adjoint, respectively, with respect to the  $L^2$  inner product. The  $L^2$  inner product of complex valued functions on  $[-1, 1]$  is defined as

$$\langle u, v \rangle_{L^2([-1, 1], \mathbb{C})} = \int_{-1}^1 \bar{u}(x) v(x) \, dx.$$

Since we will be working, for the most part, with the domain  $[-1, 1]$  and with complex valued functions, we will write  $\langle u, v \rangle$  to denote  $\langle u, v \rangle_{L^2([-1, 1], \mathbb{C})}$ .

Using integration by parts, it can be seen that

$$\langle \partial_x u, v \rangle = \int_{-1}^1 \bar{u}'(x) v(x) \, dx = \bar{u}(x) v(x) \Big|_{-1}^1 - \int_{-1}^1 \bar{u}(x) v'(x) \, dx = -\langle u, \partial_x v \rangle$$

so that  $\partial_x^* = -\partial_x$  and, consequently,  $(\partial_x^2)^* = (\partial_x^*)^2 = (-\partial_x)^2 = \partial_x^2$ . Here  $\mathcal{A}^*$  stands for the adjoint of the operator  $\mathcal{A}$ , defined as the operator that satisfies

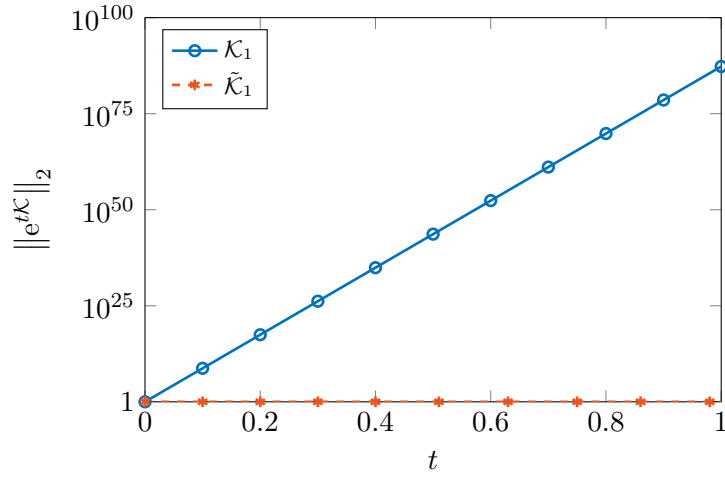
$$\langle \mathcal{A}u, v \rangle = \langle u, \mathcal{A}^*v \rangle$$

for all  $u, v$  in  $L^2([-1, 1], \mathbb{C})$ .

We say that  $\partial_x^2$  is self-adjoint or Hermitian since it is its own adjoint, while an operator like  $\partial_x$  whose adjoint is its additive inverse is called a skew-adjoint or skew-Hermitian operator. The crucial property of skew-Hermiticity of  $\partial_x$  is not shared by the one-sided difference method (2.1) since the matrix  $\mathcal{K}_1$  is not skew-symmetric. The loss of symmetry reflects in the blow up of the exponential of  $\mathcal{K}_1$ , seen in Figure 2.3. Such concerns of symmetry will play a dominant role in the development of our methods.

Corresponding to the  $L^2$  inner product, we also introduce the discrete  $\ell^2$  inner product on the grid,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^* \mathbf{v} \Delta x = \sum_{n=-N}^N \bar{u}_n v_n \Delta x$$



**Figure 2.3:** The exponential of  $K_1$  blows up even for small values of  $t$  while the exponential of  $\tilde{K}_1$  is of unit norm as expected out of a reasonable discretisation of  $\partial_x$  on a periodic grid.

where  $\mathbf{u}^*$  is the conjugate transpose of  $\mathbf{u}$ ,

$$\mathbf{u}^* = \overline{\mathbf{u}}^T.$$

The  $L^2$  norm and the  $\ell^2$  norm are obtained from this inner product since every inner product  $\langle \cdot, \cdot \rangle$  defines a norm  $\|u\| = \sqrt{\langle u, u \rangle}$ . In general, the  $L^p$  norm is defined as

$$\|u\|_p = \|u\|_{L^p} = \left( \int_{-1}^1 |u(x)|^p dx \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty,$$

and the  $\ell^p$  norm as

$$\|\mathbf{u}\|_p = \|\mathbf{u}\|_{\ell^p} = \left( \sum_n |u_n|^p \Delta x \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty.$$

For  $p = \infty$ , we need separate definitions,

$$\|u\|_\infty = \|u\|_{L^\infty} = \operatorname{ess\,sup}_{x \in [-1,1]} |u(x)|$$

and

$$\|\mathbf{u}\|_\infty = \|\mathbf{u}\|_{\ell^\infty} = \max_{n=-N, \dots, N} |u_n|.$$

We note that

$$\|\mathbf{u}\|_\infty \leq \sqrt{\frac{M}{2}} \|\mathbf{u}\|_2 \leq \sqrt{M} \|\mathbf{u}\|_\infty. \quad (2.4)$$

A norm on vectors induces a norm on matrices,

$$\|\mathcal{A}\|_p = \max_{\mathbf{u} \in \mathbb{C}^M} \frac{\|\mathcal{A}\mathbf{u}\|_p}{\|\mathbf{u}\|_p} = \max_{\substack{\|\mathbf{u}\|_p=1 \\ \mathbf{u} \in \mathbb{C}^M}} \|\mathcal{A}\mathbf{u}\|_p,$$

and a norm on functions induces a norm on linear operators,

$$\|\mathcal{A}\|_p = \sup_{u \in L^p(X, \mathbb{C})} \frac{\|\mathcal{A}u\|_p}{\|u\|_p} = \sup_{\substack{\|u\|_p=1 \\ u \in L^p(X, \mathbb{C})}} \|\mathcal{A}u\|_p.$$

The spectral radius  $\rho(\mathcal{A})$  of the matrix  $\mathcal{A}$  is its largest eigenvalue (by magnitude),

$$\rho(\mathcal{A}) = \max\{|\lambda_1|, \dots, |\lambda_M|\},$$

where  $\lambda_i$  are the eigenvalues of the  $M \times M$  matrix  $\mathcal{A}$ . In the case of a normal matrix  $\mathcal{A}$ , i.e.  $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$ ,  $\rho(\mathcal{A}) = \|\mathcal{A}\|_2$ . Unitary matrices, skew-Hermitian matrices and Hermitian matrices are normal and, therefore, the spectral radius and the  $\ell^2$  norms coincide. We often judge the ‘size’ of  $\mathcal{A}$  by its spectral radius (this is not to be confused with the dimensions of  $\mathcal{A}$ , which are  $M \times M$ ).

### 2.1.2 Higher order finite difference methods

To derive arbitrarily high-order differentiation matrices for  $\partial_x$ ,  $\partial_x^2$ , or higher degree differential operators, we need to ensure that the method is exact for polynomial interpolants of a sufficiently high degree. Consider the interpolation of  $u$  by Lagrangian cardinal functions at  $m$  (not necessarily equidistant) grid points  $\{x_i\}_{i=0}^m$  given by the polynomial

$$p_{m,u}(x) = \sum_{i=0}^m u(x_i) \ell_i(x),$$

with  $p_{m,u} \in \mathcal{P}_m(x)$ —the space of degree  $m$  polynomials in  $x$  with Lagrangian cardinal functions  $\{\ell_i\}_{i=0}^m$  as basis. The Lagrangian cardinal functions,

$$\ell_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^m \frac{x - x_j}{x_i - x_j}$$

satisfy  $\ell_i(x_j) = \delta_{ij}$  and give rise to orthonormal vectors—in fact, the natural basis  $\{\mathbf{e}_i\}_{i=0}^m$ —upon discretisation to the grid  $\{x_j\}_{j=1}^m$ . It should be noted that, as a consequence, the coordinate vector of  $p_{m,u}$  in the basis  $\{\ell_i\}_{i=0}^m$  is exactly the vector of values at the grid points,  $\mathbf{u} = (u(x_0), \dots, u(x_m))^T$ .

The first derivative of  $u$  at  $x_j$  can be approximated by differentiating the interpolant,

$$p'_{m,u}(x) \Big|_{x=x_j} = \sum_{i=0}^m u(x_i) \ell'_i(x) \Big|_{x=x_j} = \sum_{i=0}^m c_{j,i} u(x_i), \quad c_{j,i} = \ell'_i(x_j). \quad (2.5)$$

The values  $c_{j,i}$  give us a stencil for approximating the first derivative at the point  $x_j$  by a linear combination of finite number of neighbouring values, leading to finite difference methods. In this manner one can obtain stencils for higher derivatives, effective algorithms for which are discussed by Fornberg (1998). In this way we can discretise any linear differential operator  $\mathcal{L}$ . In particular for  $\mathcal{L} = \partial_x^2$ ,

$$p''_{m,u}(x) \Big|_{x=x_j} = \sum_{i=0}^m u(x_i) \ell''_i(x) \Big|_{x=x_j} = \sum_{i=0}^m \tilde{c}_{j,i} u(x_i), \quad \tilde{c}_{j,i} = \ell''_i(x_j), \quad (2.6)$$

and we merely need to evaluate  $\ell''_i(x_j)$ .

In principle one interpolates locally at each point  $x_j$  by a reasonably high degree polynomial, taking equal number of interpolation points to the left and the right, where possible. On an equispaced and periodic grid this gives us the same stencil for each point and leads to the central difference methods such as (2.2) and (2.3), which feature higher orders of accuracy and preserve symmetry (Fornberg 1998, Iserles 2008).

The accuracy of approximation of  $u'$  by  $p'_{m,u}$  is a consequence of the approximation properties of the underlying approximation space  $\mathcal{P}_m = \text{span}\{\ell_0, \dots, \ell_m\}$  and the linear interpolation method. Larger values of  $m$  lead to higher order finite difference methods and result in larger bandwidth (longer stencil) methods. For instance, the order eight central difference stencil for the second derivative has a bandwidth of nine,

$$\frac{1}{(\Delta x)^2} \left( -\frac{1}{560} \quad -\frac{8}{315} \quad -\frac{1}{5} \quad \frac{8}{5} \quad \left( -\frac{205}{72} \right) \quad \frac{8}{5} \quad -\frac{1}{5} \quad \frac{8}{315} \quad -\frac{1}{560} \right),$$

which represents the scheme

$$u''_n = \frac{1}{(\Delta x)^2} \left( -\frac{1}{560} u_{n-4} + \frac{8}{315} u_{n-3} - \frac{1}{5} u_{n-2} + \frac{8}{5} u_{n-1} - \frac{205}{72} u_n + \frac{8}{5} u_{n+1} - \frac{1}{5} u_{n+2} + \frac{8}{315} u_{n+3} - \frac{1}{560} u_{n+4} \right) + \mathcal{O}((\Delta x)^8). \quad (2.7)$$

Taylor analysis reveals that the error in this approximation for  $u \in C_p^\infty[-1, 1]$  is

$$\|\mathbf{u}'' - \tilde{\mathcal{K}}_2 \mathbf{u}\| \leq C(\Delta x)^8 \|\mathbf{u}^{(10)}\| + \mathcal{O}((\Delta x)^{10}), \quad (2.8)$$

where  $C$  is some constant that does not depend on  $u$ ,  $\mathbf{u}''$  is the vector of values of  $\partial_x^2 u$  on the grid, and  $u^{(10)}$  is the 10th derivative of  $u$ .

As a consequence of the bandwidth being nine, we end up with an  $M \times M$  matrix which is nine-diagonal except for the wrapping around that comes with being a circulant. The matrix–vector product  $\tilde{\mathcal{K}}_2 \mathbf{u}$  is still evaluated in  $\mathcal{O}(M)$  operations, although there is a linear growth of cost with the bandwidth.

The eigenvalues of an  $M \times M$  circulant, the first row of which is  $(c_0, c_{M-1}, c_{M-2}, \dots, c_1)$ , are given by

$$\lambda_j = c_0 + c_{M-1}\omega_j + c_{M-2}\omega_j^2 + \dots + c_1\omega_j^{M-1}, \quad \omega_j = \exp\left(i\frac{2\pi j}{M}\right),$$

where  $\omega_j$  is the  $j$ th root of unity. Noting that  $|\lambda_j| \leq \sum_{n=1}^M |c_n|$  for any  $j$ , the spectral radius of  $\tilde{\mathcal{K}}_2$  is bounded above by

$$\rho(\tilde{\mathcal{K}}_2) = \max_j |\lambda_j| \leq \sum_{n=1}^M |c_n| = \frac{2048}{315}(\Delta x)^{-2}.$$

Since the central difference matrix is Hermitian,  $\|\tilde{\mathcal{K}}_2\|_2 = \rho(\tilde{\mathcal{K}}_2)$ , and for short we say

$$\tilde{\mathcal{K}}_2 = \mathcal{O}(M^2),$$

keeping in mind  $(\Delta x)^{-2} = \mathcal{O}(M^2)$ . In general, the order  $k$  differentiation matrix  $\mathcal{K}_k$ , which approximates  $\partial_x^k$ , scales as  $\mathcal{O}(M^k)$ . For short, we abuse notation and write

$$\partial_x = \mathcal{O}(M), \quad \partial_x^k = \mathcal{O}(M^k), \quad (2.9)$$

when it is evident that  $\partial_x$  and  $\partial_x^k$  would be eventually discretised and replaced by  $\mathcal{K}$  and  $\mathcal{K}_k$ , even though  $\partial_x$  and  $\partial_x^k$  are, in truth, unbounded operators.

### Additional notation

We write  $\mathcal{D}_{a_n}$  to denote the diagonal  $M \times M$  matrix whose diagonal entries form the sequence  $\{a_n\}_{n=-N}^N$ . Since this case arises frequently, we write  $\mathcal{D}_f$  to denote the diagonal matrix corresponding to pointwise multiplication by  $f$  on the grid  $\{x_n\}_{n=-N}^N$ ,  $x_n = n\Delta x$ ,  $\Delta x = \frac{2}{2N+1}$ , where the underlying sequence is  $\{f(x_n)\}_{n=-N}^N$ .

Spectral radius of this matrix is bounded trivially as  $\rho(\mathcal{D}_f) = \max\{f(x_n)\}_{n=-N}^N \leq \|f\|_\infty$ . Evaluation of the matrix–vector product  $\mathcal{D}_f \mathbf{u}$  is carried out through trivial pointwise multiplication in  $\mathcal{O}(M)$  cost.

After spatial discretisation, the advection equation  $\partial_t u(x, t) = \partial_x u(x, t)$  results in  $\partial_t \mathbf{u}(t) = \mathcal{K}_1 \mathbf{u}$  while  $\partial_t u(x, t) = \partial_x u(x, t) + f(x)u(x, t)$  discretises as  $\partial_t \mathbf{u}(t) = \mathcal{K}_1 \mathbf{u} + \mathcal{D}_f \mathbf{u}$ .

### 2.1.3 Spectral collocation and pseudospectral methods

Collocation is the idea of approximating the solution of a partial differential equation such as the advection equation  $\partial_t u = \partial_x u$  by a linear combination of basis functions  $\{\phi_i\}_{i=1}^m$ ,

$$u(x, t) \approx \sum_{i=1}^m a_i(t) \phi_i(x),$$

substituting the approximation into the differential equation,

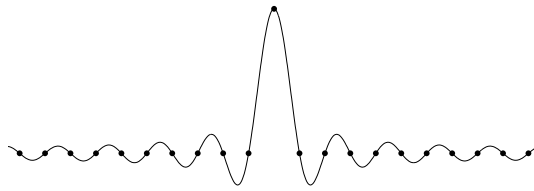
$$\sum_{i=1}^m a'_i(t) \phi_i(x) = \sum_{i=1}^m a_i(t) \phi'_i(x), \quad (2.10)$$

and solving for the coefficients  $\mathbf{a}(t)$  while ensuring that (2.10) is satisfied precisely at a set of ‘collocation’ points  $\{x_j\}_{j=1}^{\tilde{m}}$  in the domain (Kreiss & Oliger 1972, Orszag 1972). Collocation with a spectral set of basis such as the Fourier basis is called spectral collocation.

We will restrict our discussions to spectral collocation with Fourier basis on equispaced and periodic grids. On such grids, this method effectively amounts to interpolation by Fourier cardinal functions (Hesthaven, Gottlieb & Gottlieb 2007), and the development of these methods following (Fornberg 1998), therefore, turns out to be along very similar lines as the one just performed in Section 2.1.2 for finite difference methods. Here the approximation space is  $\mathcal{T}_m$ —the space of trigonometric polynomials of degree  $\leq m$ .

Fourier cardinal functions  $\{\phi_n\}_{n=-N}^N$  on the grid  $x_n = n\Delta x, n = -N, \dots, N$ , ( $\Delta x = 2/M$  where  $M = 2N + 1$ ) are given by  $\phi_n(x) = \hat{\phi}(x - x_n)$ , where

$$\hat{\phi}(x) = \frac{2}{M} \left\{ \frac{1}{2} + \cos \pi x + \cos 2\pi x + \dots + \cos N\pi x \right\}. \quad (2.11)$$



**Figure 2.4:** The Fourier cardinal function  $\hat{\phi}(x)$  on a grid with  $N = 10, M = 21$  (values at grid points are shown as dots).

Since  $\hat{\phi}(x_j) = \delta_{0j}$  we have  $\phi_n(x_j) = \delta_{nj}$  and these are seen directly to be cardinal functions. Our trigonometric interpolant to  $u$  is

$$s_{M,u}(x) = \sum_{n=-N}^N u(x_n) \phi_n(x). \quad (2.12)$$

Following the same procedure as Section 2.1.2 we arrive at a stencil for the second derivative, given by  $c_{j,n} = \phi_n''(x_j) = \hat{\phi}''(x_j - x_n) = \hat{\phi}''(x_{j-n}) =: c_{j-n}$ ,

$$s_{M,u}''(x) \Big|_{x=x_j} = \sum_{n=-N}^N u(x_n) \phi_n''(x) \Big|_{x=x_j} = \sum_{n=-N}^N c_{j,n} u(x_n), \quad c_{j,n} = \phi_n''(x_j), \quad (2.13)$$

The Fourier cardinal function is, equivalently, given by the Dirichlet kernel,

$$\hat{\phi}(x) = \frac{\sin(M\pi x/2)}{M \sin(\pi x/2)}, \quad (2.14)$$

which, upon twice differentiation and evaluation at  $x_j$ , gives us

$$c_j = \hat{\phi}''(x_j) = \frac{(-1)^{j+1} \pi^2}{2} \left( \frac{\cos(j\pi/M)}{\sin^2(j\pi/M)} \right), \quad j \neq 0. \quad (2.15)$$

For  $j = 0$ , we differentiate (2.11) to get

$$c_0 = -\frac{2}{M} \sum_{n=1}^N n^2 \pi^2 = -\pi^2 \frac{N(N+1)}{3}. \quad (2.16)$$

As the coefficients must also satisfies  $\sum_{n=-N}^N c_n = 0$ , we can instead derive  $c_0$  from this relation. The spectral collocation differentiation matrix  $(\mathcal{K}_{2,\text{SC}})_{i,j} = c_{i-j}$  is a circulant matrix just like  $\mathcal{K}_{2,\text{FD}}$ , the finite difference differentiation matrices discretising  $\partial_x^2$ . This time, however, the symbol  $\mathbf{c} = (c_{-N}, \dots, c_N)^T$  is not of a limited bandwidth and consequently  $\mathcal{K}_{2,\text{SC}}$  is a full matrix. The advantage of this approach, however, is that we inherit the spectral accuracy of the Fourier approximation space.

Since we use interpolation with cardinal functions here, the discretisation of  $u$  is again the vector of values at the grid points,  $\mathbf{u} = (u(x_{-N}), \dots, u(x_N))^T$ .  $\mathcal{K}_{2,\text{SC}}$ , being a circulant, is diagonalisable using Fourier transforms. Thus  $\mathcal{K}_{2,\text{SC}} \mathbf{u}$  is evaluated using a couple of Fast Fourier Transforms (FFTs) at an  $\mathcal{O}(M \log M)$  cost,

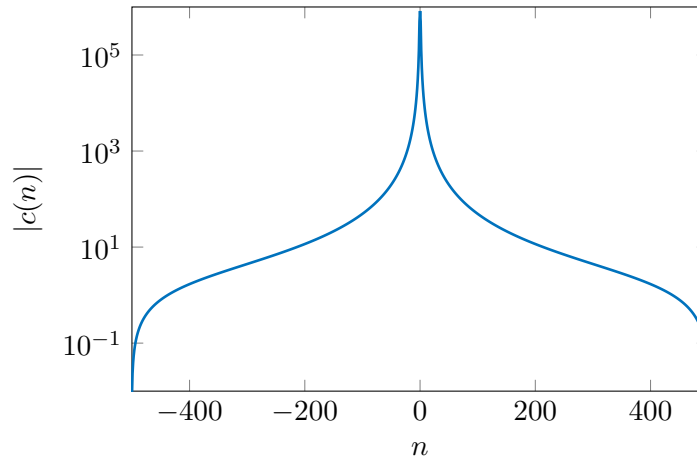
$$\mathcal{K}_{2,\text{SC}} \mathbf{u} = \mathcal{F}^{-1} \mathcal{D}_{\hat{\mathbf{c}}} \mathcal{F} \mathbf{u},$$

where  $\mathcal{F}$  is the Fourier transform matrix,  $\mathcal{F}^{-1}$  is its inverse and

$$\hat{\mathbf{c}} = \mathcal{F} \mathbf{c}$$

is the Fourier transform of the stencil vector  $\mathbf{c} = (c_{-N}, \dots, c_N)^T$ . Since the discretisation is on a grid, multiplication by a function  $f$  is once again achieved via the diagonal matrix  $\mathcal{D}_f$  and the cost of evaluating  $\mathcal{D}_f \mathbf{u}$  remains  $\mathcal{O}(M)$ .

We remind the reader that, in general, the differentiation matrix  $\mathcal{K}_2$  need not equal  $(\mathcal{K}_1)^2$ . However, in the context of equispaced and periodic grids with odd number of grid



**Figure 2.5:** The stencil  $\mathbf{c} = (c_{-N}, \dots, c_N)^T$  (plotted for  $N = 500$  here) does not decay rapidly enough away from the centre to allow for an early truncation, leading to a dense circulant matrix.

points, it holds true for spectral collocation methods. In fact, in general,  $\mathcal{K}_{k,\text{SC}} = (\mathcal{K}_{1,\text{SC}})^k$  and we may unambiguously write  $\mathcal{K}_{\text{SC}}^k$  for the degree  $k$  spectral collocation differentiation matrix.

An interesting observation by Fornberg (1998) is that one could take a very high-order finite difference scheme with a bandwidth larger than the number of grid points and, with the underlying function being periodic, we could extend the data as far off as we need to go along with the stencil coefficients. Equivalently we could wrap the stencil around and add up the overlapping coefficients. Nothing, in principle, stops us from letting the size of such a stencil go to infinity. Such a method is called a pseudospectral method. Pseudospectral methods, in this way, can be seen as infinite order finite difference schemes.

Fornberg (1998) notes that, on an equispaced and periodic grid, wrapping the stencil of an infinite order finite difference scheme in this way leads to the same the differentiation matrix as obtained in the spectral collocation method. It can be shown that, in the context of such grids, the pseudospectral and spectral collocation methods are equivalent. Since equispaced and periodic grids are what concern us for the purpose of this thesis, we will use the names pseudospectral and spectral collocation interchangeably.

#### 2.1.4 Spectral methods

Spectral methods (Orszag 1969, Orszag 1972, Hesthaven et al. 2007, Trefethen 2000) again involve approximation in the space of trigonometric polynomials,  $\mathcal{T}_M$ . Consider the



approximation

$$u(x, t) \approx \sum_{n=-N}^N \hat{u}_n(t) e^{i\pi n x} \in \mathcal{T}_M[-1, 1],$$

where the scalar coefficients  $\hat{\mathbf{u}} = (\hat{u}_{-N}, \dots, \hat{u}_N)^T = \mathcal{F}\mathbf{u}$  are the Fourier coefficients of  $\mathbf{u}$ . Substituting the approximation in a differential equation, one arrives at a differential equation in the Fourier coefficients—in other words, the frequency domain. Considering the advection equation, for instance, we require

$$\sum_{n=-N}^N \hat{u}_n(t)' e^{i\pi n x} = \sum_{n=-N}^N \hat{u}_n(t) i\pi n e^{i\pi n x}$$

to hold for each  $x \in [-1, 1]$ , whereby we can deduce the simpler differential equation

$$\hat{u}_n(t)' = i\pi n \hat{u}_n(t), \quad n \in \{-N, \dots, N\}.$$

The vector representing the approximation of the solution at time  $t$  here is the vector of Fourier coefficients,  $\hat{\mathbf{u}}(t) = (\hat{u}_{-N}(t), \dots, \hat{u}_N(t))^T$ . If the initial value for solving a differential equation is given by  $u(0)$ , one takes the FFT of the vector of values at uniform grid points  $\mathbf{u}(0)$  to arrive at the vector of Fourier coefficients  $\hat{\mathbf{u}}(0)$ , which acts as the initial value in the frequency domain.

It is evident at once that the first derivative matrix is a diagonal matrix given by the entries  $i\pi n$  and the second derivative  $\mathcal{K}_{2,S}$  (where  $S$  stands for spectral), by similar reasoning, is found to be given by diagonal entries  $(\mathcal{K}_{2,S})_{n,n} = -\pi^2 n^2$ . In our notation,

$$\mathcal{K}_{2,S} = \mathcal{D}_{-\pi^2 n^2},$$

and, in general,

$$\mathcal{K}_{k,S} = (\mathcal{K}_{1,S})^k = \mathcal{D}_{(i\pi n)^k}.$$

We remind the reader that this is a consequence of taking odd number of grid points,  $M = 2N + 1$ , which allows us to take equal number of frequencies on either side (note that  $n$  ranges from  $-N$  to  $N$ ). Since  $\mathcal{K}_{2,S}$  is a diagonal, the cost of evaluating the matrix–vector product  $\mathcal{K}_{2,S} \hat{\mathbf{u}}$  is  $\mathcal{O}(M)$ .

Pointwise multiplication of a function  $u$  by the function  $f$  in spatial domain is equivalent to convolution of Fourier coefficients of the two in the frequency domain. The matrix  $\mathcal{M}_f$  representing multiplication by  $f$ , therefore, is a circulant matrix with the Fourier coefficients of  $f$  as its symbol. Since we can easily go back and forth between spatial and frequency representations using Fourier transforms, in practice we often choose to realise

the application of  $\mathcal{M}_{f,S}$  via  $\mathcal{F}\mathcal{D}_f\mathcal{F}^{-1}$ ,

$$\mathcal{M}_f \hat{\mathbf{u}} = \mathcal{F}\mathcal{D}_f\mathcal{F}^{-1}\hat{\mathbf{u}} = \mathcal{F}\mathcal{D}_f\mathbf{u},$$

where  $\mathcal{F}$  is the Fast Fourier Transform matrix and  $\mathcal{F}^{-1}$  is its inverse—i.e. we go back into the spatial domain and perform a pointwise multiplication with  $f$ . The cost of this operation is two FFTs, each of which costs  $\mathcal{O}(M \log M)$ .

The spectral radius of  $\mathcal{M}_f$  is  $\rho(\mathcal{F}\mathcal{D}_f\mathcal{F}^{-1}) = \rho(\mathcal{D}_f) \leq \|f\|_\infty$  as  $\mathcal{F}$  is unitary and does not effect the norm. The spectral radius for  $\mathcal{K}_{2,S}$  is trivially seen to be  $N^2\pi^2$ .

### 2.1.5 Comparisons

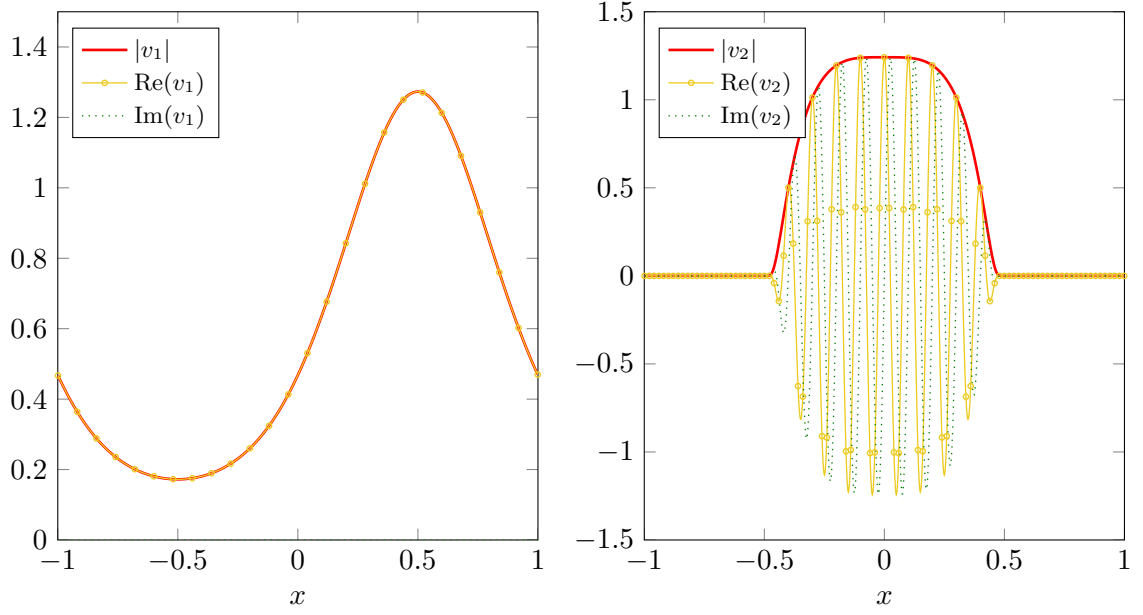
We note that the matrices discretising pointwise multiplication by  $f$  and the differential operator  $\mathcal{K}_2$  in the three cases (order eight finite differences (FD8), spectral collocation (SC) and spectral (S)) are either diagonal or circulant.

	FD8	SC	S
multiplication by $f$ structure cost of $\mathcal{D}_f\mathbf{u}$	$\mathcal{D}_f$ diagonal $\mathcal{O}(M)$	$\mathcal{D}_f$ diagonal $\mathcal{O}(M)$	$\mathcal{F}\mathcal{D}_f\mathcal{F}^{-1}$ circulant $\mathcal{O}(M \log M)$
differentiation matrix $\mathcal{K}_2$ structure cost of $\mathcal{K}_2\mathbf{u}$ accuracy of $\mathcal{K}_2\mathbf{u}$	$\mathcal{K}_{2,\text{FD8}} = \mathcal{F}^{-1}\mathcal{D}_{\hat{\mathbf{c}}_{\text{FD8}}}\mathcal{F}$ 9 diagonal circulant $\mathcal{O}(M)$ $\mathcal{O}((\Delta x)^8)$	$\mathcal{F}^{-1}\mathcal{D}_{\hat{\mathbf{c}}_{\text{SC}}}\mathcal{F}$ circulant $\mathcal{O}(M \log M)$ spectral	$\mathcal{D}_{-n^2\pi^2}$ diagonal $\mathcal{O}(M)$ spectral

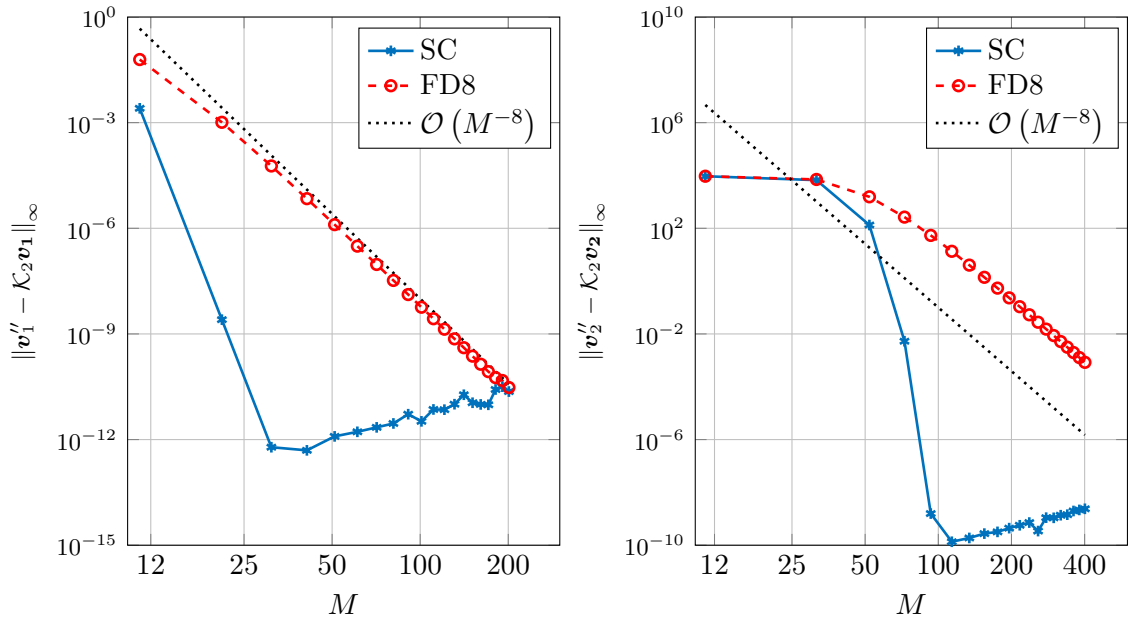
**Table 2.1:** Comparison of order eight finite differences (FD8), spectral collocation (SC) and spectral (S) methods.

This is not where the similarity ends, though. Under the grid choices we work with, it can be seen shown that  $\mathcal{F}\mathcal{K}_{k,\text{SC}}\mathcal{F}^{-1} = \mathcal{K}_{k,S}$ . We already know that  $\mathcal{F}\mathcal{D}_f\mathcal{F}^{-1} = \mathcal{M}_{f,S}$ , so that our spectral collocation methods are precisely the fourier transformations of the spectral method.

The main appeal of spectral (and spectral collocation) methods is that they exhibit *spectral convergence*: for sufficiently large  $M$  the error decays faster than  $\mathcal{O}(M^{-\alpha})$  for any  $\alpha > 0$ . The spectral accuracy of spectral collocation methods is evident in Figure 2.7, where they are seen to converge much more rapidly than the order seven central difference finite difference methods.



**Figure 2.6:** The functions  $v_1$  (left) and  $v_2$  (right) used for testing the accuracy of discretisation schemes.

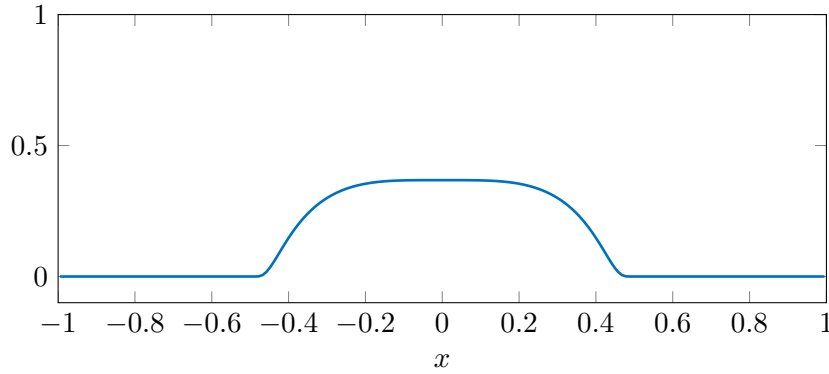


**Figure 2.7:** Error in approximating the second derivative using Spectral Collocation (SC) and Finite Difference (FD8) scheme of order eight for  $v_1$  (left) and  $v_2$  (right).

In Figure 2.6, we show the error in approximating the second derivative of the functions,

$$\begin{aligned} v_1(x) &= \exp(\sin(\pi x)), \\ v_2(x) &= \text{expbump}_p(x) \exp(i20\pi x), \end{aligned}$$

where  $\text{expbump}_p(x)$  is a periodic bump function that will occur frequently in our examples,



**Figure 2.8:** The  $\text{expbump}_p(x)$  function is a periodic bump function which is ‘active’ between  $-0.5$  and  $0.5$  and attains a maximum of  $\sim 0.37$ .

$$\text{expbump}_p(x) = \text{expbump}(1 - \sin(\pi(x + 1/2))), \quad (2.17)$$

$$\text{expbump}(x) = \begin{cases} \exp\left(-\frac{1}{1-x^2}\right) & |x| < 1, \\ 0 & |x| \geq 1. \end{cases} \quad (2.18)$$

### 2.1.6 Our discretisation choices

Spectral collocation (and spectral) methods are able to resolve high frequency oscillations much faster than finite difference methods. In Section 3.3.1 we find that finite difference methods are not particularly effective under the semiclassical regime. This has also been noted by (Bao, Jin & Markowich 2002, Jin et al. 2011). For these reasons, spectral collocation will form our weapon of choice for spatial discretisation.

In practice, we will typically work with the vector of values  $\mathbf{u} = (u(x_{-N}), \dots, u(x_N))^T$  (instead of the Fourier coefficients  $\hat{\mathbf{u}}$ ) on the equispaced periodic grid on  $[-1, 1]$ ,

$$x_n = n\Delta x, \quad n = -N, \dots, 0, \dots, N, \quad \Delta x = 2/M, \quad M = 2N + 1,$$

and will use Fourier transforms, where needed, to move between spatial and frequency

representations in order to implement differentiation operators via pointwise multiplication in the frequency domain. Thus the differential operator  $\partial_x$  will be discretised as  $\mathcal{K}$ ,

$$\mathcal{K} = \mathcal{F}^{-1} \mathcal{D}_{i\pi n} \mathcal{F},$$

to be written as

$$\partial_x \sim \mathcal{K}.$$

As noted before, knowing that

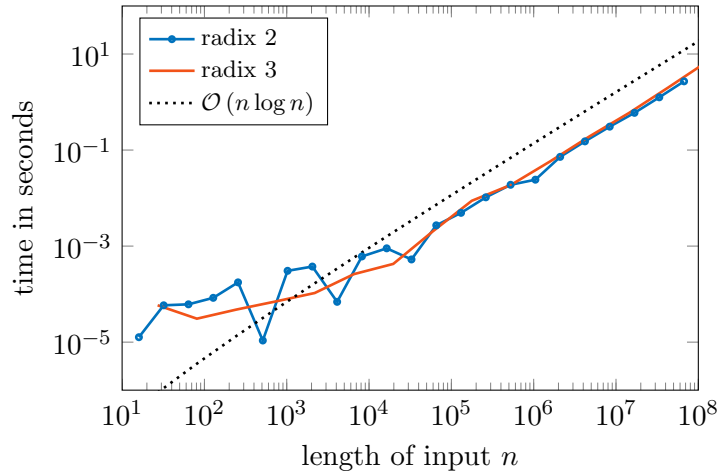
$$\mathcal{K} = \mathcal{O}(M),$$

we abuse notation and write

$$\partial_x^k = \mathcal{O}(M^k),$$

even though  $\partial_x^k$  is an unbounded operator. This is justified as long as eventual discretisation on a grid of resolution  $M$  follows.

Although an odd number of grid points are not the most efficient choice for FFTs<sup>1</sup>, which work most efficiently when the input length is in powers of 2, such grids give us the advantage of simpler differentiation matrices.



**Figure 2.9:** The difference in time taken for MATLAB’s `fft` when the input length is in powers of 2 (radix 2) and when it is in powers of 3 (radix 3) is not significant.

On such grids higher order differentiation matrices  $\partial_x^k \sim \mathcal{K}_k$  can be obtained simply via

$$\mathcal{K}_k = \mathcal{K}^k = \mathcal{F}^{-1} \mathcal{D}_{(i\pi n)^k} \mathcal{F}.$$

We note that this property helps in making notation simpler but is in no way crucial to

<sup>1</sup>Radix 2 implementations of FFTs were the most common in the early years of linear algebra and signal processing packages. However, working in a radix 3 implementation, for instance, does not make matters much worse and most modern implementations of the FFT are able to deal with such inputs reasonably efficiently (see Figure 2.9).

our analysis or efficacy of our methods.

Since we work with nodal values, multiplication by a function  $f$  is discretised as the diagonal matrix  $\mathcal{D}_f$ . We will term this method the spectral collocation method to indicate that we are working with nodal values. As we have noted previously, however, this is equivalent to the approach where we work with the vector of Fourier coefficients  $\hat{\mathbf{u}}$ ,  $\mathcal{K}$  is a diagonal and multiplication with  $f$  is diagonalised via FFTs.

The Zassenhaus splittings developed in this thesis are not necessarily tied to the choice of spectral collocation, however, and any spatial discretisation method that is able to correctly resolve oscillations and respects skew-symmetry of  $\partial_x$  can, in principle, be combined with these splittings.

## 2.2 Approximation of matrix exponentials

Once discretised in space, a partial differential equation such as  $\partial_t u = \partial_x u$  or  $\partial_t u = \partial_x^2 u$  is replaced by a system of ordinary differential equations (ODEs),

$$\partial_t \mathbf{u}(t) = \mathcal{A} \mathbf{u}(t), \quad t \geq 0, \quad \mathbf{u}(0) = \mathbf{u}^0 \in \mathbb{R}^M, \quad (2.19)$$

where  $\mathcal{A}$  is an  $M \times M$  matrix. Solving this equation typically involves discretised time-stepping. Two simple time-stepping methods for solving (2.19) are Forward Euler (FE),

$$\mathbf{u}^{n+1} = \mathbf{u}^n + h \mathcal{A} \mathbf{u}^n, \quad (2.20)$$

and Backward Euler (BE),

$$\mathbf{u}^{n+1} - h \mathcal{A} \mathbf{u}^{n+1} = \mathbf{u}^n. \quad (2.21)$$

Here  $\mathbf{u}^n$  is the numerical approximation of the exact solution of (2.19),  $\mathbf{u}(t_n)$ , at the  $n$ th time step and  $h$  is the time step such that  $t_{n+1} = t_n + h$ .

In principle, (2.19) can be solved exactly by exponentiating the matrix  $\mathcal{A}$ ,

$$\mathbf{u}(t) = e^{t\mathcal{A}} \mathbf{u}(0),$$

which can be a very effective approach when methods for efficient evaluation of the matrix exponential are available (Golub & Van Loan 1996, Higham & Al-Mohy 2010, Moler & Loan 1978). Here the matrix exponential is defined as

$$e^{\mathcal{A}} = \exp(\mathcal{A}) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathcal{A}^k. \quad (2.22)$$

When the dimension of the matrix  $\mathcal{A}$  is large (as happens when the spatial discretisation is fine and  $M$  is large), the exact evaluation of its exponential to machine precision can be

very expensive and the need for approximation methods arises. In such cases it is typical to time step,

$$\mathbf{u}^{n+1} = e^{h\mathcal{A}}\mathbf{u}^n,$$

in order to keep approximation sufficiently accurate and inexpensive. Since our motivation for approximating the exponential is to perform one time step of these iterations, methods that directly approximate matrix–vector products of the form  $\exp(h\mathcal{A})\mathbf{v}$  are also of relevance. Krylov methods, in particular, are methods of this type.

The simplest method for approximating the exponential is to truncate (2.45), the Taylor series expansion of  $\exp(z)$ . In this case we approximate  $\exp(z)$  by the analytical function  $p_m(z) = \sum_{k=0}^m \frac{1}{k!} z^k$  and approximate  $\exp(\mathcal{A}) \approx p_m(\mathcal{A})$ . These methods, called the Taylor methods, are highly sensitive to floating point rounding errors and have to be used with great care especially when  $\|\mathcal{A}\|_2$  is large. For all purposes, we will avoid using this method. In this section we encounter a few different techniques for the approximation of the matrix exponential. An excellent review article for such techniques is (Moler & Loan 1978).

### 2.2.1 Padé methods

Padé methods for exponentiation of a matrix (Golub & Van Loan 1996, Moler & Loan 1978) are rational methods based on the Padé approximation of  $\exp(z)$  around the origin. Among all the rational approximations to  $\exp(z)$  with numerator of degree  $p$  and denominator of degree  $q$ , the Padé approximation of  $\exp(z)$ ,  $R_{pq}$  are of the highest order. They are defined by  $R_{pq}(z) = D_{pq}(z)^{-1}N_{pq}(z)$  where

$$N_{pq}(z) = \sum_{j=0}^p \frac{(p+q-j)!p!}{(p+q)!j!(p-j)!} z^j,$$

and

$$D_{pq}(z) = \sum_{j=0}^q \frac{(p+q-j)!q!}{(p+q)!j!(q-j)!} (-z)^j.$$

The most interesting methods of these sort are the diagonal Padé methods,  $R_{qq}(\mathcal{A})$ . For  $p < q$ , roughly  $qM^3$  operations are required to evaluate  $R_{pq}(\mathcal{A})$  which is of order  $p+q$ , while for the same effort  $R_{qq}$  gives a method of order  $2q > p+q$ . However, for large  $q$ ,  $N_{qq}(\mathcal{A})$  starts looking like the series for  $e^{\mathcal{A}/2}$  and  $D_{qq}(\mathcal{A})$  the series for  $e^{-\mathcal{A}/2}$ . In such a case, the denominator matrix  $D_{qq}(\mathcal{A})$  can be very poorly conditioned (Moler & Loan 1978),

$$\text{cond}(D_{qq}(\mathcal{A})) \approx \text{cond}(e^{-\mathcal{A}/2}) \geq e^{(\alpha_1 - \alpha_n)/2},$$

where  $\alpha_1 \geq \dots \geq \alpha_n$  are the real parts of the eigenvalues of  $\mathcal{A}$ . This is because if  $\mathcal{A} = \mathcal{V}\Lambda\mathcal{V}^*$  is the spectral decomposition of  $\mathcal{A}$ , then  $e^{-\mathcal{A}/2} = \mathcal{V}e^{-\Lambda/2}\mathcal{V}^*$  and the condition

number would be given by the ratio of the highest and the smallest moduli of eigenvalues, which are given by the exponential of the largest and smallest real parts  $e^{\alpha_1/2}$  and  $e^{\alpha_n/2}$ .

A workaround for this problem is the scaling and squaring method which relies on the property  $e^{\mathcal{A}} = (e^{\mathcal{A}/k})^k$ . Usually, in this method  $k$  is chosen to be a power of 2, say  $k = 2^p$ . After evaluating  $e^{\mathcal{A}/k}$ , repeated squaring (in this case  $p$  times) yields  $e^{\mathcal{A}}$ . The motivation for this is to evaluate the exponential of a matrix with a smaller norm since this is likely to be more stable. A commonly used criterion for choosing  $p$  is to make it the smallest value for which  $\|\mathcal{A}\|_2/k \leq 1$ . Thus one could use a method like the diagonal Padé method for evaluating  $e^{\mathcal{A}/k}$  without being effected by the condition number, before squaring the approximation repeatedly.

This algorithm is highly effective and is, in fact, the method employed in the `expm` function of MATLAB. The overall expense of such a method, when combined with diagonal Padé methods, is  $\mathcal{O}(M^3 \log \rho)$  due to the matrix–matrix multiplications, where  $\rho = \rho(\mathcal{A})$  is the spectral radius of  $\mathcal{A}$  dictating the amount of squaring required. MATLAB’s `expm` is frequently utilised for generating reference solutions in our work—although it usually features an exorbitant cost, it can be relied upon for exponentiation accurate up to machine precision.

The Padé methods can be very effective when  $\|\mathcal{A}\|_2$  is small, for example in the case of unitary matrices  $\mathcal{U}$ , where  $\|\mathcal{U}\|_2 = 1$ . In such cases, since  $\rho(\mathcal{A}) \leq \|\mathcal{A}\|_2$ ,  $\alpha_1$  cannot not be too positive and  $\alpha_n$  cannot be too negative. In the case of skew-Hermitian matrices, these methods become even more effective since the real components of the eigenvalues vanish and the condition number is 1 (alternatively, the matrix  $e^{-\mathcal{A}/2}$  is unitary and its condition number is 1). In these cases, where the condition number does not cause much trouble, scaling and squaring is not required and one can directly use the diagonal Padé methods, whose expense is  $\mathcal{O}(M^3)$  due to the matrix–matrix multiplications involved.

### 2.2.2 Krylov subspace methods

Krylov methods are iterative methods applicable to situations where one needs to estimate matrix–vector products of the form  $\exp(\mathcal{A})\mathbf{v}$  and computing  $\exp(\mathcal{A})$  is not essential. Such methods have undergone many enhancements since the pioneering work of Tal Ezer & Kosloff (1984): in this thesis we adopt the approach in (Hochbruck & Lubich 1997).

Given an  $M \times M$  matrix  $\mathcal{A}$  and  $\mathbf{v} \in \mathbb{C}^M$ , the  $m$ th Krylov subspace is

$$\mathbf{K}_m(\mathcal{A}, \mathbf{v}) = \text{span}\{\mathbf{v}, \mathcal{A}\mathbf{v}, \mathcal{A}^2\mathbf{v}, \dots, \mathcal{A}^{m-1}\mathbf{v}\}, \quad m \in \mathbb{N}.$$

It is well known that  $\dim \mathbf{K}_{m-1}(\mathcal{A}, \mathbf{v}) \leq \dim \mathbf{K}_m(\mathcal{A}, \mathbf{v}) \leq \min\{m, M\}$  and we refer to (Golub & Van Loan 1996) for other properties of Krylov subspaces. In the Krylov subspace method,  $\exp(\mathcal{A})\mathbf{v}$  is approximated by projecting into the  $m$  dimensional Krylov subspace  $\mathbf{K}_m(\mathcal{A}, \mathbf{v})$  spanned by the orthonormal basis  $\{\mathbf{v}_i\}_{i=1}^m$ , where typically  $m \ll M$ .



The Arnoldi process	The Lanczos process
$\mathbf{v}_1 = \mathbf{v} / \ \mathbf{v}\ _2$ <b>for</b> $j = 1, \dots, m - 1$ <b>do</b> $\mathbf{t} = \mathcal{A}\mathbf{v}_j$ <b>for</b> $i = 1, \dots, j$ <b>do</b> $h_{i,j} = \mathbf{v}_i^* \mathbf{t}, \quad \mathbf{t} = \mathbf{t} - h_{i,j} \mathbf{v}_i$ <b>end for</b> $h_{j+1,j} = \ \mathbf{t}\ _2; \quad \mathbf{v}_{j+1} = \mathbf{t} / h_{j+1,j}$ <b>end for</b>	$\mathbf{v}_1 = \mathbf{v} / \ \mathbf{v}\ _2$ <b>for</b> $j = 1, \dots, m - 1$ <b>do</b> $\mathbf{t} = \mathcal{A}\mathbf{v}_j$ <b>if</b> $i > 1$ <b>then</b> $\mathbf{t} = \mathbf{t} - h_{j-1,j} \mathbf{v}_{j-1}$ <b>end if</b> $h_{j,j} = \mathbf{v}_j^* \mathbf{t}$ $\mathbf{t} = \mathbf{t} - h_{j,j} \mathbf{v}_j$ $h_{j+1,j} = \ \mathbf{t}\ _2; \quad h_{j,j+1} = -\ \mathbf{t}\ _2$ $\mathbf{v}_{j+1} = \mathbf{t} / h_{j+1,j}$ <b>end for</b>

Table 2.2: The Arnoldi and Lanczos iteration algorithms.

We approximate

$$\mathbf{e}^{\mathcal{A}} \mathbf{v} \approx \mathcal{V}_m \mathbf{e}^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v}, \quad (2.23)$$

where  $\mathcal{V}_m$  and  $\mathcal{H}_m$  are  $M \times m$  and  $m \times m$ , respectively, and the columns of  $\mathcal{V}_m$  are the orthonormal vectors  $\{\mathbf{v}_i\}_{i=1}^m$ . The matrix  $\mathcal{H}_m$  is upper Hessenberg. This approximation is motivated by two observations. The first observation is that, for a particular  $m$  and any polynomial  $p_{m-1}(z)$  of degree  $m - 1$ , the relation

$$p_{m-1}(\mathcal{A}) \mathbf{v} = \mathcal{V}_m p_{m-1}(\mathcal{H}_m) \mathcal{V}_m^* \mathbf{v}$$

is exact (Hochbruck & Lubich 1997). Secondly, a full decomposition brings us to  $\mathcal{A} = \mathcal{V}_M \mathcal{H}_M \mathcal{V}_M^*$  with  $\mathcal{V}_M$  being unitary. This allows us to express  $\exp(\mathcal{A}) = \mathcal{V}_M \exp(\mathcal{H}_M) \mathcal{V}_M^*$ . Substituting  $\mathcal{H}_M$  by  $\mathcal{H}_m$ , which is the  $m \times m$  upper left subpart of the matrix  $\mathcal{H}_M$ , amounts to projecting to a much smaller subspace. The matrices  $\mathcal{H}_m$  turn out to be very effective at approximating the spectrum of  $\mathcal{H}_M$  and, therefore, of  $\mathcal{A}$  as well. With  $m \ll M$ , we can exponentiate  $\mathcal{H}_m$  at a much lower computational cost using the MATLAB function `expm`, described in section 2.2.1.

In the particular case of skew-Hermitian matrices  $\mathcal{A} \in \mathfrak{u}_M(\mathbb{C})$ —the type of matrices we must exponentiate in the context of the symmetric Zassenhaus splittings— $\mathcal{H}_m$  turns out to be skew-Hermitian as well and, being upper Hessenberg, is reduced to a tridiagonal. This is easy to see because  $\mathcal{H}_m$  forms the upper left part of  $\mathcal{H}_M$ , which is in  $\mathfrak{u}_M(\mathbb{C})$  owing to

$$\mathcal{H}_M^* = (\mathcal{V}_M^* \mathcal{A} \mathcal{V}_M)^* = \mathcal{V}_M^* \mathcal{A}^* \mathcal{V}_M = -\mathcal{V}_M^* \mathcal{A} \mathcal{V}_M = -\mathcal{H}_M.$$

The matrices  $\mathcal{V}_m$  and  $\mathcal{H}_m$  are generated by the *Arnoldi process* which reduces to the *Lanczos process* (see Table 2.2) for skew-Hermitian matrices (Golub & Van Loan 1996, Hochbruck & Lubich 1997).

An advantage of  $\mathcal{H}_m \in \mathfrak{u}_M(\mathbb{C})$  is

$$\|\mathcal{V}_m \exp(\mathcal{H}_m) \mathbf{e}_1\|_2^2 = \mathbf{e}_1^* \exp(\mathcal{H}_m)^* \mathcal{V}_m^* \mathcal{V}_m \exp(\mathcal{H}_m) \mathbf{e}_1 = \mathbf{v}^* \mathbf{v} = \|\mathbf{v}\|_2^2,$$

which holds because  $\exp(\mathcal{H}_m)^* = \exp(\mathcal{H}_m^*) = \exp(-\mathcal{H}_m)$ . Here the unit vector  $\mathbf{e}_i$  is the  $i^{\text{th}}$  column of the  $N \times N$  identity matrix

The exponential of  $\mathcal{A} \in \mathfrak{u}_M(\mathbb{C})$  is a unitary matrix  $\exp(\mathcal{A}) \in \text{U}_M(\mathbb{C})$  and, in particular, the action of  $\exp(\mathcal{A})$  on  $\mathbf{v}$  preserves the norm of  $\mathbf{v}$ . That our numerical solution is also able to conserve norm precisely, is a very desirable aspect. This property has been used in chemical physics earlier (Park & Light 1986, Tal Ezer & Kosloff 1984). Unitary evolution, on the other hand, is not strictly preserved.

As the Hessenberg matrix obtained here is  $m \times (m - 1)$  and not the full  $\mathcal{H}_m$  required, we run the Lanczos process for an additional iteration. After  $m$  iterations the invariant

$$\mathcal{A} \mathcal{V}_m = \mathcal{V}_m \mathcal{H}_m + h_{m+1,m} \mathbf{v}_{m+1} \mathbf{e}_1^*$$

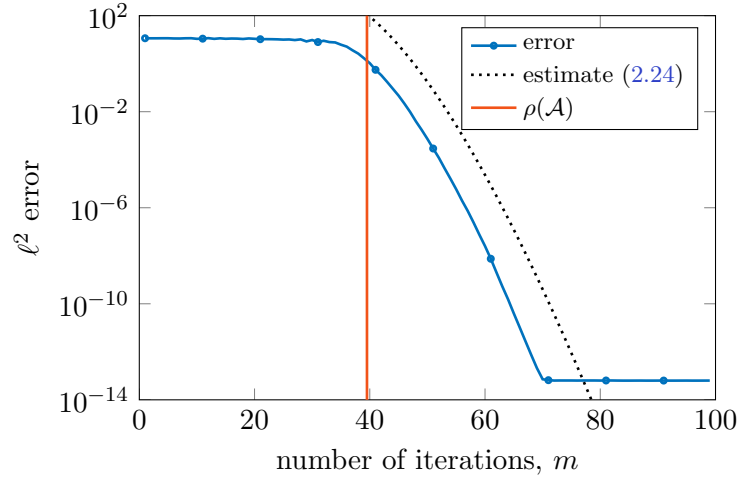
holds, where  $\mathbf{e}_i$  is the  $i^{\text{th}}$  column of the  $N \times N$  identity matrix. The additional basis vector  $\mathbf{v}_{m+1}$  can be incorporated in a slightly modified approximation but this should make only modest improvements in our case and is not explored further. Since  $\mathcal{V}_m^* \mathbf{v} = \|\mathbf{v}\|_2 \mathbf{e}_1$ , it follows that  $\exp(\mathcal{H}_m) \mathcal{V}_m^* \mathbf{v}$  is merely the first column of  $\exp(\mathcal{H}_m)$ , scaled by  $\|\mathbf{v}\|_2$ . To compute the approximation (2.23) we thus need to evaluate a *small* exponential and calculate a single matrix–vector product, at overall cost of  $\mathcal{O}(mM)$  operations.

All but the first of these iterations require a matrix–vector product of the form  $\mathcal{A} \mathbf{v}_i$ . These methods are particularly effective when  $\mathcal{A}$  is large and sparse. They can be even more effective for structured sparse matrices which we encounter here and where we often do not need to store the matrix explicitly. On the other hand,  $\exp(\mathcal{A})$  is likely to be dense even if  $\mathcal{A}$  is sparse, and computing it is undesirable since it would require large storage space. This, of course, is in addition to the undesirable  $\mathcal{O}(M^3)$  computational time associated with matrix–matrix products involved in explicitly computing  $\exp(\mathcal{A})$ . Krylov methods are found to be particularly effective on GPUs as they are well suited to parallelisation and the tighter memory constraints are not an issue.

A very tight bound on the error committed in the approximation of the exponential of  $\mathcal{A} \in \mathfrak{u}_M(\mathbb{C})$  is given by

$$\|\exp(\mathcal{A}) \mathbf{v} - \mathcal{V}_m \exp(\mathcal{H}_m) \mathcal{V}_m^* \mathbf{v}\|_2 \leq 12 \exp(-\rho^2/(4m)) \left( \frac{e\rho}{2m} \right)^m, \quad m \geq \rho, \quad (2.24)$$

where  $\rho = \rho(\mathcal{A})$  (Hochbruck & Lubich 1997). This error decay is super-linear in  $m$ . The condition  $m \geq \rho$  is, of course, not to be scoffed at and significant decrease in error starts happening only after this range, often dramatically as seen in Figure 2.10.



**Figure 2.10:** The Lanczos approximation to the matrix exponential for a random skew-Hermitian matrix  $\mathcal{A}$  starts converging rapidly after  $m \geq \rho$ .

### 2.2.3 Trotter splitting and the BCH formula

The general idea behind splitting methods is to approximate the exponential of a matrix  $\mathcal{A} = X + Y$  via exponentials of its constituents  $X$  and  $Y$ , chosen in a way that they are easier to exponentiate than exponentiating  $\mathcal{A}$  directly. While

$$e^{x+y} = e^x e^y, \quad x, y \in \mathbb{C},$$

we know that the equality

$$e^{X+Y} = e^X e^Y, \quad X, Y \in \mathbb{C}^{M \times M}$$

only holds when  $X$  and  $Y$  commute, i.e. their (matrix) commutator disappears,

$$[X, Y] = XY - YX = 0.$$

Although  $e^{X+Y} \neq e^X e^Y$  in general,  $e^X e^Y$  can, nevertheless, be considered an approximation of  $e^{X+Y}$  due to the Trotter product formula (Trotter 1959),

$$e^{X+Y} = \lim_{k \rightarrow \infty} (e^{X/k} e^{Y/k})^k.$$

Since  $e^X e^Y$  is not the exponential of  $X + Y$  in general, a natural question is “what is it the exponential of?”. In other words, we seek the matrix  $Z$  such that

$$e^Z = e^X e^Y.$$

The answer to this is provided by the matrix logarithm,

$$Z = \log(e^X e^Y).$$

This exponent is also written in the form

$$Z = \text{BCH}(X, Y),$$

where BCH stands for the *Baker–Campbell–Hausdorff formula* (Hall 2003), expressed as an infinite series of commutators of  $X$  and  $Y$ ,

$$\begin{aligned} \text{BCH}(hX, hY) &= hX + hY + \frac{1}{2}h^2[X, Y] \\ &\quad - h^3\left(\frac{1}{24}[[Y, X], X] + \frac{1}{12}[[Y, X], Y]\right) + \mathcal{O}(h^4), \end{aligned} \quad (2.25)$$

for a sufficiently small  $h$ . The coefficients occurring in this series can be evaluated algorithmically (Casas & Murua 2009, Dynkin 1947, Oteo 1991). However, for our purposes here including the first non trivial term,  $\frac{1}{2}h^2[X, Y]$ , will suffice.

Using the BCH formula,

$$e^{\text{BCH}(hX, hY)} = e^{hX} e^{hY},$$

the accuracy of the Trotter splitting can now be ascertained. We conclude that

$$e^{hX+hY} = e^{hX} e^{hY} + \mathcal{O}(h^2), \quad (2.26)$$

since we have dropped the  $\mathcal{O}(h^2)$  term  $\frac{1}{2}h^2[X, Y]$  (and smaller terms) from the BCH formula on the left hand side.

We have used the  $\mathcal{O}(h^2)$  difference in the exponents  $hX + hY$  and  $\text{BCH}(hX, hY)$  to conclude that the error in the corresponding exponentials is  $\mathcal{O}(h^2)$  without any explanation. This, however, can be made more precise. The relative perturbation in the matrix exponential of  $h\mathcal{A}$ ,

$$\phi(h) = \frac{\|e^{h(\mathcal{A}+\mathcal{E})} - e^{h\mathcal{A}}\|_2}{\|e^{h\mathcal{A}}\|_2}$$

can be bounded by

$$\phi(h) \leq h \|\mathcal{E}\|_2 e^{h(\mu(\mathcal{A}) - \eta(\mathcal{A}) - \|\mathcal{E}\|_2)}, \quad (2.27)$$

where  $\mu(\mathcal{A})$  is the largest eigenvalue of  $(\mathcal{A} + \mathcal{A}^*)/2$  (also called the logarithmic norm) and

$$\eta(\mathcal{A}) = \max\{\text{Re}(\lambda) : \lambda \text{ an eigenvalue of } \mathcal{A}\}$$

(Golub & Van Loan 1996, Moler & Loan 1978, Sheng 1994). In other words,

$$\|e^{h(\mathcal{A}+\mathcal{E})} - e^{h\mathcal{A}}\|_2 \leq h \|e^{h\mathcal{A}}\|_2 \|\mathcal{E}\|_2 e^{h(\mu(\mathcal{A}) - \eta(\mathcal{A}) - \|\mathcal{E}\|_2)} \quad (2.28)$$

or

$$\|e^{A+\mathcal{E}} - e^A\|_2 \leq \|e^A\|_2 \|\mathcal{E}\|_2 e^{\mu(A)-\eta(A)-\|\mathcal{E}\|_2} \quad (2.29)$$

for small enough  $\mathcal{A}$  and  $\mathcal{E}$ .

In particular, for  $\mathcal{A} = hX + hY$  and

$$\mathcal{E} = \text{BCH}(hX, hY) - hX - hY = \frac{1}{2}h^2[X, Y] + \mathcal{O}(h^3) = \mathcal{O}(h^2),$$

$$\exists C \geq 0, \|\mathcal{E}\| \leq Ch^2,$$

and, therefore,

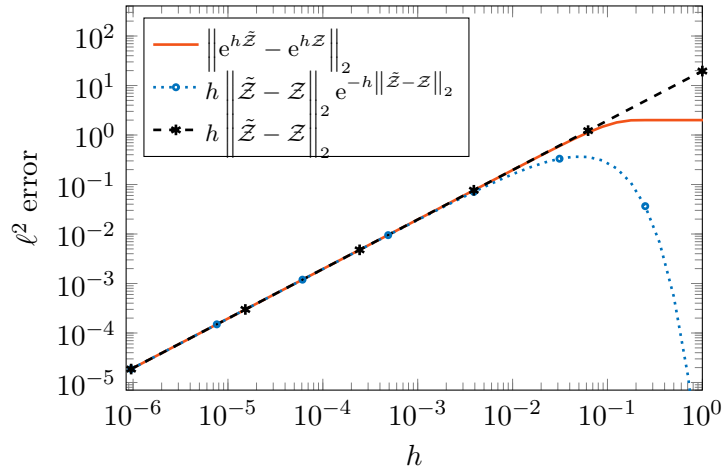
$$\begin{aligned} \left\| e^{\text{BCH}(hX, hY)} - e^{hX+hY} \right\| &\leq \left\| e^{hX+hY} \right\|_2 \|\mathcal{E}\|_2 e^{\mu(\mathcal{A})-\eta(\mathcal{A})-\|\mathcal{E}\|_2} \\ &\leq Ch^2 e^{-\|\mathcal{E}\|_2} \left\| e^{hX+hY} \right\|_2 e^{\mu(hX+hY)-\eta(hX+hY)} \\ &\leq Ch^2 \left\| e^{hX+hY} \right\|_2 e^{\mu(hX+hY)-\eta(hX+hY)}. \end{aligned} \quad (2.30)$$

For skew-Hermitian matrices  $X$  and  $Y$ , whose exponentials we will mostly be concerned with,  $\mu(\mathcal{A}) = 0$  since  $\mathcal{A}^* = -\mathcal{A}$  and  $\eta(\mathcal{A}) = 0$  since all eigenvalues are imaginary. In such cases,

$$\left\| e^{hX} \right\|_2 = \left\| e^{hY} \right\|_2 = \left\| e^{hX+hY} \right\|_2 = 1$$

since the exponential of a skew-Hermitian matrix is unitary. The error bound (2.30) now reduces to

$$\left\| e^{\text{BCH}(hX, hY)} - e^{hX+hY} \right\|_2 \leq Ch^2 e^{-\|\mathcal{E}\|_2} \leq Ch^2. \quad (2.31)$$



**Figure 2.11:** The  $\ell^2$  difference in the exponential of randomised  $100 \times 100$  skew-Hermitian matrices  $hZ$  and  $h\tilde{Z}$  compared to the error bounds of Lemma 2.2.1. Note that these bounds only hold in the asymptotic sense—i.e. when  $h$  is small enough.

**Lemma 2.2.1.** *For skew-Hermitian matrices  $\mathcal{Z}$  and  $\tilde{\mathcal{Z}}$  which are close enough,*

$$\left\| e^{\tilde{\mathcal{Z}}} - e^{\mathcal{Z}} \right\|_2 \leq \left\| \tilde{\mathcal{Z}} - \mathcal{Z} \right\|_2 e^{-\left\| \tilde{\mathcal{Z}} - \mathcal{Z} \right\|_2} \leq \left\| \tilde{\mathcal{Z}} - \mathcal{Z} \right\|_2$$

The proof for Lemma 2.2.1 follows directly from (2.29) and noting the skew-Hermiticity. In particular, if  $\left\| \tilde{\mathcal{Z}} - \mathcal{Z} \right\|_2 \leq Ch^p$ , then  $\left\| e^{\tilde{\mathcal{Z}}} - e^{\mathcal{Z}} \right\|_2 \leq Ch^p$  as well.

This observation will be inherently assumed throughout this thesis in estimating the error in the exponentials as a consequence of the error in exponents. A careful reader will note, however, that we work with unbounded operators, not matrices. The rigorous analysis that will allow us to make the relevant conclusions in case of unbounded operators will only be presented in Chapter 9.

## 2.2.4 Strang splitting and the symmetric BCH

Given complex numbers  $x$  and  $y$ , we know that the symmetric version  $e^{x+y} = e^{x/2}e^ye^{x/2}$  also holds. In the pursuit of a symmetric splitting of this sort for matrix exponentials, we seek a  $Z$  such that

$$e^Z = e^{\frac{1}{2}X}e^Ye^{\frac{1}{2}X}.$$

Such an exponent,  $Z$ , is called the symmetric BCH (sBCH) of  $X$  and  $Y$  (Casas & Murua 2009, Dynkin 1947, Hall 2003),

$$Z = \log \left( e^{\frac{1}{2}X}e^Ye^{\frac{1}{2}X} \right) = \text{sBCH}(X, Y),$$

$$e^{\frac{1}{2}X}e^Ye^{\frac{1}{2}X} = e^{\text{sBCH}(X, Y)}, \tag{2.32}$$

where

$$\begin{aligned}
 \text{sBCH}(hX, hY) &= h(X + Y) - h^3 \left( \frac{1}{24} [[Y, X], X] + \frac{1}{12} [[Y, X], Y] \right) + h^5 \left( \frac{7}{5760} [[[[Y, X], X], X], X] \right. \\
 &\quad + \frac{7}{1440} [[[[Y, X], X], X], Y] + \frac{1}{180} [[[[Y, X], X], Y], Y] \\
 &\quad + \frac{1}{720} [[[[Y, X], Y], Y], Y] + \frac{1}{480} [[[[Y, X], X], [Y, X]] \\
 &\quad - \frac{1}{360} [[[[Y, X], Y], [Y, X]]]) + h^7 \left( -\frac{31}{967680} [[[[[[Y, X], X], X], X], X], X] \right. \\
 &\quad - \frac{31}{161280} [[[[[[Y, X], X], X], X], X], Y] - \frac{13}{30240} [[[[[[Y, X], X], X], X], Y], Y] \\
 &\quad - \frac{53}{120960} [[[[[[Y, X], X], X], Y], Y], Y] - \frac{1}{5040} [[[[[[Y, X], X], Y], Y], Y], Y] \\
 &\quad - \frac{1}{30240} [[[[[[Y, X], Y], Y], Y], Y], Y] - \frac{53}{161280} [[[[[[Y, X], X], X], X], [Y, X]] \\
 &\quad - \frac{11}{12096} [[[[[[Y, X], X], X], Y], [Y, X]] - \frac{3}{4480} [[[[[[Y, X], X], Y], Y], [Y, X]] \\
 &\quad - \frac{1}{10080} [[[[[[Y, X], Y], Y], Y], [Y, X]] - \frac{1}{4032} [[[[[[Y, X], X], [Y, X]], [Y, X]] \\
 &\quad - \frac{1}{6720} [[[[[[Y, X], Y], [Y, X]], [Y, X]] - \frac{19}{80640} [[[[[[Y, X], X], X], [[Y, X], X]] \\
 &\quad - \frac{1}{10080} [[[[[[Y, X], X], Y], [[Y, X], X]] + \frac{17}{40320} [[[[[[Y, X], Y], Y], [[Y, X], X]] \\
 &\quad - \frac{53}{60480} [[[[[[Y, X], X], X], [[Y, X], Y]] - \frac{19}{13440} [[[[[[Y, X], X], Y], [[Y, X], Y]] \\
 &\quad - \frac{1}{5040} [[[[[[Y, X], Y], Y], [[Y, X], Y]]) + \mathcal{O}(h^9).
 \end{aligned}
 \tag{2.33}$$

The expansion (2.33) can be computed to an arbitrary power of  $h$  using an algorithm from (Casas & Murua 2009). Because (2.32) is palindromic, only odd powers of  $h$  feature in the expansion.

Consequently, and following the development of the Trotter splitting, we introduce the Strang splitting (Strang 1968)

$$e^{hX+hY} = e^{\frac{1}{2}hX} e^{hY} e^{\frac{1}{2}hX} + \mathcal{O}(h^3), \tag{2.34}$$

which features an  $\mathcal{O}(h^3)$  error due to the omission of the  $\mathcal{O}(h^3)$  terms occurring in the sBCH expansion (2.33). Here, once again, (2.29) and Lemma 2.2.1 come to our rescue, making it possible to analyse error by comparing the exponents. Another version of the (second order) Strang splitting is

$$e^{hX+hY} = e^{\frac{1}{2}hY} e^{hX} e^{\frac{1}{2}hY} + \mathcal{O}(h^3), \tag{2.35}$$

which can be obtained by interchanging the letters  $X$  and  $Y$ .

If a matrix  $X \in \mathbb{C}^{M \times M}$  can be split into several components,

$$X = \sum_{i=1}^m X_i, \tag{2.36}$$

the exponentials of which are easier to evaluate, we can approximate its exponential via a generalised form of the Strang splitting,

$$\exp(hX) = e^{\frac{1}{2}hX_m} e^{\frac{1}{2}hX_{m-1}} \dots e^{\frac{1}{2}hX_2} e^{hX_1} e^{\frac{1}{2}hX_2} \dots e^{\frac{1}{2}hX_{m-1}} e^{\frac{1}{2}hX_m} + \mathcal{O}(h^3). \quad (2.37)$$

### 2.2.5 Yoshida splittings

An interesting question is whether, by combining exponentials of  $X$  and  $Y$  in different ways, we can obtain splittings with greater degree of accuracy— i.e. are there coefficients  $c_i, d_i, i = 1, \dots, n$  such that

$$\prod_{i=1}^n e^{hc_i X} e^{hd_i Y} = e^{hc_1 X} e^{hd_1 Y} \dots e^{hc_n X} e^{hd_n Y} = e^{hX+hY} + \mathcal{O}(h^p) \quad (2.38)$$

for  $p > 3$ ?

This question was first answered by Yoshida (1990) in the form of a recursive procedure for developing arbitrarily high-order symmetric splittings by composing lower order methods. The technique, called the *Yoshida device*, can take any order  $2n$  splitting,  $\mathcal{S}_{2n}(h)$ , and compose it to form an order  $2n + 2$  splitting,

$$\mathcal{S}_{2n+2}(h) = \mathcal{S}_{2n}(a_{2n}h) \mathcal{S}_{2n}(b_{2n}h) \mathcal{S}_{2n}(a_{2n}h), \quad (2.39)$$

where the coefficients  $a_{2n}$  and  $b_{2n}$  are

$$a_{2n} = \frac{1}{2 - 2^{1/(2n+1)}}, \quad b_{2n} = \frac{-2^{1/(2n+1)}}{2 - 2^{1/(2n+1)}}. \quad (2.40)$$

In such a procedure, it is typical to commence from the Strang splitting (2.34) as the order two splitting,

$$\mathcal{S}_2(h) = e^{\frac{1}{2}hX} e^{hY} e^{\frac{1}{2}hX},$$

and recursively define higher order splittings. The first of these is the 4th-order Yoshida splitting  $\mathcal{S}_4(h)$ , obtained by choosing

$$a_2 = \frac{1}{2 - 2^{1/3}}, \quad b_2 = \frac{-2^{1/3}}{2 - 2^{1/3}},$$

$$\begin{aligned} \mathcal{S}_4(h) &= \mathcal{S}_2(a_2h) \mathcal{S}_2(b_2h) \mathcal{S}_2(a_2h) \\ &= \left( e^{\frac{1}{2}a_2hX} e^{ha_2Y} e^{\frac{1}{2}a_2hX} \right) \left( e^{\frac{1}{2}b_2hX} e^{hb_2Y} e^{\frac{1}{2}b_2hX} \right) \left( e^{\frac{1}{2}a_2hX} e^{ha_2Y} e^{\frac{1}{2}a_2hX} \right) \\ &= e^{\frac{1}{2}a_2hX} e^{ha_2Y} e^{\frac{1}{2}h(a_2+b_2)X} e^{hb_2Y} e^{\frac{1}{2}h(a_2+b_2)X} e^{ha_2Y} e^{\frac{1}{2}a_2hX} \\ &= e^{hX+hY} + \mathcal{O}(h^5). \end{aligned} \quad (2.41)$$



Yoshida splittings are not the only splittings of the type (2.38) and a variety of other high-order splittings of this form have been developed (Blanes, Casas & Murua 2008, McLachlan & Quispel 2002).

### Number of exponents

Let's assume that the order  $2k$  splitting,  $\mathcal{S}_{2k}(h)$ , involves  $s_i$  exponentials of  $X_i$ ,  $i = 1, \dots, m$ , and, without loss of generality, let's further assume that the outermost exponentials on both sides in  $\mathcal{S}_{2k}(h)$  are exponentials of  $X_m$ .

The recursive procedure (2.39) then dictates that in the order  $2k+2n$  Yoshida splitting  $\mathcal{S}_{2k+2n}(h)$ , we will encounter  $3^n s_i$  exponentials of  $X_i$ ,  $i = 1, \dots, m-1$ . The outermost exponentials are of  $X_m$  and two of these can be combined every time, thus we have two fewer exponent for each recursion of (2.39). A simple induction shows that the number of exponentials of  $X_m$  are  $3^n(s_m - 1) + 1$ . In general, the number of exponentials grow at the rate  $\mathcal{O}(3^n)$ .

In the case where we start with  $\mathcal{S}_2(h)$  as a Strang splitting (2.34) featuring  $X_1$  and  $X_2$  (i.e.  $m = 2$ ), we encounter  $3^{n-1}$  exponentials of  $X_1$  and  $3^{n-1} + 1$  exponentials of the outer exponent  $X_2$  for an order  $2n$  Yoshida splitting. When starting from a more general version of the Strang splitting (2.37) which features two exponentials each of  $X_2, \dots, X_m$  and one exponential of  $X_1$ , the order  $2n$  Yoshida splitting features  $3^{n-1}$  exponentials of  $X_1$ ,  $3^{n-1} + 1$  exponentials of  $X_m$  and  $2 \times 3^{n-1}$  exponentials of  $X_i, i = 2, \dots, m-1$ .

### 2.2.6 Lie groups, Lie algebras and the exponential map

It should be remarked that, although the splitting methods have been developed here in the context of matrices, the ideas extend to the splitting of the exponential map from a Lie algebra to its Lie group. We provide a very brief introduction to some of the Lie algebraic concepts required for the development of the numerical methods in this thesis and refer the reader to (Hall 2003, Khesin & Wendt 2009, Lee 2012) for more details.

**Definition 2.2.2.** *A Lie group is a smooth manifold  $G$  that is also an algebraic group where the multiplication map  $m : G \times G \rightarrow G$  and the inversion map  $i : G \rightarrow G$ ,*

$$m(g, h) = gh, \quad i(g) = g^{-1},$$

*are both smooth.*

**Example 2.2.3** (Lie Groups). Each of the following is a Lie group. The most common examples that we encounter are those of matrix Lie groups, which are manifolds in the vector space of  $m \times m$  matrices,  $M(m, \mathbb{R})$  and  $M(m, \mathbb{C})$ . Note that the matrix multiplication and inversion maps are smooth maps.

- (i) The general linear group  $GL(m, \mathbb{R})$  and the complex general linear group  $GL(m, \mathbb{C})$ , which are the sets of invertible  $m \times m$  matrices in the matrix vector spaces  $M(m, \mathbb{R})$  and  $M(m, \mathbb{C})$ , respectively, are Lie groups.
- (ii) The special linear group  $SL(m, \mathbb{R})$  and the complex special linear group  $SL(m, \mathbb{C})$  are the Lie groups of  $m \times m$  matrices with determinant 1.
- (iii) The special orthogonal group  $SO(m)$ , which is the group of real orthogonal matrices with determinant 1.
- (iv)  $U(m)$  is the group of unitary matrices in  $M(m, \mathbb{C})$ .

**Definition 2.2.4.** A Lie algebra  $\mathfrak{g}$  is a linear space equipped with a bilinear skew-symmetric operator called the Lie bracket  $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$  that satisfies the Jacobi identity. The axioms that the Lie bracket must satisfy are listed in Table 2.3.

$$\begin{aligned}
 [ax + by, z] &= a[x, z] + b[y, z], & (\text{bilinearity}) \\
 [z, ax + by] &= a[z, x] + b[z, y], & (\text{bilinearity}) \\
 [x, x] &= 0, & (\text{alternativity}) \\
 [[x, y], z] &= [x, [y, z]] + [y, [x, z]]. & (\text{Jacobi identity})
 \end{aligned}$$

**Table 2.3:** Axioms of the Lie bracket  $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$  that must hold for all  $x, y \in \mathfrak{g}$  and all  $a, b \in \mathbb{F}$  (the field over which  $\mathfrak{g}$  is a vector space).

An important property that is implied by these axioms (and can be taken as an axiom instead of alternativity in the case of fields  $\mathbb{F}$  whose characteristic is not 2) is the property of anti-commutativity of the Lie bracket,

$$[x, y] = -[y, x]. \quad (2.42)$$

The Lie algebra of the Lie group  $G$  is the tangent space at the identity of  $G$ ,  $\mathfrak{g} = T_1 G$ . Being a tangent space,  $\mathfrak{g}$  is also a vector space and its linearity is of significance in Lie group methods.

The flow  $\varphi^{[\mathcal{A}]} : G \times \mathbb{R} \rightarrow G$  of  $\mathcal{A} \in \mathfrak{g}$  is a map that describes the evolution of  $u$  from 0 to  $t$ ,

$$u(t) = \varphi^{[\mathcal{A}]}(u_0, t), \quad \forall t \geq 0.$$

under the vector field  $\mathcal{A}$ ,

$$u'(t) = \mathcal{A}u(t), \quad u(0) = u_0 \in G. \quad (2.43)$$

We note that  $\varphi^{[\mathcal{A}]}(u_0, 0) = u_0$ . If the one parameter subgroup  $\varphi^{[\mathcal{A}]}(e, \cdot)$ , generated by  $\mathcal{A}$ , exists for all  $\mathcal{A} \in \mathfrak{g}$ , where  $e$  is the group identity of  $G$ , we define the exponential map

$\exp : \mathfrak{g} \rightarrow G$  as

$$\exp(\mathcal{A}) = \varphi^{[\mathcal{A}]}(e, 1) \quad (2.44)$$

and

$$\exp(t\mathcal{A}) = \varphi^{[\mathcal{A}]}(e, t).$$

In other words, the exponential map is the integral curve of  $\mathcal{A}$  starting at the identity.

**Example 2.2.5** (Lie Algebras). Each of the following is a Lie algebra.

- (i)  $\mathfrak{gl}(n, \mathbb{R}) = M(n, \mathbb{R})$ , consisting of all  $n \times n$  matrices, is the Lie algebra of the general linear group  $GL(n, \mathbb{R})$ .
- (ii) The vector space of matrices with vanishing trace  $\mathfrak{sl}(n, \mathbb{R})$  is the Lie algebra of the special linear group  $SL(n, \mathbb{R})$ .
- (iii) The vector space of skew-symmetric matrices  $\mathfrak{so}(n, \mathbb{R})$  is the Lie algebra of the special orthogonal group  $SO(n)$ .
- (iv) The vector space of skew-Hermitian matrices  $\mathfrak{u}(n)$  is the Lie algebra of the group of unitary matrices  $U(n)$ .

The above examples are of matrix Lie groups and algebras. In these cases, the Lie bracket reduces to the matrix commutator,

$$[\mathcal{A}_1, \mathcal{A}_2] = \mathcal{A}_1\mathcal{A}_2 - \mathcal{A}_2\mathcal{A}_1,$$

and the exponential map from the Lie algebra to the Lie group,  $\exp : \mathfrak{g} \rightarrow G$ , is simply the matrix exponential,

$$\exp(\mathcal{A}) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathcal{A}^k. \quad (2.45)$$

In the case of operator algebras, such as the Lie algebra of skew-Hermitian operators, the obvious Lie bracket arises from the commutator of operators,

$$[\mathcal{A}_1, \mathcal{A}_2] = \mathcal{A}_1 \circ \mathcal{A}_2 - \mathcal{A}_2 \circ \mathcal{A}_1,$$

where  $\circ$  is operator composition and  $\mathcal{A}_i$  are linear operators.

## 2.3 The Magnus expansion

In this section we introduce methods for solving differential equations featuring time-dependent vector fields,

$$u'(t) = \mathcal{A}(t) u(t), \quad u(0) = u_0 \in G, \quad (2.46)$$

where  $G$  is a Lie group and  $\mathcal{A}(t)$  resides in its Lie algebra,  $\mathfrak{g}$ . The solution of (2.46) remains in  $G$  for all times  $t \geq 0$ ,

$$u(t) \in G, \quad t \geq 0,$$

and (2.46) is, therefore, a Lie group equation.

The Magnus expansion introduced in this section is a powerful Lie group method, which are a class of geometric integration methods (Hairer, Lubich & Wanner 2006) well suited to solving Lie group equations such as (2.46) (Iserles, Munthe-Kaas, Nørsett & Zanna 2000). This method is utilised in Chapter 7 for solving Schrödinger equations with time-dependent interaction potentials.

Lie group methods are numerical methods that exploit the properties of the Lie group and its Lie algebra, and are designed to keep the numerical solution  $u^n$  on  $G$  for all  $n \in \mathbb{N}$ , where  $u^n$  is an approximation to the true solution  $u(t_n)$  at time  $t_n$ . The broad idea of these methods is to perform the approximation in the Lie algebra  $\mathfrak{g}$ —the linearity of which is also better suited to discretisation—and use the exponential map to map back to the Lie group  $G$ .

For sufficiently small time step  $h$ , the flow can be expressed as

$$u(h) = \exp(\Theta(h))u(0),$$

where  $\Theta(h) \in \mathfrak{g}$  is a suitable function with  $\Theta(0) = 0$ . In the case where there is no time dependence in  $\mathcal{A}(t)$ ,  $\mathcal{A}(t) = \mathcal{A} \in \mathfrak{g}$  and we have  $u(t) = \exp(t\mathcal{A})u(0)$  for all  $t \in \mathbb{R}^+$ . Therefore  $\Theta(h) = h\mathcal{A}$  works in this case. In the general case, however, the solution becomes much more complicated and straightforward attempts at generalisation such as

$$u(h) = \exp\left(\int_0^h \mathcal{A}(\xi) d\xi\right)u(0)$$

do not work. In fact, this is an approximation which only holds as equality when

$$[\mathcal{A}(t_0), \mathcal{A}(t_1)] = 0, \quad \forall t_0, t_1 \in \mathbb{R}^+,$$

in which case

$$\Theta(h) = \int_0^h \mathcal{A}(\xi) d\xi$$

suffices. For general  $\mathcal{A}(t)$  that do not satisfy this criterion, the approach is to make approximations to  $\Theta$  using methods like Magnus expansions which will be discussed in this section.

If  $\tilde{\Theta} \in \mathfrak{g}$  is the approximation to  $\Theta$  at  $t_{n+1}$  and if  $u^n \in G$ , the exponential  $\exp(\tilde{\Theta})$  applied to  $u^n$  gives  $u^{n+1} = \exp(\tilde{\Theta})u^n \in G$  since  $\exp(\tilde{\Theta}) \in G$ . Thus the exact evaluation of the exponential ensures that the solution always stays on the Lie group  $G$ , in which case the error in approximation stems from the approximation of  $\Theta$ .

In Lie group methods, such as the Magnus expansion (Iserles et al. 2000), the problem thus boils down to making approximation to  $\Theta$  in  $\mathfrak{g}$  and then approximating the exponential map from  $\mathfrak{g}$  to  $G$  (Crouch & Grossman 1993, Iserles & Nørsett 1999, Munthe-Kaas 1999). We blur the boundaries by denoting the exponential map  $\exp(\Theta(h))$  by  $e^{\Theta(h)}$ , even though we might not always be working with matrix Lie algebras.

### 2.3.1 Magnus expansion

Consider the Lie group equation,

$$u'(t) = \mathcal{A}(t) u(t), \quad u(0) = u_0 \in G, \quad (2.46)$$

where  $\mathcal{A} : \mathbb{R}^+ \rightarrow \mathfrak{g}$  and, consequently,

$$u(t) \in G, \quad \forall t \geq 0.$$

For the purpose of this section it would suffice to think of  $G$  as a matrix Lie group in  $M(m, \mathbb{C})$ ,  $\mathcal{A}(t)$  as a time-dependent  $m \times m$  matrix and  $u(t) \in \mathbb{C}^m$ .

As noted earlier, the vector field described by  $\mathcal{A}(t)$  is time-dependent and the exact solution of (2.46),  $u(t)$ , can no longer be obtained by a straightforward exponentiation,

$$u(t) \neq \exp(t\mathcal{A}(t)) u_0,$$

$$u(t) \neq \exp\left(\int_0^t \mathcal{A}(\xi) d\xi\right) u_0.$$

These, however, turn out to be  $\mathcal{O}(t^2)$  and  $\mathcal{O}(t^3)$  approximations to the exact solution, respectively. Let  $\varphi_{t_2, t_1}^{[\mathcal{A}]}$  be the flow of (2.46), which describes the evolution of  $u$  from  $t_1$  to  $t_2$ ,

$$u(t_2) = \varphi_{t_2, t_1}^{[\mathcal{A}]} u(t_1), \quad \forall t_2 \geq t_1.$$

In previous notation we would have written  $\varphi^{[\mathcal{A}]}(u(t_1), t_2, t_1)$  but we will stick to the more compact notation  $\varphi_{t_2, t_1}^{[\mathcal{A}]} u(t_1)$  going forward.

The central idea of the Magnus expansion (Magnus 1954) is to seek a solution for the exact flow  $\varphi_{t_2, t_1}^{[\mathcal{A}]}$  as the exponential of some element  $\Theta(t_2, t_1)$  in the Lie algebra  $\mathfrak{g}$ ,

$$\varphi_{t_2, t_1}^{[\mathcal{A}]} = e^{\Theta(t_2, t_1)},$$

or

$$u(t_2) = e^{\Theta(t_2, t_1)} u(t_1). \quad (2.47)$$

As we will see shortly,  $\Theta(t_2, t_1)$  can be expressed as an infinite series

$$\Theta(t_2, t_1) = \sum_{k=1}^{\infty} \Theta^{[k]}(t_2, t_1),$$

where  $\Theta^{[k]}(t_2, t_1)$  features  $k$  nested integrals of grade  $k$  commutators of  $\mathcal{A}(\xi)$  at different times  $\xi \in [t_1, t_2]$ . This series is known as the Magnus expansion and its convergence is only guaranteed for sufficiently small time steps,  $|t_2 - t_1|$  (Blanes, Casas, Oteo & Ros 1999, Moan & Niesen 2008). Additionally, in practice we work with finite truncations of this series. In order to keep truncation errors small and keeping the convergence criteria in mind, it is customary to evolve the solution in small time steps  $h$ ,

$$u(t + h) = e^{\Theta(t+h, t)} u(t), \quad (2.48)$$

starting from the initial step,

$$u(h) = e^{\Theta(h, 0)} u(0). \quad (2.49)$$

The flow  $\varphi_{t+h, t}^{[\mathcal{A}]} = \exp(\Theta(t+h, t))$  is the operator that evolves the solution from  $t$  to  $t+h$ . Since  $\Theta(t+h, t)$  encodes the flow at time  $t$  under  $\mathcal{A}$ , it is recovered from  $\Theta(h, 0)$  by replacing all occurrences of  $\mathcal{A}(\zeta)$  with  $\mathcal{A}(t+\zeta)$ . For all intents and purposes, therefore, it will suffice to restrict the analysis to the first step (2.49) and the corresponding expressions for (2.48), when required, can be easily obtained by a straightforward translation of  $\mathcal{A}$ . For narrative ease, we further hide the dependence on the initial time, shortening  $\Theta(h, 0)$  to  $\Theta(h)$ .

Thus, we seek  $\Theta(h)$  such that

$$u(h) = e^{\Theta(h)} u(0), \quad (2.50)$$

is a solution of (2.46) for all  $h$ , subject to convergence of  $\Theta(h)$ . Differentiation of (2.50) together with elementary algebra (Iserles et al. 2000) shows that the exponent has to satisfy the *dexpinv equation*,

$$\Theta'(h) = \text{dexp}_{\Theta(h)}^{-1} \mathcal{A}(h) = \sum_{m=0}^{\infty} \frac{B_m}{m!} \text{ad}_{\Theta(h)}^m \mathcal{A}(h), \quad (2.51)$$

where the powers of the adjoint map,

$$\text{ad}_X(Y) = [X, Y],$$

are recursively defined by

$$\text{ad}_X^0(Y) = Y, \quad \text{ad}_X^{k+1}(Y) = [X, \text{ad}_X^k(Y)],$$

and  $B_m$  are Bernoulli numbers (Abramowitz & Stegun 1964), the first few of which are

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}, \quad B_5 = 0, \quad B_6 = \frac{1}{42}. \quad (2.52)$$

Apart from  $B_1$ , all odd-indexed Bernoulli numbers vanish.

Magnus (1954) resorted to solving the dexpinv equation (2.51) via Picard iteration,

$$\begin{aligned} \Omega_0(h) &= 0, \\ \Omega_{k+1}(h) &= \int_0^h \text{dexp}_{\Omega_k(\xi)}^{-1} \mathcal{A}(\xi) d\xi = \sum_{m=0}^{\infty} \frac{B_m}{m!} \int_0^h \text{ad}_{\Omega_k(\xi)}^m \mathcal{A}(\xi) d\xi, \end{aligned}$$

whereby  $\Theta(h) = \lim_{k \rightarrow \infty} \Omega_k(h)$ . The series  $\Omega_k(h)$  is known as the Magnus expansion.

For computational reasons, however, we adopt another version of this expansion described by Iserles & Nørsett (1999) where the operator  $\Theta(h)$  is presented as an infinite series:

$$\Theta(h) = \sum_{k=1}^{\infty} \Theta^{[k]}(h), \quad (2.53)$$

where the first few terms are

$$\begin{aligned} \Theta^{[1]}(h) &= \int_0^h \mathcal{A}(\xi_1) d\xi_1, \\ \Theta^{[2]}(h) &= -\frac{1}{2} \int_0^h \left[ \int_0^{\xi_1} \mathcal{A}(\xi_2) d\xi_2, \mathcal{A}(\xi_1) \right] d\xi_1, \\ \Theta^{[3]}(h) &= \frac{1}{12} \int_0^h \left[ \int_0^{\xi_1} \mathcal{A}(\xi_2) d\xi_2, \left[ \int_0^{\xi_1} \mathcal{A}(\xi_2) d\xi_2, \mathcal{A}(\xi_1) \right] \right] d\xi_1 \\ &\quad + \frac{1}{4} \int_0^h \left[ \int_0^{\xi_1} \left[ \int_0^{\xi_2} \mathcal{A}(\xi_3) d\xi_3, \mathcal{A}(\xi_2) \right] d\xi_2, \mathcal{A}(\xi_1) \right] d\xi_1, \\ \Theta^{[4]}(h) &= -\frac{1}{24} \int_0^h \left[ \int_0^{\xi_1} \left[ \int_0^{\xi_2} \mathcal{A}(\xi_3) d\xi_3, \left[ \int_0^{\xi_2} \mathcal{A}(\xi_3) d\xi_3, \mathcal{A}(\xi_2) \right] \right] d\xi_2, \mathcal{A}(\xi_1) \right] d\xi_1 \\ &\quad - \frac{1}{24} \int_0^h \left[ \int_0^{\xi_1} \left[ \int_0^{\xi_2} \mathcal{A}(\xi_3) d\xi_3, \mathcal{A}(\xi_2) \right] d\xi_2, \left[ \int_0^{\xi_1} \mathcal{A}(\xi_2) d\xi_2, \mathcal{A}(\xi_1) \right] \right] d\xi_1 \\ &\quad - \frac{1}{24} \int_0^h \left[ \int_0^{\xi_1} \mathcal{A}(\xi_2) d\xi_2, \left[ \int_0^{\xi_1} \left[ \int_0^{\xi_2} \mathcal{A}(\xi_3) d\xi_3, \mathcal{A}(\xi_2) \right] d\xi_2, \mathcal{A}(\xi_1) \right] \right] d\xi_1 \\ &\quad - \frac{1}{8} \int_0^h \left[ \int_0^{\xi_1} \mathcal{A}(\xi_2) d\xi_2, \left[ \int_0^{\xi_1} \mathcal{A}(\xi_2) d\xi_2, \left[ \int_0^{\xi_1} \mathcal{A}(\xi_2) d\xi_2, \mathcal{A}(\xi_1) \right] \right] \right] d\xi_1. \end{aligned}$$

### 2.3.2 Graphical representation of the Magnus expansion

The terms in the Magnus expansion can be obtained through a recursive procedure, whereby they can be easily coded as binary rooted trees. Let  $\mathcal{C}$  be a mapping from

trees to terms. We recursively define the set of trees  $\hat{\mathbb{T}}_k$  and the corresponding terms:

- (1) Let  $\hat{\mathbb{T}}_1 = \{\tau_0\}$ , and  $\mathcal{C}_{\tau_0}(h) = \mathcal{A}(h)$ .
- (2) If  $\tau_1 \in \hat{\mathbb{T}}_{m_1}$  and  $\tau_2 \in \hat{\mathbb{T}}_{m_2}$  then there exists  $\tau \in \hat{\mathbb{T}}_{m_1+m_2}$  such that

$$\mathcal{C}_{\tau}(h) = \left[ \int_0^h \mathcal{C}_{\tau_1}(\xi) d\xi, \mathcal{C}_{\tau_2}(h) \right].$$

We depict the inverse of the bijection  $\mathcal{C}$  as  $\rightsquigarrow$ , so that  $\mathcal{C}_{\tau} \rightsquigarrow \tau$  for any tree  $\tau$ . A graphical representation for these terms is obtained by representing the atomic expression  $\mathcal{A}(h)$  as a single vertex,  $\tau_0 = \bullet$ ,

$$\mathcal{A}(h) \rightsquigarrow \bullet,$$

representing the unary operator of integration by a vertical line,

$$\int_0^h \mathcal{C}_{\tau}(\xi) d\xi \rightsquigarrow \begin{array}{c} \tau \\ \vdots \\ \bullet \end{array},$$

and the Lie bracket—a binary operator—by a fork,

$$[\mathcal{C}_{\tau_1}(h), \mathcal{C}_{\tau_2}(h)] \rightsquigarrow \begin{array}{c} \tau_1 \quad \tau_2 \\ \diagdown \quad \diagup \\ \bullet \end{array}.$$

This results in a more transparent representation of Magnus expansion and what is far more comfortable for investigation of the terms. The first three sets of trees formed in this way,  $\hat{\mathbb{T}}_1$ ,  $\hat{\mathbb{T}}_2$  and  $\hat{\mathbb{T}}_3$ , are,

$$\mathcal{A}(h) \rightsquigarrow \bullet \in \hat{\mathbb{T}}_1,$$

$$\left[ \int_0^h \mathcal{A}(\xi_2) d\xi_2, \mathcal{A}(h) \right] \rightsquigarrow \begin{array}{c} \bullet \\ \vdots \\ \bullet \diagdown \quad \bullet \diagup \\ \bullet \end{array} \in \hat{\mathbb{T}}_2,$$

$$\left[ \int_0^h \mathcal{A}(\xi_2) d\xi_2, \left[ \int_0^h \mathcal{A}(\xi_2) d\xi_2, \mathcal{A}(h) \right] \right] \rightsquigarrow \begin{array}{c} \bullet \\ \vdots \\ \bullet \diagdown \quad \bullet \diagup \\ \bullet \diagdown \quad \bullet \diagup \\ \bullet \end{array} \in \hat{\mathbb{T}}_3,$$

$$\left[ \int_0^h \left[ \int_0^{\xi_2} \mathcal{A}(\xi_3) d\xi_3, \mathcal{A}(\xi_2) \right] d\xi_2, \mathcal{A}(h) \right] \rightsquigarrow \begin{array}{c} \bullet \\ \vdots \\ \bullet \diagdown \quad \bullet \diagup \\ \bullet \diagdown \quad \bullet \diagup \\ \bullet \end{array} \in \hat{\mathbb{T}}_3.$$

It should be evident that trees in  $\hat{\mathbb{T}}_k$  have  $k - 1$  vertical lines. Finally, we define the set of trees which actually appear in the Magnus expansion:  $\mathbb{T}_k$ ,

$$\tau = \begin{array}{c} \hat{\tau} \\ \vdots \\ \bullet \end{array} \in \mathbb{T}_k, \quad \mathcal{C}_{\tau}(h) = \int_0^h \mathcal{C}_{\hat{\tau}}(\xi) d\xi, \quad \forall \hat{\tau} \in \hat{\mathbb{T}}_k,$$



so that every tree  $\tau$  in  $\mathbb{T}_k$  is obtained by adding an integral to a tree  $\hat{\tau}$  from the auxiliary set  $\hat{\mathbb{T}}_k$ .

Each tree in  $\mathbb{T}_k$  possesses  $k$  integrals. For a general smooth  $\mathcal{A}$ ,  $\mathcal{A}(h) = \mathcal{O}(h^0)$ , and  $\tau \in \mathbb{T}_k$ , it immediately follows that  $\mathcal{C}_\tau(h) = \mathcal{O}(h^k)$ . We say that the tree  $\tau$  is  $\mathcal{O}(h^k)$ , for short. Similarly, every  $\tau \in \hat{\mathbb{T}}_k$  is  $\mathcal{O}(h^{k-1})$ . Each component of the Magnus expansion  $\Theta^{[k]}(h)$  is composed solely of trees from the corresponding set  $\mathbb{T}_k$ ,

$$\begin{aligned}
 \Theta^{[1]}(h) &= \text{[Diagram: a single vertical line with a dot at the top]} , \\
 \Theta^{[2]}(h) &= -\frac{1}{2} \text{[Diagram: a vertical line with a dot at the top, branching into two lines with dots at the top]} , \\
 \Theta^{[3]}(h) &= \frac{1}{12} \text{[Diagram: a vertical line with a dot at the top, branching into two lines, each of which branches into two lines with dots at the top]} + \frac{1}{4} \text{[Diagram: a vertical line with a dot at the top, branching into two lines, each of which branches into two lines with dots at the top, in a different configuration]} , \\
 \Theta^{[4]}(h) &= -\frac{1}{24} \text{[Diagram: a vertical line with a dot at the top, branching into two lines, each of which branches into two lines, each of which branches into two lines with dots at the top]} - \frac{1}{24} \text{[Diagram: another tree with 4 integrals]} - \frac{1}{24} \text{[Diagram: another tree with 4 integrals]} - \frac{1}{8} \text{[Diagram: another tree with 4 integrals]} .
 \end{aligned}$$

The set of all trees that appear in the Magnus expansion is  $\bigcup_{k \geq 1} \mathbb{T}_k$  and the Magnus expansion can be written in the form

$$\Theta(h) = \sum_{k=1}^{\infty} \Theta^{[k]}(h) = \sum_{k=1}^{\infty} \sum_{\tau \in \mathbb{T}_k} \alpha(\tau) \mathcal{C}_\tau(h), \quad h \geq 0, \quad (2.54)$$

where  $\alpha(\tau)$  is the scalar coefficient for the tree  $\tau$  which can be recursively obtained (Iserles et al. 2000, Iserles & Nørsett 1999). The reader is forewarned about certain conflict of notation: the sets  $\mathbb{T}_k$  in (Iserles et al. 2000, Iserles & Nørsett 1999) correspond to our auxiliary sets  $\hat{\mathbb{T}}_{k+1}$  and the last integral occurs explicitly in the Magnus expansion corresponding to (2.54). A few discrepancies will, therefore, be found in numbering of trees and truncations of the Magnus expansion when directly comparing with these texts.

### 2.3.3 Power truncation

A tree  $\tau$  is said to be of *power  $k$  in  $h$*  if  $k$  is the *greatest* integer such that  $\mathcal{C}_\tau(h) = \mathcal{O}(h^k)$  for every smooth  $\mathcal{A}$ . We define  $\mathbb{F}_k \subseteq \bigcup_{j \geq 1} \mathbb{T}_j$  as the set of all trees of power  $k$  in  $h$  which appear in the Magnus expansion. It is clear that  $\tau \in \mathbb{T}_k$  implies that  $\tau \in \mathbb{F}_m$  for some  $m \geq k$  since a tree with  $k$  integrals is, at the very least,  $\mathcal{O}(h^k)$ . The two sets are not

identical, however. For instance, it can be shown that the tree



belongs to  $\mathbb{T}_2$  and  $\mathbb{F}_3$ . Such a gain in power occurs wherever we encounter a pattern of the form  $\left[ \int_0^h \mathcal{C}_\tau(\xi) d\xi, \mathcal{C}_\tau(h) \right]$  with  $\tau \in \hat{\mathbb{T}}_k$ , which corresponds to the tree structure



Upon expanding  $\mathcal{C}_\tau(h)$  as  $\sum_{i=k-1}^{\infty} h^i C_i$  we find that the largest term in the above tree,  $\left[ \frac{1}{k} h^k C_{k-1}, h^{k-1} C_{k-1} \right]$ , vanishes since  $C_{k-1}$  commutes with itself (Iserles et al. 2000).

The *power truncated* Magnus expansion, which is based on truncation by the sets  $\mathbb{F}_k$ , is defined as

$$\Theta_p(h) = \sum_{k=1}^p \sum_{\tau \in \mathbb{F}_k} \alpha(\tau) \mathcal{C}_\tau(h). \quad (2.55)$$

The largest terms that have been discarded in this truncation correspond to trees from  $\mathbb{F}_{p+1}$ , which are  $\mathcal{O}(h^{p+1})$ , so that this truncated expansion incurs an error of

$$\Theta_p(h) = \Theta(h) + \mathcal{O}(h^{p+1}),$$

where  $\Theta(h)$  is the full Magnus series.

The power truncated Magnus expansion  $\Theta_4(h)$  looks no different from the regular truncated version,

$$\Theta_4(h) = \text{[tree diagram]} - \frac{1}{2} \text{[tree diagram]} + \frac{1}{12} \text{[tree diagram]} + \frac{1}{4} \text{[tree diagram]}, \quad (2.56)$$

However, in the power truncated expansion  $\Theta_5(h)$ ,

$$\Theta_5(h) = \Theta_4(h) - \frac{1}{24} \text{[tree diagram]} - \frac{1}{24} \text{[tree diagram]} - \frac{1}{8} \text{[tree diagram]}, \quad (2.57)$$

the tree



$$-\frac{1}{24} \tag{2.58}$$

drops out since it is too small when analysed correctly using Taylor analysis.

### 2.3.4 Time symmetry and gain of power

The flow of the equation (2.46),

$$\varphi_{t_2, t_1}^{[\mathcal{A}]} = e^{\Theta(t_2, t_1)},$$

satisfies

$$\varphi_{t_1, t_2}^{[\mathcal{A}]} \circ \varphi_{t_2, t_1}^{[\mathcal{A}]} = \text{id}.$$

The Magnus expansion thus satisfies

$$\Theta(t_2, t_1) = -\Theta(t_1, t_2).$$

It is advantageous to use the power truncated Magnus expansion (2.55) since they have the desirable characteristic of preserving the time symmetry. That is, their flow

$$\Phi_{t_2, t_1}^{[\mathcal{A}]} = e^{\Theta_p(t_2, t_1)},$$

also satisfies

$$\Phi_{t_1, t_2}^{[\mathcal{A}]} \circ \Phi_{t_2, t_1}^{[\mathcal{A}]} = \text{id}.$$

Consequently,

$$\Theta_p(t_2, t_1) = -\Theta_p(t_1, t_2)$$

and thus the power-truncated Magnus expansions are odd in  $h$  (Iserles et al. 2000, Iserles & Nørsett 1999, Iserles, Nørsett & Rasmussen 2001). Odd-indexed methods of this form consequently gain an extra power of  $h$ ,

$$\Theta_{2p-1}(h) = \Theta(h) + \mathcal{O}(h^{2p+1}). \tag{2.59}$$

This means that if we aim for a numerical method of order six it suffices to consider the truncation of the Magnus expansion only to the terms listed in (2.57).

### 2.3.5 Discretisation of integrals

In general it is not possible to evaluate the nested integrals of the commutators appearing in the Magnus expansion exactly. We resort to numerical quadrature methods for

approximating the integral of  $f$ ,

$$\int_0^h f(\xi) \, d\xi \approx \sum_{k=1}^n w_k f(t_k),$$

where  $w_k$  are the quadrature weights and  $t_k$  are the quadrature knots. Some of the simplest approximations are,

$$\int_0^h f(\xi) \, d\xi \approx hf(0), \quad (2.60)$$

$$\int_0^h f(\xi) \, d\xi \approx hf(h), \quad (2.61)$$

$$\int_0^h f(\xi) \, d\xi \approx hf(h/2), \quad (2.62)$$

$$\int_0^h f(\xi) \, d\xi \approx h \frac{f(0) + f(h)}{2}, \quad (2.63)$$

$$\int_0^h f(\xi) \, d\xi \approx h \frac{f(h/2 - \zeta) + f(h/2 + \zeta)}{2}, \quad \zeta \in (0, h/2). \quad (2.64)$$

By using merely two Gauss–Legendre knots,  $t_k = \frac{h}{2}(1 + k/\sqrt{3})$ ,  $k = -1, 1$ , and weights  $w_k = \frac{h}{2}$ , however, we can approximate the integral  $\int_0^h f(\xi) \, d\xi$  to  $\mathcal{O}(h^4)$  accuracy (Davis & Rabinowitz 1984, Hildebrand 1987),

$$\int_0^h f(\xi) \, d\xi = h \left( \frac{f(t_1) + f(t_{-1})}{2} \right) + \mathcal{O}(h^4). \quad (2.65)$$

This is equivalent to (2.64) with  $\zeta = \frac{h}{2\sqrt{3}}$ .

**Note:** In the context of the power truncated Magnus expansions these quadrature methods gain a power of  $h$ , becoming  $\mathcal{O}(h^5)$  in accuracy. This gain is due to the fact that such Magnus expansions are time-symmetric and, therefore, odd in  $h$  (see Section 2.3.4). Consequently, the even powers of  $h$  in its Taylor series, including error terms, vanish. Thus, any  $\mathcal{O}(h^{2n})$  accuracy quadrature method automatically becomes an  $\mathcal{O}(h^{2n+1})$  accuracy method in this context.

### Some simple Magnus expansion methods

Truncating the Magnus expansion to the first integral,

$$\Theta_1(h) = \int_0^h \mathcal{A}(\xi) \, d\xi = \Theta(h) + \mathcal{O}(h^3),$$

and using the mid point rule (2.62) for approximating the integral, the Magnus expansion reduces simply to  $h\mathcal{A}(h/2)$ . Its exponential leads to the well known exponential midpoint

rule, whose first step is

$$u^1 = \exp\left(h\mathcal{A}\left(\frac{h}{2}\right)\right)u^0.$$

Of course, as noted earlier, further steps are obtained by a simple translation of time,

$$u^{n+1} = \exp\left(h\mathcal{A}\left(t_n + \frac{h}{2}\right)\right)u^n. \quad (2.66)$$

Since (2.62) is an  $\mathcal{O}(h^3)$  approximation in the context of a power truncated Magnus expansion, the entire exponential midpoint rule is an order two method.

Pairing the Gauss–Legendre quadrature (2.65) with the first non-trivial Magnus expansion,

$$\Theta_3(h) = \int_0^h \mathcal{A}(\xi_1)d\xi_1 - \frac{1}{2} \int_0^h \left[ \int_0^{\xi_1} \mathcal{A}(\xi_2)d\xi_2, \mathcal{A}(\xi_1) \right] d\xi_1 = \Theta(h) + \mathcal{O}(h^5),$$

we arrive at the approximation of the Magnus expansion

$$\Theta_3(h) \rightsquigarrow \frac{h}{2}(\mathcal{A}(t_{-1}) + \mathcal{A}(t_1)) + \frac{\sqrt{3}h^2}{12}[\mathcal{A}(t_{-1}), \mathcal{A}(t_1)],$$

which results in the fourth order method

$$u^{n+1} = \exp\left(\frac{h}{2}(\mathcal{A}(t_{n,-1}) + \mathcal{A}(t_{n,1})) + \frac{\sqrt{3}h^2}{12}[\mathcal{A}(t_{n,-1}), \mathcal{A}(t_{n,1})]\right)u^n, \quad (2.67)$$

where  $t_{n,k} = t_n + \frac{h}{2}(1 + k/\sqrt{3})$ ,  $k = -1, 1$ . Here,  $\Theta_3(h) \rightsquigarrow \Omega$  indicates the temporal discretisation of the integrals appearing in  $\Theta_3$  using quadrature methods.

The approximation of these exponentials, of course, is an important aspect that cannot be ignored. Typically a combination of methods available for approximating matrix exponentials are applied in conjunction. Where splitting methods are employed, it makes sense to combine the order two Strang splitting with the exponential midpoint rule (2.66) and the fourth order Yoshida splitting with (2.67). Krylov methods such as Lanczos iterations provide viable alternatives, leading to Magnus–Lanczos methods, however each iteration requires the approximation of  $\Theta_3(h)\mathbf{u}$  which can prove to be expensive.

These methods have been studied extensively (Blanes, Casas, Oteo & Ros 2009, Blanes, Casas & Ros 2000) and have found various applications in the solution of Schrödinger equations (Hochbruck & Lubich 2003, Tal-Ezer, Kosloff & Cerjan 1992). They have also been found to be effective for handling stochastic ODEs (Lord, Malham & Wiese 2008).

### Multidimensional quadratures

Consider the interpolant of  $\mathcal{A}(t)$  at a set of knots  $\{t_k\}_{k=1}^n$ ,

$$\tilde{\mathcal{A}}(t) = \sum_{k=1}^n \ell_k(t) \mathcal{A}(t_k), \quad (2.68)$$

where  $\ell_k(t)$  are the Lagrange interpolation functions,

$$\ell_i(t) = \prod_{\substack{j=1 \\ j \neq k}}^n \frac{t - t_j}{t_k - t_j}, \quad k = 1, \dots, n.$$

Then  $\int_0^h \mathcal{A}(\xi) d\xi$  can be approximated by the integral of the interpolant,

$$\int_0^h \tilde{\mathcal{A}}(\xi) d\xi = \sum_{k=1}^n A(t_k) \int_0^h \ell_k(\xi) d\xi. \quad (2.69)$$

Thus,

$$w_k = \int_0^h \ell_k(\xi) d\xi$$

are the weights in the quadrature formula,

$$\int_0^h \mathcal{A}(\xi) d\xi \approx \sum_{k=1}^n w_k A(t_k). \quad (2.70)$$

When the knots  $\{t_k\}_{k=1}^n$  are the roots of Gauss–Legendre polynomials, we end up with the Gauss–Legendre quadrature which features  $\mathcal{O}(h^{2n+1})$  error for  $n$  knots<sup>2</sup> (Hildebrand 1987).

Even though a multi-dimensional integral is typically expensive to evaluate, the special form of the terms occurring in the Mangus expansion lends them to simple and efficient approximation via one-dimensional techniques (Iserles et al. 2000, Iserles & Nørsett 1999). Specifically, this is due to the multilinear nature of the nested commutators and the fact that they only involve  $\mathcal{A}$ . Having found the interpolant of  $\mathcal{A}$  using (2.70), a functional of the sort

$$\int_0^h \int_0^\xi L(\mathcal{A}(\xi), \mathcal{A}(\zeta)) d\zeta d\xi,$$

for instance, where  $L$  is some bilinear functional, can once again, be approximated by substituting the interpolant (2.69),

$$\begin{aligned} \int_0^h \int_0^\xi L(\mathcal{A}(\xi), \mathcal{A}(\zeta)) d\zeta d\xi &\approx \sum_{i=1}^n \sum_{j=1}^n L(\mathcal{A}(t_i), \mathcal{A}(t_j)) \int_0^h \int_0^\xi \ell_i(\xi) \ell_j(\zeta) d\zeta d\xi \\ &= \sum_{i=1}^n \sum_{j=1}^n L(\mathcal{A}(t_i), \mathcal{A}(t_j)) w_{i,j}, \end{aligned}$$

where  $w_{i,j}$  are the weights. Once the weights are known via the integral of the basis

---

<sup>2</sup>These methods have an accuracy of  $\mathcal{O}(h^{2n})$ , in general. However, as noted earlier in this section, due to time-symmetry, these methods always have an odd order of accuracy in the context of power-truncated Magnus expansions and thus gain one order of accuracy.

functions, we only need to combine the values of  $\mathcal{A}$  at  $n$  knots. Unlike typical two-dimensional integrals, therefore, the cost (in terms of evaluations of  $\mathcal{A}$ ) does not grow as  $n^2$ . In this way, the multi-dimensional integrals appearing in the Magnus expansion are also approximated using the evaluations of  $\mathcal{A}$  at  $t_1, \dots, t_n$ . A more detailed discourse about the discretisation of these multivariate integrals using simple univariate quadrature rules can be found in (Iserles et al. 2000).

### The self-adjoint basis of Munthe–Kaas & Owren

We will evaluate  $\mathcal{A}$  at the Gauss–Legendre quadrature points (for  $n = 3$ , these are  $t_k = \frac{h}{2} \left(1 + k\sqrt{\frac{3}{5}}\right)$ ,  $k = -1, 0, 1$ ), which leads to a far less costly quadrature.

Nevertheless, higher order Magnus expansions derived in a direct and naïve way can feature far too many commutators. Many techniques for reducing the number of terms for high-order Magnus expansions have been explored (Blanes et al. 2009). A particular strategy that we will resort to is the highly optimised version of the Magnus expansion following discretisation at the Gauss–Legendre quadrature developed by Munthe–Kaas & Owren (1999). To obtain order six approximation, all the effort of approximation of the Magnus expansion boils down to the following formula

$$\begin{aligned} \Theta_5^{[M]}(h) = & J_1 + \frac{1}{12}J_3 - \frac{1}{12}[J_1, J_2] + \frac{1}{240}[J_2, J_3] + \frac{1}{360}[J_1, [J_1, J_3]] \\ & - \frac{1}{240}[J_2, [J_1, J_2]] + \frac{1}{720}[J_1, [J_1, [J_1, J_2]]] \end{aligned} \quad (2.71)$$

where

$$\begin{aligned} J_1 &= h\mathcal{A}(t_0), \\ J_2 &= \frac{\sqrt{15}}{3}h(\mathcal{A}(t_1) - \mathcal{A}(t_{-1})), \\ J_3 &= \frac{10}{3}h(\mathcal{A}(t_1) - 2\mathcal{A}(t_0) + \mathcal{A}(t_{-1})), \end{aligned} \quad (2.72)$$

are the so called *self-adjoint basis*. These turn out to be centred finite difference approximations of  $\partial_t^j \mathcal{A}(t)$  at  $t = \frac{h}{2}$  for  $j = 0, 1, 2$  on the grid  $t_k = \frac{h}{2} \left(1 + k\sqrt{\frac{3}{5}}\right)$ ,  $k = -1, 0, 1$ .

See (Iserles et al. 2000) and (Blanes et al. 2009) for comprehensive information and ways to approximate the Magnus expansion using different quadrature rules and to higher orders. The former could be relevant if the time-dependent potential is only known at certain grid-points, as might be the case in some control setups.

## Numerical results

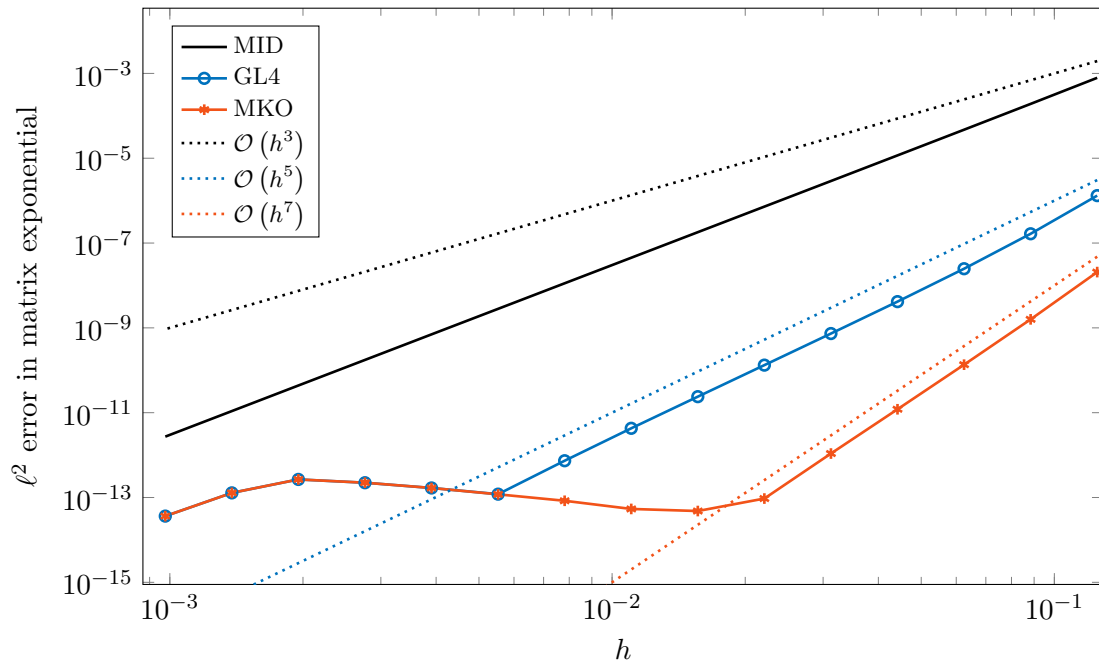
To summarise, the order two, order four and order six (discretised) Magnus expansions discussed earlier are

$$\Theta(h) = h\mathcal{A}\left(\frac{h}{2}\right) + \mathcal{O}(h^3), \quad (2.73)$$

$$= \frac{h}{2}(\mathcal{A}(t_{-1}) + \mathcal{A}(t_1)) + \frac{\sqrt{3}h^2}{12}[\mathcal{A}(t_{-1}), \mathcal{A}(t_1)] + \mathcal{O}(h^5), \quad (2.74)$$

$$= J_1 + \frac{1}{12}J_3 - \frac{1}{12}[J_1, J_2] + \frac{1}{240}[J_2, J_3] + \frac{1}{360}[J_1, [J_1, J_3]] \\ - \frac{1}{240}[J_2, [J_1, J_2]] + \frac{1}{720}[J_1, [J_1, [J_1, J_2]]] + \mathcal{O}(h^7). \quad (2.75)$$

where  $t_k = \frac{h}{2}(1 + k/\sqrt{3})$ ,  $k = -1, 1$ , and  $J_k$  are as defined in (2.72). The numerical performance of these schemes is compared in Figure 2.12.



**Figure 2.12:** The errors for the order two exponential midpoint rule (MID), the fourth order Gauss–Legendre scheme (GL4) and the sixth order Munth-Kass & Owren scheme (MKO) are compared for  $h \rightarrow 0$ . Reference solution is obtained by brute force using the midpoint rule with a very fine time step. Here  $\mathcal{A}(t) = t^3 \cos(t)X^2 + t^4 X + t^6 \cos(X) + \sin(t) \sin(X)$ , where  $X$  is a randomised  $100 \times 100$  matrix with norm  $\approx 5$ .



## Chapter 3

# Schrödinger equations and their numerical solutions

### 3.1 The time-dependent Schrödinger equation

Schrödinger equations are fundamental equations of quantum mechanics (Griffiths 2004). They describe the state of a quantum system in the wave picture of Erwin Schrödinger. The large family of Schrödinger equations can be broadly divided into two categories: the time-independent Schrödinger equations (TISEs) of the form

$$Hu(\mathbf{x}) = Eu(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \quad (3.1)$$

and the time-dependent Schrödinger equations (TDSEs),

$$i\hbar\partial_t u(\mathbf{x}, t) = Hu(\mathbf{x}, t), \quad u(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \quad t \geq 0. \quad (3.2)$$

The wave-function,  $u$ , is a complex-valued wave that resides in a Hilbert space such as  $L^2(\mathbb{R}^d, \mathbb{C})$ —the space of square integrable complex-valued functions over  $\mathbb{R}^d$ . It describes the state of the quantum system.  $H$  is the Hamiltonian operator which models the dynamics of the quantum system and governs the evolution of the wave-function  $u$ . Depending on the system under consideration, the Hamiltonian takes various forms, resulting in many different variants of the Schrödinger equation.

A common form of the Hamiltonian is

$$H = \frac{\hat{\mathbf{p}}^2}{2m} + V(\mathbf{x}) = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{x}), \quad (3.3)$$

where  $\hat{\mathbf{p}} = i\hbar\nabla$  is the momentum operator and  $V$  is a real valued function acting as the potential operator. The parameter  $\hbar$ , the *reduced Planck constant*, is truly minute,  $\hbar \approx 1.05457168 \times 10^{-34}$  Joule secs. The mass of the underlying particle,  $m$ , is a small

quantity, although substantially larger than  $\hbar$ . The Hamiltonian (3.3) has a kinetic part,  $-\hbar^2 \Delta / (2m)$ , and a potential part,  $V(\mathbf{x})$ . The notation here is not coincidental—using the classical idea of momentum,  $p = mv$ , this form results in the total energy  $\frac{1}{2}mv^2 + V(\mathbf{x})$ .

The time-independent equation (3.1) is an eigenvalue problem. It is typical to denote the eigenvalues of the Hamiltonian as  $E$  since these are the energy levels of the quantum systems. The solutions of this partial differential equation allow us to deduce the structure of a quantum system—the ground state energy, absorption spectrum and orbital shapes for an atom, for instance.

The time-dependent equation (3.4), on the other hand, is an initial value problem whose solutions allow us to predict the dynamics of a quantum system. Every solution  $u_E(\mathbf{x})$  of the time-independent problem,

$$Hu_E = Eu_E,$$

also gives us a solution of the time-dependent problem since the wave-function

$$v(\mathbf{x}, t) = e^{-itE} u_E(\mathbf{x})$$

satisfies the corresponding time-dependent equation (3.4) with initial condition  $v(\mathbf{x}, 0) = u_E(\mathbf{x})$ .

When the potential  $V$  in (3.4) is dependent on  $u$ , we end up with non-linear Schrödinger (NLS) equations such as the cubic Schrödinger equations,

$$i\hbar \partial_t u(\mathbf{x}, t) = -\frac{\hbar^2}{2m} \Delta u(\mathbf{x}, t) \pm |u(\mathbf{x}, t)|^2 u(\mathbf{x}, t), \quad \mathbf{x} \in \mathbb{R}^d, \quad t \geq 0.$$

We will, however, be concerned only with linear versions where the potential is independent of  $u$ ,

$$i\hbar \partial_t u(\mathbf{x}, t) = -\frac{\hbar^2}{2m} \Delta u(\mathbf{x}, t) + V(\mathbf{x})u(\mathbf{x}, t), \quad \mathbf{x} \in \mathbb{R}^d, \quad t \geq 0. \quad (3.4)$$

If  $u(\mathbf{x}, t)$  and  $v(\mathbf{x}, t)$  are solutions of (3.4) (with initial conditions  $u(\mathbf{x}, 0)$  and  $v(\mathbf{x}, 0)$ , respectively), so is  $au(\mathbf{x}, t) + bv(\mathbf{x}, t)$  for any  $a, b \in \mathbb{C}$  (with initial conditions  $au(\mathbf{x}, 0) + bv(\mathbf{x}, 0)$ ). This is called the principle of superposition which can be easily explained by looking at the formal solution of (3.4), written using the exponential map,

$$u(\mathbf{x}, t) = \exp(-itH)u(\mathbf{x}, 0). \quad (3.5)$$

The Hamiltonian  $H$  is a linear self-adjoint operator ( $H^* = H$ ) on the Hilbert space  $\mathcal{H} = L^2(\mathbb{R}^d, \mathbb{C})$  under the  $L^2$  inner product,

$$\langle Hu, v \rangle = \langle u, Hv \rangle, \quad \forall u, v \in L^2(\mathbb{R}^d, \mathbb{C}).$$

$itH$  is a skew-Hermitian operator and its exponential is therefore a unitary operator. Since  $\exp(itH)$  is the evolution operator (or the flow), its unitarity implies the *unitary evolution* of the wave-function,

$$\langle u(\mathbf{x}, t), v(\mathbf{x}, t) \rangle = \langle u(\mathbf{x}, 0), v(\mathbf{x}, 0) \rangle, \quad t \geq 0, \quad (3.6)$$

where  $u(\mathbf{x}, t) = \exp(-itH)u(\mathbf{x}, 0)$  is the solution of the TDSE with  $u(\mathbf{x}, 0)$  as the initial condition and  $v(\mathbf{x}, t) = \exp(-itH)v(\mathbf{x}, 0)$  is the solution with  $v(\mathbf{x}, 0)$  as the initial condition. The proof for this relation follows by noting that  $\exp(\mathcal{A})^* = \exp(\mathcal{A}^*)$  and thus,

$$\begin{aligned} \langle u(\mathbf{x}, t), v(\mathbf{x}, t) \rangle &= \langle \exp(-itH)u(\mathbf{x}, 0), \exp(-itH)v(\mathbf{x}, 0) \rangle \\ &= \langle u(\mathbf{x}, 0), \exp(itH^*) \exp(-itH)v(\mathbf{x}, 0) \rangle \\ &= \langle u(\mathbf{x}, 0), v(\mathbf{x}, 0) \rangle. \end{aligned}$$

Taking  $v = u$  in (3.6), we find that  $\|u(t)\|_2 = \|u(0)\|_2$  holds for all  $t$ . This means that the  $L^2$  norm of the wave-function does not change under evolution. It is typical to normalise the wave-function so that  $\|u(t)\|_2 = \|u(0)\|_2 = 1$ , whereby  $|u(\mathbf{x}, t)|^2$  can be interpreted as a probability density function. This probability density function—called the amplitude of the wave-function—describes the probability of finding a particle at the location  $\mathbf{x}$  at time  $t$ . In its general form, the unitary evolution described by the invariant (3.6) says that the *angle* between the two wave-function  $u$  and  $v$  does not change under evolution by the same Hamiltonian.

## 3.2 The semiclassical parameter

The reduced Planck's constant,  $\hbar \approx 1.054 \times 10^{-34} \text{ J} \cdot \text{s}$ , appearing in (3.4) is a truly minute number which would typically result in considerable difficulties as far as numerical solutions are concerned. When working in atomic units, however,  $\hbar = 1$ . In these units the typical length scales are  $10^{-11} \text{ m}$ , while typical time and mass scales are  $10^{-17} \text{ s}$  and  $10^{-30} \text{ kg}$ , respectively.

In other words, when working in the atomic units, simulations over a spatial domain of  $[-1, 1]$  and temporal domain of  $[0, 1]$  are simulations for  $10^{-17} \text{ s}$  over a domain of size  $2 \times 10^{-11} \text{ m}$ , which is incredibly small. On the other extreme, working in the standard units (where  $\hbar \approx 1.054 \times 10^{-34} \text{ J} \cdot \text{s}$ ), spatial domains of  $[-1, 1]$  and temporal domains of  $[0, 1]$  correspond to simulations lasting for a second over a range of two metres, which is far too large.

The semiclassical limit, where the Planck's constant approaches zero, is very interesting from a theoretical point of view since it ought to explain the emergence of Newtonian

mechanics from quantum mechanics as we approach macroscopic scales. We will, however, concern ourselves with spatio-temporal windows which are somewhere in between, albeit much closer to the molecular and atomic scales than macroscopic scales. In these cases we arrive at the Schrödinger equation (TDSE) under the semiclassical scaling,

$$i\varepsilon\partial_t u(x, t) = H(x)u(x, t), \quad x \in \mathbb{R}, \quad t \geq 0, \quad u(x, 0) = u_0(x), \quad (3.7)$$

where the Hamiltonian is

$$H(x) = -\varepsilon^2\partial_x^2 + V(x), \quad (3.8)$$

and the factor  $2m$  has been conveniently absorbed in  $\varepsilon$  through scaling. The standard form in which we will write the semiclassical Schrödinger equation is,

$$\partial_t u(x, t) = -i\varepsilon^{-1}H(x)u(x, t),$$

or, effectively,

$$\partial_t u(x, t) = i\varepsilon\partial_x^2 u(x, t) - i\varepsilon^{-1}V(x)u(x, t), \quad x \in \mathbb{R}, \quad t \geq 0, \quad u(x, 0) = u_0(x). \quad (3.9)$$

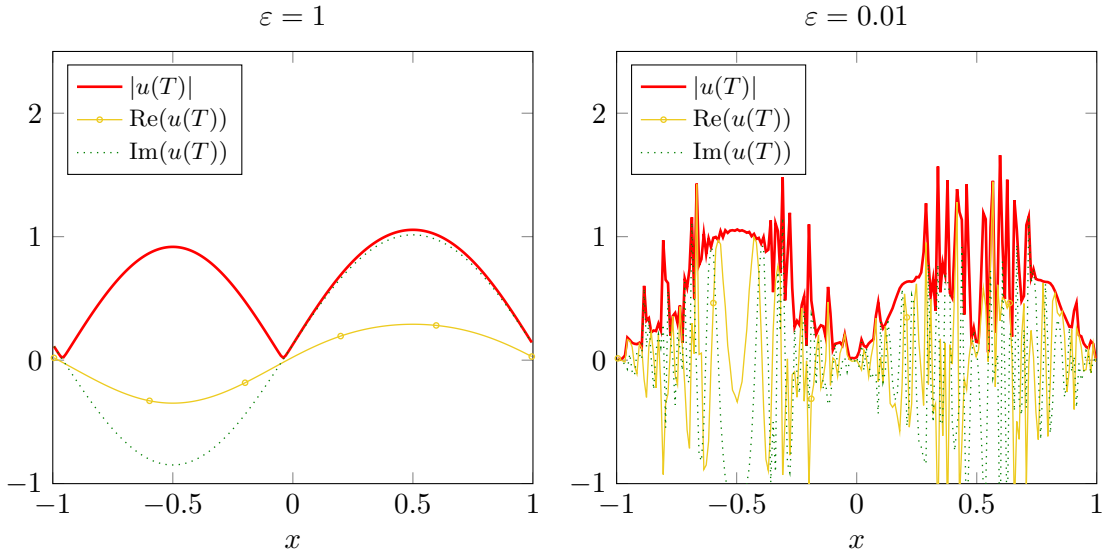
As noted previously, we are typically concerned with simulation over the spatial domain  $[-1, 1]$  and the temporal domain  $[0, 1]$ . The *semiclassical* parameter  $\varepsilon$  plays a role similar to the reduced Planck's constant. Although it is very small,  $0 < \varepsilon \ll 1$ , the semiclassical parameter is considerably larger than the Planck's constant and the range  $10^{-8} \leq \varepsilon \leq 10^{-2}$  isn't unrealistic.

The equation (3.9) also arises out of the *Born–Oppenheimer approximation* of the molecular Schrödinger equation. In this case it describes the evolution of the nuclei in the potential energy surface (PES) of the electrons,  $V(x)$ . The semiclassical parameter,  $\varepsilon$ , here, is the square root of the ratio of the mass of an electron and the heaviest nucleus. Even working in the atomic units,  $\varepsilon$  can be fairly small. Depending on the molecule under consideration,  $\varepsilon$  can be in the range  $10^{-4} \leq \varepsilon \leq 10^{-2}$ .

The two effects can be combined—when larger spatio-temporal windows are required in the simulation of the motion of nuclei following the Born–Oppenheimer approximation,  $\varepsilon$  can decrease further in size as before. However, a fact to keep in mind is that the Born–Oppenheimer approximation itself entails an error of  $\mathcal{O}((1+t)\varepsilon)$  and there is a limit to the temporal windows over which it stays valid. An excellent introduction to the derivation of the semiclassical Schrödinger equation, starting from the molecular Schrödinger equation, using the Born–Oppenheimer approximation (while keeping numerical methods in mind) can be found in (Lubich 2008).

### 3.2.1 Rapid oscillations

Regardless of the origin of the semiclassical parameter in the semiclassical Schrödinger equation (3.9), the solutions exhibit high oscillations that increase in the semiclassical limit as  $\varepsilon \rightarrow 0$ . This can be variously interpreted as the consequence of *zooming out* as we consider larger spatio-temporal domains, as an increase in momentum (which corresponds to high oscillations in the wavefunction) due to the scaling of mass, or a combination of the two.



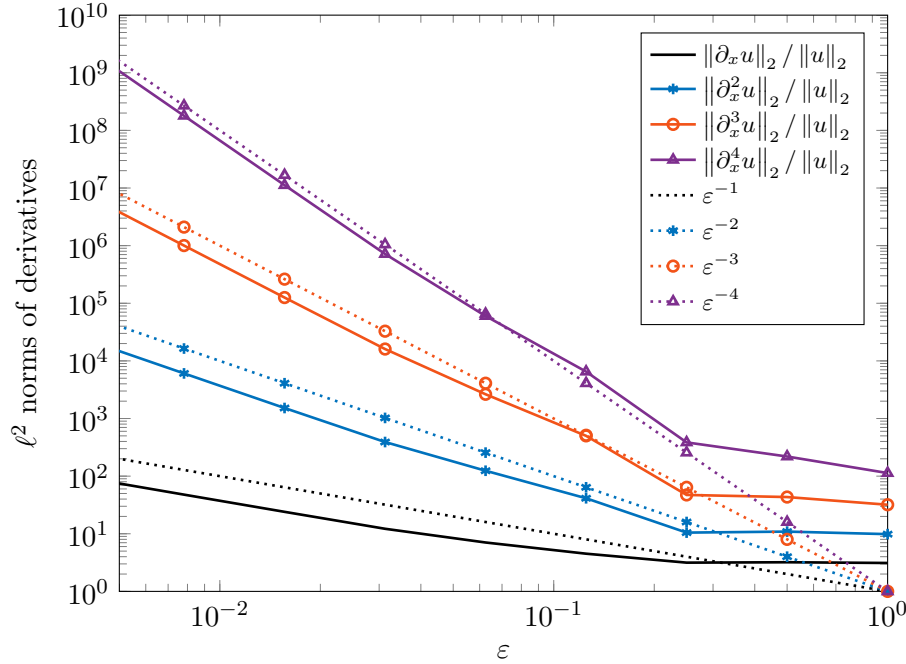
**Figure 3.1:** For small  $\varepsilon$ , even when starting from highly smooth initial data, we quickly get to a highly oscillatory solution at  $T = 1$ . The initial wave-function used here is  $u(x, 0) = \sin(\pi x)$ , which is made to evolve under the potential function  $V(x) = \exp(\sin(\pi x))$ . The solution at time  $T = 1$  is shown for two values of  $\varepsilon$ :  $\varepsilon = 1$  on the left and  $\varepsilon = 0.01$  on the right.

The rapid (spatial) oscillations caused by the small semiclassical parameter are evident in Figure 3.1. It turns out that the oscillations in the solution grow as  $\mathcal{O}(\varepsilon^{-1})$  in both space and time. This is evident from Figure 3.2, where we observe the growth of the  $k$ th derivative of the solution at time  $T = 1$  and find that  $\|\partial_x^k u\|_2 = \mathcal{O}(\varepsilon^{-k})$ . The solution at time  $T = 1$  in these figures is found by discretising in space at a fine enough resolution and directly exponentiating  $-\mathrm{i}\varepsilon^{-1}TH$  via MATLAB's `expm` function.

(Bao et al. 2002, Jin et al. 2011) note that the solution of the semiclassical equation develops oscillations with wavelength of  $\mathcal{O}(\varepsilon)$  in both space and time, and a grid resolution of  $\Delta x = \mathcal{O}(\varepsilon)$  is, therefore, essential in the discretisation in order to resolve these oscillations correctly. This observation is made rigorous in Section 9.1. Following the assumptions in (Bao et al. 2002, Jin et al. 2011) and the conclusions of Section 9.1, we assume that there exist  $C_k > 0$ , independent of  $\varepsilon$ , for every  $k \geq 0$  such that the spatial

derivatives of the solution of (3.9) grow as

$$\left\| \partial_x^k u(t) \right\|_2 \leq C_k \varepsilon^{-k}, \quad k \in \mathbb{Z}^+, \quad t \geq 0. \quad (3.10)$$



**Figure 3.2:** The  $k$ th spatial derivative of  $u(x, T)$  at time  $T = 1$  grows as  $\mathcal{O}(\varepsilon^{-k})$ . Here, once again, we start from the initial wavefunction  $u(x, 0) = \sin(\pi x)$ , evolving under the potential function  $V(x) = \exp(\sin(\pi x))$ .

### 3.3 Traditional methods for the semiclassical TDSE

The numerical solution of the semiclassical time-dependent Schrödinger equation presents several enduring challenges. Various alternatives for solving this equation have been analysed by Jin et al. (2011).

As remarked earlier, we are interested in solving this equation on a spatial domain  $[-1, 1]$  and temporal domain  $[0, T]$  where, typically, we would be interested in  $T = 1$ . Moreover, as noted by (Bao et al. 2002, Jin et al. 2011) and shown in Section 9.1, the solutions of the semiclassical Schrödinger equation are highly oscillatory and we need to resolve oscillations with wavelengths of the order  $\mathcal{O}(\varepsilon)$ . Finite difference methods, in this case, require a much higher grid resolution (Markowich, Pietra & Pohl 1999) than spectral methods (Bao et al. 2002). In the spatial discretisation of these equations, it is therefore typical to resort to spectral methods such as spectral collocation while imposing periodic boundaries at  $[-1, 1]$ .

Thus the equation we wish to solve is the linear one-dimensional semiclassical Schrödinger

equation over the torus,

$$\partial_t u(x, t) = i\varepsilon \partial_x^2 u(x, t) - i\varepsilon^{-1} V(x) u(x, t), \quad x \in [-1, 1], \quad t \geq 0, \quad (3.9)$$

where  $u$  and  $V$  are both periodic with period 2 and the wave-function  $u = u(x, t)$  is given with a periodic initial condition  $u(x, 0) = u_0(x)$ . The solution, in general, resides in the Hilbert space  $L^2([-1, 1], \mathbb{C})$  but we make further assumptions, assuming that  $u \in C_p^\infty([-1, 1], \mathbb{C})$  and  $V \in C_p^\infty([-1, 1], \mathbb{R})$  are both smooth and periodic with period 2, and that the derivatives of  $u$  are bounded by

$$\left\| \partial_x^k u(t) \right\|_2 \leq C_k \varepsilon^{-k}, \quad k \in \mathbb{Z}^+, \quad t \geq 0, \quad (3.10)$$

while  $V$  does not depend on  $\varepsilon$  in any way. Following (Jin et al. 2011) we assume that  $M = \mathcal{O}(\varepsilon^{-1})$  grid points are sufficient (and in general necessary) when approximating via spectral collocation<sup>1</sup>.

### 3.3.1 Discretisation

As we have just noted, reasonable approximation of the solution necessitates a grid resolution of  $\Delta x = \mathcal{O}(\varepsilon)$  or  $M = \mathcal{O}(\varepsilon^{-1})$  grid points at the very least for resolving the spatial oscillations. Consequently, keeping in mind the scaling (2.9), the differentiation matrix  $\mathcal{K}$  scales as  $\mathcal{O}(\varepsilon^{-1})$  and in general,

$$\mathcal{K}^k = \mathcal{O}(\varepsilon^{-k}). \quad (3.11)$$

A quick look at the error bounds for finite differences reveals a major stumbling block for these methods. The order eight finite difference method (2.7) discussed in Section 2.1.2 features an error

$$\left\| \partial_x^2 \mathbf{u} - \tilde{\mathcal{K}}_{2,FD} \mathbf{u} \right\| \leq C(\Delta x)^8 \left\| u^{(10)} \right\| + \mathcal{O}((\Delta x)^{10}), \quad (2.8)$$

where  $u^{(10)}$  is the 10th derivative of  $u$  and  $\Delta x = 2/M$  is the grid resolution. From here it could be considered obvious that the error in this approximation is  $\mathcal{O}((\Delta x)^8) = \mathcal{O}(\varepsilon^8)$ . However, this misses the crucial point that, under the semiclassical scaling,  $u$  is highly oscillatory and  $u^{(10)}$  can hardly be considered  $\mathcal{O}(1)$  with respect to  $\varepsilon$ . In fact, due to  $\mathcal{O}(\varepsilon^{-1})$  oscillations in  $u$ ,  $u^{(10)}$  is closer to  $\mathcal{O}(\varepsilon^{-10})$  as we have assumed in (3.10) and as the trend in Figure 3.2 suggests. Even multiplication by  $(\Delta x)^8$  in (2.8) only manages

<sup>1</sup>The scaling of spatial discretisation is revisited in Section 9.3.3, where we note that the actual degree of discretisation required is slightly higher. However, due to the analysis in Section 9.1, the size of exponents as well as the error of our methods remains very much in line with the results we obtain by working under the assumption  $M = \mathcal{O}(\varepsilon^{-1})$ . The difference only appears in the cost of each FFT. Thus we stick with this assumption till Chapter 9.

to bring the overall error to  $\mathcal{O}(\varepsilon^{-2})$  which is as large as the norm of the discretised Laplacian—the term we are trying to approximate!

It is possible to coax and cajole finite differences to work with  $M = \mathcal{O}(\varepsilon^{-\rho})$  for  $\rho > 1$ , but the  $\mathcal{O}(\varepsilon^{-1} \log \varepsilon^{-1})$  cost of spectral collocation cannot be improved upon in this way and the superior accuracy of these methods means that spectral collocation remains the method of choice.

### 3.3.2 Direct exponentiation

A similar trend appears in temporal evolution—under the semiclassical scaling, finite difference time stepping schemes such as Crank–Nicholson have been found to be unable to compete with spectral approximation followed by approximation of the exponential via exponential splittings such as the Strang splitting (Jin et al. 2011).

As we have seen previously in Section 3.1, the solution of the semiclassical Schrödinger equation (3.9) can be written formally as

$$u(x, t) = \exp(-it\mathbf{H}/\varepsilon) u(x, 0). \quad (3.12)$$

One approach for solving the Schrödinger equation, therefore, could be to perform spatial discretisation whereby the discretised Hamiltonian  $\mathbf{H}$  could be exponentiated directly in order to find a solution,

$$\mathbf{u}(t) = \exp(-it\mathbf{H}/\varepsilon) \mathbf{u}(0). \quad (3.13)$$

Such a solution would approximate the exact solution to the degree of accuracy of the spatial discretisation. A spectral discretisation, for instance, would result in an exponentially accurate approximation of the solution. Note that the exponential map in (3.13) is given by the matrix exponential.

More specifically, following spatial discretisation via spectral collocation, the Schrödinger equation (3.9) reduces to a system of ODEs,

$$\mathbf{u}'(t) = (i\varepsilon\mathcal{K}^2 - i\varepsilon^{-1}\mathcal{D}_V) \mathbf{u}(t), \quad t \geq 0, \quad \mathbf{u}(0) = \mathbf{u}_0, \quad (3.14)$$

where  $\mathbf{u}(t) \in \mathbb{C}^M$  is a vector of values that represent the approximate solution at time  $t$ ,  $M$  is the number of degrees of freedom (number of grid points or fourier modes) in the spatial discretisation,  $\mathcal{K}^2$  is the second order differentiation matrix approximating the infinite-dimensional operator  $\partial_x^2$ , and  $\mathcal{D}_V$  is the matrix that represents multiplication by the potential.

The exact solution of this equation can be written via direct exponentiation (3.13),

$$\mathbf{u}(t) = \exp(it\varepsilon\mathcal{K}^2 - it\varepsilon^{-1}\mathcal{D}_V) \mathbf{u}_0. \quad (3.15)$$



A natural temptation is to approximate this by any of the many methods to compute the matrix exponential discussed in Section 2.2. We recall that  $\mathcal{K}^2$  and  $\mathcal{D}_V$  are discretised as Hermitian matrices, and the exponent in (3.15) is, consequently, skew-Hermitian. Direct exponentiation via diagonal Padé methods of Section 2.2.1 requires  $\mathcal{O}(M^3)$  operations. This is in contrast to the  $\mathcal{O}(M^3 \log M)$  cost of MATLAB's `expm` which involves scaling and squaring—rendered redundant in this case since the exponent is skew-Hermitian.

Such an approach is able to preserve the unitary evolution of the wave function which is crucial, both, because of its physical significance and since it implies stability of the numerical method. Nevertheless, these methods are out of consideration since a cost of  $\mathcal{O}(\varepsilon^{-3})$  is prohibitive. This observation will form the starting point of exponential splitting methods which attempt to break down the problem of exponentiating the discretised Hamiltonian into the problem of exponentiating matrices that are much easier to exponentiate due to their structure or their size.

### 3.3.3 Lanczos iterations

Park & Light (1986) proposed Lanczos iterations for solving the Schrödinger equation. Under spectral collocation,  $\mathcal{K}^2$  is a Hermitian matrix with spectral radius  $\rho(\mathcal{K}^2) = \mathcal{O}(M^2) = \mathcal{O}(\varepsilon^{-2})$ . The exponent in (3.15) or (3.13), therefore, ends up being a skew-Hermitian matrix with a very large spectral radius,

$$\rho(-it\mathbf{H}/\varepsilon) = \rho(it\varepsilon\mathcal{K}^2 - it\varepsilon^{-1}\mathcal{D}_V) \leq t\varepsilon\rho(\mathcal{K}^2) + t\varepsilon^{-1}\|V\|_\infty = \mathcal{O}(t\varepsilon^{-1}).$$

We recall that for substantial error decrease via Lanczos iterations, the error bound (2.24) requires  $m \geq \rho = \mathcal{O}(t\varepsilon^{-1})$  iterations. If we consider large time steps,  $t = \mathcal{O}(1)$ , the spectral radius is indeed very large and we require  $\mathcal{O}(\varepsilon^{-1})$  Lanczos iterations. Each of these iterations require the evaluation of a matrix–vector product of the form  $(it\varepsilon\mathcal{K}^2 - it\varepsilon^{-1}\mathcal{D}_V)\mathbf{v}$ . While  $\mathcal{D}_V\mathbf{v}$  is evaluated in  $\mathcal{O}(M) = \mathcal{O}(\varepsilon^{-1})$  operations via point-wise multiplication,  $\mathcal{K}^2\mathbf{v}$  is evaluated via two FFTs following Section 2.1.6,

$$\mathcal{K}^2\mathbf{v} = \mathcal{F}^{-1}\mathcal{D}_{-\pi^2 n^2}\mathcal{F}\mathbf{v},$$

at a cost of  $\mathcal{O}(\varepsilon^{-1} \log \varepsilon^{-1})$ . The overall cost of the Lanczos iterations is, therefore  $\mathcal{O}(\varepsilon^{-2} \log \varepsilon^{-1})$ . This still leaves us with an upper Hessenberg matrix of size  $m \times m = \mathcal{O}(\varepsilon^{-1}) \times \mathcal{O}(\varepsilon^{-1})$  that needs to be exponentiated. The only methods that seem suited for this task are the diagonal Padé methods which entail a cost of  $\mathcal{O}(m^3) = \mathcal{O}(\varepsilon^{-3})$ . Using Lanczos iterations in this way does not make matters better!

To reduce the spectral radius of the exponent, and thereby the exorbitant expense of the Lanczos iterations, we resort to propagating (3.15) with small time steps  $t = h \ll 1$ ,

$$\mathbf{u}^{n+1} = e^{ih(\varepsilon\mathcal{K}^2 - \varepsilon^{-1}\mathcal{D}_V)}\mathbf{u}^n, \quad n \in \mathbb{Z}_+,$$

where

$$\mathbf{u}^n \approx \mathbf{u}(nh).$$

However, in order to reduce the spectral radius sufficiently, we require exceedingly small time step  $h$ . With  $h = \mathcal{O}(\varepsilon)$ , for instance, we require constant number of iterations  $m = \mathcal{O}(1)$  that do not depend on  $\varepsilon$ . The cost of exponentiating the upper Hessenberg matrix is also  $\mathcal{O}(1)$  in  $\varepsilon$  and the overall cost per time step is  $\mathcal{O}(\varepsilon^{-1} \log \varepsilon^{-1})$ . However, since we now need  $\mathcal{O}(\varepsilon^{-1})$  time steps, the overall cost of simulations over  $[0, T]$  is  $\mathcal{O}(\varepsilon^{-2} \log \varepsilon^{-1})$ .

Using the intermediate value,  $h = \mathcal{O}(\sqrt{\varepsilon})$ , we end up at the same cost. We now need  $m = \mathcal{O}(\varepsilon^{-1/2})$  and the cost of exponentiating the Hessenberg matrix amounts to  $\mathcal{O}(\varepsilon^{-3/2})$ . This is dominated by the overall cost of  $\mathcal{O}(\varepsilon^{-3/2} \log \varepsilon^{-1})$  for the Lanczos iterations. The number of time steps required are now  $\mathcal{O}(\varepsilon^{-1/2})$  and thus the cost of simulation over  $[0, T]$  is once again  $\mathcal{O}(\varepsilon^{-2} \log \varepsilon^{-1})$ .

### 3.3.4 Splitting methods

The alternative is to separate scales by means of an exponential splitting method such as the Trotter splitting (2.26),

$$\exp(ih\varepsilon\mathcal{K}^2 - ih\varepsilon^{-1}\mathcal{D}_V) = \exp(ih\varepsilon\mathcal{K}^2) \exp(-ih\varepsilon^{-1}\mathcal{D}_V) + \mathcal{O}(h^2),$$

or the Strang splittings (2.34) and (2.35),

$$\exp(ih\varepsilon\mathcal{K}^2 - ih\varepsilon^{-1}\mathcal{D}_V) = \exp\left(\frac{1}{2}ih\varepsilon\mathcal{K}^2\right) \exp(-ih\varepsilon^{-1}\mathcal{D}_V) \exp\left(\frac{1}{2}ih\varepsilon\mathcal{K}^2\right) + \mathcal{O}(h^3), \quad (3.16)$$

$$\exp(ih\varepsilon\mathcal{K}^2 - ih\varepsilon^{-1}\mathcal{D}_V) = \exp\left(-\frac{1}{2}ih\varepsilon^{-1}\mathcal{D}_V\right) \exp\left(\frac{1}{2}ih\varepsilon\mathcal{K}^2\right) \exp\left(-\frac{1}{2}ih\varepsilon^{-1}\mathcal{D}_V\right) + \mathcal{O}(h^3). \quad (3.17)$$

This has the clear virtue of separating scales (powers of  $\varepsilon$ ) and components of different structures—as we have seen in Section 2.1.6, under the choice of spectral collocation  $\mathcal{K}^2$  is circulant and  $\mathcal{D}_V$  is a diagonal.

$$\mathcal{K}^2 = \mathcal{F}^{-1} \mathcal{D}_{-\pi^2 n^2} \mathcal{F}$$

is diagonalised via Fourier transforms, whereby the exponential of the outer component in (3.16) is evaluated to machine precision using two FFTs,

$$e^{\frac{1}{2}ih\varepsilon\mathcal{K}^2} = \mathcal{F}^{-1} \mathcal{D}_{\exp(-\frac{1}{2}ih\pi^2 n^2)} \mathcal{F},$$

at  $\mathcal{O}(M \log M) = \mathcal{O}(\varepsilon^{-1} \log \varepsilon^{-1})$  cost, while the inner component, being diagonal, is exponentiated directly,

$$e^{-ih\varepsilon^{-1}\mathcal{D}_V} = \mathcal{D}_{\exp(-ih\varepsilon^{-1}V)}.$$

Strang splittings are the most standard splitting methods used for the solution of Schrödinger equations under the semiclassical scaling (Jin et al. 2011). Although the accuracy of the Strang splittings is usually derived for ODEs and exponential splittings of matrices, the error analysis can be extended for PDEs such as the Schrödinger equation which involve unbounded operators (Jahnke & Lubich 2000) or by following the analysis in Section 9.4.

Yet the order of approximation seems to be unacceptably low when we take into account the large constant that is hiding within the  $\mathcal{O}(h^3)$  accuracy of these methods—we find that this constant scales in inverse powers of the small semiclassical parameter,  $\varepsilon \ll 1$ . The error analysis for the Strang splitting (2.34), which was done by considering the size of the largest commutators discarded from the sBCH formula (2.33), shows that the highest order terms we have discarded are

$$ih^3(\tfrac{1}{24}\varepsilon[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{K}^2] + \tfrac{1}{12}\varepsilon^{-1}[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{D}_V]).$$

Going by

$$\|[X, Y]\| = \|XY - YX\| \leq 2\|X\|\|Y\|$$

and keeping in mind the fact that  $\|\mathcal{K}^2\|_2 = \mathcal{O}(\varepsilon^{-2})$ , the discarded terms ought to be  $\mathcal{O}(h^3\varepsilon^{-3})$ . Thus the error of the Strang splitting, analysed in this way, is much higher and the hidden constant  $\varepsilon^{-3}$  is truly enormous. The consequence for this splitting is that we need  $h \ll \varepsilon \ll 1$  in order to achieve reasonable accuracy.

The standard high-order generalisation of the Strang splitting bears the form

$$e^{i\alpha_1 h \varepsilon \mathcal{K}^2} e^{i\beta_1 h \varepsilon^{-1} \mathcal{D}_V} e^{i\alpha_2 h \varepsilon \mathcal{K}^2} \dots e^{i\alpha_r h \varepsilon \mathcal{K}^2} e^{i\beta_r h \varepsilon^{-1} \mathcal{D}_V} e^{i\alpha_r h \varepsilon \mathcal{K}^2} \dots e^{i\alpha_2 h \varepsilon \mathcal{K}^2} e^{i\beta_1 h \varepsilon^{-1} \mathcal{D}_V} e^{i\alpha_1 h \varepsilon \mathcal{K}^2}.$$

The palindromic form of this splitting (it reads the same from the left and from the right), which is referred to as symmetric splitting in much of the literature, is not accidental, since it guarantees higher order. The coefficients  $\alpha_i$  and  $\beta_i$  are typically chosen to ensure either higher order (because of palindromy, the order is always even) or smaller error constants or both (Blanes, Casas & Murua 2006, McLachlan & Quispel 2002).

In Section 2.2.5, we have discussed the most standard approach for achieving arbitrarily high-orders of accuracy—the Yoshida device. The cost for these splittings, however, grows exponentially in the order desired, as we have seen previously in Section 2.2.5. In this thesis we seek exponential splittings whose costs grow polynomially in the order desired, ending up with Zassenhaus splittings which feature a quadratic growth in costs.

Similar analysis of commutator sizes extended to an order  $2n$  Yoshida splitting shows

that the error expected should be  $\mathcal{O}(h^{2n+1}\varepsilon^{-(2n+1)})$ , which imposes the same restriction upon us (in terms of the time step) as the Strang splitting. This is very unlike the case of the regular Schrödinger equation where  $\varepsilon = 1$  and a severe depression of the time step isn't required.

Deeper analysis performed in Section 6.1.4 and Section 6.7 using the algebra of undiscretised operators, however, improves upon these error estimates, which are found to be overly pessimistic, and the constant in the error term is shown to grow as  $\mathcal{O}(\varepsilon^{-1})$ . However, in order to analysis this rigourously and in an elegant fashion, we first develop the algebraic theory of the Lie algebra generated by the undiscretised operators,  $\partial_x^2$  and  $V$ , which is introduced first in a rudimentary form in Chapter 4, and which forms the central theme of Chapter 5. Rigorous analysis of these bounds in the semiclassical regime are discussed in Section 9.1.

### 3.4 Time-dependent potentials

Historically, the time-independent Schrödinger equation (TISE) preceded the development of the time-dependent Schrödinger equation (TDSE). Although the most well known form of the TDSE is with the Hamiltonian (3.3), Briggs & Rost (2001) note that Schrödinger considered time dependence in the TDSE as arising from time-dependent potentials, in a classical treatment of the external environment. One of the first applications of the TDSE that Schrödinger considered was to the interaction of an atom with a classical electric field, resulting in a time-dependent potential,  $V(t)$ . Time-dependent potentials, in this way, can be considered to have a fundamental relevance to the TDSE, resulting in time-dependent Hamiltonians of the form,

$$H(t) = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{x}, t). \quad (3.18)$$

The one-dimensional, linear, semiclassical, time-dependent Schrödinger equation for a single particle moving in a time-dependent electric field is

$$i\varepsilon\partial_t u(x, t) = H(x, t)u(x, t), \quad x \in \mathbb{R}, \quad t \geq 0,$$

where the Hamiltonian now features a time-dependent electric potential,

$$H(x, t) = -\varepsilon^2\partial_x^2 + V(x, t),$$

and the wave-function  $u = u(x, t)$  is given with an initial condition  $u(x, 0) = u_0(x)$ . We write this equation in our standard form as

$$\partial_t u(x, t) = i\varepsilon\partial_x^2 u(x, t) - i\varepsilon^{-1}V(x, t)u(x, t), \quad x \in \mathbb{R}, \quad t \geq 0. \quad (3.19)$$

As typical, we impose periodic boundaries at  $x = \pm 1$ ,

$$\partial_t u(x, t) = i(\varepsilon \partial_x^2 - \varepsilon^{-1} V(x, t)) u(x, t), \quad x \in [-1, 1], \quad t \geq 0, \quad (3.20)$$

and assume that the initial conditions are periodic and smooth, i.e. the potential  $V(\cdot, t)$  lives in  $C_p^\infty([-1, 1]; \mathbb{R})$ , the space of smooth real-valued functions with period 2, and  $u(\cdot, 0)$  lives in  $C_p^\infty([-1, 1]; \mathbb{C})$ , the space of smooth complex-valued functions with period 2. In addition, we assume that  $u$  has oscillations of  $\mathcal{O}(\varepsilon)$  wavelength in space and  $M = \mathcal{O}(\varepsilon^{-1})$  grid points are sufficient (and generally necessary) for correct resolution of the oscillations.

### 3.4.1 Magnus expansions

Considering (3.20) as a PDE evolving in a Hilbert space, say  $\mathcal{H} = L_2([-1, 1], \mathbb{C})$  (the Hilbert space of complex-valued square-integrable functions over  $[-1, 1]$ ), and suppressing the dependence on  $x$ ,

$$\partial_t u(t) = (i\varepsilon \partial_x^2 - i\varepsilon^{-1} V(t)) u(t), \quad u(0) = u_0, \quad (3.21)$$

is seen to be of the ‘ODE-like’ form

$$\partial_t u(t) = \mathcal{A}(t) u(t), \quad u(0) = u_0, \quad (3.22)$$

with  $\mathcal{A}(t) = -iH(t)/\varepsilon = i\varepsilon \partial_x^2 - i\varepsilon^{-1} V(t)$  acting as a scaled version of the Hamiltonian,

$$H(t) = -\varepsilon^2 \partial_x^2 + V(t).$$

The form of (3.22) is the same as (2.46) and, despite  $\mathcal{A}(t)$  being an infinite-dimensional and unbounded operator, a Magnus expansion can be developed along the same lines,

$$u(h) = e^{\Theta(h)} u(0), \quad (3.23)$$

where  $u$  is the undiscretised wave-function and  $\Theta(h)$  is an infinite-dimensional operator that features nested integrals of nested commutators of the operator  $\mathcal{A}(\xi)$ . The application of Magnus expansion based methods for the solution of PDEs such as the Schrödinger equation, which feature unbounded operators, have been analysed by (Hochbruck & Lubich 2003). The algebraic structure of the Magnus expansion has been elaborated in greater detail in Section 2.3.

The real-valued potential,  $V$ , and the Laplacian,  $\partial_x^2$ , are Hermitian operators on  $\mathcal{H}$ , which makes  $\mathcal{A} = -iH/\varepsilon$  a skew-Hermitian operator on  $\mathcal{H}$ . The time-independent Hamiltonians generate a one-parameter unitary group,  $\{e^{-itH/\varepsilon} : t \geq 0\}$ , where the flow resides. However, here we have a time-dependent Hamiltonian. Such Hamiltonians do not generate a one-parameter unitary group. The solution for Schrödinger equations featuring

time-dependent potentials is obtained via the Magnus expansion,  $\Theta(s, t)$ . The Magnus expansion (2.53) respects Lie algebra structures (Iserles et al. 2000), so that  $\Theta(s, t)$  is a skew-Hermitian operator as well. The flow of (3.22),  $\varphi_{s,t}^{[A]}$ , therefore, resides in the two parameter unitary group  $\{e^{\Theta(s,t)} : s, t \geq 0\}$ .

As we have noted previously, unitary evolution of the wave-function,  $u(t)$ , under this flow is a central aspect of quantum mechanics. Preservation of this property under discretisation is a desirable aspect of any numerical method which naturally comes about when working in the appropriate Lie-algebraic framework such as the Magnus expansion. As we will see in Section 6.9, unitarity automatically guarantees stability of a consistent numerical scheme.

Traditionally, the first step in approximating (3.21) is spatial discretisation along the lines of (3.14),

$$\mathbf{u}'(t) = i(\varepsilon \mathcal{K}^2 - \varepsilon^{-1} \mathcal{D}_{V(\cdot, t)}) \mathbf{u}(t), \quad t \geq 0, \quad (3.24)$$

where the vector  $\mathbf{u}(t) \in \mathbb{C}^M$  represents an approximation to the solution at time  $t$ ,  $\mathbf{u}(0)$  is derived from the initial conditions, while  $\mathcal{K}^2$  and  $\mathcal{D}_{V(\cdot, t)}$  are  $M \times M$  matrices which represent (discretisation of) second derivative and a multiplication by the interaction potential  $V(\cdot, t)$ , respectively.

Although a direct exponentiation of the scaled Hamiltonian is no longer possible, the system of ODEs (3.24) can be solved via the Magnus expansion,

$$\mathbf{u}(h) = e^{\Theta(h)} \mathbf{u}(0),$$

where  $\Theta(h) \in \mathfrak{u}(M)$  is a time-dependent  $M \times M$  skew-Hermitian matrix obtained as an infinite series  $\sum_{k=1}^{\infty} \Theta^{[k]}(h)$  with each  $\Theta^{[k]}(h)$  composed of  $k$  nested integrals and commutators of the matrices  $i\varepsilon \mathcal{K}^2$  and  $i\varepsilon^{-1} \mathcal{D}_V$ .

Magnus based methods have been effectively utilised in computational chemistry for solving TDSEs with time-dependent Hamiltonians (Tal Ezer & Kosloff 1984). Following Section 2.3.5, the lowest order truncation of the Magnus expansion is

$$\Theta_1(h) = \int_0^h \mathcal{A}(\xi) d\xi = \Theta(h) + \mathcal{O}(h^3),$$

whose approximation using the mid point rule (2.62) results in the exponential mid point method,

$$u^1 = \exp\left(h\mathcal{A}\left(\frac{h}{2}\right)\right) u^0.$$

In the case of the semiclassical Schrödinger equation, this translates to

$$u^1 = \exp\left(i\varepsilon h \mathcal{K}^2 - i\varepsilon^{-1} h \mathcal{D}_{V(h/2)}\right) u^0,$$

which requires the evaluation of the potential at the midpoint of the interval  $[0, h]$ . Since

this is an order two method in  $h$ , it can be combined with a Strang splitting to obtain the order two method

$$u^1 = \exp\left(\frac{1}{2}i\varepsilon h\mathcal{K}^2\right) \exp\left(-i\varepsilon^{-1}h\mathcal{D}_{V(h/2)}\right) \exp\left(\frac{1}{2}i\varepsilon h\mathcal{K}^2\right) u^0. \quad (3.25)$$

The first non-trivial Magnus based method encountered in Section 2.3.5 was the fourth order Magnus expansion,

$$\Theta_3(h) = \int_0^h \mathcal{A}(\xi_1) d\xi_1 - \frac{1}{2} \int_0^h \left[ \int_0^{\xi_1} \mathcal{A}(\xi_2) d\xi_2, \mathcal{A}(\xi_1) \right] d\xi_1 = \Theta(h) + \mathcal{O}(h^5),$$

which is combined with Gauss–Legendre quadrature (2.65), resulting in the method

$$u^1 = \exp\left(\frac{h}{2}(\mathcal{A}(t_{-1}) + \mathcal{A}(t_1)) + \frac{\sqrt{3}h^2}{12}[\mathcal{A}(t_{-1}), \mathcal{A}(t_1)]\right) u^0,$$

where  $t_k = \frac{h}{2}(1 + k/\sqrt{3})$ ,  $k = -1, 1$ . For the semiclassical Schrödinger equation, this translates to

$$\begin{aligned} \Theta_3(h) &\rightsquigarrow \frac{h}{2}(\mathcal{A}(t_{-1}) + \mathcal{A}(t_1)) + \frac{\sqrt{3}h^2}{12}[\mathcal{A}(t_{-1}), \mathcal{A}(t_1)] \\ &= i\varepsilon h\mathcal{K}^2 - i\varepsilon^{-1}h\bar{V} + h^2 \frac{\sqrt{3}h^2}{12}[\mathcal{K}^2, \tilde{V}], \end{aligned}$$

where  $\bar{V} = \frac{V(t_{-1}) + V(t_1)}{2}$  and  $\tilde{V} = V(t_{-1}) - V(t_1)$ . The exponential of  $\Theta_3$  needs to be evaluated up to an accuracy of  $\mathcal{O}(h^5)$  to make the fourth order Magnus expansion worthwhile. The second order Strang splitting,

$$e^{\Theta_3(h)} = e^{\frac{1}{2}i\varepsilon h\mathcal{K}^2} e^{-\frac{1}{2}i\varepsilon^{-1}h\mathcal{D}_{\bar{V}}} e^{h^2 \frac{\sqrt{3}h^2}{12}[\mathcal{K}^2, \mathcal{D}_{\tilde{V}}]} e^{-\frac{1}{2}i\varepsilon^{-1}h\mathcal{D}_{\bar{V}}} e^{\frac{1}{2}i\varepsilon h\mathcal{K}^2} + \mathcal{O}(h^3),$$

therefore, does not quite suffice. The lowest order splitting that would do is a fourth order splitting such as the fourth order Yoshida splitting obtained by composing three order-two Strang splittings.

The first issue that we face here is the fact that the number of terms in the Yoshida splitting grow rapidly as we go towards higher accuracy Magnus expansions—for the fourth order Magnus expansion, the Yoshida splitting requires the evaluation of 13 exponentials. Secondly, unlike the case of Schrödinger equations with time-independent potentials, the exponentials involved are not merely of  $\mathcal{K}^2$  and  $\mathcal{D}_V$ , which possess a nice structure that makes them amenable to effective exponentiation, but also of the commutator term  $[\mathcal{K}^2, \mathcal{D}_{\tilde{V}}]$  which can be problematic to exponentiate for many reasons.

The exponent featuring  $[\mathcal{K}^2, \mathcal{D}_{\tilde{V}}]$  can not be efficiently diagonalised. Even after solving the matrix commutator  $[\mathcal{K}^2, \mathcal{D}_{\tilde{V}}]$  at a cost of  $\mathcal{O}(M^2 \log M)$  using FFTs, we end up at an  $M \times M$  matrix whose exponentiation via diagonal Padé methods costs  $\mathcal{O}(M^3)$  operations. As far as Lanczos based methods are concerned, we could get away without solving the

matrix commutator since  $[\mathcal{K}^2, \mathcal{D}_{\tilde{V}}]\mathbf{v}$  in each Lanczos iteration could be evaluated using four FFTs,

$$\mathcal{F}^{-1}\mathcal{D}_{-\pi^2 n^2}\mathcal{F}\mathcal{D}_{\tilde{V}}\mathbf{v} - \mathcal{D}_{\tilde{V}}\mathcal{F}^{-1}\mathcal{D}_{-\pi^2 n^2}\mathcal{F}\mathbf{v}.$$

The number of Lanczos iterations required, however, seems very large when we estimate

$$\|[\mathcal{K}^2, \mathcal{D}_{\tilde{V}}]\|_2 \leq 2 \|\mathcal{D}_{\tilde{V}}\|_2 \|\mathcal{K}^2\|_2 \leq C \frac{2}{\sqrt{3}} h \|\partial_t V\|_\infty \varepsilon^{-2},$$

whereby the spectral radius of the exponent seems to be  $\mathcal{O}(h^3 \varepsilon^{-2})$ . This analysis suggests that we require very small time-steps for effective exponentiation via Lanczos methods.

An alternative approach is to skip exponential splittings altogether and directly utilise Lanczos iterations to exponentiate the truncated Magnus expansion  $\Theta_3$ . We recall that the cost of exponentiating the time-independent Hamiltonian turned out to be  $\mathcal{O}(\varepsilon^{-2} \log \varepsilon^{-1})$ . This certainly cannot be improved upon for the case of time-dependent Hamiltonians. Here, we additionally encounter commutators of the form  $[\mathcal{K}^2, \mathcal{D}_{\tilde{V}}]$  which seem to make the spectral radius larger (except when we take very small time steps) and make the computation of the matrix–vector product  $\Theta_3 \mathbf{v}$  more expensive since  $[\mathcal{K}^2, \mathcal{D}_{\tilde{V}}]\mathbf{v}$  requires four more FFTs.

Additionally, we have to take into account the fact that the  $\mathcal{O}(h^3)$  accuracy of the Strang splitting, the  $\mathcal{O}(h^5)$  accuracy of the Yoshida splitting and even the  $\mathcal{O}(h^3)$  and  $\mathcal{O}(h^5)$  accuracies of the two truncated Magnus expansions,  $\Theta_1$  and  $\Theta_3$ , hide large constants in powers of  $\varepsilon^{-1}$ . Direct analysis of commutator sizes using

$$\|[X, Y]\| = \|XY - YX\| \leq 2 \|X\| \|Y\|$$

suggests that these hidden constants are very high negative powers of  $\varepsilon$ . However, these estimates, much like the previously mentioned hidden error constant in the Strang splitting for the semiclassical Schrödinger equation, turn out to be highly pessimistic.

Among others, Hochbruck & Lubich (2003) have previously noted that such commutator bounds are indeed pessimistic (even in the case of Schrödinger equations with  $\varepsilon = 1$ ) and that a more reasonable bound is presented by estimates of the form

$$\|[\mathcal{K}^2, \mathcal{D}_V]\|_2 \leq C_V \|\mathcal{K}\|_2,$$

where  $C_V$  is some constant depending on  $V$ . In Chapter 4 we will encounter this phenomenon in a slightly different way when working in the Lie algebra generated by  $\partial_x^2$  and  $V$ —this leads to the systematic reduction of spectral radius for nested commutators.

It is also worth considering the convergence of the Magnus expansion in the context of the Schrödinger equation. Iserles & Nørsett (1999) showed that the Magnus expansions for ODEs converge so long as  $h \|\mathcal{A}\| \rightarrow 0$ . Moan & Niesen (2008) improved upon this bound and showed convergence of the Magnus expansion for  $\int_0^h \|\mathcal{A}(\xi)\| \, d\xi < \pi$ . Hochbruck &



Lubich (2003) note that in the case of PDEs such as the Schrödinger equation (under  $\varepsilon = 1$ ), where we encounter unbounded operators, the Magnus expansion does not meet the convergence criteria. Nevertheless, in practice, Magnus expansion based methods have been found to be effective for much larger time steps than those suggested by the convergence bounds of (Moan & Niesen 2008).

This phenomenon has been explained by Hochbruck & Lubich (2003), who conclude that even for  $h \|\mathcal{K}\|_2 \leq c$ , where  $c$  is some constant, the fourth order Magnus expansion based methods are indeed able to achieve order four accuracy. This should be compared to the constraint  $h \|\mathcal{K}^2\|_2 < c$  implied by (Moan & Niesen 2008), which needs to be met for the convergence of the Magnus expansion, beyond which the order of accuracy should make sense (to be fair, though, these more restrictive convergence constraints were obtained for a general  $\mathcal{A}(t)$  and milder constraints might well be possible for a specific case).

For the case of the semiclassical Schrödinger equation,  $\|\mathcal{A}\|_2 = \mathcal{O}(\varepsilon^{-1})$  and the convergence constraints imposed by (Moan & Niesen 2008) translate to very small time steps of size  $h = \mathcal{O}(\varepsilon)$ . As such, a direct application of the constraint  $h \|\mathcal{K}\|_2 \leq c$  of (Hochbruck & Lubich 2003) does not seem to improve upon the time step restriction. It seems plausible but not immediately obvious that appropriately accounting for powers of  $\varepsilon$  in the analysis of (Hochbruck & Lubich 2003) will allow milder constraints.

Moreover, the error bound obtained by Hochbruck & Lubich (2003) suggest that the fourth order Magnus expansion results in an error,

$$\|u^n - u(t_n)\|_2 \leq Ch^5 t_n \max_{0 \leq t \leq t_n} \|\partial_x^4 u(t)\|_2,$$

which was noted by the authors as being worrisome due to the presence of  $\partial_x^4 u(t)$ . In the semiclassical case, this scales as  $\mathcal{O}(\varepsilon^{-4})$  due to (3.10), leading to an error of  $\mathcal{O}(h^5 \varepsilon^{-4})$  and necessitating very small time steps for achieving reasonable accuracy. A translation of these results to the semiclassical regime, therefore, needs to be done more carefully.

In Chapter 7 and Chapter 8 we develop Magnus–Zassenhaus schemes where we combine truncated Magnus expansions with Zassenhaus splittings. It turns out that, unlike the case of Yoshida splittings, the number of exponents and the overall cost of the Zassenhaus splittings does not increase much when combined with higher order Magnus expansions. Moreover, the error for the fourth order Magnus–Zassenhaus scheme turns out to be  $\mathcal{O}(h^5 \varepsilon^{-1})$ , allowing higher accuracies with less restrictive time steps. This analysis is carried out in Section 9.3 and is summed up in Theorem 9.3.5 and Theorem 9.3.6. In particular, this involves proving

$$\|[\mathcal{A}(\tau_k), [\dots, [\mathcal{A}(\tau_1), \mathcal{A}(\tau_0)] \dots]]\|_2 = \mathcal{O}(\varepsilon^{-1}), \quad \tau_i \in [0, T], \quad i = 0, \dots, k, \quad k \in \mathbb{Z}^+, \quad (3.26)$$

which results from observations made in Chapter 4, 5 and 6, summarised in Lemma 6.6.1 and formally proven in Corollary 9.1.6.



## Chapter 4

# Commutator-free asymptotic splittings

In Chapter 3 we saw that the semiclassical Schrödinger equation,

$$\partial_t u(x, t) = i\varepsilon \partial_x^2 u(x, t) - i\varepsilon^{-1} V(x) u(x, t), \quad x \in [-1, 1], \quad t \geq 0, \quad u(x, 0) = u_0(x), \quad (3.9)$$

is discretised via spectral collocation as

$$\partial_t \mathbf{u}(t) = (i\varepsilon \mathcal{K}^2 - i\varepsilon^{-1} \mathcal{D}_V) \mathbf{u}(t), \quad t \geq 0, \quad \mathbf{u}(0) = \mathbf{u}_0. \quad (3.14)$$

The exact solution of (3.14) is obtained via an exponential,

$$\mathbf{u}(h) = \exp(ih\varepsilon \mathcal{K}^2 - ih\varepsilon^{-1} \mathcal{D}_V) \mathbf{u}_0, \quad (3.15)$$

which is approximated to order two accuracy using the Strang splitting,

$$e^{ih\varepsilon \mathcal{K}^2 - ih\varepsilon^{-1} \mathcal{D}_V} = e^{\frac{1}{2}ih\varepsilon \mathcal{K}^2} e^{-ih\varepsilon^{-1} \mathcal{D}_V} e^{\frac{1}{2}ih\varepsilon \mathcal{K}^2} + \mathcal{O}(h^3). \quad (3.16)$$

Higher order splittings of the form

$$e^{i\alpha_1 h \varepsilon \mathcal{K}^2} e^{-i\beta_1 h \varepsilon^{-1} \mathcal{D}_V} e^{i\alpha_2 h \varepsilon \mathcal{K}^2} \dots e^{i\alpha_r h \varepsilon \mathcal{K}^2} e^{-i\beta_r h \varepsilon^{-1} \mathcal{D}_V} e^{i\alpha_r h \varepsilon \mathcal{K}^2} \dots e^{i\alpha_2 h \varepsilon \mathcal{K}^2} e^{-i\beta_1 h \varepsilon^{-1} \mathcal{D}_V} e^{i\alpha_1 h \varepsilon \mathcal{K}^2}$$

retain the main virtues of (3.16), namely separation of scales and the ease of computation of individual exponentials. However, an inordinately large number of exponentials is typically required to attain significant order. The simplest means toward a high-order splitting, the *Yoshida method* (McLachlan & Quispel 2002, Yoshida 1990), calls for  $r = 3^{p-1}$  (which translates to  $2 \times 3^{p-1} + 1$  exponentials) to attain order  $2p$ . Consequently, the cost of these methods grows exponentially with the order of the method, as seen in Section 2.2.5.

Our aim is to develop alternative high-order splitting schemes that require far fewer

exponentials to attain a given order. In particular, we wish the number of exponentials to grow linearly, rather than exponentially, with order. Moreover, once the number of exponentials becomes large, ideally we do not want all of them to fit into the same scale but wish for them to become increasingly smaller. Such a splitting can be termed an *asymptotic splitting*.

In this chapter we lay the groundwork for a family of exponential splittings called the *symmetric Zassenhaus schemes* which posses these favourable features. More specifically, we will work towards exponential splittings of the form

$$e^{ih(\varepsilon\mathcal{K}^2 - \varepsilon^{-1}\mathcal{D}_V)} = e^{\frac{1}{2}W^{[0]}} e^{\frac{1}{2}W^{[1]}} \dots e^{\frac{1}{2}W^{[s]}} e^{\mathcal{W}^{[s+1]}} e^{\frac{1}{2}W^{[s]}} \dots e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[0]}} + \mathcal{O}(\varepsilon^{2s+2}), \quad (4.1)$$

where

$$\begin{aligned} W^{[0]} &= W^{[0]}(h, \varepsilon, \mathcal{K}) = \mathcal{O}(\varepsilon^0), \\ W^{[k]} &= W^{[k]}(h, \varepsilon, \mathcal{K}) = \mathcal{O}(\varepsilon^{2k-2}), \quad k = 1, \dots, s, \\ \mathcal{W}^{[s+1]} &= \mathcal{W}^{[s+1]}(h, \varepsilon, \mathcal{K}) = \mathcal{O}(\varepsilon^{2s}), \end{aligned}$$

and variations on this theme. As always,  $\mathcal{K}$  and  $\mathcal{D}_V$  are matrices that approximate the differential operator and multiplication by the potential  $V$ , respectively. There are a number of critical differences between (4.1) and standard exponential splittings.

Firstly, we quantify the error not in terms of the step-size  $h$  but of the small semi-classical parameter  $\varepsilon$ . There are three small quantities at play: the parameter  $\varepsilon$  intrinsic to the semiclassical Schrödinger equation (3.9), and the two parameters of numerical discretisation— $h$  and  $\Delta x$ . By letting power laws govern the relationship between  $\varepsilon$  and the choices of  $h$  and  $\Delta x$ , we express the error in the single quantity  $\varepsilon$ .

Secondly, the number of individual terms in (4.1) is remarkably small and it grows *linearly* with  $s$ —compare with the exponential growth, as a function of order, in the number of components of familiar splittings. The reason is that the arguments of the exponentials in (4.1) decay increasingly more rapidly in  $\varepsilon$ .

Thirdly, each of these exponentials can be computed fairly easily. Some of the exponents are diagonal matrices, whereby computing the exponential is trivial. Other are circulants and can be computed with FFT. Finally, because of the minute spectral radius of the arguments for sufficiently large  $k$ , the remaining exponentials can be evaluated up to the requisite power of  $\varepsilon$  using a *very* low-dimensional Krylov subspace method. All in all, the cost of these splittings ends up being quadratic in the desired order in contrast with the exponential cost of the Yoshida method.

The asymptotic splitting (4.1) is possible because we have deliberately breached the consensus in the design of exponential splittings: the terms  $W^{[k]}$  and  $\mathcal{W}^{[s+1]}$  arise from computation of nested commutators. The use of commutators is usually frowned upon

because of their cost, and also because they are believed to increase in norm. Powerful tools like Zassenhaus splitting were historically avoided in splitting methods due to the large computational cost of nested commutators. However, as we demonstrate in Section 4.2, choosing the correct, infinite-dimensional Lie algebra in case of the Schrödinger vector field, these commutators lose their unwelcome features and enable the derivation of effective, asymptotic splittings.

The first idea is to forego the standard step of semidiscretising (3.9) to (3.14) before splitting the exponential: we semidiscretise only once the splitting has been done! Thus, the entire narrative takes place within the Lie algebra of  $\partial_x^2$  and  $V$ ,  $\mathfrak{g} = \text{LA}\{\partial_x^2, V\}$ , where  $\partial_x$  is the differential operator and  $V$  is the operation of multiplying with the interaction potential: since we have not yet discretised, both are infinite-dimensional linear operators. We demonstrate in Section 4.2 that  $\mathfrak{g}$  can be embedded in a larger Lie algebra  $\mathfrak{S}$ , where commutation has simple, straightforward interpretation. To all intents and purposes, commutators are replaced by simple linear combinations of powers of  $\partial_x$ . Moreover—and this is what lets all this procedure work in a beneficial manner—these are smaller powers of  $\partial_x$  than naïvely expected.

Our splittings build on the standard, non-symmetric, Zassenhaus splittings (Oteo 1991) by simplifying commutators of undiscretised operators in the Lie algebra  $\mathfrak{S}$  and correctly accounting for powers of  $\varepsilon$ . Although the underlying algebra is time consuming, it need be done just once.

In Section 4.1 we introduce the standard Zassenhaus splittings and explore their extension to symmetrised splittings under the semiclassical scaling. At this stage the Zassenhaus splittings are not commutator-free. We remedy this in Section 4.2 by introducing the rules of the Lie algebra  $\mathfrak{S}$ , which contains the Lie algebra generated by  $V$  and  $\partial_x^2$ . However, the structures that emerge from these algebraic computations lose the crucial property of preserving skew-Hermiticity under discretisation. We discuss this problem and a work-around in Section 4.3. A formal investigation of the Lie algebra of the symmetrised differential operators responsible for maintaining skew-Hermiticity and stability is carried out in Chapter 5. We only resume the full construction of the symmetric Zassenhaus splittings for the semiclassical Schrödinger equation in Chapter 6, having established the necessary theoretical results in Chapter 5.

## 4.1 Zassenhaus splitting

Unless  $X$  and  $Y$  commute,  $e^{h(X+Y)} = e^{hX}e^{hY}$  generally does not hold. As we saw earlier in Section 2.2.4, the discrepancy in these two formulas is quantified by the BCH formula

$$e^{hX}e^{hY} = e^{\text{BCH}(hX, hY)}. \quad (2.25)$$

An alternative way of quantifying the discrepancy is the *Zassenhaus splitting* (Oteo 1991),

$$e^{h(X+Y)} = e^{tX} e^{hY} e^{h^2 U_2(X,Y)} e^{h^3 U_3(X,Y)} e^{h^4 U_4(X,Y)} \dots, \quad (4.2)$$

where

$$\begin{aligned} U_2(X, Y) &= \frac{1}{2}[Y, X], \\ U_3(X, Y) &= \frac{1}{3}[[Y, X], Y] + \frac{1}{6}[[Y, X], X], \\ U_4(X, Y) &= \frac{1}{24}[[[Y, X], X], X] + \frac{1}{8}[[[Y, X], X], Y] + \frac{1}{8}[[[Y, X], Y], Y], \end{aligned}$$

and the focus is shifted to  $e^{h(X+Y)}$  as the primary object of interest instead of the splitting  $e^{hX}e^{hY}$ . More terms in the regular Zassenhaus splitting can be generated using the (non-symmetric) BCH formula.

#### 4.1.1 Symmetrisation of Zassenhaus

The splitting (4.2) is not well known and seldom used in computation. The natural temptation is to *symmetrise* it and consider a palindromic splitting of the form

$$e^{h(X+Y)} = \dots e^{h^5 Q_5(X,Y)} e^{h^3 Q_3(X,Y)} e^{\frac{1}{2}hX} e^{hY} e^{\frac{1}{2}hX} e^{h^3 Q_3(X,Y)} e^{h^5 Q_5(X,Y)} \dots, \quad (4.3)$$

where we can deduce by inspection of the sBCH formula (2.33),

$$\begin{aligned} \text{sBCH}(hX, hY) &= h(X + Y) - h^3 \left( \frac{1}{24}[[Y, X], X] + \frac{1}{12}[[Y, X], Y] \right) + h^5 \left( \frac{7}{5760}[[[Y, X], X], X] \right. \\ &\quad + \frac{7}{1440}[[[Y, X], X], Y] + \frac{1}{180}[[[Y, X], X], Y], Y] \\ &\quad + \frac{1}{720}[[[Y, X], Y], Y], Y] + \frac{1}{480}[[[Y, X], X], [Y, X]] \\ &\quad \left. - \frac{1}{360}[[[Y, X], Y], [Y, X]] \right) + \mathcal{O}(h^7), \end{aligned}$$

that, for example,

$$Q_3(X, Y) = \frac{1}{48}[[Y, X], X] + \frac{1}{24}[[Y, X], Y].$$

Rather than engaging in increasingly tedious calculations to compute  $Q_5$ , we replace (4.3) by a more computation-friendly splitting. We commence from the symmetric BCH formula (2.33),

$$e^{-\frac{1}{2}hX} e^{h(X+Y)} e^{-\frac{1}{2}hX} = e^{\text{sBCH}(-hX, h(X+Y))},$$

which we rewrite in the form

$$e^{h(X+Y)} = e^{\frac{1}{2}hX} e^{\text{sBCH}(-hX, h(X+Y))} e^{\frac{1}{2}hX}. \quad (4.4)$$

It follows from (2.33) that

$$\text{sBCH}(-hX, h(X+Y)) = hY + \mathcal{O}(h^3),$$

and we note that we have extracted the outer term  $hX$  from the exponent  $h(X+Y)$  in (4.4). We iterate (4.4) to further split the resulting central exponent and continue to symmetrically pull out the lowest order terms, one by one, until the central exponent reaches the desired high order,

$$\begin{aligned} \exp(h(X+Y)) &= e^{\frac{1}{2}hX} \text{sBCH}(-hX, h(X+Y)) e^{\frac{1}{2}hX} \\ &= e^{\frac{1}{2}hX} e^{\frac{1}{2}hY} \text{sBCH}(-hY, \text{sBCH}(-hX, h(X+Y))) e^{\frac{1}{2}hY} e^{\frac{1}{2}hX} = \dots \end{aligned}$$

Notice that by pulling terms out, we essentially subtract a term and add higher order corrections. It is important to observe that the order of the central exponent (in terms of  $h$ ) given by the sBCH formula in (4.4) is never decreased by this procedure<sup>1</sup>. With the notation

$$\mathcal{W}^{[k+1]} = \text{sBCH}(-W^{[k]}, \mathcal{W}^{[k]}), \quad \mathcal{W}^{[0]} = h(X+Y), \quad (4.5)$$

the result after  $s$  steps can be written as

$$\exp(h(X+Y)) = e^{\frac{1}{2}W^{[0]}} e^{\frac{1}{2}W^{[1]}} \dots e^{\frac{1}{2}W^{[s]}} e^{W^{[s+1]}} e^{\frac{1}{2}W^{[s]}} \dots e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[0]}}.$$

We emphasise that, in principle, we can freely choose the elements  $W^{[k]}$  that we want to extract. This choice can afford a great deal of flexibility—it could be based on some structural property that allows for trivial exponentiation of  $W^{[k]}$  when extracted separately, a small spectral radius which makes the term amenable to effective exponentiation through Krylov subspace methods, a combination of both, or some other criteria.

So long as the terms are decreasing in size, the convergence of the procedure is guaranteed. This can lead to many variants of the splitting, some of which could prove to have more favourable properties than others. Moreover, this flexibility will allow us to commence our splitting from more complicated exponents such as the Magnus expansion.

The first and the most obvious approach is to choose  $W^{[k]} = \mathcal{O}(h^{2k-1})$  for  $k > 0$  and  $W^{[0]} = \mathcal{O}(h)$ , which yields a separation of powers analogous to (4.3) and leads to an asymptotic splitting in  $h$  with  $s$  stages, where  $\mathcal{W}^{[s+1]} = W^{[s+1]} + \mathcal{O}(h^{2s+3})$ . This leads to a symmetrised Zassenhaus splitting of order  $2s+2$  in  $h$ . Working in this way, we require very few stages to obtain a high-order splitting: in fact, the number of stages required can be shown to grow linearly in the order.

<sup>1</sup>unless a non-existing term is subtracted and thus newly introduced instead of removed.

### 4.1.2 Zassenhaus under semiclassical scaling

The symmetrised Zassenhaus splitting procedure discussed above, however, is still a few steps away from (4.1).

Firstly, while such an asymptotic splitting in powers of  $h$  would prove useful for the analysis of ODEs, for (3.9) we also need to contend with the semiclassical parameter  $\varepsilon$  as well as the size of differentiation matrices resulting from semidiscretisation. In order to use the symmetric Zassenhaus algorithm for developing a splitting along the lines of (4.1), therefore, we need to correctly account for the size of exponents and the error in the currency of  $\varepsilon$ .

Secondly, starting from the discretised Schrödinger equation (3.14), the symmetric Zassenhaus splitting results in exponents  $W^{[k]}$  that feature nested commutators of matrices such as  $[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{K}^2]$  (as a natural consequence of the sBCH expansion) that cannot be simplified and are very expensive to exponentiate unless the time step is sufficiently depressed,  $h \ll 1$ . In order to develop an effective splitting along the lines of (4.1), therefore, it becomes crucial to solve commutators and find techniques for reducing their (spectral) size. Techniques for accomplishing this are explored in Section 4.2.1.

As always, taking into account the  $\mathcal{O}(\varepsilon^{-1})$  oscillations characteristic of the semiclassical Schrödinger equation, the spatial discretisation choice is governed by

$$\Delta x = \mathcal{O}(\varepsilon), \quad M = \mathcal{O}(\varepsilon^{-1}).$$

While being aware that  $\partial_x$  is an unbounded infinite-dimensional operator, we abuse notation and use the shorthand

$$\partial_x = \mathcal{O}(\varepsilon^{-1})$$

since  $\partial_x$  is eventually replaced by the differentiation matrix  $\mathcal{K}$  which scales<sup>2</sup> as  $\mathcal{O}(M) = \mathcal{O}(\varepsilon^{-1})$ . Similarly, we use the shorthand

$$\partial_x^k = \mathcal{O}(\varepsilon^{-k}).$$

Additionally, in order to express errors in terms of  $\varepsilon$ , we let the time step be governed by the power law<sup>3</sup>,

$$h = \mathcal{O}(\varepsilon^\sigma), \quad \sigma > 0.$$

Larger values of  $\sigma$  correspond to extremely small time steps and, for this reason, methods that work effectively despite smaller  $\sigma$  are preferable.

---

<sup>2</sup> In Section 9.3.3 we find that we require a slightly faster scaling of  $M$  with  $\varepsilon$ . Nevertheless, the analysis of the size of  $\partial_x$  and its powers that we carry out holds and, in fact, can be made more rigorous due to the observations of Section 9.1. Thus, we can safely postpone the discussion of this anomaly till Chapter 9.

<sup>3</sup> As will become evident in the analysis presented in Section 9.3 and Section 9.4, we can, in fact, analyse our methods without scaling  $h$  with  $\varepsilon$ . However, the scaling proves very convenient in judging the size of terms in a uniform manner.



As a consequence of using the sBCH series, the  $\mathcal{O}(h^3)$  terms in the symmetrised Zassenhaus splittings derived using the algorithm (4.5) of Section 4.1.1 end up being of the form  $\varepsilon h^3[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{K}^2]$  and  $\varepsilon^{-1} h^3[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{D}_V]$  in the case of the semiclassical Schrödinger equation, where a splitting of (3.15) is required. Taking the scaling  $\mathcal{K}^2 = \mathcal{O}(\varepsilon^{-2})$  into account, these are estimated to be  $\mathcal{O}(h^3 \varepsilon^{-3}) = \mathcal{O}(\varepsilon^{3(\sigma-1)})$ , just as we noted in the error analysis of Strang splitting in Section 3.3.4. The  $\mathcal{O}(h^5)$  terms, which arise from grade five commutators, are, similarly,  $\mathcal{O}(\varepsilon^{5(\sigma-1)})$ . Following this analysis, it seems that  $\sigma > 1$  is crucial for convergence and certainly for reasonable accuracy. The only way to achieve an asymptotic splitting of the kind (4.1), in fact, seems to be to choose  $\sigma = 2$ . In contrast, the range of  $\sigma$  we are interested in achieving is  $\sigma \leq 1$ . Once again, this follows the conclusions we drew about Strang and Yoshida splittings in Section 3.3.4.

In Section 4.2.2 we find that these estimates are overly pessimistic. Keeping the scaling laws in mind, we will seek splittings of the kind

$$e^{ih(\varepsilon \mathcal{K}^2 - \varepsilon^{-1} \mathcal{D}_V)} = e^{\frac{1}{2}W^{[0]}} e^{\frac{1}{2}W^{[1]}} \dots e^{\frac{1}{2}W^{[s]}} e^{\mathcal{W}^{[s+1]}} e^{\frac{1}{2}W^{[s]}} \dots e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[0]}} + \mathcal{O}\left(\varepsilon^{(2s+3)\sigma-1}\right), \quad (4.6)$$

where

$$\begin{aligned} W^{[0]} &= W^{[0]}(h, \varepsilon, \mathcal{K}) = \mathcal{O}(\varepsilon^{\sigma-1}), \\ W^{[k]} &= W^{[k]}(h, \varepsilon, \mathcal{K}) = \mathcal{O}\left(\varepsilon^{(2k-1)\sigma-1}\right), \quad k = 1, \dots, s, \\ \mathcal{W}^{[s+1]} &= \mathcal{W}^{[s+1]}(h, \varepsilon, \mathcal{K}) = \mathcal{O}\left(\varepsilon^{(2s+1)\sigma-1}\right). \end{aligned}$$

We denote the splitting (4.6) as  $\mathcal{Z}_{s,\sigma}$ . In the first instance, we will consider the scaling  $\sigma = 1$  or  $h = \mathcal{O}(\varepsilon)$ , which is the choice that leads us to (4.1).

## 4.2 Splitting the undiscretised Hamiltonian

We recall that the formal solution of

$$\partial_t u(x, t) = -i\varepsilon^{-1} H(x) u(x, t), \quad x \in [-1, 1], \quad t \geq 0, \quad u(x, 0) = u_0(x), \quad (3.2)$$

with the Hamiltonian,

$$H = -\varepsilon^2 \partial_x^2 + V(x), \quad (3.8)$$

is

$$u(x, t) = e^{-itH/\varepsilon} u(x, 0). \quad (3.12)$$

In Chapter 3, we commence our analysis by discretising the Schrödinger equation, effectively replacing the problem of exponentiating the Hamiltonian operator (3.8) by the problem of approximating the exponential of its discretised form using exponential split-

ting schemes. In this section we attempt to perform a splitting directly on the exponential (3.12) of the undiscretised Hamiltonian (3.8).

The vector field in the semiclassical Schrödinger equation (3.2) is a linear combination of the action of two operators,  $\partial_x^2$  and multiplication by the interaction potential  $V$ . Since the development of exponential splittings entails nested commutation, the focus of our interest is on the Lie algebra generated by  $V$  and  $\partial_x^2$ ,

$$\mathfrak{g} = \text{LA}\{V, \partial_x^2\},$$

i.e. the linear-space closure of all nested commutators generated by  $\partial_x^2$  and  $V$ . The elements of  $\mathfrak{g}$  are operators, acting on sufficiently smooth functions including the initial value of (3.2): for the purpose of this thesis and for simplicity sake we assume that the solution of (3.2) is a periodic function in  $C_p^\infty([-1, 1], \mathbb{C})$ , but our results should extend to functions of lower smoothness.

#### 4.2.1 Solving commutators

To compute commutators we need, in principle, to describe their action on functions, e.g.

$$[V, \partial_x^2]u = V(\partial_x^2 u) - \partial_x^2(Vu) = -(\partial_x^2 V)u - 2(\partial_x V)\partial_x u$$

implies that  $[V, \partial_x^2] = -(\partial_x^2 V) - 2(\partial_x V)\partial_x$ . This algebra necessitates knowing derivatives of the interaction potential, which are assumed for the scope of this chapter to be given exactly but in practice can be obtained via differentiation matrices. The derivatives of  $V$  appearing in our splitting need to be known only to a certain accuracy, and finite difference methods of fairly reasonable orders suffice since  $V$  is assumed to be independent of  $\varepsilon$  and not highly oscillatory. It must be noted that these derivatives, if not given exactly, need be derived only once and the overhead is bearable.

In Table 4.1, we list all nested commutators of  $V$  and  $\partial_x^2$  up to grade 7 that appear in the sBCH formula. Here the ‘grade’ of a commutator refers to the number of ‘letters’  $V$  and  $\partial_x^2$  in it, while  $\chi_j$  is the coefficient of this commutator in the symmetric BCH formula, (2.33). There are a number of ways of forming such a basis, e.g. the Hall basis, Lyndon basis and Dynkin basis (Reutenauer 1993): here we will use the most popular one, the Hall basis (Murua 2010).

As it turns out, the 41 commutators listed in Table 4.1 are the ones we would need to solve when seeking a splitting with  $\mathcal{O}(\varepsilon^{9\sigma-1})$  error. For our purposes, where the highest order splitting discussed will feature an  $\mathcal{O}(\varepsilon^{7\sigma-1})$  error, much fewer commutators will suffice (upto grade 5).

**Table 4.1:** The terms of the Hall basis of  $\mathfrak{g}$  of grade  $\leq 7$ .

$j$	Nested commutator	$\chi_j$	grade
$S_1$	$\partial_x^2$	1	1
$S_2$	$V$	1	1
$S_3$	$[V, \partial_x^2]$	0	2
$S_4$	$[[V, \partial_x^2], \partial_x^2]$	$-\frac{1}{24}$	3
$S_5$	$[[V, \partial_x^2], V]$	$-\frac{1}{12}$	3
$S_6$	$[[[V, \partial_x^2], \partial_x^2], \partial_x^2]$	0	4
$S_7$	$[[[V, \partial_x^2], \partial_x^2], V]$	0	4
$S_8$	$[[[V, \partial_x^2], V], V]$	0	4
$S_9$	$[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], \partial_x^2]$	$\frac{7}{5760}$	5
$S_{10}$	$[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], V]$	$\frac{7}{1440}$	5
$S_{11}$	$[[[[V, \partial_x^2], \partial_x^2], V], V]$	$\frac{1}{180}$	5
$S_{12}$	$[[[[V, \partial_x^2], V], V], V]$	$\frac{1}{720}$	5
$S_{13}$	$[[[V, \partial_x^2], \partial_x^2], [V, \partial_x^2]]$	$\frac{1}{480}$	5
$S_{14}$	$[[[V, \partial_x^2], V], [V, \partial_x^2]]$	$-\frac{1}{360}$	5
$S_{15}$	$[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], \partial_x^2], \partial_x^2]$	0	6
$S_{16}$	$[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], \partial_x^2], V]$	0	6
$S_{17}$	$[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], V], V]$	0	6
$S_{18}$	$[[[[[V, \partial_x^2], \partial_x^2], V], V], V]$	0	6
$S_{19}$	$[[[[[V, \partial_x^2], V], V], V], V]$	0	6
$S_{20}$	$[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], [V, \partial_x^2]]$	0	6
$S_{21}$	$[[[[[V, \partial_x^2], \partial_x^2], V], [V, \partial_x^2]]$	0	6
$S_{22}$	$[[[[[V, \partial_x^2], V], V], [V, \partial_x^2]]$	0	6
$S_{23}$	$[[[[V, \partial_x^2], V], [[V, \partial_x^2], \partial_x^2]]$	0	6
$S_{24}$	$[[[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], \partial_x^2], \partial_x^2], \partial_x^2]$	$-\frac{31}{967680}$	7
$S_{25}$	$[[[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], \partial_x^2], \partial_x^2], V]$	$-\frac{31}{161280}$	7
$S_{26}$	$[[[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], \partial_x^2], V], V]$	$-\frac{13}{30240}$	7
$S_{27}$	$[[[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], V], V], V]$	$-\frac{53}{120960}$	7
$S_{28}$	$[[[[[[[V, \partial_x^2], \partial_x^2], V], V], V], V]$	$-\frac{1}{5040}$	7
$S_{29}$	$[[[[[[[V, \partial_x^2], V], V], V], V], V]$	$-\frac{1}{30240}$	7
$S_{30}$	$[[[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], \partial_x^2], [V, \partial_x^2]]$	$-\frac{53}{161280}$	7
$S_{31}$	$[[[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], V], [V, \partial_x^2]]$	$-\frac{11}{12096}$	7
$S_{32}$	$[[[[[[[V, \partial_x^2], \partial_x^2], V], V], [V, \partial_x^2]]$	$-\frac{3}{4480}$	7

Table 4.1 (contd)

$S_{33}$	$[[[[[V, \partial_x^2], V], V], V], [V, \partial_x^2]]$	$-\frac{1}{10080}$	7
$S_{34}$	$[[[[[V, \partial_x^2], \partial_x^2], [V, \partial_x^2]], [V, \partial_x^2]]$	$-\frac{1}{4032}$	7
$S_{35}$	$[[[[[V, \partial_x^2], V], [V, \partial_x^2]], [V, \partial_x^2]]$	$-\frac{1}{6720}$	7
$S_{36}$	$[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], [[V, \partial_x^2], \partial_x^2]]$	$-\frac{19}{80640}$	7
$S_{37}$	$[[[[[V, \partial_x^2], \partial_x^2], V], [[V, \partial_x^2], \partial_x^2]]$	$-\frac{1}{10080}$	7
$S_{38}$	$[[[[[V, \partial_x^2], V], V], [[V, \partial_x^2], \partial_x^2]]$	$-\frac{17}{40320}$	7
$S_{39}$	$[[[[[V, \partial_x^2], \partial_x^2], \partial_x^2], [[V, \partial_x^2], V]]$	$-\frac{53}{60480}$	7
$S_{40}$	$[[[[[V, \partial_x^2], \partial_x^2], V], [[V, \partial_x^2], V]]$	$-\frac{19}{13440}$	7
$S_{41}$	$[[[[[V, \partial_x^2], V], V], [[V, \partial_x^2], V]]$	$-\frac{1}{5040}$	7

Computing the first few commutators explicitly, we find

$$\begin{aligned}
S_3 &= -(\partial_x^2 V) - 2(\partial_x V)\partial_x, \\
S_4 &= (\partial_x^4 V) + 4(\partial_x^3 V)\partial_x + 4(\partial_x^2 V)\partial_x^2, \\
S_5 &= -2(\partial_x V)^2, \\
S_6 &= -(\partial_x^6 V) - 6(\partial_x^5 V)\partial_x - 12(\partial_x^4 V)\partial_x^2 - 8(\partial_x^3 V)\partial_x^3, \\
S_7 &= 4[(\partial_x V)(\partial_x^3 V) + (\partial_x^2 V)^2] + 8(\partial_x V)(\partial_x^2 V)\partial_x, \\
S_8 &= 0.
\end{aligned}$$

#### 4.2.2 The Lie algebra $\mathfrak{S}$ and reduction of height

We note that the terms  $S_1, \dots, S_8$  all belong to the set

$$\mathfrak{S} = \left\{ \sum_{k=0}^n f_k(x) \partial_x^k : n \in \mathbb{Z}_+, f_0, \dots, f_n \in C_p^\infty([-1, 1], \mathbb{R}) \right\}.$$

It is trivial to observe that  $\mathfrak{S}$  is itself a Lie algebra. Since  $\partial_x^2$  and  $V$  reside in this Lie algebra, so does the Lie algebra they generate,

$$\mathfrak{g} = \text{LA}\{V, \partial_x^2\} \subseteq \mathfrak{S}.$$

There are numerous cancellations, similar to  $S_8 = 0$ , because of the special structure induced by the letters  $\partial_x^2$  and  $V(x)$ . Nevertheless, for our exposition it is more appropriate to directly operate in the larger Lie-algebra  $\mathfrak{S}$ , where all cancellations will be taken care

of by simple computation of the commutators according to

$$\left[ \sum_{i=0}^n f_i(x) \partial_x^i, \sum_{j=0}^m g_j(x) \partial_x^j \right] = \sum_{i=0}^n \sum_{j=0}^m \sum_{\ell=0}^i \binom{i}{\ell} f_i(x) \left( \partial_x^{i-\ell} g_j(x) \right) \partial_x^{\ell+j} - \sum_{j=0}^m \sum_{i=0}^n \sum_{\ell=0}^j \binom{j}{\ell} g_j(x) \left( \partial_x^{j-\ell} f_i(x) \right) \partial_x^{\ell+i}. \quad (4.7)$$

Once we begin to simplify commutators of terms in the Lie algebra  $\mathfrak{S}$  using the above simplification procedure, we encounter systematic reduction in the degree of differential operators, which is summarised in the observation of *height reduction*. This observation is what justifies working directly in the undiscretised operators—it leads to a reduction in the spectral radius of terms, eventually allowing an effective asymptotic splitting.

**Definition 4.2.1.** *The height of a term in the Lie algebra  $\mathfrak{S}$  is defined as the degree of the highest-degree differential operator occurring in the term,*

$$\text{ht} \left( \sum_{k=0}^n f_k(x) \partial_x^k \right) = n.$$

We note that the height of  $(\partial_x^3 V) \partial_x + V \partial_x^2$ , defined in this way, is two, not three, since  $\partial_x^3$  is not acting as an operator but instead describes the third derivative of  $V$ . The term  $0 \in \mathfrak{S}$  is assigned a height of  $-1$ , making it the only term with negative height.

We extend the notion of height to commutators in the Lie algebra  $\mathfrak{g}$  by defining it as the height of the term in  $\mathfrak{S}$  to which the commutator reduces upon applying the simplification rule (4.7). Additionally, we may extend the notion of height to other algebras including the algebra generated by  $\partial_x^2$  and  $V$  under operatorial composition  $\circ$  and the Jordan algebra generated by them (which is the algebra generated by them under the Jordan product,  $A \bullet B = \frac{1}{2}(A \circ B + B \circ A)$ ).

**Lemma 4.2.2** (Height reduction in  $\mathfrak{S}$ ). *For all  $C_1, C_2 \in \mathfrak{S}$ , such that  $C_1, C_2 \neq 0$ ,*

$$\text{ht}([C_1, C_2]) \leq \text{ht}(C_1) + \text{ht}(C_2) - 1. \quad (4.8)$$

*Proof.* Let  $\text{ht}(C_1) = n$  and  $\text{ht}(C_2) = m$ , so that for some  $f_i$  and  $g_j$ ,  $C_1 = \sum_{i=0}^n f_i(x) \partial_x^i$  and  $C_2 = \sum_{j=0}^m g_j(x) \partial_x^j$ . The commutator  $[C_1, C_2]$  is then described exactly by (4.7), where the terms featuring  $\partial_x^{n+m}$  in the two summations cancel out. The highest degree differential operator, therefore, has a degree not exceeding  $n + m - 1$ .  $\square$

We note that in general the terms corresponding to  $\partial_x^{n+m-1}$  do not cancel out and, with the exception of some special cases, (4.8) holds as an equality.

The observation of height reduction allows us to study a systematic decrease in the highest degree differential operator occurring in a term. Upon eventual spatial discretisa-

tion the differential operators occurring in each expression are replaced by matrices, the spectral radius of which is largely governed by the highest degree differential operator. The notion of height provides a direct estimate of the spectral radius in this way. For instance, the term with height  $n$ ,

$$\text{ht} \left( \sum_{k=0}^n f_k(x) \partial_x^k \right) = n,$$

is  $\mathcal{O}(\varepsilon^{-n})$  in size (provided  $f_k = \mathcal{O}(\varepsilon^0)$ ) since  $\partial_x^k = \mathcal{O}(\varepsilon^{-k})$ . To be more precise, once such a term is discretised, it yields a matrix whose spectral radius is  $\mathcal{O}(\varepsilon^{-n})$ ,

$$\rho \left( \sum_{k=0}^n \mathcal{D}_{f_k} \mathcal{K}^k \right) = \mathcal{O}(\varepsilon^{-n}).$$

In particular, due to height reduction,

$$\text{ht}([V, \partial_x^2], \partial_x^2) = 2, \quad \text{ht}([V, \partial_x^2], V) = 0.$$

Consequently, once these commutators are simplified in  $\mathfrak{S}$ ,

$$[[V, \partial_x^2], \partial_x^2] = (\partial_x^4 V) + 4(\partial_x^3 V) \partial_x + 4(\partial_x^2 V) \partial_x^2, \quad [[V, \partial_x^2], V] = -2(\partial_x V)^2,$$

and subsequently discretised,

$$[[V, \partial_x^2], \partial_x^2] \rightsquigarrow 4\mathcal{D}_{\partial_x^2 V} \mathcal{K}^2 + 4\mathcal{D}_{\partial_x^3 V} \mathcal{K} + \mathcal{D}_{\partial_x^4 V}, \quad [[V, \partial_x^2], V] \rightsquigarrow -2\mathcal{D}_{(\partial_x V)^2},$$

they result in matrices that are  $\mathcal{O}(\varepsilon^{-2})$  and  $\mathcal{O}(\varepsilon^0)$  in size, respectively.

Due to the spectral accuracy of our discretisation, the direct discretisations of these commutators,  $[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{K}^2]$  and  $[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{D}_V]$ , are not very far away from the discretisations of the simplified commutators either (see Figure 4.1, for instance). Therefore, these commutators are also  $\mathcal{O}(\varepsilon^{-2})$  and  $\mathcal{O}(\varepsilon^0)$ , respectively, not  $\mathcal{O}(\varepsilon^{-4})$  and  $\mathcal{O}(\varepsilon^{-2})$  as might be estimated by using standard bounds for commutators.

In Section 3.3.4, we analysed the error for the Strang splitting by considering the size of nested commutators of the form  $\varepsilon h^3 [[\mathcal{D}_V, \mathcal{K}^2], \mathcal{K}^2]$  and  $\varepsilon^{-1} h^3 [[\mathcal{D}_V, \mathcal{K}^2], \mathcal{D}_V]$  that are discarded from the sBCH. Using our latest estimates, however, these turn out to be  $\mathcal{O}(h^3 \varepsilon^{-1})$ , not  $\mathcal{O}(h^3 \varepsilon^{-3})$  as pessimistically estimated previously. Similar terms appear when using the symmetrised Zassenhaus algorithm of Section 4.1.1 to split the discretised Hamiltonian. As a consequence of the reduction of the negative powers of  $\varepsilon$ , it turns out that we do not need  $\sigma = 2$  for achieving asymptotic splittings of the kind (4.1) and  $\sigma = 1$  will, in fact, suffice. These observations are presented in a more generalised form in Chapter 6 and will apply to Yoshida splittings and Magnus expansions in addition to

symmetric Zassenhaus splittings.

### 4.2.3 The first stage of Zassenhaus splitting

Having established the techniques for solving commutators and noting the correct scaling laws, we commence upon the first stage of the symmetric Zassenhaus splitting. We aim to devise an asymptotic splitting scheme of the form (4.1) (which results from (4.6) under the scaling  $\sigma = 1$ ) with  $s = 2$ , i.e. bearing the error of  $\mathcal{O}(\varepsilon^6)$ . Given that  $\varepsilon > 0$  is very small, this presents a method which is very accurate—arguably, of higher accuracy than required in standard numerical computations.

We will expand the commutators in powers of  $\varepsilon$  and successively remove them from the core of our expansion, aiming for  $W^{[k]} = \mathcal{O}(\varepsilon^{2k-2})$  except for  $W^{[0]}$  which will be  $\mathcal{O}(\varepsilon^0)$ . Note that  $ih\varepsilon\partial_x^2 = \mathcal{O}(\varepsilon^0)$  and  $-ih\varepsilon^{-1}V = \mathcal{O}(\varepsilon^0)$  under  $h = \mathcal{O}(\varepsilon)$ , or more generally

$$h^\ell \varepsilon^m f(x) \partial_x^n = \mathcal{O}(\varepsilon^{\ell+m-n}), \quad \varepsilon \rightarrow 0. \quad (4.9)$$

We can now commence the algorithm (4.5), setting

$$\mathcal{W}^{[0]} = ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V, \quad W^{[0]} = -ih\varepsilon^{-1}V.$$

With the help of (2.33), we compute the commutators in  $\mathcal{W}^{[1]} = \text{sBCH}(-W^{[0]}, \mathcal{W}^{[0]})$  according to (4.7). This task faces us with long and tedious algebra, but can, however be automatised with a computer algebra program. It is worth pointing out, that all simplifications, such as  $[[[V, \partial_x^2], V], V] = 0$  are automatically performed once we work in the larger Lie algebra  $\mathfrak{S}$  with differential operators and scalar functions. Likewise, there is no need for a tedious representation of expansion elements in, say the Hall basis.

Substituting  $W^{[0]}$  and  $\mathcal{W}^{[0]}$ , and aggregating terms of the same order of magnitude, we obtain

$$\begin{aligned} \mathcal{W}^{[1]} = & \overbrace{ih\varepsilon\partial_x^2}^{\mathcal{O}(\varepsilon^0)} - \overbrace{\frac{1}{12}ih^3\varepsilon^{-1}(\partial_x V)^2 - \frac{1}{3}ih^3\varepsilon(\partial_x^2 V)\partial_x^2}^{\mathcal{O}(\varepsilon^2)} - \overbrace{\frac{1}{3}ih^3\varepsilon(\partial_x^3 V)\partial_x}^{\mathcal{O}(\varepsilon^3)} \\ & - \overbrace{\frac{1}{60}ih^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 - \frac{1}{12}ih^3\varepsilon(\partial_x^4 V)}^{\mathcal{O}(\varepsilon^4)} \\ & + \overbrace{ih^5\varepsilon\{\frac{4}{45}(\partial_x^2 V)^2 - \frac{1}{90}(\partial_x^3 V)(\partial_x V)\}\partial_x^2 - \frac{1}{45}ih^5\varepsilon^{-3}(\partial_x^4 V)\partial_x^4}^{\mathcal{O}(\varepsilon^4)} \\ & + \overbrace{ih^5\varepsilon\{\frac{1}{6}(\partial_x^3 V)(\partial_x^2 V) - \frac{1}{90}(\partial_x^4 V)(\partial_x V)\}\partial_x}^{\mathcal{O}(\varepsilon^5)} \\ & - \overbrace{\frac{2}{45}ih^5\varepsilon^{-3}(\partial_x^5 V)\partial_x^3}^{\mathcal{O}(\varepsilon^5)} + \mathcal{O}(\varepsilon^6). \end{aligned} \quad (4.10)$$

### 4.3 Loss of stability

Unfortunately, we run into a problem at the very first stage: (4.10) contains terms of size  $\mathcal{O}(\varepsilon^3)$  and  $\mathcal{O}(\varepsilon^5)$  that are both due to the presence of odd powers of  $\partial_x$ . This presence is worrisome for an important reason, namely *stability*. Both  $\partial_x^2$  and multiplication by  $V$  are Hermitian operators, therefore  $i\hbar(\varepsilon\partial_x^2 - \varepsilon^{-1}V)$  is a skew-Hermitian operator: its exponential is thus unitary (see Chapter 3). This survives under eventual discretisation, because any reasonable approximation of  $\partial_x^2$  preserves Hermitian structure. However,  $\partial_x$  (and, in general, odd powers of  $\partial_x$ ) is a skew-Hermitian operator, hence  $i\partial_x$  is Hermitian and so are its reasonable approximations.

The introduction of odd powers of  $\partial_x$  seems fraught with loss of unitarity and stability. However, even the terms  $-\frac{1}{3}i\hbar^3\varepsilon(\partial_x^2V)\partial_x^2$  and  $-\frac{1}{45}i\hbar^5\varepsilon^{-3}(\partial_x^4V)\partial_x^4$  that do feature even derivatives seem to be cause for concern—they are neither Hermitian nor skew-Hermitian. We should clarify that, having arisen from nested commutators of skew-Hermitian operators, the operator  $\mathcal{W}^{[1]}$  ought to be skew-Hermitian (since the linear space of skew-Hermitian operators is a Lie algebra), and it indeed is! The trouble is two-fold.

Firstly, when such an operator is discretised in the usual way, the occurrences of the differential operator are replaced by corresponding differentiation matrices. While any reasonable discretisation strategy does maintain skew-Hermiticity of  $\partial_x$  under discretisation, it does not preserve the skew-Hermiticity of all skew-Hermitian operators such as  $\mathcal{W}^{[1]}$ .

This problem is easily demonstrated by using a simple example. In Section 4.2.1 we commence by noting

$$[\partial_x^2, V] = 2(\partial_x V)\partial_x + (\partial_x^2 V).$$

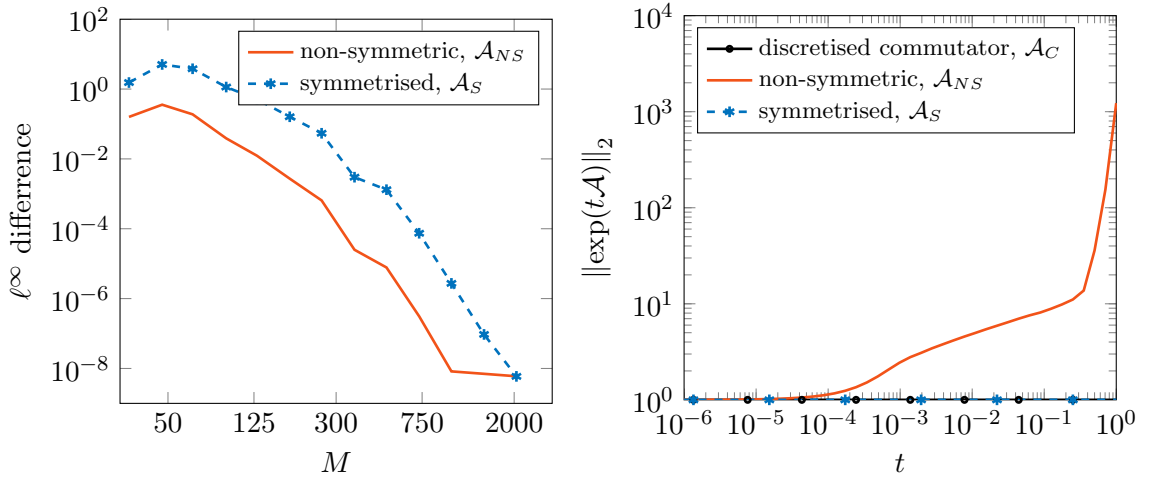
When the differential operators are replaced by differentiation matrices, these two forms, which are equivalent in the undiscretised case, lead to two different matrices,

$$\mathcal{A}_C = [\mathcal{K}^2, \mathcal{D}_V], \quad \mathcal{A}_{NS} = 2\mathcal{D}_{\partial_x V}\mathcal{K} + \mathcal{D}_{\partial_x^2 V}.$$

Here  $\mathcal{A}_C$  is the form where the commutator is directly discretised and  $\mathcal{A}_{NS}$  is the non-symmetric form where the Lie bracket has first been simplified using the rules of Section 4.2.1. In Figure 4.1, we find that, while the skew-Hermiticity of  $[\partial_x^2, V]$  is preserved when discretised directly to  $\mathcal{A}_C$ , the non-symmetric form  $\mathcal{A}_{NS}$  does not preserve this crucial feature and consequently its exponential blows up. This is despite the fact that the action of  $\mathcal{A}_C$  and  $\mathcal{A}_{NS}$  when applied to the wave-function  $u = \exp(\sin(\pi x))$  is not all that different.

As it turns out, the strategy for replacing odd derivatives pursued in the next section





**Figure 4.1:**  $\ell^\infty$  difference in the computation of  $\mathcal{A}u$  as compared to  $\mathcal{A}_C u$  for different grid resolutions, where  $u = \exp(\sin(\pi x))$  and  $V(x) = \text{expbump}_p(x)$ . This shows that both  $\mathcal{A}_{NS}$  and  $\mathcal{A}_S$  are accurate approximations and, in fact,  $\mathcal{A}_{NS}$  is somewhat better (*left*). The exponential  $\exp(t\mathcal{A}_{NS})$  blows up due to the loss of skew-Hermiticity, leading to loss of stability (*right*).

suffices for restoring skew-Hermiticity. The theoretical reasons for the success of this strategy, however, are deeper and are pursued in a much more abstract way in Chapter 5. The symmetrised form obtained by following this strategy,

$$[\partial_x^2, V] = (\partial_x V)\partial_x + \partial_x \circ (\partial_x V),$$

results in a symmetrised discretisation,

$$\mathcal{A}_S = \mathcal{D}_{\partial_x V} \mathcal{K} + \mathcal{K} \mathcal{D}_{\partial_x V},$$

which is skew-Hermitian and whose exponential doesn't blow up.

Secondly, in the Zassenhaus splitting algorithm we will discard at each stage terms that are smaller than our error tolerance ( $\mathcal{O}(\varepsilon^6)$  in the example considered here). Discarding certain terms that might otherwise be small, however, can again lead to loss of skew-Hermiticity. In light of this, we would require a more reliable way of knowing when a term can be discarded.

The approach pursued in the following section is a way to overcome this hurdle of instability that threatens to derail the usability of our splitting schemes.

### 4.3.1 Restoring skew-Hermiticity

Let  $f$  be a  $C^1$  function. The starting point for our current construction is the operatorial identity

$$f(x)\partial_x = -\frac{1}{2} \left[ \int_{x_0}^x f(\xi) d\xi \right] \partial_x^2 - \frac{1}{2} \partial_x f(x) + \frac{1}{2} \partial_x^2 \left[ \int_{x_0}^x f(\xi) d\xi \cdot \right], \quad (4.11)$$

where  $x_0$  is arbitrary: its direct proof is trivial. Here

$$\partial_x^k[f \cdot] = \partial_x^k \circ f,$$

and its action on  $u$  is  $\partial_x^k(fu)$ . Note that, while we have  $\partial_x$  on the left, the right-hand side features  $\partial_x^0$  and  $\partial_x^2$ , both even powers of the differentiation operator. The challenge is thus to generalise (4.11) and express  $f(x)\partial_x^{2s+1}$ ,  $s \in \mathbb{Z}_+$ , solely by means of even derivatives.

**Theorem 4.3.1.** *Let  $s \in \mathbb{Z}_+$ , define the real sequence  $\{\beta_k\}_{k \geq 0}$  by*

$$\sum_{k=0}^{\infty} \frac{(-1)^k \beta_k}{(2k+1)!} T^k = \frac{1}{T} \left( 1 - \frac{T^{1/2}}{\sinh T^{1/2}} \right)$$

and set

$$Q_k(x) = (-1)^{s-k+1} \beta_{s-k} \binom{2s+1}{2k} \partial_x^{2s-2k+1} f(x), \quad k = 0, 1, \dots, s, \quad (4.12)$$

$$Q_{s+1}(x) = \frac{1}{2s+2} \int_{x_0}^x f(\xi) d\xi, \quad (4.13)$$

$$P_k(x) = - \sum_{\ell=k}^{s+1} \binom{2\ell}{2k} \partial_x^{2\ell-2k} Q_\ell(x), \quad k = 1, 2, \dots, s+1. \quad (4.14)$$

Then

$$f(x)\partial_x^{2s+1} = \sum_{k=0}^{s+1} P_k(x)\partial_x^{2k} + \sum_{k=0}^{s+1} \partial_x^{2k} [Q_k(x) \cdot]. \quad (4.15)$$

*Proof.* We act on the second sum on the right of (4.15) with the Leibnitz rule, whereby

$$\begin{aligned} f\partial_x^{2s+1} &= \sum_{k=1}^{s+1} P_k \partial_x^{2k} + \sum_{\ell=0}^{s+1} \sum_{k=0}^{2\ell} \binom{2\ell}{k} (\partial_x^{2\ell-k} Q_\ell) \partial_x^k \\ &= \sum_{k=1}^{s+1} P_k \partial_x^{2k} + \sum_{k=0}^{s+1} \left[ \sum_{\ell=k}^{s+1} \binom{2\ell}{2k} (\partial_x^{2(\ell-k)} Q_\ell) \right] \partial_x^{2k} \\ &\quad + \sum_{k=0}^s \left[ \sum_{\ell=k+1}^{s+1} \binom{2\ell}{2k+1} (\partial_x^{2(\ell-k)-1} Q_\ell) \right] \partial_x^{2k+1}. \end{aligned}$$

Equating powers of  $\partial_x$  on both sides, we obtain (4.13), (4.14) and the equations

$$\sum_{\ell=k+1}^{s+1} \binom{2\ell}{2k+1} \partial_x^{2(\ell-k)-1} Q_\ell = 0, \quad k = s-1, s-2, \dots, 0. \quad (4.16)$$

Our contention is that there exist coefficients  $\{\beta_k\}_{k \geq 0}$  such that (4.12) is true. Indeed, substituting (4.12) in (4.16) yields, after simple algebra, the triangular linear system

$$\sum_{\ell=k+1}^s (-1)^{s-\ell} \binom{2s-2k}{2s+1-2\ell} \beta_{s-\ell} = \frac{1}{2s-2k+1}, \quad k = 0, 1, \dots, s-1.$$

We deduce that

$$\sum_{\ell=0}^{k-1} (-1)^\ell \binom{2k}{2\ell+1} \beta_\ell = \frac{1}{2k+1}, \quad k \in \mathbb{N}.$$

Finally, we multiply the last equation by  $T^{k-1}/(2k)!$  and sum up for  $k \in \mathbb{N}$ . On the left we have

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{1}{(2k)!} \sum_{\ell=0}^{k-1} (-1)^\ell \binom{2k}{2\ell+1} \beta_\ell T^{k-1} &= \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \beta_\ell}{(2\ell+1)!} \sum_{k=\ell+1}^{\infty} \frac{T^{k-1}}{(2k-2\ell-1)!} \\ &= \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \beta_\ell}{(2\ell+1)!} T^\ell \sum_{k=0}^{\infty} \frac{T^k}{(2k+1)!} \\ &= \frac{\sinh T^{1/2}}{T^{1/2}} \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \beta_\ell}{(2\ell+1)!} T^\ell, \end{aligned}$$

while on the right we obtain

$$\sum_{k=1}^{\infty} \frac{T^{k-1}}{(2k+1)!} = \frac{1}{T} \left( \frac{\sinh T^{1/2}}{T^{1/2}} - 1 \right).$$

This confirms (4.12) and completes the proof.  $\square$

First few values are  $\beta_0 = \frac{1}{6}$ ,  $\beta_1 = \frac{7}{60}$ ,  $\beta_2 = \frac{31}{126}$ ,  $\beta_3 = \frac{127}{120}$ ,  $\beta_4 = \frac{511}{66}$ ,  $\beta_5 = \frac{1414477}{16380}$  and  $\beta_6 = \frac{8191}{6}$ . Since

$$\frac{te^{xt}}{(e^t - 1)} = \sum_{k=0}^{\infty} \frac{\mathcal{B}_k(x)}{k!} t^k,$$

where  $\mathcal{B}_k$  is the  $k$ th *Bernoulli polynomial*, it is easy to confirm that

$$\beta_k = \frac{(-1)^{k+1} 2^{2k+1} \mathcal{B}_{2k+2}(\frac{1}{2})}{k+1}, \quad k \in \mathbb{Z}_+.$$

Practically, just

$$\begin{aligned}
f\partial_x &= -\frac{1}{2} \left[ \int_0^x f(\xi) d\xi \right] \partial_x^2 - \frac{1}{2} \partial_x f + \frac{1}{2} \partial_x^2 \left[ \int_0^x f(\xi) d\xi \cdot \right], \\
f\partial_x^3 &= -(\partial_x f) \partial_x^2 - \frac{1}{4} \left[ \int_0^x f(\xi) d\xi \right] \partial_x^4 + \frac{1}{4} \partial_x^3 f - \frac{1}{2} \partial_x^2 [(\partial_x f) \cdot] + \frac{1}{4} \partial_x^4 \left[ \int_0^x f(\xi) d\xi \cdot \right], \\
f\partial_x^5 &= \frac{4}{3} (\partial_x^3 f) \partial_x^2 - \frac{5}{3} (\partial_x f) \partial_x^4 - \frac{1}{6} \left[ \int_0^x f(\xi) d\xi \right] \partial_x^6 - \frac{1}{2} \partial_x^5 f + \frac{7}{6} \partial_x^2 [(\partial_x^3 f) \cdot] \\
&\quad - \frac{5}{6} \partial_x^4 [(\partial_x f) \cdot] + \frac{1}{6} \partial_x^6 \left[ \int_0^x f(\xi) d\xi \cdot \right],
\end{aligned}$$

are likely to be needed in our computations.

These replacement rules might seem concerning from the point of view of height—the operator  $f\partial_x^{2s+1}$  on the left hand side has the height  $2s+1$  but its replacement involves a term whose height is  $2s+2$ . It can be proven that the replacement rules only need to be applied to the second largest or smaller terms in any commutator appearing the sBCH, whereby the effective height in the context that concerns us is not changed. However, the proof is unwieldy, and the more abstract and rigourous approach pursued in Chapter 5 handles the concepts of height and skew-Hermiticity more elegantly, obviating the need for such a proof.

### 4.3.2 Some structural observations

Using (4.11) and its generalisations to replace all the occurrences of  $\partial_x$  and  $\partial_x^3$  in (4.10), we express  $\mathcal{W}^{[1]}$  in the form

$$\begin{aligned}
\mathcal{W}^{[1]} &= \overbrace{i\hbar\varepsilon\partial_x^2}^{\mathcal{O}(\varepsilon^0)} - \overbrace{\frac{1}{12}i\hbar^3\varepsilon^{-1}(\partial_x V)^2 - \frac{1}{6}i\hbar^3\varepsilon \{(\partial_x^2 V)\partial_x^2 + \partial_x^2[(\partial_x^2 V) \cdot]\}}^{\mathcal{O}(\varepsilon^2)} \\
&\quad - \overbrace{\frac{1}{60}i\hbar^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 + \frac{1}{12}i\hbar^3\varepsilon(\partial_x^4 V)}^{\mathcal{O}(\varepsilon^4)} \\
&\quad + \overbrace{\frac{1}{180}i\hbar^5\varepsilon \{8(\partial_x^2 V)^2\partial_x^2 + 8\partial_x^2[(\partial_x^2 V)^2 \cdot] - (\partial_x^3 V)(\partial_x V)\partial_x^2 - \partial_x^2[(\partial_x^3 V)(\partial_x V) \cdot]\}}^{\mathcal{O}(\varepsilon^4)} \\
&\quad - \overbrace{\frac{1}{90}i\hbar^5\varepsilon^{-3} \{(\partial_x^4 V)\partial_x^4 + \partial_x^4[(\partial_x^4 V) \cdot]\}}^{\mathcal{O}(\varepsilon^4)} + \mathcal{O}(\varepsilon^6).
\end{aligned}$$

Once appropriate odd and even differential operators are replaced, operators of the form  $f\partial_x^k + \partial_x^k[f \cdot]$  start appearing ubiquitously in our analysis. Far from being unique to our computation of  $\mathcal{W}^{[1]}$ , these are characteristic of the Lie algebra generated by  $\partial_x^2$  and  $V$ . Proving this observation rigourously is a different ball game that requires a fundamental investigation of the Lie algebra of the symmetrised differential operators,  $f\partial_x^k + \partial_x^k[f \cdot]$ ,

which forms the central theme of Chapter 5.

At this stage we take a break from Zassenhaus splittings to analyse these symmetric structures in more detail. The results of this investigation will prove to be indispensable for designing and analysing our splittings. In Chapter 6 we resume the development of Zassenhaus splittings in the language of the algebraic structures of Chapter 5.



## Chapter 5

# A Lie algebra of symmetrised differential operators

In the pursuit of stable commutator-free Zassenhaus splittings in Chapter 4, we encountered terms of the form  $f\partial_x^k + \partial_x^k[f \cdot]$ . As evident from Figure 4.1, symmetrised differential operators such as  $(\partial_x V)\partial_x + \partial_x[(\partial_x V) \cdot]$  are crucial for ensuring skew-Hermiticity of exponents upon discretisation, the loss of which leads to exponential blow up and numerical instability.

A linear differential operator such as

$$f_2(x)\partial_x^2 + f_1(x)\partial_x + f_0(x) = f_2(x) \circ \partial_x^2 + f_1(x) \circ \partial_x + f_0(x),$$

where  $\circ$  is operator composition, can be viewed as a polynomial in  $\partial_x$  with function coefficients  $f_0, f_1, f_2$ . The symmetrised differential operators such as

$$\mathbf{i} (f_4\partial_x^4 + \partial_x^4[f_4 \cdot]) + \mathbf{i} (f_2\partial_x^2 + \partial_x^2[f_2 \cdot]) + \mathbf{i}f_0 = \mathbf{i} (f_4 \circ \partial_x^4 + \partial_x^4 \circ f_4) + \mathbf{i} (f_2 \circ \partial_x^2 + \partial_x^2 \circ f_2) + \mathbf{i}f_0,$$

which we encountered in the pursuit of stability in Section 4.3.2 can similarly be viewed as polynomials in  $\partial_x$  with function coefficients with one difference—instead of the product operator being operator composition,  $\circ$ , these are polynomials under the Jordan product,  $\bullet$ , which is defined by

$$\mathcal{L}_1 \bullet \mathcal{L}_2 = \frac{1}{2} (\mathcal{L}_1 \circ \mathcal{L}_2 + \mathcal{L}_2 \circ \mathcal{L}_1).$$

With this notation in place, the symmetrised differential operators we encountered in Section 4.3.2 can be written in the form

$$2\mathbf{i}f_4 \bullet \partial_x^4 + 2\mathbf{i}f_2 \bullet \partial_x^2 + \mathbf{i}f_0.$$

Such an operator can be called a ‘Jordan polynomial’ (i.e. a polynomial under the Jordan product,  $\bullet$ ) in  $\partial_x$  with function coefficients  $\mathbf{i}f_0, 0, 2\mathbf{i}f_2, 0, 2\mathbf{i}f_4$ .

The symmetric form of the Jordan product naturally results in the Jordan polynomials being symmetric. Once we discretise

$$f \rightsquigarrow \mathcal{D}_f, \quad \partial_x^k \rightsquigarrow \mathcal{K}^k,$$

these symmetrised differential operators discretise to

$$2if_4 \bullet \partial_x^4 + 2if_2 \bullet \partial_x^2 + if_0 \rightsquigarrow 2i\mathcal{D}_{f_4} \bullet \mathcal{K}^4 + 2i\mathcal{D}_{f_2} \bullet \mathcal{K}^2 + i\mathcal{D}_{f_0}.$$

It is easy to verify that each

$$2i\mathcal{D}_f \bullet \mathcal{K}^{2k} = i \left( \mathcal{D}_f \mathcal{K}^{2k} + \mathcal{K}^{2k} \mathcal{D}_f \right)$$

is skew-Hermitian, and its exponential is unitary.

That the coefficients of  $\partial_x$  and  $\partial_x^3$  don't appear here is crucial—if we also had  $if_3 \bullet \partial_x^3$  in Section 4.3.2, for instance, such a term would result in a discretisation which mixes Hermitian terms (arising from discretisation of  $if_3 \bullet \partial_x^3$ ) with the skew-Hermitian terms arising from discretisation of  $2if_4 \bullet \partial_x^4 + 2if_2 \bullet \partial_x^2 + if_0$ . This would result in loss of unitarity and stability—precisely what we are trying to avoid. This ‘non-mixing property’ will be an important result of this chapter.

We find that these symmetrised differential operators constitute a Lie algebra—i.e. the commutator of two symmetrised differential operators is also a symmetrised differential operator. The characterisation of this Lie algebra forms the motivation of this chapter. In Section 5.3, we will find that this algebra possesses a  $\mathbb{Z}_2$ -graded structure that amounts to the ‘non-mixing property’ we alluded to. Along with the in-built symmetry of the symmetrised differential operators, this non-mixing property proves crucial for the unitary evolution of wave-functions as well as the numerical stability of Zassenhaus splittings.

We also note the property of height reduction in Section 5.3 which is used in Section 6.1.4 to show that commutators of differential operators are much smaller than naïvely analysed using commutator bounds. Additionally, the property of height reduction will allow us to formally prove the quadratic growth in costs for Zassenhaus splittings in Section 6.7.

In Section 4.2, we had noted that the structure of the Lie algebra generated by  $V$  and  $\partial_x^2$ ,

$$\mathfrak{g} = \text{LA}\{V, \partial_x^2\},$$

is of great interest to us since the terms in the sBCH expansion feature nested commutators of  $V$  and  $\partial_x^2$  all of which reside in  $\mathfrak{g}$  by definition. A suitable characterisation of this Lie algebra using symmetrised differential operators, therefore, can prove very helpful—it will allow us to study the nested commutators of  $V$  and  $\partial_x^2$  directly in terms of symmetrised differential operators. Instead of characterising  $\mathfrak{g}$ , however, we will attempt the



characterisation of a bigger Lie algebra,

$$\text{LA}(\mathcal{G} \cup \mathcal{P}(\partial_x)),$$

generated by all functions in the function space  $\mathcal{G}$  and all polynomials of the differential operator  $\partial_x$  with constant coefficients. This characterisation will be done in Section 5.3.1.

These algebraic characterisations, however, will be introduced in a more abstract form—in the context of an associative algebra  $\mathcal{A}$  with a commutative subalgebra  $\mathcal{C}$  and its Lie idealiser  $\mathcal{I}$ . The algebraic structures of the Zassenhaus schemes for the Schrödinger equation are seen to be special cases of the Lie algebras introduced here.

The abstract formulation pursued in this chapter should also prove of crucial importance in extending the use of Zassenhaus algorithms to other equations of quantum mechanics and possibly beyond. Some of these applications are considered briefly in Sections 10.4.2 and 10.4.3.

In Section 5.1 we introduce the abstract context in which our algebra of Jordan polynomials  $\mathfrak{G}$  is defined. In Section 5.2, we show that  $\mathfrak{G}$  is an associative algebra. While it follows immediately that  $\mathfrak{G}$  is also a Lie algebra, in Section 5.3 we use the results of Sections 5.2.1 and 5.2.2 to demonstrate that it possesses a very interesting structure, including a  $\mathbb{Z}_2$ -grading (the non-mixing property), which has ramifications for the unitarity and stability of numerical methods developed in Chapter 6. The tables of coefficients in Section 5.4 should aid direct applications of the results of this chapter.

## 5.1 Notations

Consider a commutative algebra  $\mathcal{C}$  which is a subalgebra of the unital associative algebra  $(\mathcal{A}, \cdot, +)$  over the field of rational numbers  $\mathbb{Q}$ . The commutator on the associative product,

$$[a, b] = a \cdot b - b \cdot a,$$

acts as the canonical Lie product while the anticommutator ,

$$a \bullet b = \frac{1}{2} (a \cdot b + b \cdot a),$$

acts as a Jordan product.  $\mathcal{A}$ , along with the Lie (Jordan) product, forms a Lie (Jordan) algebra which we can identify with  $\mathcal{A}$  again. This is true since for every  $a, b \in \mathcal{A}$ , the products  $a \cdot b$  and  $b \cdot a$  are in  $\mathcal{A}$  and so is their sum or difference.

In the application of the results of this chapter to the Schrödinger equation,  $\mathcal{A}$  will be the algebra of linear differential operators (with  $\cdot = \circ$  being operator composition) and  $\mathcal{C}$  will be the commutative algebra of multiplication operators. This special case forms our primary motivation for studying these algebraic structures in this thesis. For the purpose

of our numerical methods, not much will be lost, therefore, if the reader were to assume throughout that  $\mathcal{C} = C_p^\infty([-1, 1], \mathbb{R})$  and  $\mathcal{A} = (\text{End}(\mathcal{C}), \circ, +)$ —an assumption that is made more precise in Section 6.1.3.

In particular, for two linear differential operators  $\mathcal{L}_1, \mathcal{L}_2$ , the commutator is

$$[\mathcal{L}_1, \mathcal{L}_2] = \mathcal{L}_1 \circ \mathcal{L}_2 - \mathcal{L}_2 \circ \mathcal{L}_1,$$

as usual, and its action on  $u$  is

$$[\mathcal{L}_1, \mathcal{L}_2](u) = \mathcal{L}_1(\mathcal{L}_2(u)) - \mathcal{L}_2(\mathcal{L}_1(u)).$$

The Jordan product, similarly, is

$$\mathcal{L}_1 \bullet \mathcal{L}_2 = \frac{1}{2} (\mathcal{L}_1 \circ \mathcal{L}_2 + \mathcal{L}_2 \circ \mathcal{L}_1).$$

As an example, for  $\mathcal{L}_1 = f$  and  $\mathcal{L}_2 = \partial_x^k$ , we have

$$f \bullet \partial_x^k = \frac{1}{2} \left( f \circ \partial_x^k + \partial_x^k \circ f \right),$$

which can be considered a monomial in  $\partial_x$  of degree  $k$  (a polynomial with only the degree  $k$  term). We will frequently study the behaviour of polynomials by studying the behaviour of such monomials, which are simpler.

**Note:** As we have seen earlier in Section 4.3.2, such symmetrised differential operators seem to appear naturally when we try to regain stability by replacing certain differential operators using the rules of Section 4.3.1. In this chapter we are attempting to find an algebraic proof for the appearance of these symmetrised differential operators.

Since all the properties that we will prove in this chapter are due to the algebraic properties of  $\partial_x$ , they can be studied in a greater degree of abstraction. The algebraic generalisation of the derivative  $\partial_x$  that we are interested in is a ‘derivation’. A map  $\delta : \mathcal{R} \rightarrow \mathcal{R}$  is called a derivation on the algebra  $\mathcal{R}$  if

$$\delta(a \cdot b) = \delta(a) \cdot b + a \cdot \delta(b), \quad \forall a, b \in \mathcal{R}.$$

Essentially, a derivation is a map that follows the Leibniz rule (chain rule) and therefore behaves like differentiation, algebraically speaking. The differential operator  $\partial_x$  is a derivation on  $C_p^\infty$ , for instance.

We will also need to exploit the observation that the commutator of  $\partial_x$  with any function  $f \in C_p^\infty$ ,

$$[\partial_x, f] = (\partial_x f),$$

is again a function in the same space—in fact, it is the derivative of  $f$ . We seek a generalisation of  $\partial_x$  to an element, say,  $d$  such that

$$[d, f] = \text{ad}_d(f) = D(f),$$

where  $D = \text{ad}_d$  is a derivation on the function space ( $d$ , on the other hand, need not be a derivation). It is crucial, of course, that  $D(f)$  is in the function space. Any such pair  $\{d, D\}$  will do for the purpose of our generalisation.

**Note:** The Lie idealiser of  $\mathcal{C}$  in  $\mathcal{A}$ ,

$$\mathcal{I} = \{d \in \mathcal{A} : [d, \mathcal{C}] \subseteq \mathcal{C}\}.$$

is the largest subalgebra of  $\mathcal{A}$  in which  $\mathcal{C}$  is a Lie ideal. Essentially, therefore, any element of the Lie idealiser of  $\mathcal{C}$  will suffice for our purposes as a suitable choice of  $d$ .

For  $d \in \mathcal{C}$ , however,  $D = 0$ . Such a case will be considered trivial and will not be of much interest.

With this generalisation of  $\partial_x$  in place, we generalise the Jordan monomials

$$f \bullet \partial_x^k = \frac{1}{2} \left( f \circ \partial_x^k + \partial_x^k \circ f \right)$$

in the following definition.

**Definition 5.1.1** (Jordan monomials in  $d$ ). *For  $x \in \mathcal{C}$ ,  $d \in \mathcal{I}$  and a non-negative integer  $k$ , we introduce the notation*

$$\langle x \rangle_k^d = x \bullet d^k = \frac{1}{2} (x \cdot d^k + d^k \cdot x),$$

*for the degree  $k$  Jordan monomial in  $d$  with coefficient  $x \in \mathcal{C}$ .*

**Note:** The power  $d^k$  is defined unambiguously as

$$d^{k+1} = d \bullet d^k = d \cdot d^k, \quad d^0 = 1_{\mathcal{A}}.$$

Once we have defined monomials, the definition of polynomials follows trivially. The space of (Jordan) monomials of degree  $k$  will be denoted as  $\mathfrak{J}_k^d$  while the space of all (Jordan) polynomials in  $d$  will be denoted as  $\mathfrak{G}^d$ . These are formalised in the following definition.

**Definition 5.1.2** (Jordan polynomials in  $d$ ). *For any  $d \in \mathcal{I}$  and non-negative integer  $k$ , we denote the linear space of all degree  $k$  Jordan monomials in  $d$  with coefficients in  $\mathcal{C}$  as*

$$\mathfrak{F}_k^d = \left\{ \langle x \rangle_k^d : x \in \mathcal{C} \right\},$$

obtaining the linear space of all Jordan polynomials in  $d$  with coefficients in  $\mathcal{C}$  as a direct sum of these linear spaces,

$$\mathfrak{G}^d = \bigoplus_{k \in \mathbb{Z}_+} \mathfrak{F}_k^d.$$

In this chapter we will only use a single non-trivial element of  $\mathcal{J}$  which is not in  $\mathcal{C}$ ,  $d \in \mathcal{J} \setminus \mathcal{C}$ , and can drop the superscript  $d$ , writing  $\langle x \rangle_k$ ,  $\mathfrak{F}_k$  and  $\mathfrak{G}$  in place of  $\langle x \rangle_k^d$ ,  $\mathfrak{F}_k^d$  and  $\mathfrak{G}^d$ , respectively.

Since  $D = \text{ad}_d$  is a derivation on  $\mathcal{A}$  (and  $\mathcal{C}$ ),  $D(a \cdot b) = Da \cdot b + a \cdot Db$ , and it distributes binomially on  $\mathcal{A}$  (and  $\mathcal{C}$ ),

$$D^k(a \cdot b) = \sum_{i=0}^k \binom{k}{i} D^i a \cdot D^{k-i} b. \quad (5.1)$$

Starting from  $d \cdot x = Dx + x \cdot d$ , a simple inductive procedure leads to a similar binomial identity,

$$d^k \cdot x = \sum_{i=0}^k \binom{k}{i} D^i x \cdot d^{k-i}. \quad (5.2)$$

These are trivial generalisations of the consequences of the Leibniz that are immediately obvious and very well known for the case  $d = \partial_x$ .

**Note:** In this chapter we typically use  $a, b, c$  for elements of the associative algebra  $\mathcal{A}$ ,  $x, y, z$  for elements of the commutative subalgebra  $\mathcal{C}$  and reserve  $d$  for elements of the Lie idealiser  $\mathcal{J}$ , except where noted otherwise. For elements  $x, y$  of the commutative algebra  $\mathcal{C}$ , we write  $xy$  to denote  $x \cdot y$ , dropping the explicit use of the multiplication operator. The letters  $k, l, m, n, p, r, s, i$  are used for non-negative integers.

**Note:** For all practical purposes, a reader who feels more comfortable with concrete examples might consider  $d = \partial_x$  and  $D = \partial_x$  for the rest of this chapter, considering elements of the algebra  $\mathcal{A}$  (usually denoted  $a, b, c$ ) as linear differential operators and elements of the commutative algebra  $\mathcal{C}$  (usually denoted  $x, y, z$ ) as functions (or multiplicative operators).

## 5.2 Associative algebra of Jordan polynomials

We find that when we compose two symmetrised differential operators such as  $f_1 \bullet \partial_x + f_0$  and  $g_2 \bullet \partial_x^2 + g_1 \bullet \partial_x + g_0$ , using operator composition,  $\circ$ , we can express the result as another symmetrised differential operator (of degree three, in this case). Note carefully

that the product used for defining the polynomials,  $\bullet$ , is different from the product,  $\circ$ , we are applying for multiplying them.

This result is summarised in Lemma 5.2.1 under the general case where we use  $d$  and  $D$  instead of  $\partial_x$ , as usual, and recalling that terms like  $f_1 \bullet \partial_x + f_0$  are written in the form  $\langle f_1 \rangle_1 + \langle f_0 \rangle_0$ , following the notation introduced earlier. The result follows by straightforward application of the Leibniz rule, followed by a recursive procedure for finding the hypothesised terms. Thus the proof is constructive in nature.

**Lemma 5.2.1.** *The linear space of Jordan polynomials,  $\mathfrak{G}$ , is an associative algebra under the associative product ‘ $\cdot$ ’,*

$$\langle x \rangle_k \cdot \langle y \rangle_l = \sum_{n=0}^{k+l} \langle z_n \rangle_{k+l-n}, \quad (5.3)$$

where the terms

$$z_n = \sum_{i=0}^n \pi_{n,i}^{k,l} D^i x D^{n-i} y \quad (5.4)$$

are in  $\mathcal{C}$  and  $\pi_{n,i}^{k,l} \in \mathbb{Q}$ .

In other words, when we multiply two Jordan monomials using the product  $\cdot$ , we get a Jordan polynomial whose degree is equal to the sum of the degrees of the monomials. Thus, it follows trivially, that when we multiply two polynomials, we again get a polynomial whose degree is equal to the sum of the degrees of polynomials. Once we know the coefficients  $\pi$ , which can be computed using the recursive procedure (5.11), we have a simple procedure by which we could multiply such polynomials.

Knowing how to multiply these polynomials is important since we eventually want to be able to solve the commutators of the symmetrised differential operators. This naturally involves being able to multiply two such operators.

**Note:** It is interesting to see that monomials of all degrees from 0 to  $k + l$  appear in the polynomial formed as a product of the two monomials. Thus, even if we were to take the monomials  $\langle f \rangle_1$  and  $\langle g \rangle_1$ , which are both skew-Hermitian operators in the context of our application ( $d = \partial_x$ ), their product will consist of a skew-Hermitian term of the form  $\langle h_1 \rangle_1$  but will also have the Hermitian terms  $\langle h_2 \rangle_2$  and  $\langle h_0 \rangle_0$ . This is not surprising—the product of two skew-Hermitian operators might end up being neither Hermitian nor skew-Hermitian.

However, when we compute commutators, it would become important that terms such as  $\langle h_2 \rangle_2$  and  $\langle h_0 \rangle_0$  don’t appear since the commutator of two skew-Hermitian operators must be skew-Hermitian. Moreover, the main motivation from a computational point of view is that the resulting term should readily lead to a discretisation that is also skew-Hermitian.

These results, which are highly desirable, will be sought in Section 5.3 in Theorem 5.3.2. The non-mixing property, which ensures that we don't mix Hermitian and skew-Hermitian terms was alluded to in the introduction of this chapter, and is proven by showing that the coefficients  $\pi_{n,i}^{k,l}$  possess a nice symmetry result in Lemma 5.3.1. In order to do so, however, we first try to find a more explicit form of the coefficients in Section 5.2.1 than evident from the recursive procedure (5.11) in the proof that follows, and by finding the generating function of these coefficients in Section 5.2.2.

*Proof.* A proof for Lemma 5.2.1 follows by expanding (5.3) using the binomial identity (5.2) and comparing powers of  $d$ . At first, we expand the left hand side,

$$\begin{aligned}
4 \langle x \rangle_k \cdot \langle y \rangle_l &= (x \cdot d^k + d^k \cdot x) \cdot 2 \langle y \rangle_l \\
&= x \cdot (d^k \cdot y) \cdot d^l + x \cdot (d^{k+l} \cdot y) + (d^k \cdot x) \cdot 2 \langle y \rangle_l \\
&= x \cdot \left( \sum_{i=0}^k \binom{k}{i} D^i y \cdot d^{k-i} \right) \cdot d^l + x \cdot \sum_{i=0}^{k+l} \binom{k+l}{i} D^i y \cdot d^{k+l-i} \\
&\quad + \left( \sum_{i=0}^k \binom{k}{i} D^i x \cdot d^{k-i} \right) \cdot (y \cdot d^l + d^l \cdot y) \\
&= \sum_{i=0}^{k+l} \left[ \binom{k}{i} + \binom{k+l}{i} \right] (x \cdot D^i y) \cdot d^{k+l-i} \\
&\quad + \sum_{i=0}^k \sum_{j=0}^{k+l-i} \binom{k}{i} \left[ \binom{k-i}{j} + \binom{k+l-i}{j} \right] (D^i x \cdot D^j y) \cdot d^{k+l-i-j}.
\end{aligned}$$

The right hand side of (5.3) is expanded in similar fashion,

$$4 \sum_{n=0}^{k+l} \langle z_n \rangle_{k+l-n} = 2 \sum_{n=0}^{k+l} z_n \cdot d^{k+l-n} + 2 \sum_{n=0}^{k+l} \sum_{i=0}^{k+l-n} \binom{k+l-n}{i} D^i z_n \cdot d^{k+l-n-i}.$$

A sequence of terms  $z_n$  for which all terms accompanying  $d^{k+l-a}$ ,  $a \in \{0, \dots, k+l\}$ , match on both sides will certainly satisfy (5.3) although it need not be the unique solution. Note that, in a break from convention, we have used the letter  $a$  as a non-negative integer. This leads to the relation

$$\begin{aligned}
2z_a + 2 \sum_{n=0}^{k+l} \binom{k+l-n}{a-n} D^{a-n} z_n &= \left[ \binom{k}{a} + \binom{k+l}{a} \right] (x \cdot D^a y) \\
&\quad + \sum_{i=0}^a \binom{k}{i} \left[ \binom{k-i}{a-i} + \binom{k+l-i}{a-i} \right] (D^i x \cdot D^{a-i} y),
\end{aligned} \tag{5.5}$$

which can easily be fashioned into a recursive procedure for solving  $z_a$ ,

$$z_a = -\frac{1}{2} \sum_{n=0}^{a-1} \binom{k+l-n}{a-n} D^{a-n} z_n + \frac{1}{4} \left[ \binom{k}{a} + \binom{k+l}{a} \right] (x D^a y) \quad (5.6)$$

$$+ \frac{1}{4} \sum_{i=0}^a \binom{k}{a} \left[ \binom{k-i}{a-i} + \binom{k+l-i}{a-i} \right] (D^i x D^{a-i} y),$$

starting from  $z_0 = xy$  (found by substituting  $a = 0$ ). We remind the reader that  $a \leq k+l$ , and  $\binom{n}{k} = 0$  for  $k < 0$  if  $n > 0$ .

The second part of the lemma states that  $z_n$  of the form (5.4) satisfy (5.3) for some  $\pi_{n,i}^{k,l} \in \mathbb{Q}$ . After substituting this form in (5.5), the left side is comprised of

$$2 \sum_{i=0}^a \pi_{a,i}^{k,l} D^i x D^{a-i} y \quad (5.7)$$

and

$$\begin{aligned} 2 \sum_{n=0}^{k+l} \binom{k+l-n}{a-n} D^{a-n} \left( \sum_{i=0}^n \pi_{n,i}^{k,l} D^i x D^{n-i} y \right) \\ = 2 \sum_{n=0}^{k+l} \binom{k+l-n}{a-n} \sum_{i=0}^n \pi_{n,i}^{k,l} \sum_{j=0}^{a-n} \binom{a-n}{j} D^{i+j} x D^{a-i-j} y, \end{aligned} \quad (5.8)$$

where the inner term  $D^{a-n} \cdot (D^i x D^{n-i} y)$  has been expanded using (5.1).

We now equate terms accompanying  $D^p x D^{a-p} y$  in (5.5), noting that this would give a solution which need not be unique. We arrive at equations of the form

$$R_{a,p}^{k,l} = L_{a,p}^{k,l}, \quad p \in \{0, \dots, a\}, \quad a \in \{0, \dots, k+l\},$$

where

$$R_{a,p}^{k,l} = 2\pi_{a,p}^{k,l} + 2 \sum_{n=0}^a \binom{k+l-n}{a-n} \sum_{i=0}^n \pi_{n,i}^{k,l} \binom{a-n}{p-i}, \quad (5.9)$$

$$\begin{aligned} L_{a,p}^{k,l} &= \delta_{p,0} \left( \binom{k}{a} + \binom{k+l}{a} \right) \\ &\quad + \binom{k}{p} \left( \binom{k-p}{a-p} + \binom{k+l-p}{a-p} \right), \end{aligned} \quad (5.10)$$

and  $\delta_{i,j}$  is the Kronecker delta function. The fact that a recursive procedure for finding  $\pi$ s can be designed,

$$\pi_{a,p}^{k,l} = \frac{1}{4} \left( L_{a,p}^{k,l} - 2 \sum_{n=0}^{a-1} \binom{k+l-n}{a-n} \sum_{i=0}^n \pi_{n,i}^{k,l} \binom{a-n}{p-i} \right), \quad (5.11)$$

starting from  $\pi_{0,0}^{k,l} = 1$ , is the proof that such  $\pi_{n,i}^{k,l} \in \mathbb{Q}$  exist, whereby  $z_n \in \mathcal{C}$  of the form (5.4) satisfies (5.3). Hence  $\mathfrak{G}$  is an associative algebra with the prescribed properties (5.3) and (5.4).  $\square$

### 5.2.1 Explicit form of the coefficients

The recursive procedure (5.11) suffices for the proof of Lemma 5.2.1 and for expanding products of Jordan polynomials using a symbolic computation algorithm. Coefficients  $\pi_{n,i}^{k,l}$  for a few combinations of  $k$  and  $l$ , which can be derived using this recursive procedure, are listed in Table 5.1 in Section 5.4.

Nevertheless, an explicit form for these coefficients remains highly desirable since the observations which lead us to the structural properties (such as the non-mixing property) of the Lie algebra in Section 5.3—the main result of this chapter—are not immediately evident in the recursive form.

**Lemma 5.2.2.** *The explicit form of the coefficients is given by*

$$\pi_{n,i}^{k,l} = \frac{1}{2} \sum_{s=0}^{k+l} \sum_{j=0}^{k+l} Q_{(n,i),(s,j)}^{k+l} L_{s,j}^{k,l}, \quad (5.12)$$

where

$$Q_{(n,i),(s,j)}^q = \begin{cases} \delta_{n,s} \delta_{i,j} - \frac{P_{n-s+1}}{n-s+1} \binom{q-s}{n-s} \binom{n-s}{i-j} & n \geq s, i \geq j, \\ 0 & \text{otherwise,} \end{cases} \quad (5.13)$$

and  $P_r$  are defined in terms of the Bernoulli numbers  $B_r$  (Abramowitz & Stegun 1964),

$$P_r = (-1)^r (2^r - 1) B_r.$$

*Proof.* Let

$$S_{(a,p),(n,i)}^q = \begin{cases} \delta_{a,n} \delta_{p,i} + \binom{q-n}{a-n} \binom{a-n}{p-i} & a \geq n, p \geq i, \\ 0 & \text{otherwise.} \end{cases} \quad (5.14)$$

Note that, in a break from convention, we have used the letters  $a$  and  $b$  as non-negative integers in this proof. The explicit form of  $\pi$ s in (5.12) allows us to express  $R_{a,p}^{k,l}$  in (5.9) as

$$\begin{aligned} R_{a,p}^{k,l} &= 2 \sum_{n=0}^{k+l} \sum_{i=0}^{k+l} S_{(a,p),(n,i)}^{k+l} \pi_{n,i}^{k,l} \\ &= \sum_{n=0}^{k+l} \sum_{i=0}^{k+l} S_{(a,p),(n,i)}^{k+l} \sum_{s=0}^{k+l} \sum_{j=0}^{k+l} Q_{(n,i),(s,j)}^{k+l} L_{s,j}^{k,l} \\ &= \sum_{s=0}^{k+l} \sum_{j=0}^{k+l} \left[ \sum_{n=0}^{k+l} \sum_{i=0}^{k+l} S_{(a,p),(n,i)}^{k+l} Q_{(n,i),(s,j)}^{k+l} \right] L_{s,j}^{k,l}. \end{aligned}$$



To prove Lemma 5.2.2, we need to prove that  $Q$  in (5.12) is such that  $R_{a,p}^{k,l} = L_{a,p}^{k,l}$  is satisfied. This certainly holds if

$$\sum_{n=0}^q \sum_{i=0}^q S_{(a,p),(n,i)}^q Q_{(n,i),(s,j)}^q = \delta_{a,s} \delta_{p,j}, \quad (5.15)$$

holds for any  $q$  and  $a, s, p, j \in \{0, \dots, q\}$ . To prove this we note that  $S$  and  $Q$ , and therefore their product  $SQ$ , are lower triangular. Thus we may concern ourselves solely with the case  $0 \leq s \leq a \leq q$  and  $0 \leq j \leq p \leq q$ . Denoting  $(SQ)_{(a,p),(s,j)}^q$  as  $T$  for brevity,

$$T = \sum_{n=0}^q \sum_{i=0}^q \left[ \delta_{a,n} \delta_{p,i} + \binom{q-n}{a-n} \binom{a-n}{p-i} \right] \left[ \delta_{n,s} \delta_{i,j} - \frac{P_{n-s+1}}{n-s+1} \binom{q-s}{n-s} \binom{n-s}{i-j} \right]$$

can be separated into four parts,

$$\begin{aligned} T_1 &= \sum_{n=0}^q \sum_{i=0}^q \delta_{a,n} \delta_{p,i} \delta_{n,s} \delta_{i,j} = \delta_{a,s} \delta_{p,j}, \\ T_2 &= - \sum_{n=0}^q \sum_{i=0}^q \delta_{a,n} \delta_{p,i} \frac{P_{n-s+1}}{n-s+1} \binom{q-s}{n-s} \binom{n-s}{i-j} \\ &= - \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j}, \\ T_3 &= \sum_{n=0}^q \sum_{i=0}^q \delta_{n,s} \delta_{i,j} \binom{q-n}{a-n} \binom{a-n}{p-i} = \binom{q-s}{a-s} \binom{a-s}{p-j}, \\ T_4 &= - \sum_{n=0}^q \sum_{i=0}^q \binom{q-n}{a-n} \binom{a-n}{p-i} \frac{P_{n-s+1}}{n-s+1} \binom{q-s}{n-s} \binom{n-s}{i-j}. \end{aligned}$$

In the case of  $T_4$  we note that the binomial coefficients vanish except for  $s \leq n \leq a$  and  $j \leq i \leq p$ , when expanding them leads to the expression

$$T_4 = - \frac{(q-s)!}{(q-a)!} \sum_{n=s}^a \frac{P_{n-s+1}}{n-s+1} \sum_{i=j}^p \frac{1}{(p-i)!(a-n-p+i)!(i-j)!(n-s-i+j)!}.$$

In order to reduce this expression, will need the following identities,

$$\sum_{i=j}^p \frac{1}{(p-i)!(a-n-p+i)!(i-j)!(n-s-i+j)!} = \frac{1}{(a-n)!(n-s)!} \binom{a-s}{p-j}, \quad (5.16)$$

$$\sum_{n=0}^b \binom{b+1}{n+1} (2^{n+1} - 1) B_{n+1} = -\frac{1}{2} \delta_{b,0} - \delta_{b>0} P_{b+1}, \quad (5.17)$$

and

$$\sum_{n=0}^b \frac{P_{n+1}}{(n+1)!(b-n)!} = \frac{1}{b!} - \frac{1}{2}\delta_{b,0} - \frac{P_{b+1}}{(b+1)!}\delta_{b>0}, \quad (5.18)$$

where  $\delta_{b>0}$  is 1 if  $b > 0$  and 0 otherwise. Using these identities (proved at the end of this section),

$$\begin{aligned} T_4 &\stackrel{(5.16)}{=} -\frac{(q-s)!(a-s)}{(q-a)!(p-j)} \sum_{n=s}^a \frac{P_{n-s+1}}{(n-s+1)!(a-n)!} \\ &= -\frac{(q-s)!(a-s)}{(q-a)!(p-j)} \sum_{n=0}^{a-s} \frac{P_{n+1}}{(n+1)!((a-s)-n)!} \\ &\stackrel{(5.18)}{=} -\frac{(q-s)!(a-s)}{(q-a)!(p-j)} \left[ \frac{1}{(a-s)!} - \frac{1}{2}\delta_{a,s} - \frac{P_{a-s+1}}{(a-s+1)!}\delta_{a>s} \right] \\ &= -\binom{q-s}{a-s} \binom{a-s}{p-j} + \frac{1}{2}\delta_{a,s}\delta_{p,j} + \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j} \delta_{a>s} \end{aligned}$$

This allows us to complete the proof of Lemma 5.2.2 by proving (5.15),

$$\begin{aligned} T &= \delta_{a,s}\delta_{p,j} - \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j} + \binom{q-s}{a-s} \binom{a-s}{p-j} \\ &\quad - \binom{q-s}{a-s} \binom{a-s}{p-j} + \frac{1}{2}\delta_{a,s}\delta_{p,j} + \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j} \delta_{a>s} \\ &= \frac{3}{2}\delta_{a,s}\delta_{p,j} - (1 - \delta_{a>s}) \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j} \\ &= \frac{3}{2}\delta_{a,s}\delta_{p,j} - \delta_{a \leq s} \frac{P_{a-s+1}}{a-s+1} \binom{q-s}{a-s} \binom{a-s}{p-j} \\ &= \frac{3}{2}\delta_{a,s}\delta_{p,j} - \delta_{a,s}\delta_{p,j}P_1 = \delta_{a,s}\delta_{p,j}, \end{aligned}$$

in proving which we have used  $P_1 = 1/2$  and the fact that  $SQ$  is lower triangular (thus the only case of  $a \leq s$  that we need to consider is  $s = a$ ). This completes the proof of Lemma 5.2.2.  $\square$

Proofs for the identities (5.16), (5.17) and (5.18) used in the proof of Lemma 5.2.2 are derived here.

*Proof.* (5.16)

$$\begin{aligned} \sum_{i=j}^p \frac{(a-n)!(n-s)!}{(p-i)!(a-n-p+i)!(i-j)!(n-s-i+j)!} &= \sum_{i=j}^p \binom{a-n}{p-i} \binom{n-s}{i-j} \\ &= \sum_{i=0}^{p-j} \binom{a-n}{(p-j)-i} \binom{n-s}{i} = \binom{a-s}{p-j}, \end{aligned}$$

since

$$\sum_{i=0}^k \binom{n}{k-i} \binom{m}{i} = \binom{n+m}{k}.$$

□

*Proof.* (5.17)

$$\begin{aligned} \sum_{n=0}^b \binom{b+1}{n+1} (2^{n+1} - 1) B_{n+1} &= \sum_{n=0}^b \binom{b+1}{n} (2^{b-n+1} - 1) B_{b-n+1} \\ &= \sum_{n=0}^{b+1} \binom{b+1}{n} (2^{b-n+1} - 1) B_{b-n+1} - 0 \\ &= \left[ 2^{b+1} \mathcal{B}_{b+1}(1/2) - \mathcal{B}_{b+1}(1) \right] \\ &= \left[ 2 - 2^{b+1} - (-1)^{b+1} \right] B_{b+1}, \end{aligned} \quad (5.19)$$

where  $\mathcal{B}_k(x)$  are the Bernoulli polynomials (Abramowitz & Stegun 1964),

$$\mathcal{B}_k(x) = \sum_{i=0}^k \binom{k}{i} B_{k-i} x^i,$$

whose values at 1 and 1/2 are,

$$\mathcal{B}_k(1) = (-1)^k B_k, \quad \mathcal{B}_k(1/2) = (2^{1-k} - 1) B_k.$$

For  $b = 0$ , the expression (5.19) evaluates to  $-1/2$ , while for  $b > 0$ ,  $(-1)^{b+1}$  can be replaced by 1 since  $B_{b+1}$  vanishes for all cases when  $(-1)^{b+1}$  is negative. Thus, for  $b > 0$ , (5.19) evaluates to  $(1 - 2^{b+1}) B_{b+1}$ . Using the same logic, we may multiply it by  $(-1)^{b+1}$  to get  $-P_{b+1}$ , completing the proof. □

*Proof.* (5.18)

$$\begin{aligned} \sum_{n=0}^b \frac{P_{n+1}}{(n+1)!(b-n)!} &= \frac{1}{(b+1)!} \sum_{n=0}^b \binom{b+1}{n+1} (-1)^{n+1} (2^{n+1} - 1) B_{n+1} \\ &= \frac{1}{(b+1)!} \left[ \sum_{n=0}^b \binom{b+1}{n+1} (2^{n+1} - 1) B_{n+1} - 2(b+1) B_1 \right] \\ &= \frac{1}{b!} + \frac{1}{(b+1)!} \sum_{n=0}^b \binom{b+1}{n+1} (2^{n+1} - 1) B_{n+1} \\ &\stackrel{(5.17)}{=} \frac{1}{b!} - \frac{1}{2} \delta_{b,0} - \frac{P_{b+1}}{(b+1)!} \delta_{b>0}. \end{aligned}$$

Where we have used the fact that, except for the  $n = 0$  case, all negative occurrences of

$(-1)^{n+1}$  vanish since  $B_{n+1}$  vanishes. □

### 5.2.2 The generating function for the coefficients

The symmetry of the coefficients  $\pi_{n,i}^{k,l}$  that we require in order for the non-mixing property to hold will be proven via the generating function of these coefficients, which is developed in this section. We will break the notational convention of this chapter in this section, using  $u, w, x$  and  $y$  as variables in which the formal series of the generating function is specified.

**Lemma 5.2.3.** *The generating function,*

$$h(u, w, y, x) = \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \pi_{n,i}^{k,l}, \quad (5.20)$$

for the coefficients  $\pi_{n,i}^{k,l}$  appearing in (5.3) is

$$h(u, w, y, x) = \frac{\exp((wy - uxy)/2)}{1 - (w + u)} \frac{\cosh(uy/2) \cosh(wxy/2)}{\cosh(y(u + w)(1 + x)/2)}. \quad (5.21)$$

*Proof.* We wish to find an explicit form for the generating function

$$h(u, w, y, x) = \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \pi_{n,i}^{k,l}. \quad (5.20)$$

We start from the result of Lemma 5.2.2, substituting (5.12) in (5.20) and splitting  $\pi_{n,i}^{k,l}$  into eight parts for convenience,

$$\pi_{n,i}^{k,l} = \frac{1}{2} \sum_{j=1}^8 p_j(k, l, n, i), \quad (5.22)$$

where

$$\begin{aligned} p_1(k, l, n, i) &= \delta_{i,0} \binom{k}{n}, & p_2(k, l, n, i) &= \delta_{i,0} \binom{k+l}{n}, \\ p_3(k, l, n, i) &= \binom{k}{i} \binom{k-i}{n-i}, & p_4(k, l, n, i) &= \binom{k}{i} \binom{k+l-i}{n-i}, \\ p_5(k, l, n, i) &= - \sum_{r=0}^n \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i} \binom{k}{n-r}, \\ p_6(k, l, n, i) &= - \sum_{r=0}^n \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i} \binom{k+l}{n-r}, \\ p_7(k, l, n, i) &= - \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i-j} \binom{k}{j} \binom{k-j}{n-r-j}, \end{aligned}$$

$$p_8(k, l, n, i) = - \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i-j} \binom{k}{j} \binom{k+l-j}{n-r-j},$$

are obtained after a change of variables,  $r = n - s$ , and noting  $\binom{0}{i-j} = \delta_{i,j}$  and the fact that  $\binom{n}{k}$  vanishes for  $k < 0$  when  $n \geq 0$ . We will simplify the corresponding parts of  $h(u, w, y, x)$ ,

$$h_j(u, w, y, x) = \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i p_j(k, l, n, i),$$

combining them to find an expression for the generating function

$$h(u, w, y, x) = \sum_{j=1}^8 h_j(u, w, y, x)/2.$$

In this pursuit, we will repeatedly use a few results,

$$\sum_{r=0}^{\infty} \frac{B_r}{r!} x^r = \frac{x}{e^x - 1}, \quad (5.23)$$

which is well known,

$$\begin{aligned} \sum_{r=0}^{\infty} \frac{P_{r+1}}{(r+1)!} x^r &= \frac{1}{x} \sum_{r=0}^{\infty} \frac{P_r}{r!} x^r = \frac{1}{x} \sum_{r=0}^{\infty} \frac{(-1)^r (2^r - 1) B_r}{r!} x^r \\ &= \frac{1}{x} \sum_{r=0}^{\infty} \frac{B_r}{r!} (-2x)^r - \frac{1}{x} \sum_{r=0}^{\infty} \frac{B_r}{r!} (-x)^r \\ &= - \left( \frac{2}{e^{-2x} + 1} - \frac{1}{e^{-x} + 1} \right) = \frac{e^x}{e^x + 1}, \end{aligned} \quad (5.24)$$

where we use the fact that  $P_0 = 0$ , and

$$\begin{aligned} \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+r}{l} u^l w^{k+n-l} &= \sum_{k=0}^{\infty} w^{k+n} \sum_{l=0}^{\infty} \binom{k+r}{l} \left( \frac{u}{w} \right)^l = \sum_{k=0}^{\infty} w^{k+n} \left( 1 + \frac{u}{w} \right)^k + r \\ &= w^{n-r} (w+u)^r \sum_{k=0}^{\infty} (w+u)^k = \frac{w^{n-r} (w+u)^r}{1 - (w+u)}. \end{aligned} \quad (5.25)$$

Two standard tricks for exchanging summations that we exploit are

$$\sum_{n=0}^{\infty} \sum_{r=0}^n \alpha_{n,r} = \sum_{r=0}^{\infty} \sum_{n=0}^{\infty} \alpha_{n+r,r}, \quad (5.26)$$

$$\sum_{k=0}^{\infty} \sum_{n=0}^{k+l} \alpha_{n,k,l} = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \alpha_{n,k+n-l,l}. \quad (5.27)$$

With these tools, we proceed to seek expressions for the generating functions  $h_j$ , writing  $h_j$  as shorthand for  $h_j(u, w, y, x)$ .

$$\begin{aligned} h_1 &= \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \delta_{i,0} \binom{k}{n} \\ &= \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} \binom{k+l-n}{l} \frac{y^n}{n!} = \sum_{n=0}^{\infty} \frac{y^n}{n!} \left[ \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k}{l} u^l w^{k+n-l} \right] \\ &= \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} \frac{(yw)^n}{n!} = \frac{e^{yw}}{1-(w+u)}, \\ h_2 &= \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \delta_{i,0} \binom{k+l}{n} \\ &= \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} \binom{k+l}{l} \frac{y^n}{n!} = \sum_{n=0}^{\infty} \frac{y^n}{n!} \left[ \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+n}{l} u^l w^{k+n-l} \right] \\ &= \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} \frac{(y(w+u))^n}{n!} = \frac{e^{y(w+u)}}{1-(w+u)}, \\ h_3 &= \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \binom{k}{i} \binom{k-i}{n-i} \\ &= \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} \frac{y^n}{n!} \binom{k+l-n}{l} \left[ \sum_{i=0}^n \binom{n}{i} x^i \right] \\ &= \sum_{n=0}^{\infty} \frac{(y(1+x))^n}{n!} \left[ \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k}{l} u^l w^{k+n-l} \right] \\ &= \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} \frac{(yw(1+x))^n}{n!} = \frac{e^{yw(1+x)}}{1-(w+u)}, \end{aligned}$$

$$\begin{aligned}
 h_4 &= \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \binom{k}{i} \binom{k+l-i}{n-i} \\
 &= \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} \frac{y^n}{n!} \sum_{i=0}^n \binom{n}{i} x^i \binom{k+l-i}{l} \\
 &= \sum_{n=0}^{\infty} \frac{y^n}{n!} \sum_{i=0}^n \binom{n}{i} x^i \left[ \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+n-i}{l} u^l w^{k+n-l} \right] \\
 &= \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} \frac{y^n}{n!} \sum_{i=0}^n \binom{n}{i} x^i w^i (w+u)^{n-i} \\
 &= \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} \frac{(y((w+u)+xw))^n}{n!} = \frac{e^{y(w+u)+xyw}}{1-(w+u)}, \\
 h_5 &= - \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \sum_{r=0}^n \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i} \binom{k}{n-r} \\
 &= - \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} y^n \sum_{r=0}^n \sum_{i=0}^r x^i \frac{P_{r+1}}{(r+1)!(n-r)!} \binom{k+l-n+r}{l} \binom{r}{i} \\
 &= - \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \frac{P_{r+1}}{(r+1)!(n-r)!} \left[ \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+r}{l} u^l w^{k+n-l} \right] \left[ \sum_{i=0}^r \binom{r}{i} x^i \right],
 \end{aligned}$$

where we have changed summation limits on  $n$  using (5.27), and for  $i$  since  $r \leq n$  and  $\binom{r}{i} = 0$  for  $i > r \geq 0$ . Using (5.25) and changing limits again using (5.26),

$$\begin{aligned}
 h_5 &= - \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \frac{P_{r+1}}{(r+1)!(n-r)!} w^{n-r} ((w+u)(1+x))^r \\
 &= - \frac{1}{1-(w+u)} \left[ \sum_{n=0}^{\infty} \frac{(yw)^n}{n!} \right] \sum_{r=0}^{\infty} \frac{P_{r+1}}{(r+1)!} (y(w+u)(1+x))^r \\
 &= - \frac{e^{yw}}{1-(w+u) e^z + 1},
 \end{aligned}$$

where we have defined  $z = y(w + u)(1 + x)$  for convenience.

$$\begin{aligned}
 h_6 &= - \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i \sum_{r=0}^n \frac{P_{r+1}}{r+1} \binom{k+l-n+r}{r} \binom{r}{i} \binom{k+l}{n-r} \\
 &= - \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} y^n \sum_{r=0}^n \frac{P_{r+1}}{(r+1)!(n-r)!} \binom{k+l}{l} \left[ \sum_{i=0}^r \binom{r}{i} x^i \right] \\
 &= - \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \frac{P_{r+1}}{(r+1)!(n-r)!} (1+x)^r \left[ \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+n}{l} u^l w^{k+n-l} \right] \\
 &= - \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} (y(w+u))^n \sum_{r=0}^n \frac{P_{r+1}}{(r+1)!(n-r)!} (1+x)^r \\
 &= - \frac{1}{1-(w+u)} \left[ \sum_{n=0}^{\infty} \frac{(y(w+u))^n}{n!} \right] \sum_{r=0}^{\infty} \frac{P_{r+1}}{(r+1)!} (y(w+u)(1+x))^r \\
 &= - \frac{e^{y(w+u)}}{1-(w+u)} \frac{e^z}{e^z + 1}.
 \end{aligned}$$

After manipulating binomial coefficients, we change summation limits (5.27) and use (5.25),

$$\begin{aligned}
 h_7 &= - \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} y^n \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{(r+1)!j!(n-r-j)!} \binom{k+l-n+r}{l} \left[ \sum_{i=0}^n \binom{r}{i-j} x^i \right] \\
 &= - \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{(r+1)!j!(n-r-j)!} \left[ \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+r}{l} u^l w^{k+n-l} \right] \left[ \sum_{i=j}^{r+j} \binom{r}{i-j} x^i \right] \\
 &= - \frac{1}{1-(w+u)} \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{(r+1)!j!(n-r-j)!} w^{n-r} (w+u)^r x^j (1+x)^r,
 \end{aligned}$$

since  $\binom{r}{i-j}$  vanishes unless  $j \leq i \leq r+j$  and  $r+j \leq n$ . Exchanging limits twice using (5.26), once for  $r$  and once for  $j$ ,

$$\begin{aligned}
 h_7 &= - \frac{1}{1-(w+u)} \sum_{r=0}^{\infty} \sum_{n=0}^{\infty} (yw)^n \sum_{j=0}^n \frac{P_{r+1}}{(r+1)!j!(n-j)!} z^r x^j \\
 &= - \frac{1}{1-(w+u)} \left[ \sum_{n=0}^{\infty} \frac{(yw)^n}{n!} \right] \left[ \sum_{r=0}^{\infty} \frac{P_{r+1}}{(r+1)!} z^r \right] \left[ \sum_{j=0}^{\infty} \frac{(ywx)^j}{j!} \right] \\
 &= - \frac{e^{yw(1+x)}}{1-(w+u)} \frac{e^z}{e^z + 1}.
 \end{aligned}$$



Along similar lines we manipulate

$$\begin{aligned}
 h_8 &= \sum_{l=0}^{\infty} u^l \sum_{k=0}^{\infty} w^k \sum_{n=0}^{k+l} y^n \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{-P_{r+1}}{(r+1)!j!(n-r-j)!} \binom{k+l-j}{l} \left[ \sum_{i=0}^n \binom{r}{i-j} x^i \right] \\
 &= \sum_{n=0}^{\infty} \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{-P_{r+1} y^n}{(r+1)!j!(n-r-j)!} \left[ \sum_{i=j}^{r+j} \binom{r}{i-j} x^i \right] \left[ \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+n-j}{l} u^l w^{k+n-l} \right] \\
 &= -\frac{1}{1-(w+u)} \sum_{n=0}^{\infty} y^n \sum_{r=0}^n \sum_{j=0}^{n-r} \frac{P_{r+1}}{(r+1)!j!(n-r-j)!} x^j (1+x)^r w^j (w+u)^{n-j},
 \end{aligned}$$

since  $\binom{r}{i-j}$  vanishes unless  $j \leq i \leq r+j$  and  $r+j \leq n$ . Once again, exchanging limits twice using (5.26), once for  $r$  and once for  $j$ ,

$$\begin{aligned}
 h_8 &= -\frac{1}{1-(w+u)} \sum_{r=0}^{\infty} \sum_{n=0}^{\infty} y^{n+r} \sum_{j=0}^n \frac{P_{r+1}}{(r+1)!j!(n-j)!} x^j (1+x)^r w^j (w+u)^{n+r-j} \\
 &= -\frac{1}{1-(w+u)} \left[ \sum_{n=0}^{\infty} \frac{(y(w+u))^n}{n!} \right] \left[ \sum_{r=0}^{\infty} \frac{P_{r+1}}{(r+1)!} z^r \right] \left[ \sum_{j=0}^{\infty} \frac{(xyw)^j}{j!} \right] \\
 &= -\frac{e^{y(w+u)+xyw}}{1-(w+u)} \frac{e^z}{e^z+1}.
 \end{aligned}$$

These expressions are then combined to form the full generating function

$$\begin{aligned}
 2h(u, w, y, x) &= \frac{e^{yw}}{1-(w+u)} + \frac{e^{y(w+u)}}{1-(w+u)} + \frac{e^{yw(1+x)}}{1-(w+u)} + \frac{e^{y(w+u)+xyw}}{1-(w+u)} \\
 &\quad - \frac{e^{yw}}{1-(w+u)} \frac{e^z}{e^z+1} - \frac{e^{y(w+u)}}{1-(w+u)} \frac{e^z}{e^z+1} \\
 &\quad - \frac{e^{yw(1+x)}}{1-(w+u)} \frac{e^z}{e^z+1} - \frac{e^{y(w+u)+xyw}}{1-(w+u)} \frac{e^z}{e^z+1} \\
 &= \frac{e^{yw} + e^{y(w+u)} + e^{yw(1+x)} + e^{y(w+u)+xyw}}{1-(w+u)} \left[ 1 - \frac{e^z}{e^z+1} \right] \\
 &= \frac{1}{1-(w+u)} \left[ \frac{e^{yw} + e^{y(w+u)} + e^{yw(1+x)} + e^{y(w+u)+xyw}}{e^z+1} \right].
 \end{aligned}$$

For further simplification, consider  $g = 2h(u, w, 2y, x)(1-(w+u))$ , letting  $a = uy, b =$

$wy, c = uxy$  and  $d = wxy$ ,

$$\begin{aligned}
g &= \frac{e^{2yw} + e^{2y(w+u)} + e^{2yw(1+x)} + e^{2y(w+u)+2xyw}}{e^{2yw+2ywx+2yu+2yux} + 1} \\
&= \frac{e^{2b} + e^{2a+2b} + e^{2b+2d} + e^{2a+2b+2d}}{e^{2a+2b+2d+2c} + 1} \\
&= \frac{e^{-a+b-c-d} + e^{a+b-c-d} + e^{-a+b-c+d} + e^{a+b-c+d}}{e^{a+b+c+d} + e^{-a-b-c-d}} \\
&= e^{b-c} \frac{e^{-a-d} + e^{a-d} + e^{-a+d} + e^{a+d}}{e^{a+b+c+d} + e^{-a-b-c-d}} \\
&= e^{b-c} \frac{(e^{-a} + e^a)(e^{-d} + e^d)}{e^{a+b+c+d} + e^{-a-b-c-d}} \\
&= 2e^{b-c} \frac{\cosh(a) \cosh(d)}{\cosh(a+b+c+d)}.
\end{aligned}$$

This brings us to the desired form of the generating function,

$$h(u, w, y, x) = \frac{\exp((wy - uxy)/2)}{1 - (w + u)} \frac{\cosh(uy/2) \cosh(wxy/2)}{\cosh(y(u + w)(1 + x)/2)}, \quad (5.21)$$

and completes the proof of Lemma 5.2.3.  $\square$

### 5.3 Lie algebra of Jordan polynomials

Recall that, as a consequence of Lemma 5.2.1 discussed in Section 5.2, we have a procedure for multiplying two Jordan monomials (and, therefore, also for multiplying two Jordan polynomials). This involved a recursive procedure (5.11) for finding the coefficients  $\pi_{n,i}^{k,l}$ . As we had noted, once we know how to compose two symmetrised differential operators (Jordan polynomials with  $d = \partial_x$ ),  $\mathcal{L}_1$  and  $\mathcal{L}_2$ , we also have a procedure for computing their commutators since

$$[\mathcal{L}_1, \mathcal{L}_2] = \mathcal{L}_1 \circ \mathcal{L}_2 - \mathcal{L}_2 \circ \mathcal{L}_1.$$

However, a major motivation for working with the symmetrised differential operators is to ensure that Hermitian and skew-Hermitian terms do not mix. This non-mixing property was not obvious from the results of Lemma 5.2.1. However, the results of Section 5.2.1 and Section 5.2.2 will allow us to show the symmetry property of the coefficients in Lemma 5.3.1, which leads to the  $\mathbb{Z}_2$ -graded structure of our Lie algebra. This property, which can be gleaned from Theorem 5.3.2, is the crucial property of non-mixing that we seek.

As we have already stated, since Jordan polynomials in  $d$  with coefficients in  $\mathcal{C}$  form an associative algebra  $(\mathfrak{G}, \cdot, +)$ , it is immediately obvious that they also form a Lie algebra.

From (5.3), we know that the commutator of the monomials can be expanded to

$$[\langle x \rangle_k, \langle y \rangle_l] = \sum_{n=0}^{k+l} \sum_{i=0}^n \mu_{n,i}^{k,l} \langle D^i x D^{n-i} y \rangle_{k+l-n}, \quad (5.28)$$

where  $\mu_{n,i}^{k,l} = \pi_{n,i}^{k,l} - \pi_{n,n-i}^{l,k}$ . However, as we have suspected all along, the structure of  $\mathfrak{G}$  turns out to be more interesting than this due to a symmetry of the coefficients,  $\pi_{n,i}^{k,l}$ .

**Lemma 5.3.1.** *The coefficients possess the symmetry,*

$$\pi_{n,i}^{k,l} = (-1)^n \pi_{n,n-i}^{l,k}.$$

*Proof.* We show that the generating function for  $(-1)^n \pi_{n,n-i}^{l,k}$  coincides with the generating function of  $\pi_{n,i}^{k,l}$ ,

$$\begin{aligned} g(u, w, y, x) &= \sum_{l=0}^{\infty} \frac{u^l}{l!} \sum_{k=0}^{\infty} \frac{w^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! y^n \sum_{i=0}^n x^i (-1)^n \pi_{n,n-i}^{l,k} \\ &= \sum_{l=0}^{\infty} \frac{w^l}{l!} \sum_{k=0}^{\infty} \frac{u^k}{k!} \sum_{n=0}^{k+l} (k+l-n)! (-y)^n \sum_{i=0}^n x^{n-i} \pi_{n,i}^{k,l} \\ &= h(w, u, -xy, 1/x) \\ &= \frac{\exp((-uxy + wy)/2)}{1 - (w + u)} \frac{\cosh(-wxy/2) \cosh(-uy/2)}{\cosh(-yx(u + w)(1 + 1/x)/2)} \\ &= \frac{\exp((wy - uxy)/2)}{1 - (w + u)} \frac{\cosh(wxy/2) \cosh(uy/2)}{\cosh(yx(u + w)(1 + 1/x)/2)} \\ &= h(u, w, y, x). \end{aligned}$$

□

**Theorem 5.3.2.** *Commutators in the Lie algebra of Jordan polynomials  $(\mathfrak{G}, [\cdot, \cdot], +)$  can be solved explicitly using the rule*

$$[\langle x \rangle_k, \langle y \rangle_l] = \sum_{n=0}^{\frac{k+l-1}{2}} \sum_{i=0}^{2n+1} \lambda_{n,i}^{k,l} \langle D^i x D^{2n+1-i} y \rangle_{k+l-2n-1}, \quad (5.29)$$

where  $\lambda_{n,i}^{k,l} = 2\pi_{2n+1,i}^{k,l} \in \mathbb{Q}$ .

*Proof.* The even indexed coefficients  $\mu_{2n,i}^{k,l}$  in (5.28) vanish due to Lemma 5.3.1,  $\mu_{2n,i}^{k,l} = \pi_{2n,i}^{k,l} - (-1)^{2n} \pi_{2n,i}^{l,k} = 0$ , while  $\mu_{2n+1,i}^{k,l} = \pi_{2n+1,i}^{k,l} - (-1)^{2n+1} \pi_{2n+1,i}^{l,k} = 2\pi_{2n+1,i}^{k,l}$ . We conveniently rename  $\mu_{2n+1,i}^{k,l}$  as  $\lambda_{n,i}^{k,l}$ . □

Theorem 5.3.2 is the central result of this chapter. The following consequences of this theorem will be discussed in the rest of this section and exploited further in Section 6.1.3.

- (i) Commutators of symmetrised differential operators can be expressed as other symmetrised differential operators, and we have a procedure based on the coefficients  $\lambda_{n,i}^{k,l}$  that allows us to do this in a symbolic algebra program. Moreover, due to (5.12) these coefficients are known explicitly. Thus if we were to start with symmetrised differential operators and form nested commutators, we will still be able to write any such nested commutator as a symmetrised differential operator.

This is the property allowing us to simplify nested commutators in Zassenhaus splittings and Magnus expansions to symmetrised differential operators.

- (ii) The  $\langle \cdot \rangle_{k+l}$  term that appears when composing two symmetrised differential operators of the form  $\langle \cdot \rangle_k$  and  $\langle \cdot \rangle_l$  cancels out when we, instead, consider their commutator. Thus the highest possible degree differential operator in the commutator vanishes.

Due to highly oscillatory solutions in the semiclassical regime, high derivatives of the wavefunction are very large. Thus, this property of ‘height reduction’ significantly reduces the size of nested commutators, leading to much smaller errors, milder constraints and lower costs of our numerical schemes.

- (iii) Either  $\langle \cdot \rangle_{2m+1}$  or  $\langle \cdot \rangle_{2m}$  terms, but not both, appear in the commutator of two monomials—this is the property of non-mixing that proves crucial for separating Hermitian and skew-Hermitian terms, leading to unitarity and stability of our methods.

We will now consider the above properties of this Lie algebra in more details.

### Height reduction

The linear space of Jordan polynomials of degree  $\leq n$  is written as

$$\mathfrak{G}_n = \bigoplus_{k \leq n} \mathfrak{F}_k.$$

The algebra  $\mathfrak{G} = \lim_{n \rightarrow \infty} \mathfrak{G}_n$  can be seen as a filtered Lie algebra since

$$[\mathfrak{G}_k, \mathfrak{G}_l] \subseteq \mathfrak{G}_{k+l}.$$

In other words, when we take the commutator of a degree  $k$  symmetrised differential operator and a degree  $l$  symmetrised differential operator, we expect the result to be another symmetrised differential operator of degree  $k + l$ , just like the product of these operators. As we see shortly, however, this turns out not to be the case—instead the commutator is a symmetrised differential operator of degree  $k + l - 1$ .

**Definition 5.3.3 (Height).** *The height of a Jordan polynomial  $W \in \mathfrak{G}$  is defined as its degree and written as  $\text{ht}(W)$ —it is the smallest integer  $n$  such that  $W \in \mathfrak{G}_n$ . We define*

the height of 0 (the zero Jordan polynomial) to be  $-1$  as per common convention, this being the only term with negative height.

Effectively, the height of a general term in  $\mathfrak{G}_n$  (a degree  $n$  Jordan polynomial) is

$$\text{ht} \left( \sum_{k=0}^n \langle x_k \rangle_k \right) = n,$$

provided  $x_n$  does not vanish. A degree  $k$  Jordan monomial  $\langle x \rangle_k$  has height  $k$ .

**Lemma 5.3.4 (Height Reduction for Commutators of Jordan Monomials).**

$$\text{ht} ([\langle x \rangle_k, \langle y \rangle_l]) \leq k + l - 1. \quad (5.30)$$

*Proof.* The largest degree monomial on the right hand side of (5.29) has the index  $k+l-1$ , which occurs for  $n=0$ . As a consequence, the height of the right hand side of (5.29) is  $k+l-1$ .  $\square$

We note that, as a consequence of Theorem 5.3.2 and Lemma 5.3.4,

$$[\mathfrak{G}_k, \mathfrak{G}_l] \subseteq \mathfrak{G}_{k+l-1},$$

while keeping in mind that  $\mathfrak{G}_{k+l-1} \subseteq \mathfrak{G}_{k+l}$ , so that there is no contradiction. Thus, the commutator of two symmetrised differential operators is a symmetrised differential operator of degree one lower than the sum of their degrees.

**Corollary 5.3.5 (Height Reduction).** *For any commutator  $C$  featuring the ‘letters’  $W_i \in \mathfrak{G}_{k_i}$ ,  $i = 1, \dots, n$ ,*

$$\text{ht} (C(W_1, \dots, W_n)) \leq \sum_{i=1}^n k_i - n + 1. \quad (5.31)$$

### Non-mixing property

The non-mixing property of this Lie algebra is that it doesn’t *mix* terms of certain forms. We define

$$\mathfrak{e} = \bigoplus_{k \geq 0} \mathfrak{F}_{2k}, \quad \mathfrak{o} = \bigoplus_{k \geq 0} \mathfrak{F}_{2k+1},$$

so that  $\mathfrak{G} = \mathfrak{o} \oplus \mathfrak{e}$ .

In the special case of  $d = \partial_x$ , which interests us while solving the Schrödinger equation, the even indexed terms  $\mathfrak{e}$  correspond to Hermitian operators and the odd indexed terms  $\mathfrak{o}$  correspond to skew-Hermitian operators. The non-mixing property means that the commutator of two monomial symmetrised differential operators can be written as another symmetrised differential operator which is either Hermitian or skew-Hermitian but doesn’t

combine Hermitian and skew-Hermitian terms. Moreover, the discretisation of such a differential operator also results in a Hermitian or skew-Hermitian matrix, appropriately, without mixing terms.

The following relations are evident from (5.29),

$$\begin{aligned} [\mathfrak{e}, \mathfrak{e}] &\subseteq \mathfrak{o}, & [\mathfrak{o}, \mathfrak{o}] &\subseteq \mathfrak{o}, \\ [\mathfrak{e}, \mathfrak{o}] &\subseteq \mathfrak{e}, & [\mathfrak{o}, \mathfrak{e}] &\subseteq \mathfrak{e}. \end{aligned} \quad (5.32)$$

For instance, the commutator of two terms in  $\mathfrak{e}$ , such as  $[\langle f \rangle_{2k}, \langle g \rangle_{2l}]$ , is expressed as a linear combination of terms of the form  $\langle \cdot \rangle_{2k+2l-2n-1}$  which have odd indices and are, therefore, in  $\mathfrak{o}$ .

As a consequence of (5.32), for instance,

$$[[[\mathfrak{e}, \mathfrak{e}], \mathfrak{e}], \mathfrak{o}], \mathfrak{e}] \subseteq \mathfrak{o}, \quad [[[[\mathfrak{e}, \mathfrak{o}], \mathfrak{e}], \mathfrak{e}], \mathfrak{o}] \subseteq \mathfrak{e}. \quad (5.33)$$

Thus, if each letter  $W_i$  is either in  $\mathfrak{e}$  or in  $\mathfrak{o}$ , the commutator  $C(W_1, \dots, W_n)$  is either in  $\mathfrak{e}$  or  $\mathfrak{o}$  and does not mix terms from the two. Moreover, commutators with the same letters will fall in the same space,  $\mathfrak{e}$  or  $\mathfrak{o}$ . For example,

$$[[[[\mathfrak{e}, \mathfrak{o}], \mathfrak{e}], \mathfrak{e}], \mathfrak{e}] \subseteq \mathfrak{o}, \quad [[[[\mathfrak{e}, \mathfrak{o}], \mathfrak{o}], \mathfrak{e}], \mathfrak{e}] \subseteq \mathfrak{e}, \quad (5.34)$$

are obtained merely by exchanging the positions of  $\mathfrak{o}$  and  $\mathfrak{e}$  in the commutators (5.33). As already stated earlier, this non-mixing property has implications for our Zassenhaus schemes discussed in Chapter 6 where this property translates into a matrix being either Hermitian or skew-Hermitian but not a mix of the two, which would have resulted in unfavourable structures like the ones appearing in Chapter 4.

**Note:** Formally, the property (5.32) says that  $\mathfrak{G}$  is a  $\mathbb{Z}_2$ -graded Lie algebra with the grade 0 component  $\mathfrak{o}$  and grade 1 component  $\mathfrak{e}$ . A graded Lie algebra  $\mathfrak{s}$  is a Lie algebra which is a direct sum of vector spaces  $\mathfrak{s}_k$ ,

$$\mathfrak{s} = \bigoplus_{k \in \mathbb{Z}} \mathfrak{s}_k,$$

such that the Lie bracket respects the gradation,

$$[\mathfrak{s}_k, \mathfrak{s}_l] \subseteq \mathfrak{s}_{k+l}.$$

A  $\mathbb{Z}_n$ -graded Lie algebra is similar, except the index is in  $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$  instead of  $\mathbb{Z}$ ,

$$\mathfrak{s} = \bigoplus_{k \in \mathbb{Z}_n} \mathfrak{s}_k,$$

and

$$[\mathfrak{s}_k, \mathfrak{s}_l] \subseteq \mathfrak{s}_{k \oplus_n l},$$

where  $k \oplus_n l = (k + l) \bmod n$  is the modular addition.

In light of this, (5.33) and (5.34) can be easily understood by considering the grade:  $[[[[\mathfrak{e}, \mathfrak{e}], \mathfrak{e}], \mathfrak{o}], \mathfrak{e}]$  has the grade  $1 \oplus 1 \oplus 1 \oplus 0 \oplus 1 = 4 \bmod 2 = 0$ , thus residing in the grade 0 component,  $\mathfrak{o}$ .

**Note:** The linear space  $\mathfrak{o}$  is a Lie algebra in its own right, while structures of the form  $\mathfrak{e}$  are also called Lie triple systems which are closed under double commutation:

$$[\mathfrak{e}, [\mathfrak{e}, \mathfrak{e}]] \subseteq \mathfrak{e}.$$

These notions are closely related to Lie groups and symmetric spaces and have found applications in linear algebra methods such as the generalised polar decomposition (Munthe-Kaas, Quispel & Zanna 2001).

### 5.3.1 Characterisation of the Lie algebra

As anticipated, there is a close relation between  $\mathfrak{G}$  and the Lie algebra generated by  $\mathcal{C}$  and polynomials in  $d$  with constant coefficients, i.e. the linear space closure of all their commutators. Recall that we were interested in studying the linear space of all nested commutators since commutators of all grades appear in the sBCH series, which underlies the Zassenhaus algorithm. Thus, the fact that these commutators can all be written in terms of the symmetrised differential operators means that, as a consequence of Lemma 5.3.6, the exponents in our Zassenhaus splittings can always be written as symmetrised differential operators and their properties directly become relevant for our schemes.

**Lemma 5.3.6.** *The Lie algebra generated by  $\mathcal{C}$  and  $\mathcal{P}(d)$  (the ring of polynomials in  $d$  with constant coefficients) is contained in  $\mathfrak{G}$ ,*

$$\mathfrak{f} = \text{LA}(\mathcal{C} \cup \mathcal{P}(d)) \subseteq \mathfrak{G}.$$

*The two are identical if an inverse of the mapping  $D = \text{ad}_d : \mathcal{C} \rightarrow \mathcal{C}$  exists.*

*Proof.* The containment is not difficult to prove. Since  $d^k = \langle 1 \rangle_k$ , it is contained in  $\mathfrak{F}_k$  and therefore  $\mathcal{P}(d) \subseteq \mathfrak{G}$ . The algebra  $\mathcal{C}$  is also contained in  $\mathfrak{G}$  since every  $x \in \mathcal{C}$  can be written in the form  $x = \langle x \rangle_0$ . The Lie algebra  $\mathfrak{f}$  generated by  $\mathcal{C} \cup \mathcal{P}(d)$  is the intersection of all Lie algebras containing  $\mathcal{C}$  and  $\mathcal{P}(d)$  and is therefore contained in the Lie algebra  $\mathfrak{G}$ .

The two algebras  $\mathfrak{f}$  and  $\mathfrak{G}$  are identical if inverse of the map  $D$  exists. Any term in  $\mathfrak{F}_0$  is of the form  $\langle y \rangle_0$  for some  $y \in \mathcal{C}$  and therefore trivially resides in  $\mathfrak{f}$ . Take any  $x \in \mathcal{C}$  and

note that  $d^2 \in \mathcal{P}(d)$ . These are both contained in  $\mathfrak{f}$  and by definition of being a Lie algebra  $[d^2, x] = [\langle 1 \rangle_2, \langle x \rangle_0] = \lambda_{0,0}^{2,0} \langle D^0(1)D^1(x) \rangle_1 + \lambda_{0,1}^{2,0} \langle D^1(1)D^0(x) \rangle_1 = 2 \langle D(x) \rangle_1$  resides in  $\mathfrak{f}$  as well. Consequently, any  $\langle y \rangle_1 \in \mathfrak{F}_1$  can be expressed as  $\frac{1}{2}[d^2, D^{-1}(y)]$  so long as the inverse of  $D$  exists. Thus we have  $\mathfrak{F}_0, \mathfrak{F}_1 \subseteq \mathfrak{f}$ . These two cases form the base case of our induction argument.

Assume that for all  $k < n$ ,  $\mathfrak{F}_{2k}$  and  $\mathfrak{F}_{2k+1}$  are contained in  $\mathfrak{f}$ . For completing the inductive proof, consider

$$\begin{aligned} [d^{2n+1}, x] &= [\langle 1 \rangle_{2n+1}, \langle x \rangle_0] = \sum_{s=0}^n \sum_{i=0}^{2s+1} \lambda_{s,i}^{2n+1,0} \langle D^i(1)D^{2s+1-i}(x) \rangle_{2n-2s} \\ &= \sum_{s=0}^n \lambda_{s,0}^{2n+1,0} \langle D^{2s+1}(x) \rangle_{2n-2s} \\ &= \lambda_{0,0}^{2n+1,0} \langle D(x) \rangle_{2n} + \sum_{s=1}^n \lambda_{s,0}^{2n+1,0} \langle D^{2s+1}(x) \rangle_{2n-2s}. \end{aligned}$$

Setting  $x = D^{-1}(y) \in \mathcal{C}$ ,

$$\langle y \rangle_{2n} = \frac{1}{\lambda_{0,0}^{2n+1,0}} [d^{2n+1}, D^{-1}(y)] - \sum_{s=1}^n \frac{\lambda_{s,0}^{2n+1,0}}{\lambda_{0,0}^{2n+1,0}} \langle D^{2s}(y) \rangle_{2n-2s}.$$

The first term on the right hand side is in  $\mathfrak{f}$  since  $d^{2n+1} \in \mathcal{P}(d)$ , while the terms in the summation fall in  $\mathfrak{f}$  due to the induction hypothesis. Thus, we find that  $\langle y \rangle_{2n} \in \mathfrak{f}$  for any  $y \in \mathcal{C}$ . A similar proof shows that  $\langle y \rangle_{2n+1}$  also resides in  $\mathfrak{f}$ , proving that  $\mathfrak{F}_k \subseteq \mathfrak{f}$  for every  $k$ . Since  $\mathfrak{f}$  is also a linear space, the direct sum of these spaces,  $\mathfrak{G} = \bigoplus_{k \in \mathbb{Z}_+} \mathfrak{F}_k$ , is also contained in it. This completes our proof,  $\mathfrak{f} = \mathfrak{G}$ .  $\square$

We note that, in general, the Lie algebra  $\mathfrak{f}$  is not a subalgebra of the Lie idealiser  $\mathcal{I}$  since  $d^2$  need not be in  $\mathcal{I}$  for every  $d \in \mathcal{I}$ . We also remind the reader that

$$\mathfrak{g} = \text{LA}\{V, \partial_x^2\} \subseteq \text{LA}(\mathcal{C} \cup \mathcal{P}(d)) = \mathfrak{f}.$$

Of less immediate and practical interest to us is the fact that the Jordan polynomials in  $d$  also form a Jordan algebra,  $(\mathfrak{G}, \bullet, +)$ . This follows directly and trivially from the fact that they form an associate algebra  $(\mathfrak{G}, \cdot, +)$  (Lemma 5.2.1). More interestingly, this Jordan algebra is also  $\mathbb{Z}_2$ -graded since, along similar lines to (5.28),

$$\langle x \rangle_k \bullet \langle y \rangle_l = \sum_{n=0}^{k+l} \sum_{i=0}^n \gamma_{n,i}^{k,l} \langle D^i x D^{n-i} y \rangle_{k+l-n}, \quad (5.35)$$

where  $\gamma_{n,i}^{k,l} = (\pi_{n,i}^{k,l} + \pi_{n,n-i}^{l,k})/2$  vanishes for odd values of  $n$  due to Lemma 5.3.1, while  $\gamma_{2n,i}^{k,l} = \pi_{2n,i}^{k,l}$  survives. Corresponding observations about the Jordan algebra generated by



$\mathcal{C}$  and  $\mathcal{P}(d)$  can also be made along similar lines, however a property corresponding to height reduction of Corollary 5.3.5 does not follow.

## 5.4 Tables of coefficients

The coefficients  $\pi_{n,i}^{k,l}$  appearing in Lemma 5.2.1 for  $k+l$  ranging between 1 and 4 are presented in Table 5.1. The case  $\langle x \rangle_0 \cdot \langle y \rangle_0 = \langle xy \rangle_0$  is trivial and not listed. From

$(k, l)$	$n$	$\pi_{n,0}^{k,l}$	$\pi_{n,1}^{k,l}$	$\pi_{n,2}^{k,l}$	$\pi_{n,3}^{k,l}$	$\pi_{n,4}^{k,l}$
(1, 0)	0	1				
	1	1/2	0			
(2, 0)	0	1				
	1	1	0			
	2	0	-1/2	0		
(1, 1)	0	1				
	1	1/2	-1/2			
	2	-1/4	-3/4	-1/4		
(3, 0)	0	1				
	1	3/2	0			
	2	0	-3/2	0		
	3	-1/4	-3/4	0	0	
(2, 1)	0	1				
	1	1	-1/2			
	2	-1/2	2	-1/2		
	3	-1/4	-1/2	0	0	
(4, 0)	0	1				
	1	2	0			
	2	0	-3	0		
	3	-1	-3	0	0	
	4	0	1/2	3/2	1/2	0
(3, 1)	0	1				
	1	3/2	-1/2			
	2	-3/4	-15/4	-3/4		
	3	-1	-9/4	0	0	
	4	1/8	1	15/8	7/8	1/8
(2, 2)	0	1				
	1	1	-1			
	2	-1	-4	-1		
	3	-1/2	-1	1	1/2	
	4	1/4	5/4	9/4	5/4	1/4

**Table 5.1:** A table of the coefficients  $\pi_{n,i}^{k,l}$ ,  $n \in \{0, \dots, k+l\}$ ,  $i \in \{0, \dots, n\}$ , which appear in Lemma 5.2.1.

Lemma 5.3.1 we know that  $\pi_{n,i}^{k,l} = (-1)^n \pi_{n,n-i}^{l,k}$ , so that specifying the rows (1, 3) and

$(0, 4)$ , for instance, would be redundant.

In Table 5.2 we present the coefficients  $\lambda_{n,i}^{k,l}$  appearing in Theorem 5.3.2 for  $k+l$  ranging between 1 and 6, while noting that the relation  $\lambda_{n,i}^{k,l} = -\lambda_{n,2n+1-i}^{l,k}$  makes redundant the need to specify coefficients when  $k$  and  $l$  exchange values. The values of the coefficients for  $(k, l) = (0, 2)$  can be inferred from the row  $(k, l) = (2, 0)$ , for instance. Since  $\lambda_{n,i}^{k,l} = 2\pi_{2n+1,i}^{k,l}$ , the first eight rows can be read directly by doubling the corresponding rows in Table 5.1. Note that  $[\langle x \rangle_0, \langle y \rangle_0] = 0$ , and the case  $k+l = 0$  doesn't merit a mention in the table.

Substituting the coefficients from Table 5.2, row  $(k, l) = (2, 1)$  into (6.2) we find,

$$[\langle x \rangle_2, \langle y \rangle_1] = \langle -\frac{1}{2}x D^3 y - Dx D^2 y \rangle_0 + \langle 2x Dy - Dx y \rangle_2,$$

while from the row  $(k, l) = (3, 2)$  we can compute

$$\begin{aligned} [\langle x \rangle_3, \langle y \rangle_2] &= \langle 3x Dy - 2Dx y \rangle_0 \\ &+ \langle -(7/2)x D^3 y - (15/2)Dx D^2 y + 3D^2 x Dy + (3/2)D^3 x y \rangle_1 \\ &+ \langle (3/4)x D^5 y + 3Dx D^4 y + (7/2)D^2 x D^3 y - D^4 x Dy - (1/4)D^5 x y \rangle_2. \end{aligned}$$

Note that these brackets are linear,  $\langle \alpha x + \beta y \rangle_k = \alpha \langle x \rangle_k + \beta \langle y \rangle_k$ .

## 5.5 A finite-dimensional example

While the applications of these algebraic ideas in this thesis, which are first explored in Section 6.1, will be based in infinite-dimensional algebras and function spaces, it is also possible to construct finite-dimensional examples of such structures. Consider the commutative subalgebra of  $2n \times 2n$  matrices,  $\mathcal{A} = M(2n, \mathbb{R})$ ,

$$\mathcal{C} = \left\{ \begin{pmatrix} aI_n & A \\ O_n & aI_n \end{pmatrix} : a \in \mathbb{R}, A \in M(n, \mathbb{R}) \right\},$$

where  $I_n$  is the  $n \times n$  identity matrix and  $O_n$  is the  $n \times n$  zero matrix. Consider any

$$d = \begin{pmatrix} E_{11} & E_{12} \\ O_n & E_{22} \end{pmatrix},$$

where  $E_{ij} \in M(n, \mathbb{R})$ . Then, for every  $x \in \mathcal{C}$ ,

$$[d, x] = \begin{pmatrix} O_n & E_{11}A - AE_{22} \\ O_n & O_n \end{pmatrix} \in \mathcal{C}.$$

Thus,  $d$  is in the Lie idealiser of  $\mathcal{C}$ ,  $D = ad_d$  is a derivation on  $\mathcal{C}$ , and, consequently the Lie algebra

$$\mathfrak{G}^d = \bigoplus_{k \in \mathbb{Z}_+} \{ \langle x \rangle_k^d = \frac{1}{2}(xd^k + d^kx) : x \in \mathcal{C} \}$$

is a  $\mathbb{Z}_2$ -graded matrix Lie algebra with height reduction where commutators can be solved directly.

Similar structures should be found in other commutative subalgebras of  $M(n, \mathbb{R})$ . Applications of these structural observations to linear algebra algorithms are yet to be explored.

$(k, l)$	$n$	$\lambda_{n,0}^{k,l}$	$\lambda_{n,1}^{k,l}$	$\lambda_{n,2}^{k,l}$	$\lambda_{n,3}^{k,l}$	$\lambda_{n,4}^{k,l}$	$\lambda_{n,5}^{k,l}$
(1, 0)	0	1	0				
(2, 0)	0	2	0				
(1, 1)	0	1	-1				
(3, 0)	0	3	0				
	1	-1/2	-3/2	0	0		
(2, 1)	0	2	-1				
	1	-1/2	-1	0	0		
(4, 0)	0	4	0				
	1	-2	6	0	0		
(3, 1)	0	3	-1				
	1	-2	-9/2	0	0		
(2, 2)	0	2	-2				
	1	-1	-2	2	1		
(5, 0)	0	5	0				
	1	-5	-15	0	0		
	2	1	5	15/2	5/2	0	0
(4, 1)	0	4	-1				
	1	-5	-12	0	0		
	2	1	9/2	6	2	0	0
(3, 2)	0	3	-2				
	1	-7/2	-15/2	3	3/2		
	2	3/4	3	7/2	0	-1	-1/4
(6, 0)	0	6	0				
	1	-10	-30	0	0		
	2	6	30	45	15	0	0
(5, 1)	0	5	-1				
	1	-10	-25	0	0		
	2	6	55/2	75/2	25/2	0	0
(4, 2)	0	4	-2				
	1	-8	-18	4	2		
	2	5	21	26	4	-4	-1
(3, 3)	0	3	-3				
	1	-5	-21/2	21/2	5		
	2	3	12	21/2	-21/2	-12	-3

**Table 5.2:** A table of the coefficients  $\lambda_{n,i}^{k,l}$ ,  $n \in \{0, \dots, (k+l-1)/2\}$ ,  $i \in \{0, \dots, 2n+1\}$ , which appear in Theorem 5.3.2.

## Chapter 6

# Symmetric Zassenhaus splittings

Having laid down most of the groundwork for symmetric Zassenhaus splittings for the semiclassical Schrödinger equation in Chapter 4, we encountered the problem of instability caused by loss of skew-Hermiticity. As it turned out, replacement of odd derivatives in Section 4.3 restored the skew-Hermiticity in the first stage of the Zassenhaus splitting we considered in Section 4.2.3. However, this does not constitute a proof of the effectiveness of the procedure in restoring skew-Hermiticity in arbitrary settings.

In Chapter 5 we investigated the Lie algebra of symmetrised differential operators that start appearing in our workings following replacement of derivatives, albeit in a highly general and abstract manner. In this chapter we will resume the development of the symmetric Zassenhaus splittings, while working directly in the language of these symmetrised differential operators. We will see how the structural properties of these algebras translate into the remarkable features of symmetric Zassenhaus splittings for semiclassical Schrödinger equations. In particular, these structural properties allow us to devise arbitrarily high-order, stable, commutator-free, asymptotic splittings which benefit from height reduction and conservation of unitary evolution, and whose costs grow quadratically with the order.

In Section 6.1, we translate the results of Chapter 5 into a language that is relevant for the Schrödinger equation. The first symmetric Zassenhaus splitting is fully developed in the language of these symmetrised differential operators in Section 6.2, following the algorithms introduced in Chapter 4. A variant splitting is presented in Section 6.3, demonstrating the versatility of the symmetric Zassenhaus algorithm. At this stage the splittings are still in operatorial form. The computation of the exponentials featuring in these schemes, following discretisation, is discussed in Section 6.4.

In Section 6.5, we consider Zassenhaus splittings for  $h = \mathcal{O}(\varepsilon^\sigma)$  for  $\sigma \leq 1$ . In Section 6.6, 6.7 and 6.8 we consider some consequences of height reduction on the termination of the Zassenhaus algorithm and the cost of these splittings—both local as well as global. In Section 6.9, we formally prove the numerical stability of these schemes before presenting numerical results in Section 6.10.

## 6.1 Working with symmetrised differential operators

Applications of the algebras of Chapter 5 in solving partial differential equations such as the time-dependent Schrödinger equation will arise from considering function spaces as commutative algebras and the algebra of linear differential operators acting on these spaces as our associative algebra. With this motivation, we will restrict our attention to cases where the commutative algebra  $\mathcal{C}$  is isomorphic to a function space.

In Section 6.1.1 and Section 6.1.2, we will pursue a very formal and pedantic notation in order to see how the results of the previous chapter translate to the case of the Schrödinger equation. However, since much of this intuition has already been given in the previous chapter and since the following narrative is highly pedantic, not much will be lost by simply recalling the results of Section 5.3 and considering the case of  $d = \partial_x$  as a special case. A reader may, therefore, safely assume that the symmetrised differential operators of the form

$$\langle f \rangle_k = \frac{1}{2} \left( f \circ \partial_x^k + \partial_x^k \circ f \right)$$

inherit all the nice structural properties discussed in Section 5.3 and proceed directly to Section 6.1.3.

### 6.1.1 The most general case of interest

Let  $\mathcal{C}$  be isomorphic to  $\mathcal{G}$  and  $\Psi$  be the isomorphism between them. Anticipating the case where  $\mathcal{G}$  is a function space, we will use  $f, g$  to denote its elements. For  $f \in \mathcal{G}$ , we use the notation  $\Psi_f$  and  $\Psi(f)$  interchangeably for the corresponding element in  $\mathcal{C}$ .

We recall that for an element  $d \in \mathcal{J}$  in the Lie idealiser of  $\mathcal{C}$ ,  $D = \text{ad}_d$  is a derivation on  $\mathcal{C}$ . For any  $f \in \mathcal{G}$ ,

$$D(\Psi_f) = [d, \Psi_f] \in \mathcal{C}$$

so that

$$\tilde{d}(f) := \Psi^{-1}(D(\Psi_f)) \in \mathcal{G}.$$

Since  $D$  is a derivation on  $\mathcal{C}$ , and  $\mathcal{C}$  is isomorphic to  $\mathcal{G}$ ,  $\tilde{d}$  is a derivation on  $\mathcal{G}$ ,

$$\begin{aligned} \tilde{d}(fg) &= \Psi^{-1}(D(\Psi_{fg})) \\ &= \Psi^{-1}(D(\Psi_f \Psi_g)) \\ &= \Psi^{-1}(D(\Psi_f) \Psi_g + \Psi_f D(\Psi_g)) \\ &= \Psi^{-1}(D(\Psi_f))g + f \Psi^{-1}(D(\Psi_g)) \\ &= \tilde{d}(f)g + f \tilde{d}(g). \end{aligned}$$

We say that  $\tilde{d}$  is a derivation induced by  $d$ . The element of the Lie idealiser that we will require in the following sections will be the differential operator  $d = \partial_x$ , which will end up

coinciding with the induced derivation  $\tilde{d}$ .

The isomorphism  $\Psi$  that will concern us while solving the Schrödinger equation will be  $\mathcal{M}$ , the left multiplication map. In the case of the Wigner equation, the isomorphism  $\Psi$  is more complicated and maps functions to pseudo-differential operators. The study of the Wigner equation is beyond the scope of this thesis. However, we consider the algebraic behaviour of the operators appearing in this equation in Section 10.4.3, which has formed the motivation for presenting the discourse leading up to here in a more abstract manner.

### 6.1.2 A more specific case

Let  $\mathcal{G} \subseteq \mathcal{H}$  be real and complex valued function spaces, respectively. Endomorphisms on  $\mathcal{H}$  form an associative algebra  $\mathcal{A} = (\text{End}(\mathcal{H}), \circ, +)$  where  $\circ$  is operator composition. For convenience, we consider  $\mathcal{G} = C_p^\infty([-1, 1], \mathbb{R})$  and  $\mathcal{H} = C_p^\infty([-1, 1], \mathbb{C})$ , the space of smooth periodic functions over  $[-1, 1]$  with values in  $\mathbb{R}$  and  $\mathbb{C}$ , respectively. Let  $\mathcal{M} : \mathcal{G} \rightarrow \text{End}(\mathcal{H})$  be the left multiplication map ,

$$\mathcal{M}_f(g) = fg \in \mathcal{H}, \quad f \in \mathcal{G}, g \in \mathcal{H}.$$

We write  $\mathcal{M}_f$  to denote the operator  $\mathcal{M}(f)$  whose action is that of multiplying by  $f$ . It is not uncommon to denote the map  $\mathcal{M}_f$  simply by  $f$ —in fact, this is precisely the way in which we have used  $V$  as an operator of multiplying by the potential.

The image of  $\mathcal{G}$  under  $\mathcal{M}$ ,

$$\mathcal{C} = \mathcal{M}(\mathcal{G}),$$

is a commutative algebra of multiplication operators, with  $\mathcal{M}$  acting as an isomorphism between  $\mathcal{G}$  and  $\mathcal{C}$ . The partial differentiation operator  $\partial_x$  is an element of the Lie idealiser of  $\mathcal{C}$  in the associative operator algebra  $(\text{End}(\mathcal{H}), \circ, +)$ ,

$$[\partial_x, \mathcal{M}_f] = \partial_x \circ \mathcal{M}_f - \mathcal{M}_f \circ \partial_x = \mathcal{M}_{\partial_x f} + \mathcal{M}_f \circ \partial_x - \mathcal{M}_f \circ \partial_x = \mathcal{M}_{\partial_x f} \in \mathcal{C},$$

where the commutator is the canonical Lie product as usual. Following our terminology, the operator  $\partial_x \in \text{End}(\mathcal{H})$  induces the derivation  $\tilde{\partial}_x = \partial_x$  on  $\mathcal{G}$ . Here the induced derivation  $\tilde{d}$  overlaps with the element  $d$ , but this need not be the case in general. To be very precise,  $\tilde{\partial}_x$  is an operator on  $\mathcal{G}$  while  $\partial_x$  is an operator on  $\mathcal{H}$ , but the two coincide on  $\mathcal{G}$ .

Having chosen  $\mathcal{A} = (\text{End}(\mathcal{H}), \circ, +)$ ,  $\mathcal{C} = \mathcal{M}(\mathcal{G})$  and  $d = \partial_x$ , we can define the linear spaces of Jordan monomials in  $\partial_x$  with coefficients in  $\mathcal{C}$ ,

$$\mathfrak{J}_k^{\partial_x} = \left\{ \langle \mathcal{M}_f \rangle_k^{\partial_x} : f \in \mathcal{G} \right\},$$

along the same line as Chapter 5, obtaining the linear space of Jordan polynomials in  $\partial_x$

as the direct sum of these linear spaces,

$$\mathfrak{G}^{\partial_x} = \bigoplus_{k \in \mathbb{Z}_+} \mathfrak{F}_k^{\partial_x}.$$

As usual, since we consider only one element of  $\mathcal{J}$ ,  $d = \partial_x$ , we drop the superscript and write  $\mathfrak{F}_k$ ,  $\mathfrak{G}$  and  $\langle \mathcal{M}_f \rangle_k$  instead of  $\mathfrak{F}_k^{\partial_x}$ ,  $\mathfrak{G}^{\partial_x}$  and  $\langle \mathcal{M}_f \rangle_k^{\partial_x}$ , respectively.

Results of Chapter 5 imply that  $\mathfrak{G}$  is a  $\mathbb{Z}_2$ -graded Lie algebra and Lemma 5.3.6 directly allows us to conclude that

$$\text{LA}(\mathcal{M}(\mathcal{G}) \cup \mathcal{P}(\partial_x)) = \mathfrak{G}.$$

Here  $D^{-1}$  is isomorphic to  $\tilde{d}^{-1}$ , which is the inverse of differentiation on the space  $\mathcal{G}$ . Assuming  $\mathcal{G}$  is closed under integration, therefore, an appropriate inverse of  $D$  exists and the Lie algebra generated by multiplicative operators  $\mathcal{M}(\mathcal{G})$  and polynomials (with constant coefficients) in the differential operators,  $\mathcal{P}(\partial_x)$ , is characterised completely by  $\mathfrak{G}$ , the Lie algebra of Jordan polynomials in  $\partial_x$  with coefficients in  $\mathcal{G}$ .

### 6.1.3 Schrödinger equation in the language of $\mathfrak{G}$

The symmetrised differential operators are written using the notation of the monomials,

$$\langle f \rangle_k = f \bullet \partial_x^k = \frac{1}{2} \left( f \circ \partial_x^k + \partial_x^k \circ f \right).$$

In this notation

$$\langle 1 \rangle_k = \partial_x^k, \quad \langle f \rangle_0 = f,$$

where the unit 1 in  $\langle 1 \rangle_k$  is the constant function over  $\mathcal{G}$  with value 1.

In our new notation, the semiclassical Schrödinger equation (3.9),

$$\partial_t u = i\varepsilon \partial_x^2 u - i\varepsilon^{-1} V u,$$

can be written in the form

$$\partial_t u = i(\varepsilon \langle 1 \rangle_2 - \varepsilon^{-1} \langle V \rangle_0) u. \tag{6.1}$$

Recall that  $\mathfrak{G}$  is the space of all symmetrised differential operators and  $\mathfrak{F}_k$  constitutes all the monomials of degree  $k$ . We recall other useful definitions from Chapter 5,

$$\mathfrak{G}_n = \bigoplus_{k \leq n} \mathfrak{F}_k,$$



which shows  $\mathfrak{G} = \lim_{n \rightarrow \infty} \mathfrak{G}_n$  as a filtered Lie algebra, and

$$\mathfrak{e} = \bigoplus_{k \geq 0} \mathfrak{F}_{2k}, \quad \mathfrak{o} = \bigoplus_{k \geq 0} \mathfrak{F}_{2k+1},$$

which form the grade 1 and grade 0 components, respectively, of  $\mathfrak{G} = \mathfrak{o} \oplus \mathfrak{e}$  seen as a  $\mathbb{Z}_2$ -graded Lie algebra. As we had discussed in the previous chapters, and which is easy to verify, for a real valued  $f$ , the monomial  $\langle f \rangle_k$  is symmetric if  $k$  is even and skew-symmetric otherwise. We define the direct sum of the skew-Hermitian spaces  $\mathfrak{o}$  and  $\mathfrak{i}\mathfrak{e}$ ,

$$\mathfrak{H} = \mathfrak{o} \oplus \mathfrak{i}\mathfrak{e} = \bigoplus_{k \in \mathbb{Z}_+} \{\mathfrak{i}^{k+1} \langle f \rangle_k : f \in \mathcal{G}\}.$$

Components of  $\mathfrak{H}$  are, naturally, skew-Hermitian operators. The structure of  $\mathfrak{o}$  and  $\mathfrak{e}$  means that  $\mathfrak{H}$  also turns out to be a  $\mathbb{Z}_2$ -graded Lie algebra that inherits all the favourable features of  $\mathfrak{G}$ .

Our computations will be performed in  $\mathfrak{H}$ , whereby we trivially preserve skew-Hermiticity. However, since the properties of height reduction (where the height (Definition 5.3.3) is the degree of the symmetrised differential operator) and solution of commutators are inherited from  $\mathfrak{G}$ , where such considerations are the dominating factor we can ignore the multiplication by the imaginary unit and the analysis can be carried out in  $\mathfrak{G}$ .

#### 6.1.4 Consequences for exponential splitting schemes

Following the strategy pursued in Chapter 4, we can find the solution of the Schrödinger equation (6.1), written in the symmetrised differential operator form, by formally exponentiating the Hamiltonian without discretisation,

$$u(h) = \exp(\mathfrak{i}h\varepsilon \langle 1 \rangle_2 - \mathfrak{i}h\varepsilon^{-1} \langle V \rangle_0) u(0).$$

We may now perform splittings directly on this undiscretised Hamiltonian. In Chapter 4 we found that, working in the Lie algebra  $\mathfrak{G}$ , we could simplify commutators and achieve height reduction. However, we encountered instability when the elements of  $\mathfrak{G}$  were discretised in a straightforward manner. In this subsection we briefly consider how the Jordan polynomials in  $\partial_x$  will allow us to simplify commutators, achieve height reduction, while maintaining stability under straightforward discretisation. These observations are pursued in more details in Section 6.2, 6.6 and 6.9, respectively.

#### Simplification of commutators

Methods such as the symmetric Zassenhaus splitting of Section 4.1.1, which result in the appearance of nested matrix commutators such as  $[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{K}^2]$ , feature corresponding

commutators of operators  $\langle 1 \rangle_2$  and  $\langle V \rangle_0$ , such as  $[[\langle V \rangle_0, \langle 1 \rangle_2], \langle 1 \rangle_2]$  when splitting the undiscretised Hamiltonian. Since the two operators reside in our Lie algebra,

$$\langle 1 \rangle_2, \langle V \rangle_0 \in \mathfrak{G} = \bigoplus_{k \in \mathbb{Z}_+} \{ \langle f \rangle_k : f \in \mathcal{G} \},$$

(i.e. they are symmetrised differential operators), so do all of their commutators. Moreover, as a consequence of Theorem 5.3.2, commutators in this Lie algebra can be solved explicitly to other symmetrised differential operators using the rule

$$[\langle f \rangle_k, \langle g \rangle_l] = \sum_{n=0}^{\frac{k+l-1}{2}} \sum_{i=0}^{2n+1} \lambda_{n,i}^{k,l} \langle (\partial_x^i f)(\partial_x^{2n+1-i} g) \rangle_{k+l-2n-1}, \quad (6.2)$$

where  $\lambda$ s are the same as in Chapter 5. This allows us to design commutator-free methods by explicitly working out nested commutators such as  $[[\langle V \rangle_0, \langle 1 \rangle_2], \langle 1 \rangle_2]$  and  $[[\langle V \rangle_0, \langle 1 \rangle_2], \langle V \rangle_0]$  appearing in the sBCH of  $ih\varepsilon \langle 1 \rangle_2$  and  $-ih\varepsilon^{-1} \langle V \rangle_0$  using the rules,

$$\begin{aligned} [\langle f \rangle_2, \langle g \rangle_1] &= -\langle (\partial_x f)(\partial_x^2 g) + \tfrac{1}{2}f(\partial_x^3 g) \rangle_0 + \langle 2f(\partial_x g) - (\partial_x f)g \rangle_2, \\ [\langle f \rangle_2, \langle g \rangle_0] &= 2\langle f(\partial_x g) \rangle_1, \\ [\langle f \rangle_1, \langle g \rangle_0] &= \langle f(\partial_x g) \rangle_0, \end{aligned}$$

which can be read off Table 5.2 in Section 5.4. We easily deduce

$$\begin{aligned} [\langle V \rangle_0, \langle 1 \rangle_2] &= -2\langle \partial_x V \rangle_1, \\ [[\langle V \rangle_0, \langle 1 \rangle_2], \langle V \rangle_0] &= -2\langle (\partial_x V)^2 \rangle_0, \\ [[\langle V \rangle_0, \langle 1 \rangle_2], \langle 1 \rangle_2] &= -\langle \partial_x^4 V \rangle_0 + 4\langle \partial_x^2 V \rangle_2. \end{aligned}$$

The sBCH of  $ih\varepsilon \langle 1 \rangle_2$  and  $-ih\varepsilon^{-1} \langle V \rangle_0$  up to and including grade three terms using the above simplifications is

$$ih\varepsilon \langle 1 \rangle_2 - ih\varepsilon^{-1} \langle V \rangle_0 - \tfrac{1}{6}ih^3\varepsilon \langle \partial_x^2 V \rangle_2 + \tfrac{1}{24}ih^3\varepsilon \langle \partial_x^4 V \rangle_0 - \tfrac{1}{6}h^3\varepsilon^{-1} \langle (\partial_x V)^2 \rangle_0.$$

Such simplifications are used in Section 6.2 to derive the symmetric Zassenhaus splitting of the first kind.

### Height reduction

We recall that we need  $M = \mathcal{O}(\varepsilon^{-1})$  grid points to resolve spatial oscillations. As we have discussed previously, the spectral radius of the differentiation matrix  $\mathcal{K}$  scales as  $\mathcal{O}(M)$ ,

i.e.  $\mathcal{O}(\varepsilon^{-1})$ . Keeping eventual discretisation in mind and noting that

$$\langle f \rangle_k \rightsquigarrow \frac{1}{2}(\mathcal{D}_f \mathcal{K}^k + \mathcal{K}^k \mathcal{D}_f) = \mathcal{O}(\varepsilon^{-k}),$$

we use the shorthand<sup>1</sup>

$$\langle f \rangle_k = \mathcal{O}(\varepsilon^{-k}).$$

It is now easy to see how the height (or degree) of a Jordan polynomial  $W \in \mathfrak{G}_n$  acts as a proxy for the spectral radius of the eventual discretisation,

$$\text{ht}(W) = n \implies W = \mathcal{O}(\varepsilon^{-n}).$$

As we discussed in Section 4.2.2, naïve estimates using typical commutator bounds suggest that the commutator  $[[\mathcal{D}_V, \mathcal{K}^2], \mathcal{K}^2]$  should scale as  $\mathcal{O}(\varepsilon^{-4})$  since the matrix  $\mathcal{K}^2$  scales as  $\mathcal{O}(\varepsilon^{-2})$ . However, the property of height reduction in  $\mathfrak{G}$  (Corollary 5.3.5) follows the observations of height reduction in  $\mathfrak{S}$  (Lemma 4.2.2) made in Section 4.2.2, and the undiscretised commutator  $[[\langle V \rangle_0, \langle 1 \rangle_2], \langle 1 \rangle_2]$  scales as  $\mathcal{O}(\varepsilon^{-2})$ .

In Section 6.6 and 6.7 we see how the property of height reduction plays a role in the finite termination of the Zassenhaus algorithm and leads to the quadratic growth of costs of the resulting splittings.

### Stability

The non-mixing property is an important consequence of Theorem 5.3.2. Due to this property, all terms in the sBCH of  $i\hbar\varepsilon \langle 1 \rangle_2$  and  $-i\hbar\varepsilon^{-1} \langle V \rangle_0$  reside in the Lie algebra

$$\mathfrak{H} = \mathfrak{o} \oplus i\mathfrak{c} = \bigoplus_{k \in \mathbb{Z}_+} \{i^{k+1} \langle f \rangle_k : f \in \mathcal{G}\},$$

whose elements are skew-Hermitian operators. A careful reader would note that if it were not for  $[\mathfrak{o}, \mathfrak{o}] \subseteq \mathfrak{o}$ , we might have featured terms such as the Hermitian term  $\langle f \rangle_2$  (note the lack of a leading imaginary unit) when expanding commutators of elements in  $\mathfrak{H}$ , for instance.

Upon discretisation using spectral collocation methods, these elements of  $\mathfrak{H}$  are replaced by

$$i^{k+1} \langle f \rangle_k \rightsquigarrow \frac{1}{2}i^{k+1}(\mathcal{D}_f \mathcal{K}^k + \mathcal{K}^k \mathcal{D}_f),$$

where  $\mathcal{D}_f$  is a diagonal matrix and  $\mathcal{K}$  a skew-symmetric circulant differentiation matrix. Consequently, the discretised forms of elements of  $\mathfrak{H}$  are skew-Hermitian matrices, the exponentials of which are unitary matrices.

Once Zassenhaus splittings are carried out in this algebra, they feature exponentials

<sup>1</sup>In Section 9.1 we will give meaning to this notation without reference to discretisation.

of terms in  $\mathfrak{H}$ , thereby guaranteeing unitary evolution of the wave-function and resulting in unconditional stability of these numerical methods. Concerns of stability are fully addressed in Section 6.9 once the details of the symmetric Zassenhaus splittings have been fully fleshed out.

### 6.1.5 Relation to other algebras

The gain of powers of  $\varepsilon$  with commutation (a consequence of height reduction) has also been observed by Gaim & Lasser (2014) using Moyal brackets in the phase space. However, the analysis is not as easily generalised as the algebraic approach here. Moreover, the structures involved are not symmetrised and are reminiscent of the algebra  $\mathfrak{S}$  of Chapter 4.

These ideas are closely related to the Weyl algebra (Coutinho 1997, Dixmier 1968) which is the universal enveloping algebra of the Lie algebra of Heisenberg groups. The univariate Weyl algebra can be written in the form

$$\mathfrak{W} = \left\{ \sum_{k=0}^n p_k(x) \partial_x^k : p_k \in \mathcal{P}(x), k = 0, \dots, n, n \in \mathbb{Z}_+ \right\}.$$

These are special cases of the more general form discussed in (Bader, Iserles, Kropielnicka & Singh 2014)

$$\mathfrak{S} = \left\{ \sum_{k=0}^n f_k(x) \partial_x^k : f_k \in \mathcal{G}, k = 0, \dots, n, n \in \mathbb{Z}_+ \right\},$$

under the choice of polynomials in  $x$ ,  $\mathcal{P}(x)$ , as the function space  $\mathcal{G}$ .

We note that  $\mathfrak{S}$  is contained within  $\mathfrak{G}$  as an associative algebra since  $f_k, \partial_x^k \in \mathfrak{G}$ , while, in the other direction,  $\mathfrak{G} \subseteq \mathfrak{S}$  is evident using the Leibniz rule. Thus, the two are identical. The analysis in Chapter 4 is carried out in this algebra and we note that height reduction has also been proven for  $\mathfrak{S}$  in Lemma 4.2.2. However, we remind the reader that the elements of  $\mathfrak{S}$  are not symmetrised. Their non-symmetry led to instability when discretised in a straightforward manner in Zassenhaus splittings (see Figure 4.1). It was this very drawback that forced us to seek more symmetrised structures in the form of Jordan polynomials whose direct discretisation preserves skew-Hermiticity.

Recalling the discourse in Section 4.3, one could argue that any nested commutator of skew-Hermitian operators in  $\mathfrak{S}$  should be skew-Hermitian. As we had noted, this is indeed true prior to discretisation. However, unless some highly specialised differentiation matrices can be constructed, the discretised version of the commutator simplified in  $\mathfrak{S}$  possesses no such structure and, as seen in Figure 4.1, this leads to an exponential blow up and numerical instability even in the simplest case.

What is more, even prior to discretisation we prefer to discard terms smaller than a certain size (while analysing in powers of  $\varepsilon$ ). Working in  $\mathfrak{S}$  instead of  $\mathfrak{G}$ , it becomes

difficult to discern which components of a term such as  $\sum_{k=0}^n f_k(x) \partial_x^k$  can be discarded and which need to be kept despite their small size in order to preserve the skew-Hermiticity of the undiscretised operator.

There is a significant advantage to working in the symmetrised form of  $\mathfrak{G}$  since  $\mathfrak{S}$  lacks an obvious  $\mathbb{Z}_2$ -graded structure. The  $\mathbb{Z}_2$  grading on  $\mathfrak{G}$  and the symmetric nature of its elements ensures that elements of  $\mathfrak{H}$  are skew-Hermitian after discretisation. This proves crucial for devising stable numerical schemes such as the symmetric Zassenhaus schemes once we start utilising nested commutators. In contrast, lacking a  $\mathbb{Z}_2$  grading, a clean separation of terms does not occur in the form  $\mathfrak{S}$ .

## 6.2 The first Zassenhaus splitting

All necessary tools for developing symmetric Zassenhaus splittings for the semiclassical Schrödinger equation are now available and we dedicate this section to illustrating how to compute the splitting (4.1), whose development we had left incomplete at the end of Chapter 4.

We present here the identities of  $\mathfrak{G}$  which suffice for simplifying all commutators appearing in this chapter,

$$\begin{aligned}
 [\langle f \rangle_4, \langle g \rangle_0] &= 4 \langle f(\partial_x g) \rangle_3 - 2 \langle 3(\partial_x f)(\partial_x^2 g) + f(\partial_x^3 g) \rangle_1, \\
 [\langle f \rangle_3, \langle g \rangle_0] &= 3 \langle f(\partial_x g) \rangle_2 - \frac{1}{2} \langle 3(\partial_x f)(\partial_x^2 g) + f(\partial_x^3 g) \rangle_0, \\
 [\langle f \rangle_2, \langle g \rangle_2] &= 2 \langle f(\partial_x g) - (\partial_x f)g \rangle_3 + \langle 2(\partial_x^2 f)(\partial_x g) - 2(\partial_x f)(\partial_x^2 g) + (\partial_x^3 f)g - f(\partial_x^3 g) \rangle_1, \\
 [\langle f \rangle_2, \langle g \rangle_1] &= \langle 2f(\partial_x g) - (\partial_x f)g \rangle_2 - \frac{1}{2} \langle 2(\partial_x f)(\partial_x^2 g) + f(\partial_x^3 g) \rangle_0, \\
 [\langle f \rangle_2, \langle g \rangle_0] &= 2 \langle f(\partial_x g) \rangle_1, \\
 [\langle f \rangle_1, \langle g \rangle_1] &= \langle f(\partial_x g) - (\partial_x f)g \rangle_1, \\
 [\langle f \rangle_1, \langle g \rangle_0] &= \langle f(\partial_x g) \rangle_0.
 \end{aligned} \tag{6.3}$$

These can be read off Table 5.2.

As a matter of convenience, we will often write  $f$  in place of  $\langle f \rangle_0$  and  $\partial_x^k$  instead of  $\langle 1 \rangle_k$ . This will help a reader immediately identify the components that can be exponentiated directly (in the case of  $f$ ) or via FFTs (in the case of  $\partial_x^k$ ).

We recall that we were working under the scaling  $\sigma = 1$  or  $h = \mathcal{O}(\varepsilon)$ , and commence with the splitting using the symmetric Zassenhaus algorithm in Table 6.1.

We have started the algorithm with

$$W^{[0]} = -i h \varepsilon^{-1} V.$$

---

**Symmetric Zassenhaus Algorithm**

---

$\mathcal{W}^{[0]} = i\hbar(\varepsilon\partial_x^2 - \varepsilon^{-1}V)$   
 $W^{[0]} = -i\hbar\varepsilon^{-1}V$   
 $s := 0$   
**do**  
     $s := s + 1$   
    compute  $\mathcal{W}^{[s]} = \text{sBCH}(-W^{[s-1]}, \mathcal{W}^{[s-1]})$   
    expand result in powers of  $\varepsilon$   
    discard terms of size  $\mathcal{O}(\varepsilon^{2s_{\max}+2})$   
    define  $W^{[s]} = \mathcal{O}(\varepsilon^{2s-2})$ , s.t.  $W^{[s]} - \mathcal{W}^{[s]} = \mathcal{O}(\varepsilon^{2s})$   
**while**  $s < \text{desired order } s_{\max}$

Resulting method:

$$e^{\mathcal{W}^{[0]}} = e^{W^{[0]}/2} e^{W^{[1]}/2} \dots e^{\mathcal{W}^{[s_{\max}]}} \dots e^{W^{[1]}/2} e^{W^{[0]}/2} + \mathcal{O}(\varepsilon^{2s_{\max}+2})$$


---

**Table 6.1:** The symmetric Zassenhaus splitting algorithm

Simplifying using the rules (6.3), we find that

$$\begin{aligned}
\mathcal{W}^{[1]} = & \overbrace{i\hbar\varepsilon\partial_x^2}^{\mathcal{O}(\varepsilon^0)} - \overbrace{\frac{1}{12}i\hbar^3\varepsilon^{-1}(\partial_x V)^2 - \frac{1}{3}i\hbar^3\varepsilon\langle\partial_x^2 V\rangle_2}^{\mathcal{O}(\varepsilon^2)} \\
& - \overbrace{\frac{1}{60}i\hbar^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 + \frac{1}{12}i\hbar^3\varepsilon(\partial_x^4 V)}^{\mathcal{O}(\varepsilon^4)} \\
& + \overbrace{\frac{1}{90}i\hbar^5\varepsilon\langle 8(\partial_x^2 V)^2 - (\partial_x^3 V)(\partial_x V) \rangle_2}^{\mathcal{O}(\varepsilon^4)} \\
& - \overbrace{\frac{1}{45}i\hbar^5\varepsilon^{-3}\langle\partial_x^4 V\rangle_4}^{\mathcal{O}(\varepsilon^4)} + \mathcal{O}(\varepsilon^6).
\end{aligned}$$

To progress to the second stage, we choose to eliminate the lowest  $\varepsilon$ -order term,

$$W^{[1]} = i\hbar\varepsilon\partial_x^2,$$

from  $\mathcal{W}^{[1]}$  by computing the sBCH

$$\begin{aligned}
 \mathcal{W}^{[2]} &= \text{sBCH}(-W^{[1]}, \mathcal{W}^{[1]}) \\
 &= -W^{[1]} + \mathcal{W}^{[1]} - \frac{1}{24}[[\mathcal{W}^{[1]}, W^{[1]}], W^{[1]}] - \frac{1}{12}[[\mathcal{W}^{[1]}, W^{[1]}], \mathcal{W}^{[1]}] + \mathcal{O}(\varepsilon^6). \\
 &= \overbrace{-\frac{1}{12}ih^3\varepsilon^{-1}(\partial_x V)^2 - \frac{1}{3}ih^3\varepsilon\langle\partial_x^2 V\rangle_2}^{\mathcal{O}(\varepsilon^2)} \\
 &\quad - \overbrace{\frac{1}{60}ih^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 + \frac{1}{12}ih^3\varepsilon(\partial_x^4 V)}^{\mathcal{O}(\varepsilon^4)} \\
 &\quad + \overbrace{\frac{1}{60}ih^5\varepsilon\langle 7(\partial_x^2 V)^2 + (\partial_x^3 V)(\partial_x V)\rangle_2}^{\mathcal{O}(\varepsilon^4)} \\
 &\quad + \overbrace{\frac{1}{30}ih^5\varepsilon^{-3}\langle\partial_x^4 V\rangle_4}^{\mathcal{O}(\varepsilon^4)} + \mathcal{O}(\varepsilon^6).
 \end{aligned}$$

In the next iteration, we pull out the  $\mathcal{O}(\varepsilon^2)$  term,

$$W^{[2]} = -\frac{1}{12}ih^3\varepsilon^{-1}(\partial_x V)^2 - \frac{1}{3}ih^3\varepsilon\langle\partial_x^2 V\rangle_2$$

and need to compute  $\mathcal{W}^{[3]}$ . The  $\mathcal{O}(\varepsilon^2)$  terms in  $\mathcal{W}^{[2]}$  are precisely  $W^{[2]}$  and, consequently, they cancel out in the commutator  $[W^{[2]}, \mathcal{W}^{[2]}]$ . The terms in  $\mathcal{W}^{[2]}$  whose commutation with  $W^{[2]}$  does not trivially vanish are  $\mathcal{O}(\varepsilon^4)$  in size. As a consequence of height reduction, the commutator gains a power of  $\varepsilon$  and  $[W^{[2]}, \mathcal{W}^{[2]}] = \mathcal{O}(\varepsilon^7)$ . Since this term is smaller than our error tolerance, commutators can be disregarded to obtain  $\mathcal{W}^{[3]}$  simply as the  $\mathcal{O}(\varepsilon^4)$  components of  $\mathcal{W}^{[2]}$ . The asymptotic splitting is therefore

$$\mathcal{Z}_{2,1}^{[1]} = e^{\frac{1}{2}W^{[0]}} e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[2]}} e^{\mathcal{W}^{[3]}} e^{\frac{1}{2}W^{[2]}} e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[0]}}, \quad (6.4)$$

where

$$\begin{aligned}
 W^{[0]} &= -ih\varepsilon^{-1}V = \mathcal{O}(\varepsilon^0), \\
 W^{[1]} &= ih\varepsilon\partial_x^2 = \mathcal{O}(\varepsilon^0), \\
 W^{[2]} &= -\frac{1}{12}ih^3\varepsilon^{-1}(\partial_x V)^2 - \frac{1}{3}ih^3\varepsilon\langle\partial_x^2 V\rangle_2 = \mathcal{O}(\varepsilon^2), \\
 \mathcal{W}^{[3]} &= -\frac{1}{60}ih^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 + \frac{1}{12}ih^3\varepsilon(\partial_x^4 V) \\
 &\quad + \frac{1}{60}ih^5\varepsilon\langle 7(\partial_x^2 V)^2 + (\partial_x^3 V)(\partial_x V)\rangle_2 \\
 &\quad + \frac{1}{30}ih^5\varepsilon^{-3}\langle\partial_x^4 V\rangle_4 = \mathcal{O}(\varepsilon^4).
 \end{aligned} \quad (6.5)$$

The notation  $\mathcal{Z}_{2,1}^{[1]}$  is mostly self-explanatory: the splitting (4.6) is denoted as  $\mathcal{Z}_{s,\sigma}$  and the numbers 2, 1 in the subscript mean  $s = 2$ ,  $\sigma = 1$ . The superscript  $^{[1]}$  stands for an

asymptotic splitting *of the first kind*: in Section 6.3 we consider an alternative splitting (with the initial term  $W^{[0]}$  equalling  $ih\varepsilon\partial_x^2$ ), which we designate as an asymptotic splitting *of the second kind*.

### 6.3 Zassenhaus splitting of the second kind

The motivation for splitting  $\mathcal{W}^{[0]} = ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V$  is down to the structural differences in  $-ih\varepsilon^{-1}V$  and  $ih\varepsilon\partial_x^2$  which makes it easy to exponentiate either separately. There is, however, no reason why we must commence with  $W^{[0]} = -ih\varepsilon^{-1}V$ . In this Section we start with the term  $ih\varepsilon\partial_x^2$ , instead, and arrive at a variant of the splitting  $\mathcal{Z}_{2,1}^{[1]}$ .

Revisiting the narrative of Section 6.2, we start from

$$W^{[0]} = ih\varepsilon\partial_x^2, \quad \mathcal{W}^{[0]} = ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V.$$

This results in

$$\mathcal{W}^{[1]} = \text{sBCH}(-W^{[0]}, \mathcal{W}^{[0]}) = \sum_{j=0}^{\infty} \mathcal{W}_j^{[1]}, \quad \text{where} \quad \mathcal{W}_j^{[1]} = \mathcal{O}(\varepsilon^{2j}),$$

and

$$\begin{aligned} \mathcal{W}_0^{[1]} &= -ih\varepsilon^{-1}V, \\ \mathcal{W}_1^{[1]} &= \frac{1}{6}ih^3\varepsilon^{-1}(\partial_x V)^2 + \frac{1}{6}ih^3\varepsilon\langle\partial_x^2 V\rangle_2, \\ \mathcal{W}_2^{[1]} &= -\frac{1}{24}ih^3\varepsilon(\partial_x^4 V) - \frac{2}{45}ih^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 \\ &\quad + \frac{1}{30}ih^5\varepsilon\langle(\partial_x^2 V)^2 - 2(\partial_x^3 V)(\partial_x V)\rangle_2 \\ &\quad - \frac{1}{120}ih^5\varepsilon^3\langle\partial_x^4 V\rangle_4. \end{aligned}$$

We next remove  $W^{[1]} = \mathcal{W}_0^{[1]} = \mathcal{O}(\varepsilon^0)$  and obtain, with the shorthand  $X = -W^{[1]}$ ,  $Y = \mathcal{W}^{[1]}$ ,

$$\mathcal{W}^{[2]} = \text{sBCH}(X, Y) = X + Y - \frac{1}{24}[[Y, X], X] - \frac{1}{12}[[Y, X], Y] + \mathcal{O}(\varepsilon^6) = \sum_{j=1}^{\infty} \mathcal{W}_j^{[2]},$$

where

$$\begin{aligned} \mathcal{W}_1^{[2]} &= \frac{1}{6}ih^3\varepsilon^{-1}(\partial_x V)^2 + \frac{1}{6}ih^3\varepsilon\langle\partial_x^2 V\rangle_2, \\ \mathcal{W}_2^{[2]} &= -\frac{1}{24}ih^3\varepsilon(\partial_x^4 V) - \frac{7}{120}ih^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 \\ &\quad + \frac{1}{30}ih^5\varepsilon\langle(\partial_x^2 V)^2 - 2(\partial_x^3 V)(\partial_x V)\rangle_2 \\ &\quad - \frac{1}{120}ih^5\varepsilon^3\langle\partial_x^4 V\rangle_4. \end{aligned}$$



Finally,

$$W^{[2]} = \mathcal{W}_1^{[2]}, \quad \mathcal{W}_3 = \mathcal{W}_2^{[2]}.$$

The outcome is the splitting

$$\mathcal{Z}_{2,1}^{[2]} = e^{\frac{1}{2}W^{[0]}} e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[2]}} e^{\mathcal{W}^{[3]}} e^{\frac{1}{2}W^{[2]}} e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[0]}}, \quad (6.6)$$

where

$$W^{[0]} = i\hbar\varepsilon\partial_x^2 = \mathcal{O}(\varepsilon^0), \quad (6.7)$$

$$W^{[1]} = -i\hbar\varepsilon^{-1}V = \mathcal{O}(\varepsilon^0),$$

$$W^{[2]} = \frac{1}{6}i\hbar^3\varepsilon^{-1}(\partial_x V)^2 + \frac{1}{6}i\hbar^3\varepsilon\langle\partial_x^2 V\rangle_2 = \mathcal{O}(\varepsilon^2),$$

$$\begin{aligned} \mathcal{W}^{[3]} = & -\frac{1}{24}i\hbar^3\varepsilon(\partial_x^4 V) - \frac{7}{120}i\hbar^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 \\ & + \frac{1}{30}i\hbar^5\varepsilon\langle(\partial_x^2 V)^2 - 2(\partial_x^3 V)(\partial_x V)\rangle_2 \\ & - \frac{1}{120}i\hbar^5\varepsilon^3\langle\partial_x^4 V\rangle_4 = \mathcal{O}(\varepsilon^4). \end{aligned} \quad (6.8)$$

## 6.4 Computation of exponentials

Once we replace derivatives by differentiation matrices, the evaluation of a single time step  $\mathbf{u}^{n+1} = (\tilde{\mathcal{Z}}_{2,1}^{[2]})\mathbf{u}^n$  of the Zassenhaus splitting (6.6), where  $\tilde{\mathcal{Z}}_{2,1}^{[2]}$  follows from  $\mathcal{Z}_{2,1}^{[2]}$  by spatial discretisation of all exponents, requires in principle seven exponentials.

We note that  $g(\mathcal{D}_f) = \mathcal{D}_{g(f)}$  for any analytic function  $g$ , while diagonalisation of  $\mathcal{K}$  via Fourier transforms,  $\mathcal{K} = \mathcal{F}^{-1}\mathcal{D}_{in\pi}\mathcal{F}$ , implies that  $g(\mathcal{K}) = \mathcal{F}^{-1}\mathcal{D}_{g(in\pi)}\mathcal{F}$ . Consequently, the exponential of  $W^{[0]} = i\hbar\varepsilon\partial_x^2$  is evaluated using two FFTs,

$$e^{\frac{1}{2}W^{[0]}}u \rightsquigarrow \mathcal{F}^{-1}\mathcal{D}_{\exp(-i\hbar\varepsilon n^2\pi^2/2)}\mathcal{F}u,$$

in  $\mathcal{O}(M \log M)$  operations and  $W^{[1]}$  is exponentiated trivially in  $\mathcal{O}(M)$  operations,

$$e^{\frac{1}{2}W^{[1]}}u = \exp\left(-\frac{1}{2}i\varepsilon^{-1}V\right)u \rightsquigarrow \mathcal{D}_{\exp(-\frac{1}{2}i\varepsilon^{-1}V)}u,$$

as we had seen in Section 3.3.4. Inexpensive computation of the exponentials of these terms is an important factor because  $\tilde{W}^{[0]}$  and  $\tilde{W}^{[1]}$ , which follow from the spatial discretisation of  $W^{[0]}$  and  $W^{[1]}$ , respectively, are (spectrally) the largest matrices present. All other exponents in  $\tilde{\mathcal{Z}}_{2,1}^{[2]}$  are  $\mathcal{O}(\varepsilon^2)$  or smaller and, as will be clear later in this section, their computation with Lanczos iterations is very affordable.

Note that the palindromic property allows us to further reduce the number of exponentials if no output at intermediate steps is required. This so-called First-Same-As-Last

(FSAL) property effectively yields a method

$$\tilde{\mathcal{Z}}(\alpha)_{2,1}^{[2]} = e^{\frac{1}{2}\tilde{W}^{[1]}} e^{\frac{1}{2}\tilde{W}^{[2]}} e^{\tilde{\mathcal{W}}^{[3]}} e^{\frac{1}{2}\tilde{W}^{[2]}} e^{\frac{1}{2}\tilde{W}^{[1]}} e^{\frac{1}{2}\alpha\tilde{W}^{[0]}}, \quad (6.9)$$

where the first step has to be calculated with  $\alpha = 1$ , and further steps with  $\alpha = 2$ . Whenever output is required, we apply  $e^{\frac{1}{2}\tilde{W}^{[0]}}$ , and initialise the method by letting  $\alpha = 1$  for the next step. All in all, we only need to compute six exponentials each step, two of which are diagonal matrices, one is circulant and the remaining three can be approximated cheaply by Krylov methods.

Unlike  $W^{[0]}$  and  $W^{[1]}$ , the exponents  $W^{[2]}$  and  $\mathcal{W}^{[3]}$  in splittings (6.4) and (6.6) do not possess a structure that makes them amenable to direct exponentiation. However, they are small— $\mathcal{O}(\varepsilon^2)$  and  $\mathcal{O}(\varepsilon^4)$ , respectively—and approximating their exponentials to  $\mathcal{O}(\varepsilon^6)$  accuracy is relatively inexpensive.

Since these matrices are skew-Hermitian, we resort to Lanczos iterations of Table 2.2 for exponentiating them. As we had seen in Section 2.2.2, exponentiation using this Krylov subspace method involves approximating the exponential of a matrix  $\mathcal{A}$  by the exponential of a small  $m \times m$  upper Hessenberg matrix  $\mathcal{H}_m$  (which happens to be tridiagonal in the case of skew-Hermitian matrices and Lanczos iterations),

$$e^{\mathcal{A}}\mathbf{v} \approx \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v}. \quad (2.23)$$

Here  $m$  is the number of Lanczos iterations.

Typically these methods are effective when  $m \ll M$  and  $e^{\mathcal{H}_m}$  is inexpensive to evaluate. In such cases, the computational cost of exponentiation via these Krylov subspace methods is dominated by the cost of the iterations required for constructing the basis vectors of the Krylov subspace,  $\mathcal{V}_m$ , and the Hessenberg matrix,  $\mathcal{H}_m$ . Each Lanczos iteration involves evaluating a matrix–vector product of the form  $\mathcal{A}\mathbf{v}_j$ .

#### 6.4.1 Evaluating matrix–vector products

In symmetric Zassenhaus splittings the exponents will always be in  $\mathfrak{H}$  and, therefore, of the form

$$W = \sum_{j=0}^k i^{j+1} \langle f_j \rangle_j.$$

The matrices that require exponentiation by Krylov methods will, consequently, involve a symmetric (Jordan) multiplication of a diagonal and a circulant,

$$W = \sum_{j=0}^k i^{j+1} \langle f_j \rangle_j \rightsquigarrow \sum_{j=0}^k i^{j+1} \mathcal{D}_{f_j} \bullet \mathcal{K}^j = \frac{1}{2} \sum_{j=0}^k i^{j+1} (\mathcal{D}_{f_j} \mathcal{K}^j + \mathcal{K}^j \mathcal{D}_{f_j}) = \tilde{W}.$$

The circulants can be diagonalised using Fourier transform matrices  $\mathcal{F}$ ,

$$\frac{1}{2} \sum_{j=0}^k i^{j+1} (\mathcal{D}_{f_j} \mathcal{F}^{-1} \mathcal{D}_{(in\pi)^j} \mathcal{F} + \mathcal{F}^{-1} \mathcal{D}_{(in\pi)^j} \mathcal{F} \mathcal{D}_{f_j}).$$

A few FFTs are therefore unavoidable in each iteration, bringing the overall cost to  $\mathcal{O}(mM \log M)$  operations, where  $m$  is the number of Lanczos iterations.

When attempting to approximate each  $i^{j+1} \langle f_j \rangle_j v$  separately via

$$\frac{1}{2} i^{j+1} (\mathcal{D}_{f_j} \mathcal{F}^{-1} \mathcal{D}_{(in\pi)^j} \mathcal{F} v + \mathcal{F}^{-1} \mathcal{D}_{(in\pi)^j} \mathcal{F} \mathcal{D}_{f_j} v),$$

it seems that we would need four FFTs per  $\langle f_j \rangle_j$ , bringing the total number of FFTs for a term  $W = \sum_{j=0}^k i^{j+1} \langle f_j \rangle_j$  to  $4k$  (note that no FFT is required for  $i \langle f_0 \rangle_0$ ).

In order to evaluate  $\tilde{W} v$  efficiently, we could combine the FFTs more cleverly,

$$i \mathcal{D}_{f_0} v + \frac{1}{2} \left( \sum_{j=1}^k i^{j+1} \mathcal{D}_{f_j} \mathcal{F}^{-1} \mathcal{D}_{(in\pi)^j} \right) \mathcal{F} v + \frac{1}{2} \mathcal{F}^{-1} \left( \sum_{j=1}^k i^{j+1} \mathcal{D}_{(in\pi)^j} \mathcal{F} \mathcal{D}_{f_j} v \right),$$

bringing the cost to  $2k + 2$  FFTs.

In the case of the splittings (6.4) and (6.6), designed for the Schrödinger equation (3.9), the exponents  $W$  are of the form  $i \sum_{j=0}^k \langle f_j \rangle_{2j}$ , so that no odd-indexed term appears. In this case, the number of FFTs required would again be  $2k + 2$ —given the height this means fewer FFTs (note that the height of this term is  $2k$ , not  $k$ ). This is not the case for Zassenhaus splittings for the Magnus expansion that we encounter in Chapter 8.

### 6.4.2 Estimating Lanczos iterations

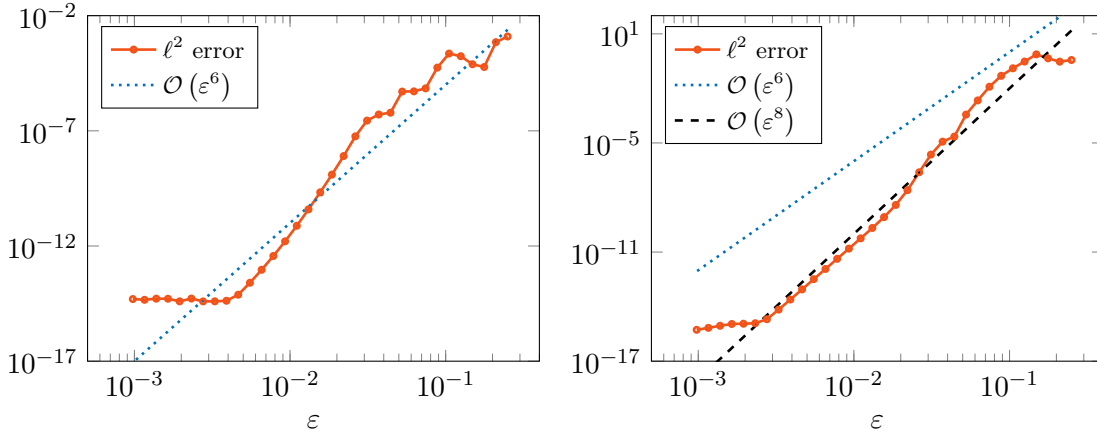
The question of an appropriate number of Lanczos iterations,  $m$ , is answered by the inequality

$$\|e^{\mathcal{A}} v - \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* v\|_2 \leq 12 e^{-\rho^2/(4m)} \left( \frac{e\rho}{2m} \right)^m, \quad m \geq \rho, \quad (2.24)$$

where  $\rho = \rho(\mathcal{A})$  is the spectral radius of  $\mathcal{A}$  (Hochbruck & Lubich 1997). We know that  $\tilde{W}^{[2]} = \mathcal{O}(\varepsilon^2)$  and assume, with very minor loss of generality, that  $\rho(\tilde{W}^{[2]}) \leq c\varepsilon^2$  for some  $c > 0$ . We thus deduce from (2.24) that

$$\|e^{\tilde{W}^{[2]}} v - \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* v\|_2 \leq 12 \left( \frac{ec}{2m} \right)^m \varepsilon^{2m}, \quad m \geq \rho,$$

and  $m = 3$  is sufficient to reduce the error to  $\mathcal{O}(\varepsilon^6)$ , in line with the error of our symmetric Zassenhaus algorithm. This is true provided that  $\rho \leq 3$ , i.e.  $\varepsilon \leq \sqrt{3/c}$ —since we expect  $\varepsilon > 0$  to be very small, this is not much in a way of restriction. Likewise,  $\tilde{W}^{[3]} = \mathcal{O}(\varepsilon^4)$



**Figure 6.1:** The error, compared to the required order  $\mathcal{O}(\varepsilon^6)$ , in computing  $e^{\tilde{W}^{[2]}}\mathbf{v}$  (left) and  $e^{\tilde{\mathcal{W}}^{[3]}}\mathbf{v}$  (right).

and the inequality  $\rho(\tilde{\mathcal{W}}^{[3]}) \leq \tilde{c}\varepsilon^4$  implies that

$$\left\| e^{\tilde{\mathcal{W}}^{[3]}}\mathbf{v} - \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v} \right\|_2 \leq 12 \left( \frac{e\tilde{c}}{2m} \right)^m \varepsilon^{4m}, \quad m \geq \rho$$

and for  $\varepsilon \leq (2/\tilde{c})^{1/4}$  we need just  $m = 2$ . Altogether, we deduce that the computation (consistent with the error of  $\mathcal{O}(\varepsilon^6)$ ) of  $e^{\tilde{W}^{[2]}}\mathbf{v}$  (twice) and  $e^{\tilde{\mathcal{W}}^{[3]}}\mathbf{v}$  in each step of (6.4) or (6.6) cost just  $\mathcal{O}(M \log M)$  operations.

### Numerical examples

Figure 6.1 presents the  $L^2$  error committed in approximating the exponentials  $e^{\tilde{W}^{[2]}}\mathbf{v}$  and  $e^{\tilde{\mathcal{W}}^{[3]}}\mathbf{v}$ , where we take

$$\phi(x) = e^{-20 \sin^2(\pi x/2)}$$

as the interaction potential  $V$  and

$$\psi(x) = e^{-4(\sin^2(5\pi x/2) + \sin^2(\pi x/2))}$$

as the wave-function  $u$ , both discretised on  $M = 2N + 1$  grid points with  $N = \lceil \varepsilon^{-1} \rceil$ .

Although we have used just  $m = 3$  for  $e^{\tilde{W}^{[2]}}\mathbf{v}$  (i.e., approximated the  $(2N+1) \times (2N+1)$  exponential by an  $3 \times 3$  one) the error is truly minuscule. Moreover, consistently with our theory (but not with conventional numerical intuition) it decreases with  $\varepsilon$ . Indeed, the sort of accuracies we obtain for small values of  $\varepsilon$  are well in excess of what is required in realistic numerical computations. In the case of  $e^{\tilde{\mathcal{W}}^{[3]}}\mathbf{v}$ , we approximate with just a  $2 \times 2$  exponential! Again, everything is consistent with our analysis.

The slope of the error bound is steeper than  $\mathcal{O}(\varepsilon^6)$  in the second figure and this should cause no surprise. The error for  $e^{\tilde{\mathcal{W}}^{[3]}}\mathbf{v}$  decays like  $\mathcal{O}(\varepsilon^8)$ , much faster than required.

## 6.5 Splittings for different scaling laws

The Zassenhaus splitting procedure used for deriving the splitting (4.6) is hardly tied to the choice of  $\sigma = 1$  and works just as well for any other choice. With minimal analysis, for instance, we can show that the relevant splitting for  $0 < \sigma \leq 1$  is

$$\mathcal{Z}_{2,\sigma}^{[2]} = e^{\frac{1}{2}W^{[0]}} e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[2]}} e^{\mathcal{W}^{[3]}} e^{\frac{1}{2}W^{[2]}} e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[0]}} = e^{ih(\varepsilon\partial_x^2 - \varepsilon^{-1}V)} + \mathcal{O}(\varepsilon^{7\sigma-1}), \quad (6.10)$$

with

$$\begin{aligned} W^{[0]} &= ih\varepsilon\partial_x^2 = \mathcal{O}(\varepsilon^{\sigma-1}), \\ W^{[1]} &= -ih\varepsilon^{-1}V = \mathcal{O}(\varepsilon^{\sigma-1}), \\ W^{[2]} &= \frac{1}{6}ih^3\varepsilon^{-1}(\partial_x V)^2 + \frac{1}{6}ih^3\varepsilon\langle\partial_x^2 V\rangle_2 = \mathcal{O}(\varepsilon^{3\sigma-1}), \\ \mathcal{W}^{[3]} &= -\frac{1}{24}ih^3\varepsilon(\partial_x^4 V) - \frac{7}{120}ih^5\varepsilon^{-1}(\partial_x^2 V)(\partial_x V)^2 \\ &\quad + \frac{1}{30}ih^5\varepsilon\langle(\partial_x^2 V)^2 - 2(\partial_x^3 V)(\partial_x V)\rangle_2 \\ &\quad - \frac{1}{120}ih^5\varepsilon^3\langle\partial_x^4 V\rangle_4 = \mathcal{O}(\varepsilon^{5\sigma-1}), \end{aligned}$$

which coincides with (6.6)—the only difference being the factor  $-\frac{1}{24}ih^3\varepsilon(\partial_x^4 V)$  in  $\mathcal{W}^{[3]}$  which can be discarded for  $\sigma \leq \frac{1}{2}$ . This is because the term is  $\mathcal{O}(\varepsilon^{3\sigma+1})$ , which is  $\mathcal{O}(\varepsilon^{5\sigma-1})$  for  $\sigma \leq 1$  and  $\mathcal{O}(\varepsilon^{7\sigma-1})$  for  $\sigma \leq \frac{1}{2}$ . Since the error in these splittings is  $\mathcal{O}(\varepsilon^{7\sigma-1})$ , we may discard this additional term in the case of  $\sigma \leq \frac{1}{2}$ .

### 6.5.1 Number of Lanczos iterations

Consider the problem of exponentiating a skew-Hermitian matrix  $\mathcal{A} = \mathcal{O}(\varepsilon^{k\sigma-1})$  in a splitting featuring error of  $\mathcal{O}(\varepsilon^{r\sigma-1})$ . Assuming  $\rho(\mathcal{A}) \leq c\varepsilon^{k\sigma-1}$ , the error in approximating  $\exp(\mathcal{A})\mathbf{v}$  by  $m$  Lanczos iterations is bounded above by

$$\|e^{\mathcal{A}}\mathbf{v} - \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v}\|_2 \leq 12 \left(\frac{ec}{2m}\right)^m \varepsilon^{m(k\sigma-1)}, \quad m \geq c\varepsilon^{k\sigma-1}. \quad (6.11)$$

For our splitting,

$$m(k\sigma - 1) \geq r\sigma - 1, \quad m \in \mathbb{Z}_+,$$

will ensure that the error in Krylov approximation of the exponential is equal to or smaller than the error incurred in the splitting procedure. Thus we need

$$m = \left\lceil \frac{r\sigma - 1}{k\sigma - 1} \right\rceil$$

Lanczos iterations.

Note that the number of Lanczos iterations required for approximation grows linearly with the order of the method  $r$  and decreases with  $k$  (i.e. is smaller for smaller exponents).

For  $r = 15$ , for instance, which corresponds to a splitting featuring  $\mathcal{O}(\varepsilon^{15\sigma-1})$  error, and  $\sigma = 1$ , the number of iterations required for progressive exponents in the Zassenhaus splitting ( $k = 3, 5, \dots, 11, 13$ ) are 7, 4, 3, 2, 2, 2, while for  $\sigma = 1/2$ , these are 13, 5, 3, 2, 2, 2.

## 6.6 Termination of the Zassenhaus algorithm

**Lemma 6.6.1.** Grade  $n$  commutators of  $i\varepsilon\langle 1 \rangle_2$  and  $-i\varepsilon^{-1}\langle V \rangle_0$  are of size  $\mathcal{O}(\varepsilon^{-1})$ .

*Proof.* Consider a grade  $n$  commutator  $C$  of  $\langle 1 \rangle_2$  and  $\langle V \rangle_0$ , featuring  $k$  occurrences of the letter  $\langle 1 \rangle_2$  and  $n - k$  occurrences of  $\langle V \rangle_0$ . Since  $\langle 1 \rangle_2 \in \mathfrak{G}_2$  and  $\langle V \rangle_0 \in \mathfrak{G}_0$ , using Corollary 5.3.5,

$$\text{ht}(C) \leq 2k - n + 1.$$

Thus  $C$  is  $\mathcal{O}(\varepsilon^{-2k+n-1})$ . However, the corresponding commutator of  $i\varepsilon\langle 1 \rangle_2$  and  $-i\varepsilon^{-1}\langle V \rangle_0$  is scaled by  $k$  occurrences of  $\varepsilon$  and  $n - k$  occurrences of  $\varepsilon^{-1}$ , bringing its size to  $\mathcal{O}(\varepsilon^{-1})$ .  $\square$

**Corollary 6.6.2.** Grade  $n$  commutators of  $ih\varepsilon\langle 1 \rangle_2$  and  $-ih\varepsilon^{-1}\langle V \rangle_0$  are of size  $\mathcal{O}(\varepsilon^{n\sigma-1})$  under the scaling  $h = \mathcal{O}(\varepsilon^\sigma)$ ,  $\sigma > 0$ .

The  $\mathcal{O}(\varepsilon^{n\sigma-1})$  scaling of Corollary 6.6.2 is in contrast to the  $\mathcal{O}(\varepsilon^{n(\sigma-1)})$  scaling seen when using naïve estimates of commutator size for matrix commutators. Whereas the latter requires  $\sigma > 1$  (which corresponds to very small time steps) for terms in sBCH to become progressively smaller, our analysis suggests that no such restrictions apply for the semiclassical Schrödinger equation. With time steps of size  $\mathcal{O}(\sqrt[4]{\varepsilon})$ , which corresponds to  $\sigma = 1/4$ , for instance, grade  $n$  commutators are  $\mathcal{O}(\varepsilon^{n/4-1})$ .

The Zassenhaus procedure terminates so long as the correction terms are progressively smaller. As we have seen, every correction term added in the Zassenhaus procedure is a commutator of a higher grade which, as a consequence of Corollary 6.6.2, is progressively smaller in size.

## 6.7 Cost of Zassenhaus splittings

As we have seen in the previous sections, the property of height reduction which underlies the asymptotic splitting of the symmetric Zassenhaus kind results in inexpensive exponentiations via Lanczos iterations.

**Note:** The advantages of height reduction also extend to error analysis for existing methods such as Yoshida splittings since commutators discarded in these methods can be assumed to be undiscretised and analysed appropriately. An order six Yoshida splitting, for instance, discards all commutators of grade seven and higher, thereby committing an error of  $\mathcal{O}(\varepsilon^{7\sigma-1})$  by the virtue of Corollary 6.6.2. This is an

improvement upon analysis that does not take height reduction into account—working with matrices and analysing naïvely one would expect the error of these methods to be  $\mathcal{O}(\varepsilon^{7(\sigma-1)})$ .

Unlike the exponential growth of costs for Yoshida splittings, however, Zassenhaus splittings feature a quadratic growth of costs which results from a systematic exploitation of the structures of  $\mathfrak{G}$  and the property of height reduction, combined with the asymptotic splitting approach.

**Theorem 6.7.1.** *Cost of Zassenhaus splittings grows quadratically in the order desired.*

*Proof.* Zassenhaus splittings (Table 6.1) recursively utilise the sBCF formula featuring odd grade commutators of  $ih\varepsilon\langle 1 \rangle_2$  and  $-ih\varepsilon^{-1}\langle V \rangle_0$ . Due to Corollary 6.6.2, any grade  $n$  commutator of  $ih\varepsilon\langle 1 \rangle_2$  and  $-ih\varepsilon^{-1}\langle V \rangle_0$  is  $\mathcal{O}(\varepsilon^{n\sigma-1})$ . Splittings with  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  error, therefore, require commutators of grades up to  $2n+1$ . Such a splitting has the form,

$$\mathcal{Z}_{n,\sigma} = e^{\frac{1}{2}W^{[0]}} \dots e^{\frac{1}{2}W^{[n]}} e^{W^{[n+1]}} e^{\frac{1}{2}W^{[n]}} \dots e^{\frac{1}{2}W^{[0]}} = e^{-ihH/\varepsilon} + \mathcal{O}\left(\varepsilon^{(2n+3)\sigma-1}\right), \quad (4.1)$$

where  $W^{[0]} = ih\varepsilon\langle 1 \rangle_2$  and  $W^{[1]} = -ih\varepsilon^{-1}\langle V \rangle_0$ , both of which are  $\mathcal{O}(\varepsilon^{\sigma-1})$ . (Note that we have renamed the central exponent  $\mathcal{W}^{[n+1]}$  to  $W^{[n+1]}$  for convenience.) For  $k \geq 2$ ,  $W^{[k]}$  is of size  $\mathcal{O}(\varepsilon^{(2k-1)\sigma-1})$  and arises from grade  $\leq 2k-1$  commutators of  $ih\varepsilon\langle 1 \rangle_2$  and  $-ih\varepsilon^{-1}\langle V \rangle_0$ .

As discussed in Section 6.4, the first (outermost) exponent,  $W^{[0]}$ , is discretised as a circulant matrix,  $ih\varepsilon\mathcal{K}^2$ , whose exponential can be evaluated using two Fast Fourier Transforms (FFTs), each of which costs  $\mathcal{O}(M \log M) = \mathcal{O}(\varepsilon^{-1} \log \varepsilon^{-1})$ . The second exponent,  $W^{[1]}$ , is discretised as a diagonal matrix,  $-ih\varepsilon^{-1}\mathcal{D}_V$ , and can be exponentiated directly.

Remaining exponents require Lanczos iterations. Since  $W^{[k]}$  is  $\mathcal{O}(\varepsilon^{(2k-1)\sigma-1})$ , we need  $\left\lceil \frac{(2n+3)\sigma-1}{(2k-1)\sigma-1} \right\rceil$  iterations for approximating its exponential to  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  accuracy, as seen in Section 6.5.1. Each of these iterations require the evaluation of matrix–vector products of the form  $W^{[k]}u$ .

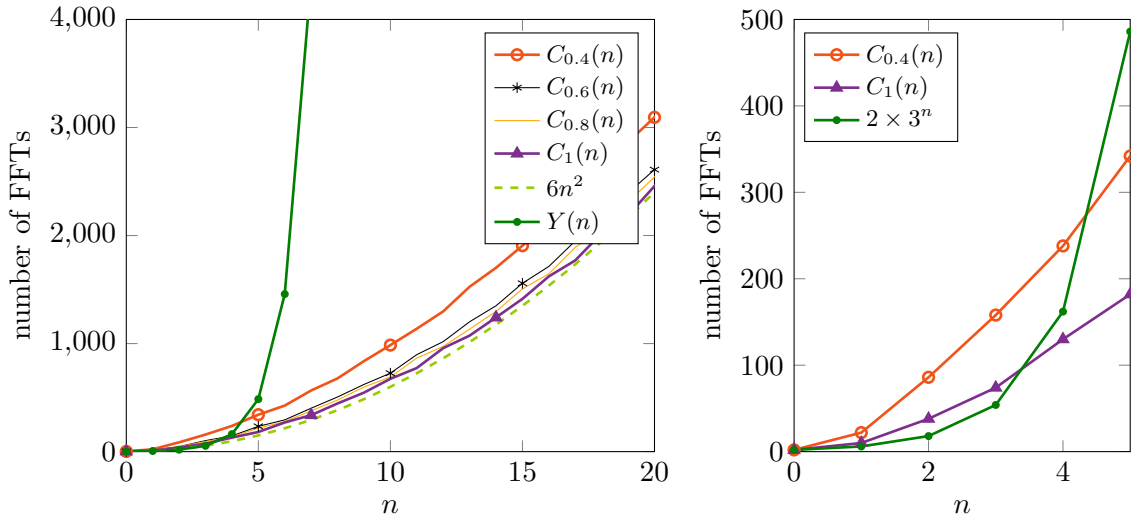
Since  $W^{[k]}$  arises from grade  $2k-1$  commutators (at most) of the terms  $\langle 1 \rangle_2 \in \mathfrak{G}_2 \cap \mathfrak{e}$  and  $\langle V \rangle_0 \in \mathfrak{G}_0 \cap \mathfrak{e}$ , it resides in the intersection of  $\mathfrak{G}_{2k-2}$  and  $\mathfrak{e}$ . Thus, it consists of terms of the form  $\langle f_0 \rangle_0, \langle f_2 \rangle_2, \langle f_4 \rangle_4, \dots, \langle f_{2k-2} \rangle_{2k-2}$ . To evaluate  $W^{[k]}u$ , we need to evaluate  $\langle f_0 \rangle_0 u, \langle f_2 \rangle_2 u, \langle f_4 \rangle_4 u, \dots, \langle f_{2k-2} \rangle_{2k-2} u$ . The first of these,  $\langle f_0 \rangle_0 u$ , is a pointwise product, while the rest require four FFTs each since  $\langle f \rangle_k$  is discretised as  $(\mathcal{D}_f \mathcal{K}^2 + \mathcal{K}^2 \mathcal{D}_f)/2$ . However, as we have seen in Section 6.4.1, there are more clever ways of combining the FFTs using which  $W^{[k]}u$  can be evaluated using  $2k$  FFTs.

The cost of the splitting is dominated by the cost of the FFT operations, each of which requires  $\mathcal{O}(\varepsilon^{-1} \log \varepsilon^{-1})$  operations. Since each evaluation of  $W^{[k]}u$  requires  $2k$  FFTs, the

number of FFTs required per time step of the splitting scheme is

$$C_\sigma(n) = 2 + 2(n+1) \left\lceil \frac{(2n+3)\sigma - 1}{(2n+1)\sigma - 1} \right\rceil + 2 \sum_{k=2}^n 2k \left\lceil \frac{(2n+3)\sigma - 1}{(2k-1)\sigma - 1} \right\rceil, \quad (6.12)$$

which grows quadratically in  $n$  (as seen in Figure 6.2). Here we have used the FASL property, combining the outermost exponentials (see Section 6.4). A comparable Yoshida splitting requires  $Y(n) = 2 \times 3^n$  FFTs (see Section 2.2.5).  $\square$



**Figure 6.2:** Number of FFTs in the Zassenhaus splitting  $\mathcal{Z}_{n,\sigma}$  of  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  accuracy for  $\sigma = 2/5, 3/5, 4/5, 1$ , compared to the  $Y(n) = 2 \times 3^n$  FFTs required for the comparable Yoshida splitting  $\mathcal{S}_{2n+2}$ . For  $n \leq 3$ , the number of FFTs for Yoshida splittings is lower than all Zassenhaus splittings.

The overall cost for an  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  splitting is  $\mathcal{O}(C_\sigma(n)\varepsilon^{-1} \log \varepsilon^{-1})$  per time step. We remark that careful choices in a splitting could allow us to reduce the exact number of FFTs required. However, we are largely concerned with asymptotic growths here.

**Note:** Since this proof does not assume a specific form of the exponents,  $W^{[k]}_S$ , and can work solely on the basis of the structure of  $\mathfrak{G}_k$  and  $\mathfrak{H}$ , and the height reduction of Corollary 5.3.5, the cost estimates derived here also apply directly to Magnus–Zassenhaus approaches of Chapter 8 (with the exception that Magnus–Zassenhaus splittings will also feature exponents with odd-indexed components of the form  $\langle f \rangle_{2k+1}$  and the exact number of FFTs are somewhat higher).

**Note:** While the quadratic growth of Zassenhaus methods is theoretically impressive, small values of  $n$  (or moderately high-order methods) are all that are required in practical applications. A careful reader would notice in Figure 6.2 the number of



FFTs in Yoshida splitting are lower than those required in a corresponding Zassenhaus splitting for small  $n$ . Featuring a similar error, therefore, Yoshida splittings are superior for time-independent Hamiltonians for moderately high orders.

**Note:** Yoshida splittings are not the most efficient and optimised of high order splittings in use. It would be worth comparing the cost with highly optimised high order splittings—those discussed in (Blanes et al. 2008), for instance. However, as we note in Figure 6.2, for moderately high orders the cost of Zassenhaus splittings exceeds that of Yoshida splittings. Thus, where Yoshida already performs better, more optimised splittings are certainly expected to outperform.

On the other hand, from a complexity point of view, the quadratic growth of Zassenhaus splittings is unlikely to be beaten by optimised splittings.

The true advantage of Zassenhaus splittings becomes evident when the Hamiltonian is a bit more complicated and features more terms, as it does in the case of Magnus expansions. A cost analysis for Magnus–Zassenhaus schemes is presented along similar lines in Section 8.5. Here Yoshida splittings are applicable but become prohibitively expensive. On the other hand, no splittings for time-dependent potentials along the lines of (Blanes et al. 2008) were known to the author at the time of writing. Keeping these factors in mind, the choice of Yoshida splittings as a reference seems reasonable.

**Example 6.7.2.** For  $n \geq 2$  and under the scaling  $\sigma = 1$  and,  $C_1(n) \sim 6n^2$ ,

$$\begin{aligned}
 C_1(n) &= 2 + 2(n+1) \left\lceil \frac{2n+2}{2n} \right\rceil + 2 \sum_{k=2}^n 2k \left\lceil \frac{2n+2}{2k-2} \right\rceil \\
 &\leq 2 + 4(n+1) + 4 \sum_{k=2}^n k \left( \frac{n+1}{k-1} + 1 \right) \\
 &= 2 + 4(n+1) + 4 \sum_{k=1}^{n-1} (k+1) \left( \frac{n+1}{k} + 1 \right) \\
 &= 2 + (n-1) + 4(n+1) + 4 \sum_{k=1}^{n-1} (n+1+k) + 4(n+1) \sum_{k=1}^{n-1} \frac{1}{k} \\
 &= 6n^2 + 3n + 1 + 4(n+1)H_{n-1} \\
 &\leq 6n^2 + 3n + 1 + 4(n+1)(\log(n-1) + \gamma) + 4 + \mathcal{O}\left(\frac{1}{n-1}\right) \\
 &\leq 6n^2 + 3n + 6 + 4(n+1)(\log(n-1) + \gamma),
 \end{aligned}$$

where  $\gamma \approx 0.577$  is the Euler gamma constant,  $H_{n-1}$  is the harmonic number and  $(n+1)H_{n-1}$

is estimated using Euler's summation formula following (Apostol 1976),

$$\begin{aligned}
(x+2)H_x &= (x+2) \sum_{k=1}^x \frac{1}{k} \\
&= (x+2) \left[ \int_1^x \frac{1}{t} dt - \int_1^x \frac{t - \lfloor t \rfloor}{t^2} dt + 1 - \frac{x - \lfloor x \rfloor}{x} \right] \\
&= (x+2) \left[ \log x + 1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt + \int_x^\infty \frac{t - \lfloor t \rfloor}{t^2} dt \right] - 2 \frac{x - \lfloor x \rfloor}{x} - (x - \lfloor x \rfloor) \\
&\leq (x+2) (\log x + \gamma) + 1 + 2 \frac{1 - (x - \lfloor x \rfloor)}{x} - (x - \lfloor x \rfloor) \\
&\leq (x+2) (\log x + \gamma) + 1 + \mathcal{O}\left(\frac{1}{x}\right).
\end{aligned}$$

Here we use the fact that

$$\int_x^\infty \frac{t - \lfloor t \rfloor}{t^2} dt \leq \int_x^\infty \frac{1}{t^2} dt = \frac{1}{x},$$

and

$$1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt = \gamma.$$

## 6.8 Scaling choices and global cost concerns

For the splitting  $\mathcal{Z}_{2,\sigma}^{[2]}$ , which is valid for  $\sigma \leq 1$ , the exponents  $\tilde{W}^{[0]}$  and  $\tilde{W}^{[1]}$  are exponentiated easily despite their  $\mathcal{O}(\varepsilon^{\sigma-1})$  size (note that for  $\sigma = \frac{1}{2}$  these exponents increase as  $\varepsilon \rightarrow 0$ ). The remaining exponents are  $\tilde{W}^{[2]} = \mathcal{O}(\varepsilon^{3\sigma-1})$  and  $\tilde{W}^{[3]} = \mathcal{O}(\varepsilon^{5\sigma-1})$ . Except for cases with  $\sigma < \frac{1}{3}$ , where the spectral radius of  $\tilde{W}^{[2]}$  increases as  $\varepsilon \rightarrow 0$ , the number of iterations required in the Krylov approximation of the exponentials is small and derived in a straightforward manner from the bound (2.24). The exponentiation of  $\tilde{W}^{[2]}$  requires 3 iterations for  $\sigma = 1$ , as we have seen, and 5 iterations for  $\sigma = \frac{1}{2}$ , while  $\tilde{W}^{[3]}$  can be exponentiated in 2, 2 and 3 iterations, respectively, in the cases  $\sigma = 1, \frac{1}{2}, \frac{1}{4}$ .

For the case  $\sigma = \frac{1}{4}$ , the spectral radius of  $\tilde{W}^{[2]}$  becomes large enough to be of concern with regards to the side condition in (2.24),  $m \geq \rho$ , where, it should be recalled,  $m$  was the number of iterations and  $\rho$  the spectral radius. In this case, as noted by Hochbruck & Lubich (1997), the error does not decrease substantially till  $m \geq \rho$ , decreasing rapidly thereafter. Re-writing (2.24) as

$$\left\| e^{\tilde{W}^{[2]}} \mathbf{v} - \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v} \right\|_2 \leq 12 \exp\left(\frac{-\rho^2 + 4m^2(1 - \log 2)}{4m}\right) \left(\frac{\rho}{m}\right)^m, \quad m \geq \rho, \quad (6.13)$$

from where, with the choice  $m \geq \alpha\rho$  for some  $\alpha > 1$ , we end up with an estimate

$$\left\| e^{\tilde{W}^{[2]}} \mathbf{v} - \mathcal{V}_m e^{\mathcal{H}_m} \mathcal{V}_m^* \mathbf{v} \right\|_2 \leq 12 \exp\left(\frac{-\rho^2 + 4m^2(1 - \log 2 - \log \alpha)}{4m}\right).$$

Exponential convergence, well in excess of what we require, can be achieved by an appropriate choice of  $\alpha$ , whereby  $\mathcal{O}(\varepsilon^{-1/4})$  iterations prove adequate. With  $\alpha = 2$ , for instance, the error term works out to roughly  $\exp(-\frac{7}{8}\rho)$ .

After  $\mathcal{O}(\varepsilon^{-1/4})$  iterations we are also left with the task of exponentiating a  $\mathcal{O}(\varepsilon^{-1/4}) \times \mathcal{O}(\varepsilon^{-1/4})$  upper Hessenberg matrix which, although large, can be exponentiated by brute force using diagonal Padé methods in  $\mathcal{O}(\varepsilon^{-3/4})$  operations. The  $\mathcal{O}(\varepsilon^{-5/4} \log \varepsilon^{-1})$  cost of the Lanczos iterations, however, overshadows the cost for approximating the exponential of the Hessenberg matrix.

The cost of the exponential  $e^{\tilde{W}^{[2]}} \mathbf{v}$  dominates the cost of each time-step of the splitting under  $\sigma = \frac{1}{4}$ , making the overall (global) cost  $\mathcal{O}(\varepsilon^{-3/2} \log \varepsilon^{-1})$  for solving the equation over a fixed time interval  $[0, T]$ . This is no less than the cost of the more accurate  $\sigma = \frac{1}{2}$  splitting. In the case of  $\sigma = 1$ —where, of course, a much smaller error is achieved with the same number of exponentials per time step—the overall cost comes to  $\mathcal{O}(\varepsilon^{-2} \log \varepsilon^{-1})$ .

Of the three choices  $\sigma = \frac{1}{4}, \frac{1}{2}, 1$ , the  $\mathcal{O}(C_{\frac{1}{2}}(n) \varepsilon^{-3/2} \log \varepsilon^{-1})$  cost for  $\sigma = \frac{1}{2}$  is the lowest. There seems little point in considering a  $\sigma$  smaller than  $\frac{1}{3}$  where  $\tilde{W}^{[2]}$  becomes  $\mathcal{O}(1)$  or larger and the number of Lanczos iterations required for  $e^{\tilde{W}^{[2]}} \mathbf{v}$  starts increasing as  $\varepsilon \rightarrow 0$ . Even where the spectral radius does decrease with  $\varepsilon$ , a small  $\sigma$  makes the constraint  $m \geq \rho$  in (2.24) a graver concern. With  $\sigma = \frac{1}{4}$ , for instance, where  $\tilde{W}^{[3]}$  scales as  $\tilde{c}\varepsilon^{1/4}$ , this requires  $\varepsilon \leq (3/\tilde{c})^4$ . The constant  $\tilde{c}$  depends on the interaction potential and circumstances where this constraint can become a serious concern are far from inconceivable.

## 6.9 Stability

The convergence of classical methods for initial-value partial differential equations is governed by the Lax equivalence theorem: convergence equals consistency plus stability (Iserles 2008). Our method is clearly consistent but the question is whether, once derivatives are replaced by differentiation matrices, the ensuing finite-dimensional operator is stable in the sense of Lax. Within our formalism this is equivalent to

$$\lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \|(\tilde{\mathcal{Z}}_{2,1}^{[1]})^n\| < \infty, \quad (6.14)$$

where  $\tilde{\mathcal{Z}}_{2,1}^{[1]}$  is the finite-dimensional discretisation of  $\mathcal{Z}_{2,1}^{[1]}$ .

Upon discretisation of  $\partial_x$  and  $f$  as  $M \times M$  matrices  $\mathcal{K}$  and  $\mathcal{D}_f$ , respectively, where  $\mathcal{K}$  is skew-Hermitian and  $\mathcal{D}_f$  is diagonal, elements of the Lie algebra  $\mathfrak{H}$  are discretised as skew-Hermitian matrices,

$$\sum_{k=0}^n \mathbf{i}^{k+1} \langle f_k \rangle_k \rightsquigarrow \frac{1}{2} \sum_{k=0}^n \mathbf{i}^{k+1} \left( \mathcal{D}_f \mathcal{K}^k + \mathcal{K}^k \mathcal{D}_f \right),$$

which reside in  $\mathfrak{u}(M)$ . Since the computations of the Zassenhaus procedure of Section 6.2

are carried out entirely in  $\mathfrak{H}$ , it guarantees  $W^{[k]} \in \mathfrak{H}$  and thus the discretisation  $\tilde{W}^{[k]} \in \mathfrak{u}(M)$ . Consequently,  $\exp(\tilde{W}^{[k]}) \in U(M)$  and  $\tilde{\mathcal{Z}}_{2,1}^{[1]} \in U(M)$ .

The condition (6.14) is clearly implied by  $\tilde{\mathcal{Z}}_{2,1}^{[1]}$  being a unitary matrix for all (sufficiently small)  $\varepsilon > 0$ , in other words by the discretisation method being *unitary*. This has the added virtue of the discretisation method mimicking the unitarity of the infinite-dimensional operator  $\exp(ih(\varepsilon\partial_x^2 - \varepsilon^{-1}V))$  which was discussed in Chapter 3. Consequently, in that case we obtain a *geometric integrator* in the sense of (Faou 2012, Hairer et al. 2006, Lubich 2008).

We note that once we resort to Lanczos iterations (a Krylov subspace method) to exponentiate some exponents, the unitarity is lost. However,  $\ell^2$  norm is preserved under these numerical methods. Consequently,

$$\|(\tilde{\mathcal{Z}}_{2,1}^{[1]})^n \mathbf{u}\| = \|\mathbf{u}\|,$$

holds for all  $\mathbf{u}$ , and we conclude that

$$\|(\tilde{\mathcal{Z}}_{2,1}^{[1]})^n\| = 1.$$

This is sufficient to guarantee (6.14).

## 6.10 Numerical experiments

We present numerical results for two interaction-potential wave-function pairs. The wave-functions used in our experiments are  $u_1(x) = \frac{1}{100}\phi(x + 0.6)e^{i20\pi x}$  and  $u_2 = \psi$ , where  $\phi$  and  $\psi$  have been previously introduced as

$$\phi(x) = e^{-20\sin^2(\pi x/2)}, \quad \psi(x) = e^{-4(\sin^2(5\pi x/2) + \sin^2(\pi x/2))}.$$

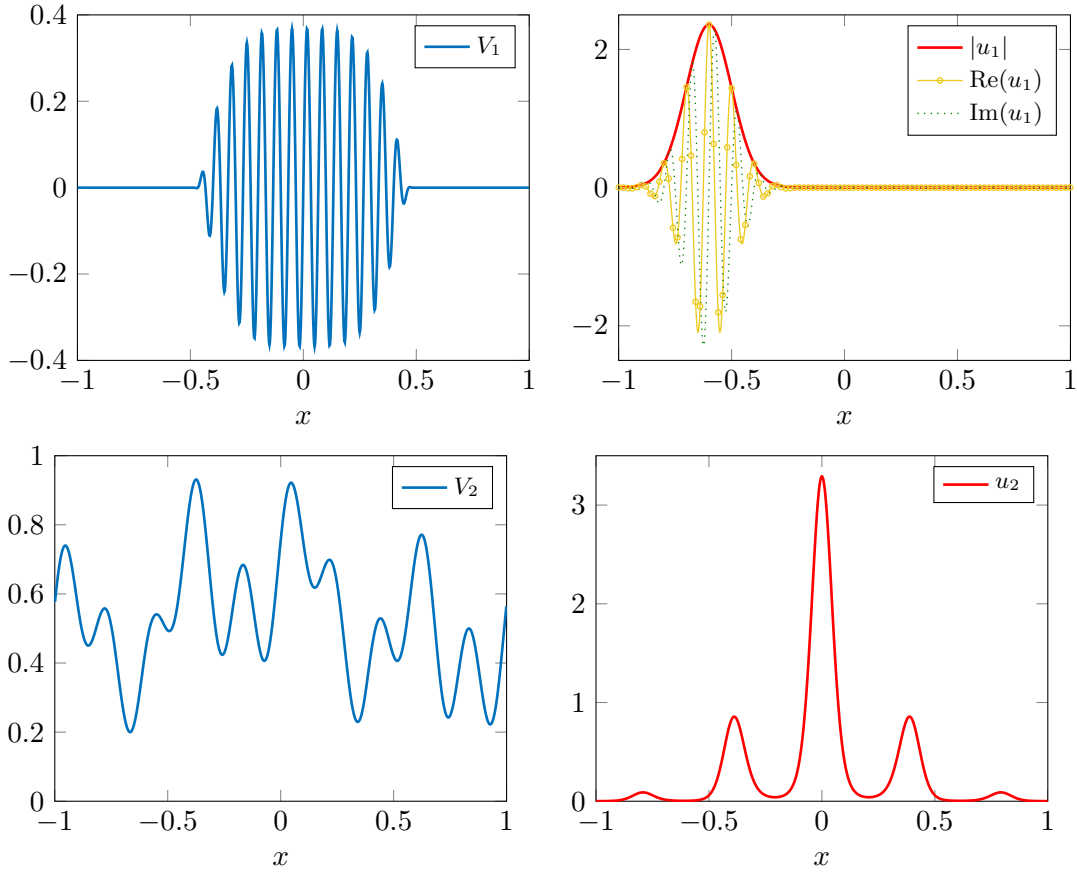
The first of these is a moderately oscillating wave-train with a periodic Gaussian envelope seen travelling to the right in free space ( $V = 0$ ). In our experiments, these move under influence of the interaction potentials

$$\begin{aligned} V_1(x) &= \text{expbump}_p(x) \sin(20\pi x), \\ V_2(x) &= \frac{1}{5} + \frac{1}{2}\text{expbump}_p\left(x + \frac{1}{10}\right) + \frac{3}{10} \left( \sin^4\left(2\pi x - \frac{24}{35}\right) + \sin^2\left(5\pi x - \frac{8}{3}\right) \right), \end{aligned}$$

where  $\text{expbump}_p$  is the periodic bump function (2.17).

Physically, the first pair shown in Fig. 6.3 (top row) is an attempt at modelling a wave-packet heading towards a periodic lattice. The second pair, Fig 6.3 (bottom row), has no physical motivation and is chosen for its complexity.

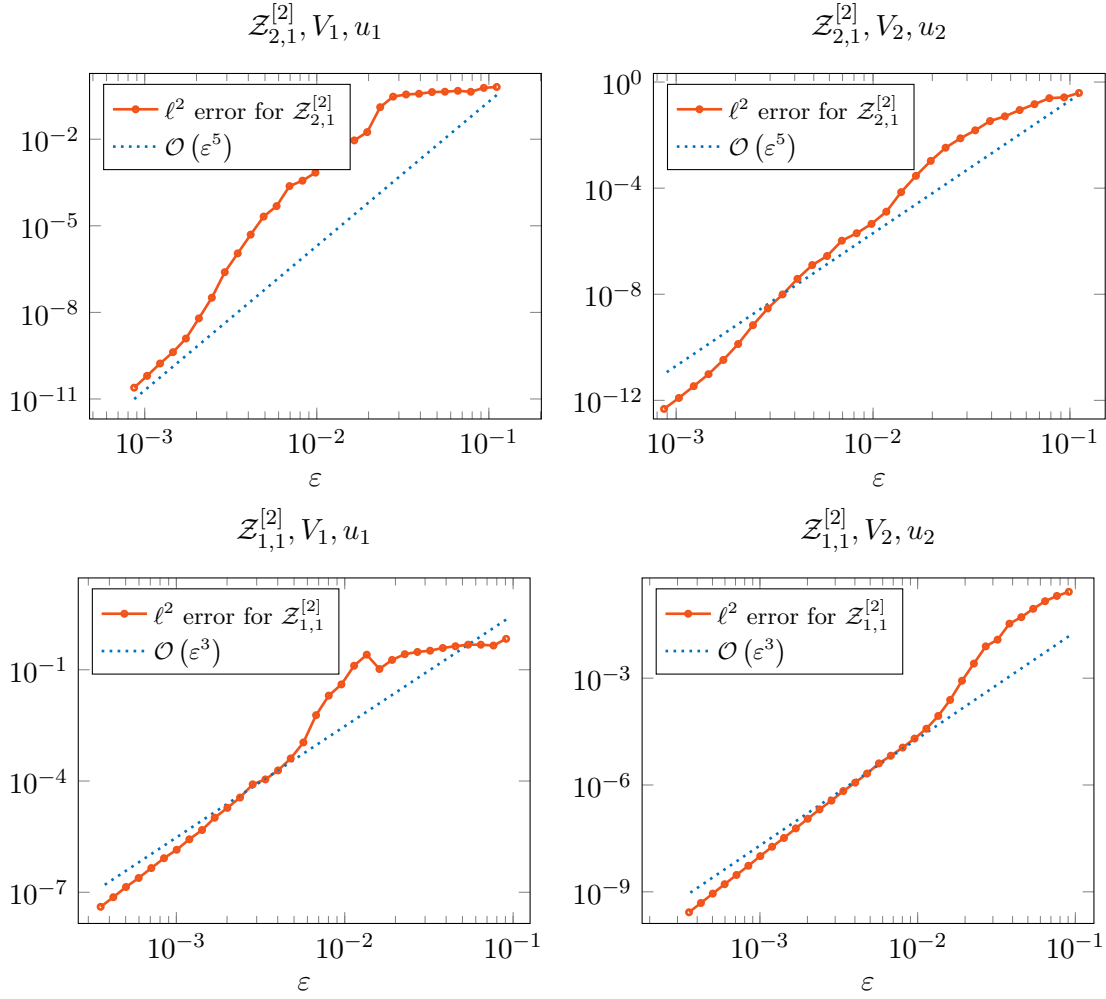
The error estimates are, of course, of an asymptotic nature and it is little surprise



**Figure 6.3:** The interaction-potential wave-function pair  $V_1, u_1$  (top row) and  $V_2, u_2$  (bottom row).

that for some cases the true nature does not emerge till very small values of  $\varepsilon$ . In the case of  $V_1$  with  $\sigma = \frac{1}{2}$  or  $\frac{1}{4}$ , for instance, where one of the terms omitted in the splitting,  $-\frac{1}{24}i\hbar^3\varepsilon(\partial_x^4 V)$ , is fairly large, we do not see a noticeable decrease till very small values of  $\varepsilon$  unless the magnitude of the interaction potential is decreased. In Figure 6.5 (top row) the error is seen to approach the asymptotic estimate at an earlier stage in the case of a smaller potential. The asymptotic bounds are very much adhered to, but here we become limited by inefficiency of the reference method—MATLAB’s `expm`—which does not allow us to go beyond moderate values of  $M = \mathcal{O}(\varepsilon^{-1})$ .

All estimates and bounds in our analysis have, of course, been derived with respect to the  $L^2$  norm. This is approximated by the  $\ell^2$  norm on the grid. Where  $\ell^\infty$  error estimates are required, noting the inequality (2.4),  $\|\mathbf{u}\|_\infty \leq \sqrt{M/2} \|\mathbf{u}\|_2$ , we expect the  $\ell^\infty$  error to be worse off than the  $\ell^2$  error by a factor of  $\sqrt{M/2} = \mathcal{O}(\varepsilon^{-1/2})$ . This is indeed seen to be the case in our experiments (Figures 6.5 and 6.6, bottom row).

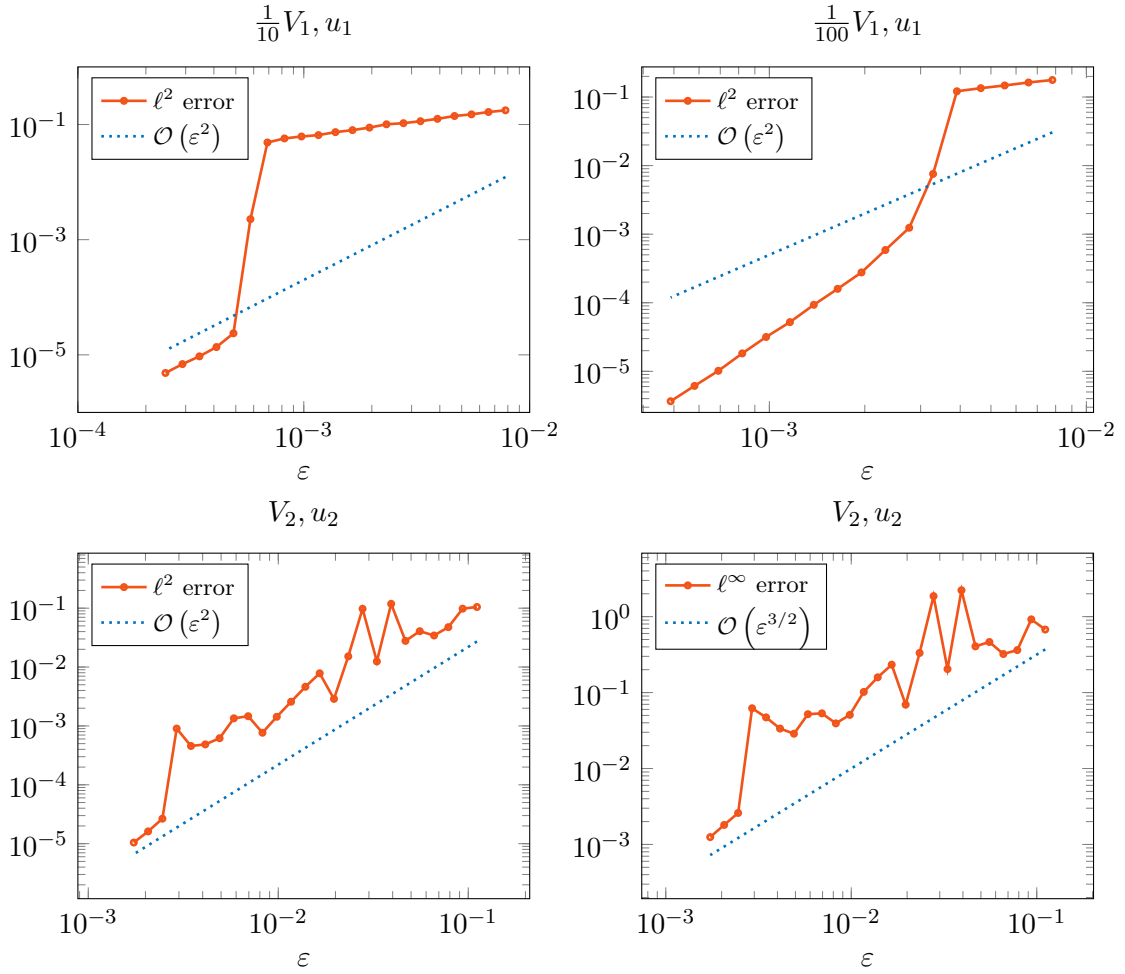


**Figure 6.4:** The  $\mathcal{Z}_{2,1}^{[2]}$  splitting (6.6) with seven exponentials (*top row*) has a global  $\ell^2$  error estimate of  $\mathcal{O}(\varepsilon^5)$  at  $T = 1$ , while the  $\mathcal{Z}_{1,1}^{[2]}$  splitting which omits  $\tilde{\mathcal{W}}^{[3]}$  uses five exponentials (*bottom row*) and has an error estimate of  $\mathcal{O}(\varepsilon^3)$ . Errors are presented with initial value combinations  $V_1, u_1$  and  $V_2, u_2$ .

### 6.10.1 Finding a reference solution

To arrive at a reference solution, we solve (3.15) with increasing spatial degrees of freedom  $M_R$  (decoupled in this case from  $\varepsilon$  and  $h$ ) till the reference numerical solution  $\mathbf{u}_R(T)$  seems to converge with respect to  $\ell^\infty$  norm, up to a certain level of tolerance.

Direct exponentiation using MATLAB's `expm` function is a naïve approach that is discussed in Section 3.3, and is seen to be inefficient to the point of impracticality for small values of  $\varepsilon$ . Generating reference solutions using this method is, therefore, an extremely resource intensive and time consuming affair. However, its accuracy can be relied upon as a reference solution. The inefficiency of this method, rather than that of the Zassenhaus splitting method, sets limits on the range of  $\varepsilon$  for which we can verify the accuracy of our



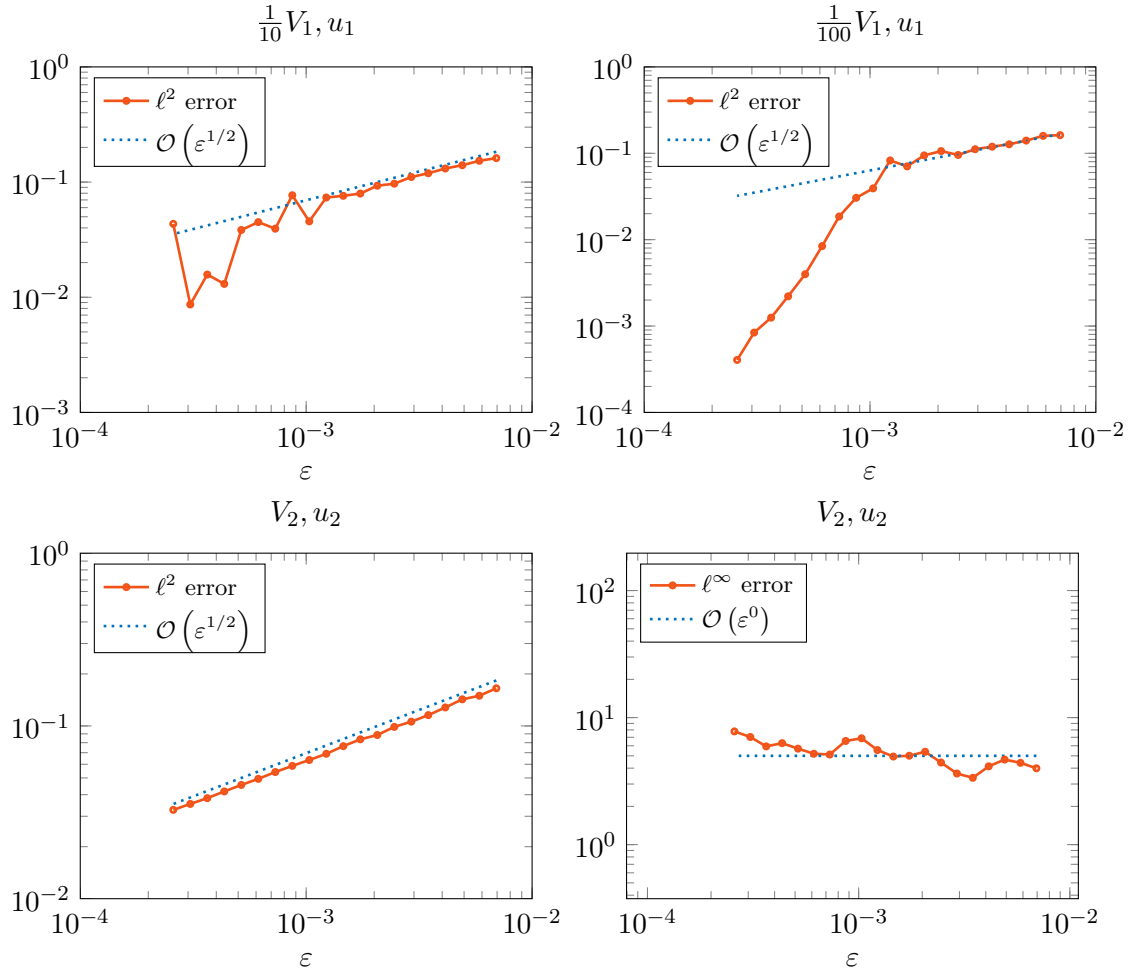
**Figure 6.5:** Global error at  $T = 1$  for  $\mathcal{Z}_{2, \frac{1}{2}}^{[2]}$  (where  $\sigma = \frac{1}{2}$ ): asymptotic nature of the error bounds is evident in the top row where asymptotic decrease is achieved faster when  $V_1$  is scaled down by a factor of  $\frac{1}{100}$ . The bottom row shows the relationship between the  $\ell^2$  error and  $\ell^\infty$  error for the initial value pair  $V_2, u_2$ .

splitting schemes.

### Choice of grid resolutions

As we make the grid finer in the pursuit of a converged reference solution, we must ensure that the coarser grid forms a subset of the finer grid. Solution obtained on the coarser grid can then be compared against the solution on the finer grid restricted to the coarse grid points.

Let  $M_1, M_2$  be two grid resolutions such that  $M_1 < M_2$ . The two grids are given by  $x_i^j = i/(N_j + 1/2)$ ,  $i = -N_j, \dots, 0, \dots, N_j$  with  $M_j = 2N_j + 1$ ,  $j = 1, 2$ . In order for



**Figure 6.6:** Global error at  $T = 1$  for  $\mathcal{Z}_{2, \frac{1}{4}}^{[2]}$  (where  $\sigma = \frac{1}{4}$ ): asymptotic nature of the error bounds is evident in the top row where asymptotic decrease is achieved faster when  $V_1$  is scaled down by a factor of  $\frac{1}{100}$ . The bottom row shows the relationship between the  $\ell^2$  error and  $\ell^\infty$  error for the initial value pair  $V_2, u_2$ .

$x_1^1$  to feature in the finer grid, there must be a  $k$  such that  $x_1^1 = x_k^2$ , or

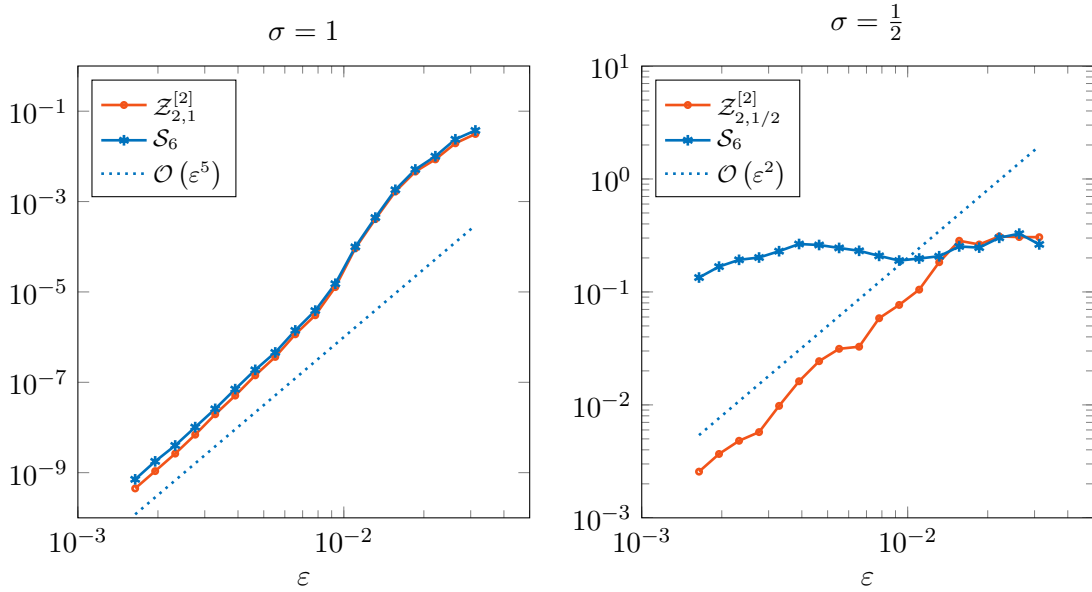
$$\frac{1}{N_1 + 1/2} = k \frac{1}{N_2 + 1/2}$$

which gives us the relation  $N_2 = kN_1 + (k-1)/2$  whereby a choice of an odd  $k$  is necessary and sufficient. For the purpose of our experiments we stick with the factor of 3 and the corresponding relation  $M_2 = 2N_2 + 1 = 2(3N_1 + 1) + 1 = 3M_1$ . For this reason, we stick to comparing discretised solutions on grids with resolutions that differ by powers of 3.



### 6.10.2 Comparison with Yoshida splittings

As we noted in Section 6.7, Yoshida splittings should also benefit from decrease in height of the commutators (that are discarded and lead to the  $\mathcal{O}(h^{2n+1})$  error estimate for the order  $2n$  Yoshida splitting  $\mathcal{S}_{2n}$ ). Consequently,  $\mathcal{S}_{2n+2}$  should feature a  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  error which corresponds to the accuracy of the Zassenhaus splitting  $\mathcal{Z}_{n,\sigma}$ . The order six Yoshida splitting  $\mathcal{S}_6$  should therefore be as accurate as  $\mathcal{Z}_{2,\sigma}^{[2]}$ .

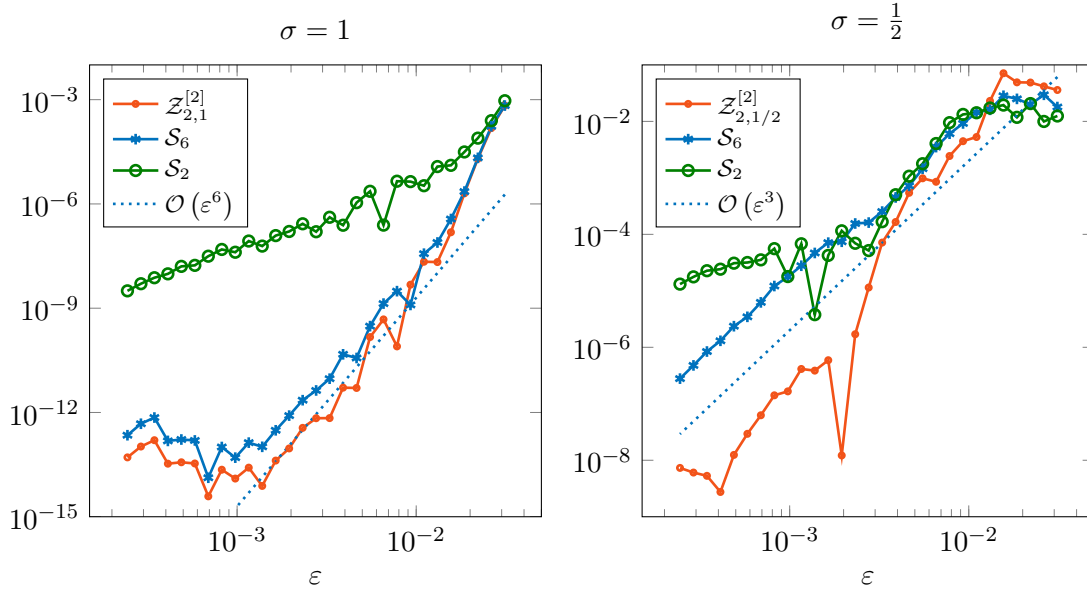


**Figure 6.7:** Global  $\ell^2$  error at  $T = 1$  for the initial value pair  $V_2, u_2$  using  $\mathcal{Z}_{2,\sigma}^{[2]}$  and the sixth order Yoshida splitting  $\mathcal{S}_6$  under  $\sigma = 1$  (left) and  $\sigma = \frac{1}{2}$  (right).

In Figure 6.7 we show the results of numerical experiments comparing the accuracy of  $\mathcal{Z}_{2,\sigma}^{[2]}$  and  $\mathcal{S}_6$  for the initial value pair  $V_2, u_2$  discussed earlier. The exact scaling used here is  $M \sim \varepsilon^{-1}$  and  $h \sim \varepsilon^\sigma$ . For  $\sigma = 1$ , the two splittings have nearly identical errors. In practice we have found that, depending on the example under consideration and the exact choice of discretisation scaling, either of them could turn out somewhat superior for  $\sigma = 1$ . However, for small values of  $\sigma$ , the Zassenhaus splittings usually perform better. In this example, however, the Yoshida splitting performs unusually bad in the case of  $\sigma = \frac{1}{2}$ . This behaviour requires further investigation.

In Figure 6.8 we show the errors in energy for the Strang splitting  $\mathcal{S}_2$ , the Yoshida splitting  $\mathcal{S}_6$ , and the Zassenhaus splitting  $\mathcal{Z}_{2,\sigma}^{[2]}$ . The initial value pair  $V_2, u_2$  and the choice of discretisation scaling is the same as discussed earlier.

**Note:** The error in energy seems to gain an entire power of  $\varepsilon$ , scaling as  $\mathcal{O}(\varepsilon^{6\sigma})$  for the Zassenhaus and Yoshida splittings and as  $\mathcal{O}(\varepsilon^{2\sigma})$  for the Strang splitting. This is in line with the observations of (Jin et al. 2011), who note that when we are only concerned about observables such as the energy, the time step in the Strang



**Figure 6.8:** Error in energy at  $T = 1$  for the initial value pair  $V_2, u_2$  using  $\mathcal{Z}_{2,\sigma}^{[2]}$ , the sixth order Yoshida splitting  $\mathcal{S}_6$  and the Strang splitting  $\mathcal{S}_2$  under  $\sigma = 1$  (left) and  $\sigma = \frac{1}{2}$  (right).

splitting can be taken independently of  $\varepsilon$  and the negative power of  $\varepsilon$  doesn't show up in the error.

Consequently the (global) accuracy of Strang splitting in the context of observables is expected to be  $\mathcal{O}(h^2)$ . This is effectively what we are witnessing here since the experiments were performed under the scaling  $h = \mathcal{O}(\varepsilon^\sigma)$ .

A similar analysis needs to be carried out for Yoshida splittings and Zassenhaus splittings. However, the evidence seems to support the hypothesis that, much like the case of the Strang splitting, the error in energy does not feature a  $\varepsilon^{-1}$  factor and, in fact, scales as  $\mathcal{O}(h^6) = \mathcal{O}(\varepsilon^{6\sigma})$ .

Numerical experiments concerning error in observables also need to be carried out while choosing  $h$  independently of  $\varepsilon$ . However, this is beyond the scope of this thesis.

## Chapter 7

# Time-dependent potentials and Magnus expansions

The one-dimensional semiclassical Schrödinger equation featuring a time-dependent electric potential introduced in Section 3.4 is

$$\partial_t u(x, t) = i(\varepsilon \partial_x^2 - \varepsilon^{-1} V(x, t)) u(x, t), \quad x \in [-1, 1], \quad t \geq 0, \quad (3.20)$$

where the wave-function  $u = u(x, t)$  is given with an initial condition  $u(x, 0) = u_0(x)$ . In Section 3.4.1, we imposed periodic boundary conditions at  $\pm 1$  to allow an effective approximation by spectral methods, arriving at the discretised equation,

$$\mathbf{u}'(t) = i(\varepsilon \mathcal{K}^2 - \varepsilon^{-1} \mathcal{D}_{V(\cdot, t)}) \mathbf{u}(t), \quad t \geq 0, \quad (3.24)$$

which is solved via the Magnus expansion,

$$\mathbf{u}(h) = e^{\Theta(h)} \mathbf{u}(0).$$

Magnus expansions such as the fourth order truncation,

$$\Theta_3(h) \rightsquigarrow i\varepsilon h \mathcal{K}^2 - i\varepsilon^{-1} h \overline{V} + h^2 \frac{\sqrt{3}h^2}{12} [\mathcal{K}^2, \tilde{V}]$$

naturally feature commutators of matrices and the approximation of their exponential via Krylov methods (which results in the Magnus–Lanczos schemes) is estimated to cost  $\mathcal{O}(\varepsilon^{-2} \log \varepsilon^{-1})$ .

In this context, Yoshida splittings also become less effective on two counts: firstly, even the lowest order method—the Strang splitting—features a large number of exponentials, leading to an even larger number of exponentials in higher order Yoshida splittings derived by composing the Strang splitting; secondly, unlike the case of time-independent Hamiltonians, there seems to be no way to avoid commutators, whose exponentials are

more expensive to compute than the exponentials of  $\mathcal{K}^2$  and  $\mathcal{D}_V$ . Moreover, no highly optimised splitting along the lines of (Blanes et al. 2008) were known to the author at the time of writing.

In contrast to the traditional approach discussed in Chapter 3, we follow the ideas pursued in Chapter 6: we postpone the spatial discretisation to the very last moment, working directly in the Lie algebra generated by  $\partial_x^2$  and  $V$ . As always, we assume throughout that the interaction potential  $V(\cdot, t)$  and the wave-function  $u(\cdot, t)$  are sufficiently smooth. For the purpose of simplicity we may assume that they belong to  $C_p^\infty([-1, 1], \mathbb{R})$  and  $C_p^\infty([-1, 1], \mathbb{C})$ —the spaces of real-valued and complex-valued smooth periodic functions over  $[-1, 1]$ , respectively—but the results should extend to functions of lower smoothness.

Consider (3.20) written in the form

$$\partial_t u(t) = \mathcal{A}(t) u(t), \quad u(0) = u_0, \quad (3.22)$$

where  $\mathcal{A}(t) = i\varepsilon \langle 1 \rangle_2 - i\varepsilon^{-1} \langle V(t) \rangle_0 \in \mathfrak{H}$ . The solution of the partial differential equation (3.22) is obtained via the Magnus expansion

$$u(t) = e^{\Theta(t)} u(0),$$

where the infinite series  $\Theta(t) = \sum_{k=1}^{\infty} \Theta_k(t)$  is also an element of the underlying Lie algebra  $\mathfrak{H}$ . The Lie algebra

$$\mathfrak{H} = \bigoplus_{k \in \mathbb{Z}_+} \{i^{k+1} \langle f \rangle_k : f \in C_p^\infty([-1, 1], \mathbb{R})\},$$

which has been introduced previously in Section 6.1.3, possesses structural properties that allow us to simplify commutators while gaining powers of  $\varepsilon$  and preserving stability. The graded version of the Magnus expansion reads

$$\begin{aligned} \Theta(t) = & \int_0^t \mathcal{A}(\xi) d\xi - \frac{1}{2} \int_0^t \int_0^{\xi_1} [\mathcal{A}(\xi_2), \mathcal{A}(\xi_1)] d\xi \\ & + \frac{1}{12} \int_0^t \int_0^{\xi_1} \int_0^{\xi_1} [\mathcal{A}(\xi_2), [\mathcal{A}(\xi_3), \mathcal{A}(\xi_1)]] d\xi \\ & + \frac{1}{4} \int_0^t \int_0^{\xi_1} \int_0^{\xi_2} [[\mathcal{A}(\xi_3), \mathcal{A}(\xi_2)], \mathcal{A}(\xi_1)] d\xi + \dots \end{aligned} \quad (7.1)$$

(A more detailed version of the Magnus expansion, (2.53), appears in Section 2.3.1.) Since  $\mathcal{A}(t) \in \mathfrak{H}$ , all commutators can be simplified to commutator-free terms in  $\mathfrak{H}$ . The skew-Hermitian nature of elements of  $\mathfrak{H}$  implies that our Magnus expansion, or any finite truncation in  $\mathfrak{H}$ , remains skew-Hermitian even after spatial discretisation. Once again, this leads to unitary evolution and stability of our Magnus based methods.

In Section 7.2 we develop truncated integral and commutator-free Magnus expansions

$\tilde{\Theta}_p^{\varepsilon[M]}$  that can achieve an arbitrarily high order, expressed in powers of the semiclassical parameter  $\varepsilon$ . We commence our analysis from an integral-free approximation of the Magnus expansion, following the approach of (Munthe-Kaas & Owren 1999) discussed in Section 2.3.5. Here the integrals appearing in the Magnus expansion are replaced by Gauss–Legendre quadratures at the outset. When the interaction potential  $V$  possesses favourable temporal behaviour, values of  $V$  at only a few temporal grid points are required for a high-order method. For  $\mathcal{O}(\varepsilon^{7\sigma-1})$  accuracy, for instance, we require merely three Gauss–Legendre knots<sup>1</sup>,  $t_k = h(1 + k\sqrt{3/5})/2, k = -1, 0, 1$ .

In another approach, presented in Section 7.3, evaluation of integrals is postponed till the very end to derive integral-preserving commutator-free Magnus expansions  $\tilde{\Theta}_p^{\varepsilon[I]}$ . The mathematical machinery we need to introduce here becomes more involved. However, we end up with a highly flexible method—not only is it possible to approximate the integrals through any quadrature method, but we may also use exact integrals for potentials that possess an analytic form. This flexibility can also allow us to tackle effectively potentials with weaker temporal regularity as well as highly oscillatory potentials of certain forms.

Deriving a commutator-free Magnus expansion, however, is only half the story—to develop an effective scheme for (3.20), we need to find efficient ways to approximate the exponential of these expansions. In Section 7.4 we discuss the cost of exponentiation via Lanczos iterations and Yoshida splittings. More effective means are only presented in Chapter 8 when we combine Magnus expansions with Zassenhaus splittings.

## 7.1 Truncation by powers of the semiclassical parameter

From Section 2.3, we recall that the Magnus expansion can be written in the form,

$$\Theta(h) = \sum_{k=1}^{\infty} \Theta^{[k]}(h) = \sum_{k=1}^{\infty} \sum_{\tau \in \mathbb{T}_k} \alpha(\tau) \mathcal{C}_{\tau}(h), \quad h \geq 0, \quad (2.54)$$

where  $\mathbb{T}_k$  is the set of trees featuring  $k$  integrals (and  $k$  commutators). As we noted in Section 2.3.4, it is more advantageous to use the *power truncated* Magnus expansion,

$$\Theta_p(h) = \sum_{k=1}^p \sum_{\tau \in \mathbb{F}_k} \alpha(\tau) \mathcal{C}_{\tau}(h), \quad (2.55)$$

where we include trees based on the sets  $\mathbb{F}_k$  (defined in terms of powers of  $h$ ) instead of the sets  $\mathbb{T}_k$  (based on number of integrals). Such a power truncated Magnus expansion has the desirable characteristic of being odd in  $h$  due to the time symmetry of its flow

<sup>1</sup>Although a Gauss–Lengendre quadrature with three knots usually has an accuracy of  $\mathcal{O}(h^6)$ , due to time-symmetry and odd nature of power-truncated Magnus expansions, all even terms in the Taylor series, including error terms, vanish. Consequently, these methods become  $\mathcal{O}(h^7)$  accuracy in our context. See Section 2.3.4 and Section 2.3.5.

(Iserles et al. 2000, Iserles & Nørsett 1999, Iserles et al. 2001). Odd-indexed methods of this form, consequently, gain an extra power of  $h$ ,

$$\Theta_{2p-1}(h) = \Theta(h) + \mathcal{O}(h^{2p+1}). \quad (2.59)$$

Once we start analysing the Magnus expansion for the Schrödinger equation in the currency of  $\varepsilon$ , however, we need to consider truncating the sets of trees by powers of  $\varepsilon$ . For this purpose we define the set  $\mathbb{E}_k$  along similar lines to the definition of  $\mathbb{F}_k$ :  $\tau \in \mathbb{E}_k$  if  $k$  is the *greatest* integer such that  $\tau = \mathcal{O}(\varepsilon^{k\sigma-1})$  for all smooth  $V(\cdot, t)$ . It turns out that the sets  $\mathbb{E}_k$  and  $\mathbb{F}_k$  are, in fact, identical under the choice of spatial discretisation scaling  $M = \mathcal{O}(\varepsilon^{-1})$ .

**Theorem 7.1.1.**  $\mathbb{E}_m = \mathbb{F}_m$

*Proof.* Consider any  $\tau \in \mathbb{F}_m$ . Clearly  $\tau \in \mathbb{T}_n$  for some  $n \leq m$ , and we know that  $\mathcal{C}_\tau(h)$  is a nested integral of a commutator  $C(\mathcal{A}(\xi_1), \dots, \mathcal{A}(\xi_n))$ , where  $\mathcal{A}(\xi_k) = i\varepsilon\partial_x^2 - i\varepsilon^{-1}V(\xi_k)$  for  $\xi_k \in \mathbb{R}$ . By linearity of the Lie bracket, this commutator expands to a linear combination of commutators of the form  $C(E_1, \dots, E_n)$ , where each  $E_k \in \{i\varepsilon\langle 1 \rangle_2\} \cup \{i\varepsilon^{-1}\langle V(\xi) \rangle_0 : \xi \in \mathbb{R}\}$ .

Not all of these commutators end up vanishing, and we restrict our attention to those with non-negative height,  $C(E_1, \dots, E_n) \neq 0$ . By Lemma 6.6.1, any commutator of  $i\varepsilon\langle 1 \rangle_2$  and  $i\varepsilon^{-1}\langle V(\xi_k) \rangle_0$  is  $\mathcal{O}(\varepsilon^{-1})$ . Thus, the largest commutator occurring in  $\tau$  is  $\mathcal{O}(\varepsilon^{-1})$  (note that considerations of time step and integration have not yet been taken into account).

Analysed in terms of the time step  $h$  alone, however, such a term is  $\mathcal{O}(h^m)$  by definition of  $\tau \in \mathbb{F}_m$ . Combining these facts, we see that the tree  $\tau$  is  $\mathcal{O}(\varepsilon^{m\sigma-1})$  in size and consequently also belongs to  $\mathbb{E}_m$ .

Every  $\tau \in \mathbb{E}_m$ , of course, must be in some  $\mathbb{F}_n$  and thus in  $\mathbb{E}_n$ . However, the definition of  $\mathbb{E}_k$  dictates that  $n = m$ . Thus the sets  $\mathbb{F}_m$  and  $\mathbb{E}_m$  must coincide.  $\square$

The consequence of Theorem 7.1.1 is that Magnus expansions truncated by  $\mathbb{E}_k$  (which interest us when analysing in powers of  $\varepsilon$ ),

$$\Theta_p^\varepsilon(h) = \sum_{k=1}^p \sum_{\tau \in \mathbb{E}_k} \alpha(\tau) \mathcal{C}_\tau^\varepsilon(h), \quad (7.2)$$

where the superscript  $\varepsilon$  explicitly acknowledges the dependence on  $\varepsilon$ , are identical to the corresponding Magnus expansions (2.55) truncated by  $\mathbb{F}_k$  (analysed in powers of  $h$ ). Thus we can commence our analysis from the regular power-truncated Magnus expansions.

**Note:** Under grid resolution scaling that differs from our choice of  $M = \mathcal{O}(\varepsilon^{-1})$ , observations such as Theorem 7.1.1 and Lemma 6.6.1 do not hold.

The largest trees that are discarded in this expansion are  $\mathcal{O}(\varepsilon^{(p+1)\sigma-1})$  since they belong to  $\mathbb{E}_{p+1}$ , and the expansion incurs an error of

$$\Theta_p^\varepsilon(h) = \Theta(h) + \mathcal{O}(\varepsilon^{(p+1)\sigma-1}). \quad (7.3)$$

Odd-indexed truncations, as expected, gain a power of  $\varepsilon^\sigma$  since  $h = \mathcal{O}(\varepsilon^\sigma)$ ,

$$\Theta_{2p-1}^\varepsilon(h) = \Theta(h) + \mathcal{O}(\varepsilon^{(2p+1)\sigma-1}). \quad (7.4)$$

We must make a clear distinction between discarding *trees* by powers of  $\varepsilon$  and discarding *terms* obtained upon simplification of these trees in  $\mathfrak{H}$ . A tree  $\tau \in \mathbb{E}_m$  is  $\mathcal{O}(\varepsilon^{m\sigma-1})$  by definition, and is included in the expansion  $\Theta_{2p-1}^\varepsilon$  if  $m \leq 2p-1$ . Upon simplifying, however, it might feature  $\mathcal{O}(\varepsilon^{2p\sigma-1})$ ,  $\mathcal{O}(\varepsilon^{(2p+1)\sigma-1})$  or smaller terms, which must, nevertheless, be retained in the expansion  $\Theta_{2p-1}^\varepsilon$  for the sake of time symmetry.

Once the desirable time symmetry properties are fully exploited, though, we discard all  $\mathcal{O}(\varepsilon^{(2p+1)\sigma-1})$  and smaller terms from  $\Theta_{2p-1}^\varepsilon$  and arrive at  $\tilde{\Theta}_{2p-1}^\varepsilon$ , which carries an error of

$$\tilde{\Theta}_{2p-1}^\varepsilon(h) = \Theta_{2p-1}^\varepsilon(h) + \mathcal{O}(\varepsilon^{(2p+1)\sigma-1}) = \Theta(h) + \mathcal{O}(\varepsilon^{(2p+1)\sigma-1}). \quad (7.5)$$

We are prohibited, however, to discard any  $\mathcal{O}(\varepsilon^{2p\sigma-1})$  terms that might originate in simplifications without compromising on error. For completeness, we also define  $\tilde{\Theta}_{2p}^\varepsilon$  as the expansion obtained after discarding all  $\mathcal{O}(\varepsilon^{(2p+1)\sigma-1})$  and smaller terms from  $\Theta_{2p}^\varepsilon$ . However, it is preferable to work solely with odd-indexed expansions due to the free gain in power.

## 7.2 Approach 1: Discretised integrals

In combining Magnus expansions with the algebraic advantages of working in the Lie algebra  $\mathfrak{H}$ , we encounter some technicalities. In this section we start with the most straightforward combination of these techniques: we commence our analysis directly from the results of (Munthe-Kaas & Owren 1999), discussed in Section 2.3.5, where the integrals occurring in the Magnus expansion are approximated using Gauss–Legendre quadratures at the outset.

Following the techniques discussed here, it should be relatively straightforward to commence the development of a numerical scheme from an alternative quadrature method for approximating the integrals in the Magnus expansion. See (Iserles et al. 2000) and (Blanes et al. 2009) for comprehensive information and ways to approximate the Magnus expansion using different quadrature rules and to higher orders. The former could be relevant if the time-dependent potential is only known at certain grid-points as might be

the case in some control setups.

### 7.2.1 The self-adjoint basis of Munthe-Kaas & Owren

As we saw in Section 2.3.5, the seemingly complicated multivariate integrals occurring in the Magnus expansion (7.1) can be computed using simple univariate quadrature rules. Among the approaches for discretising these integrals, the approach of Munthe-Kaas & Owren (1999) is particularly efficient for smooth  $\mathcal{A}$ . Following this approach, an order six approximation (which incurs an  $\mathcal{O}(h^7)$  error) is obtained by evaluating  $\mathcal{A}$  at three Gauss-Legendre quadrature points<sup>2</sup>

$$t_k = \left(\frac{1}{2} + k\frac{\sqrt{15}}{10}\right)h, \quad k = -1, 0, 1.$$

The values at the Gauss-Legendre knots are used for generating the so called *self-adjoint* basis (2.72),

$$\begin{aligned} J_1 &= h\mathcal{A}(t_0), \\ J_2 &= \frac{\sqrt{15}}{3}h(\mathcal{A}(t_1) - \mathcal{A}(t_{-1})), \\ J_3 &= \frac{10}{3}h(\mathcal{A}(t_1) - 2\mathcal{A}(t_0) + \mathcal{A}(t_{-1})). \end{aligned}$$

The power-truncated Magnus expansion up to order six expressed in terms of these self-adjoint basis is

$$\begin{aligned} \Theta_5^{[M]}(h) &= J_1 + \frac{1}{12}J_3 - \frac{1}{12}[J_1, J_2] + \frac{1}{240}[J_2, J_3] + \frac{1}{360}[J_1, [J_1, J_3]] \\ &\quad - \frac{1}{240}[J_2, [J_1, J_2]] + \frac{1}{720}[J_1, [J_1, [J_1, J_2]]]. \end{aligned} \tag{2.71}$$

We will start our analysis from this result of (Munthe-Kaas & Owren 1999), where we use the superscript  $[M]$  to distinguish this Magnus expansion from later ones discussed in Section 7.3 where integrals are preserved.

In the case of the Schrödinger equation (3.21), the time-dependent component of  $\mathcal{A}(t)$  is  $V(t)$  and in order to evaluate  $J_1, J_2, J_3$ , the values of  $V$  at the three Gauss-Legendre knots are required. As noted in Section 2.3.5, the self-adjoint basis  $J_1, J_2$  and  $J_3$  approximate the degree zero, one and two derivatives of  $\mathcal{A}$  in time, respectively, using central differences on the Gauss-Legendre knots. Foreseeing a corresponding need for the central difference

---

<sup>2</sup>Although a Gauss-Legendre quadrature with three knots usually has an accuracy of  $\mathcal{O}(h^6)$ , due to time-symmetry and odd nature of power-truncated Magnus expansions, all even terms in the Taylor series, including error terms, vanish. Consequently, these methods become  $\mathcal{O}(h^7)$  accuracy in our context. See Section 2.3.4 and Section 2.3.5.



approximations of derivatives of  $V$ , we define

$$\begin{aligned} V_0 &= V(t_0), \\ V_1 &= \frac{\sqrt{15}}{3h}(V(t_1) - V(t_{-1})), \\ V_2 &= \frac{10}{3h^2}(V(t_1) - 2V(t_0) + V(t_{-1})). \end{aligned} \quad (7.6)$$

Since  $\mathcal{A}(t) = i\varepsilon \langle 1 \rangle_2 - i\varepsilon^{-1} \langle V(t) \rangle_0$ , the self-adjoint basis of (2.72) are

$$\begin{aligned} J_1 &= i h \varepsilon \langle 1 \rangle_2 - i h \varepsilon^{-1} \langle V_0 \rangle_0, \\ J_2 &= -i h^2 \varepsilon^{-1} \langle V_1 \rangle_0, \\ J_3 &= -i h^3 \varepsilon^{-1} \langle V_2 \rangle_0. \end{aligned} \quad (7.7)$$

We note that  $J_i \in \mathfrak{H}$ . Consequently, the truncated Magnus expansion (2.71) can be expanded in  $\mathfrak{H}$  to a commutator-free form,  $\sum_k i^{k+1} \langle f_k \rangle_k$  with  $f_k \in C_p^\infty([-1, 1], \mathbb{R})$ .

### 7.2.2 Computations in the Lie algebra $\mathfrak{H}$

For the purpose of an  $\mathcal{O}(\varepsilon^{7\sigma-1})$  Magnus expansion, we only require the following identities,

$$\begin{aligned} [\langle f \rangle_2, \langle g \rangle_2] &= 2 \langle f(\partial_x g) - (\partial_x f)g \rangle_3 + \langle 2(\partial_x^2 f)(\partial_x g) - 2(\partial_x f)(\partial_x^2 g) + (\partial_x^3 f)g - f(\partial_x^3 g) \rangle_1, \\ [\langle f \rangle_2, \langle g \rangle_1] &= \langle 2f(\partial_x g) - (\partial_x f)g \rangle_2 - \frac{1}{2} \langle 2(\partial_x f)(\partial_x^2 g) + f(\partial_x^3 g) \rangle_0, \\ [\langle f \rangle_2, \langle g \rangle_0] &= 2 \langle f(\partial_x g) \rangle_1, \\ [\langle f \rangle_1, \langle g \rangle_1] &= \langle f(\partial_x g) - (\partial_x f)g \rangle_1, \\ [\langle f \rangle_1, \langle g \rangle_0] &= \langle f(\partial_x g) \rangle_0, \\ [\langle f \rangle_0, \langle g \rangle_0] &= 0, \end{aligned} \quad (7.8)$$

which can be read off Table 5.2 and which are a subset of those listed previously in (6.3).

Using (7.7) and the identities (7.8), the grade one commutators of the self-adjoint basis can be simplified as follows,

$$\begin{aligned} [J_1, J_2] &= [i h \varepsilon \langle 1 \rangle_2 - i h \varepsilon^{-1} \langle V_0 \rangle_0, -i h^2 \varepsilon^{-1} \langle V_1 \rangle_0] \\ &= h^3 [\langle 1 \rangle_2, \langle V_1 \rangle_0] = 2h^3 \langle \partial_x V_1 \rangle_1, \end{aligned} \quad (7.9)$$

$$\begin{aligned} [J_1, J_3] &= [i h \varepsilon \langle 1 \rangle_2 - i h \varepsilon^{-1} \langle V_0 \rangle_0, -i h^3 \varepsilon^{-1} \langle V_2 \rangle_0] \\ &= h^4 [\langle 1 \rangle_2, \langle V_2 \rangle_0] = 2h^4 \langle \partial_x V_2 \rangle_1, \end{aligned} \quad (7.10)$$

$$\begin{aligned} [J_2, J_3] &= [-i h^2 \varepsilon^{-1} \langle V_1 \rangle_0, i h^3 \varepsilon^{-1} \langle V_2 \rangle_0] \\ &= 0. \end{aligned} \quad (7.11)$$

Consequently, the grade two commutators appearing in (2.71) are,

$$\begin{aligned}
 [J_1, [J_1, J_3]] &= [\mathrm{i}h\varepsilon \langle 1 \rangle_2 - \mathrm{i}h\varepsilon^{-1} \langle V_0 \rangle_0, 2h^4 \langle \partial_x V_2 \rangle_1] \\
 &= 2\mathrm{i}h^5\varepsilon [\langle 1 \rangle_2, \langle \partial_x V_2 \rangle_1] + 2\mathrm{i}h^5\varepsilon^{-1} [\langle \partial_x V_2 \rangle_1, \langle V_0 \rangle_0] \\
 &= 2\mathrm{i}h^5\varepsilon (2 \langle \partial_x^2 V_2 \rangle_2 - \tfrac{1}{2} \langle \partial_x^4 V_2 \rangle_0) + 2\mathrm{i}h^5\varepsilon^{-1} \langle (\partial_x V_2)(\partial_x V_0) \rangle_0 \\
 &= 4\mathrm{i}h^5\varepsilon \langle \partial_x^2 V_2 \rangle_2 - \mathrm{i}h^5\varepsilon \langle \partial_x^4 V_2 \rangle_0 + 2\mathrm{i}h^5\varepsilon^{-1} \langle (\partial_x V_2)(\partial_x V_0) \rangle_0, \quad (7.12)
 \end{aligned}$$

$$\begin{aligned}
 [J_2, [J_1, J_2]] &= [-\mathrm{i}h^2\varepsilon^{-1} \langle V_1 \rangle_0, 2h^3 \langle \partial_x V_1 \rangle_1] \\
 &= 2\mathrm{i}h^5\varepsilon^{-1} [\langle \partial_x V_1 \rangle_1, \langle V_1 \rangle_0] \\
 &= 2\mathrm{i}h^5\varepsilon^{-1} \langle (\partial_x V_1)^2 \rangle_0, \quad (7.13)
 \end{aligned}$$

$$\begin{aligned}
 [J_1, [J_1, J_2]] &= [\mathrm{i}h\varepsilon \langle 1 \rangle_2 - \mathrm{i}h\varepsilon^{-1} \langle V_0 \rangle_0, 2h^3 \langle \partial_x V_1 \rangle_1] \\
 &= 2\mathrm{i}h^4\varepsilon [\langle 1 \rangle_2, \langle \partial_x V_1 \rangle_1] + 2\mathrm{i}h^4\varepsilon^{-1} [\langle \partial_x V_1 \rangle_1, \langle V_0 \rangle_0] \\
 &= 2\mathrm{i}h^4\varepsilon (2 \langle \partial_x^2 V_1 \rangle_2 - \tfrac{1}{2} \langle \partial_x^4 V_1 \rangle_0) + 2\mathrm{i}h^4\varepsilon^{-1} \langle (\partial_x V_1)(\partial_x V_0) \rangle_0 \\
 &= 4\mathrm{i}h^4\varepsilon \langle \partial_x^2 V_1 \rangle_2 - \mathrm{i}h^4\varepsilon \langle \partial_x^4 V_1 \rangle_0 + 2\mathrm{i}h^4\varepsilon^{-1} \langle (\partial_x V_1)(\partial_x V_0) \rangle_0. \quad (7.14)
 \end{aligned}$$

The only grade three commutator that we need is

$$\begin{aligned}
 [J_1, [J_1, [J_1, J_2]]] &= [\mathrm{i}h\varepsilon \langle 1 \rangle_2 - \mathrm{i}h\varepsilon^{-1} \langle V_0 \rangle_0, \quad (7.15) \\
 &\quad 4\mathrm{i}h^4\varepsilon \langle \partial_x^2 V_1 \rangle_2 - \mathrm{i}h^4\varepsilon \langle \partial_x^4 V_1 \rangle_0 + 2\mathrm{i}h^4\varepsilon^{-1} \langle (\partial_x V_1)(\partial_x V_0) \rangle_0] \\
 &= -4h^5\varepsilon^2 [\langle 1 \rangle_2, \langle \partial_x^2 V_1 \rangle_2] + h^5\varepsilon^2 [\langle 1 \rangle_2, \langle \partial_x^4 V_1 \rangle_0] \\
 &\quad - 2h^5 [\langle 1 \rangle_2, \langle (\partial_x V_1)(\partial_x V_0) \rangle_0] - 4h^5 [\langle \partial_x^2 V_1 \rangle_2, \langle V_0 \rangle_0] \\
 &= -4h^5\varepsilon^2 (2 \langle \partial_x^3 V_1 \rangle_3 - \langle \partial_x^5 V_1 \rangle_1) + 2h^5\varepsilon^2 \langle \partial_x^5 V_1 \rangle_1 \\
 &\quad - 4h^5 \langle (\partial_x^2 V_1)(\partial_x V_0) + (\partial_x V_1)(\partial_x^2 V_0) \rangle_1 - 8h^5 \langle (\partial_x^2 V_1)(\partial_x V_0) \rangle_1 \\
 &= -8h^5\varepsilon^2 \langle \partial_x^3 V_1 \rangle_3 + 3h^5\varepsilon^2 \langle \partial_x^5 V_1 \rangle_1 - h^5 \langle 12(\partial_x^2 V_1)(\partial_x V_0) + 4(\partial_x V_1)(\partial_x^2 V_0) \rangle_1.
 \end{aligned}$$

### 7.2.3 Commutator-free Magnus expansion

Substituting (7.9–7.15) in (2.71) gives us a Magnus expansion for (3.21) in the Lie algebra  $\mathfrak{H}$ ,

$$\begin{aligned}
 \Theta_5^{\varepsilon[M]}(h) &= J_1 + \frac{1}{12}J_3 - \frac{1}{12}[J_1, J_2] + \frac{1}{240}[J_2, J_3] + \frac{1}{360}[J_1, [J_1, J_3]] \\
 &\quad - \frac{1}{240}[J_2, [J_1, J_2]] + \frac{1}{720}[J_1, [J_1, [J_1, J_2]]] \\
 &= i h \varepsilon \langle 1 \rangle_2 - i h \varepsilon^{-1} \langle V_0 \rangle_0 - \frac{1}{12} i h^3 \varepsilon^{-1} \langle V_2 \rangle_0 - \frac{1}{6} h^3 \langle \partial_x V_1 \rangle_1 \\
 &\quad + \frac{1}{360} (4 i h^5 \varepsilon \langle \partial_x^2 V_2 \rangle_2 - i h^5 \varepsilon \langle \partial_x^4 V_2 \rangle_0 + 2 i h^5 \varepsilon^{-1} \langle (\partial_x V_2)(\partial_x V_0) \rangle_0) \\
 &\quad - \frac{1}{120} i h^5 \varepsilon^{-1} \langle (\partial_x V_1)^2 \rangle_0 + \frac{1}{720} (-8 h^5 \varepsilon^2 \langle \partial_x^3 V_1 \rangle_3 + 3 h^5 \varepsilon^2 \langle \partial_x^5 V_1 \rangle_1 \\
 &\quad \quad - h^5 \langle 12(\partial_x^2 V_1)(\partial_x V_0) + 4(\partial_x V_1)(\partial_x^2 V_0) \rangle_1) . \\
 \\
 \Theta_5^{\varepsilon[M]}(h) &= \overbrace{i h \varepsilon \partial_x^2 - i h \varepsilon^{-1} V_0}^{\mathcal{O}(\varepsilon^{\sigma-1})} - \overbrace{\frac{1}{12} i h^3 \varepsilon^{-1} V_2 - \frac{1}{6} h^3 \langle \partial_x V_1 \rangle_1}^{\mathcal{O}(\varepsilon^{3\sigma-1})} \\
 &\quad + \overbrace{\frac{1}{360} i h^5 \varepsilon^{-1} (2(\partial_x V_2)(\partial_x V_0) - 3(\partial_x V_1)^2)}^{\mathcal{O}(\varepsilon^{5\sigma-1})} - \overbrace{\frac{1}{180} h^5 \langle (\partial_x V_1)(\partial_x^2 V_0) + 3(\partial_x V_0)(\partial_x^2 V_1) \rangle_1}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\
 &\quad + \overbrace{\frac{1}{90} i h^5 \varepsilon \langle \partial_x^2 V_2 \rangle_2 - \frac{1}{90} h^5 \varepsilon^2 \langle \partial_x^3 V_1 \rangle_3}^{\mathcal{O}(\varepsilon^{5\sigma-1})} - \overbrace{\frac{1}{360} i h^5 \varepsilon \langle \partial_x^4 V_2 \rangle + \frac{1}{240} h^5 \varepsilon^2 \langle \partial_x^5 V_1 \rangle_1}^{\mathcal{O}(\varepsilon^{5\sigma+1})} \\
 &= \Theta + \mathcal{O}(\varepsilon^{7\sigma-1}) .
 \end{aligned} \tag{7.16}$$

We remind the reader that the superscript  $[M]$  is used for distinguishing this Magnus expansion from later ones discussed in Section 7.3 where integrals are preserved. For  $\sigma \leq 1$ , the last two terms in  $\Theta_5^{\varepsilon[M]}(h)$ , which are  $\mathcal{O}(\varepsilon^{5\sigma+1})$ , become  $\mathcal{O}(\varepsilon^{7\sigma-1})$  and can be

discarded. After discarding these terms, the Magnus expansion reduces to

$$\begin{aligned}
 \tilde{\Theta}_5^{\varepsilon[M]}(h) = & \overbrace{i h \varepsilon \partial_x^2 - i h \varepsilon^{-1} V_0}^{\mathcal{O}(\varepsilon^{\sigma-1})} - \overbrace{\frac{1}{12} i h^3 \varepsilon^{-1} V_2 - \frac{1}{6} h^3 \langle \partial_x V_1 \rangle_1}^{\mathcal{O}(\varepsilon^{3\sigma-1})} \\
 & + \overbrace{\frac{1}{360} i h^5 \varepsilon^{-1} \left( 2(\partial_x V_2)(\partial_x V_0) - 3(\partial_x V_1)^2 \right)}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\
 & - \overbrace{\frac{1}{180} h^5 \langle (\partial_x V_1)(\partial_x^2 V_0) + 3(\partial_x V_0)(\partial_x^2 V_1) \rangle_1}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\
 & + \overbrace{\frac{1}{90} i h^5 \varepsilon \langle \partial_x^2 V_2 \rangle_2 - \frac{1}{90} h^5 \varepsilon^2 \langle \partial_x^3 V_1 \rangle_3}^{\mathcal{O}(\varepsilon^{5\sigma-1})} = \Theta + \mathcal{O}(\varepsilon^{7\sigma-1}).
 \end{aligned} \tag{7.17}$$

Since  $\mathcal{A}(t) \in \mathfrak{H}$  and  $\mathcal{A}(t) = i(\varepsilon \langle 1 \rangle_2 - \varepsilon^{-1} \langle V(t) \rangle_0) = \mathcal{O}(\varepsilon^{-1})$ ,  $t \geq 0$ , Lemma 6.6.1 implies,

$$\|\mathcal{A}(\tau_k), [\dots, [\mathcal{A}(\tau_1), \mathcal{A}(\tau_0)] \dots]\|_2 = \mathcal{O}(\varepsilon^{-1}), \quad \tau_i \in [0, T], \quad i = 0, \dots, k, \quad k \in \mathbb{Z}^+, \tag{3.26}$$

which is responsible for the terms in the Magnus expansion being progressively smaller, even for large time steps. Due to Corollary 6.6.2, a grade  $n$  term in the Magnus expansion of  $\mathcal{A}(t)$  is  $\mathcal{O}(\varepsilon^{n\sigma-1})$ . Thus, asymptotically speaking in terms of  $\varepsilon$ , the terms in the expansion are decreasing in size with increasing  $n$  for any  $\sigma > 0$ .

### 7.3 Approach 2: Undiscretised integrals

The Magnus expansion (7.2), truncated by powers of  $\varepsilon$ , features integrals that were approximated in the previous section by Gauss–Legendre quadratures. Getting rid of the integrals in this way allowed us to simplify the development of the commutator-free Magnus expansion greatly: once we commence from the optimised Magnus expansion of (Munthe–Kaas & Owren 1999), relatively straightforward algebraic manipulations in the Lie algebra  $\mathfrak{H}$  bring us to the commutator-free Magnus expansion  $\tilde{\Theta}_5^{\varepsilon[M]}$  (7.17) in just a few steps.

While this simplifies the analysis greatly, there is no fundamental reason why we cannot perform similar algebraic manipulations while keeping the integrals intact. There are many good reasons for attempting to do so—the result of such a procedure would be a method that could be easily tailored to potential functions that can be integrated in closed form, are optimally discretised via quadrature methods other than the Gauss–Legendre quadrature (for highly-oscillatory quadrature, or uniform grids, for instance), or possess a lower degree of temporal regularity.

In this section we demonstrate how commutators in a power-truncated Magnus expan-

sion

$$\Theta_4(h) = \text{[diagram: a single vertical line with a dot at the top]} - \frac{1}{2} \text{[diagram: a V-shape with a dot at the top]} + \frac{1}{12} \text{[diagram: a more complex tree with 4 nodes]} + \frac{1}{4} \text{[diagram: another tree with 4 nodes]}, \quad (7.18)$$

where integrals haven't been discretised, can still be expanded in the algebra  $\mathfrak{G}$ .

### 7.3.1 Computing with integrals intact

We note that all integrals occurring in our work are in the temporal domain, while the differential operators are in the spatial domain. Consequently, integral and differential operators occurring out here commute,

$$\partial_x^k \left( \int_0^h V(\xi) d\xi \right) = \int_0^h (\partial_x^k V(\xi)) d\xi,$$

and

$$\left( \int_0^h V(\xi) d\xi \right) \partial_x^k = \int_0^h V(\xi) \partial_x^k d\xi,$$

for instance. In the context of the Lie algebra  $\mathfrak{G}$ ,

$$\int_0^h \langle V(\xi) \rangle_k d\xi = \left\langle \int_0^h V(\xi) d\xi \right\rangle_k$$

and it is the latter representation that we will use as the canonical form.

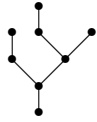
With  $A(\zeta) = i\varepsilon \langle 1 \rangle_2 - i\varepsilon^{-1} \langle V(\zeta) \rangle_0$ , the first term of the Magnus expansion is the integral

$$\text{[diagram: a single vertical line with a dot at the top]} : \int_0^h A(\zeta) d\zeta = ih\varepsilon \langle 1 \rangle_2 - i\varepsilon^{-1} \left\langle \int_0^h V(\zeta) d\zeta \right\rangle_0. \quad (7.19)$$

Other terms can be written as nested integrals of commutators  $C(\mathcal{A}_1, \dots, \mathcal{A}_n)$ ,  $\mathcal{A}_j = \mathcal{A}(\xi_j) = i\varepsilon \langle 1 \rangle_2 - i\varepsilon^{-1} \langle V(\xi_j) \rangle_0$ , which can again be worked out in the Lie algebra  $\mathfrak{G}$ . The first non-trivial tree, for instance, is simplified as

$$\begin{aligned} \text{[diagram: a tree with 3 nodes]} & : \int_0^h \left[ \int_0^\zeta \mathcal{A}(\xi) d\xi, A(\zeta) \right] d\zeta = \int_0^h \int_0^\zeta [\mathcal{A}(\xi), \mathcal{A}(\zeta)] d\xi d\zeta \\ & = \int_0^h \int_0^\zeta [i\varepsilon \langle 1 \rangle_2 - i\varepsilon^{-1} \langle V(\xi) \rangle_0, i\varepsilon \langle 1 \rangle_2 - i\varepsilon^{-1} \langle V(\zeta) \rangle_0] d\xi d\zeta \\ & = \int_0^h \int_0^\zeta [\langle 1 \rangle_2, \langle V(\zeta) \rangle_0] + [\langle V(\xi) \rangle_0, \langle 1 \rangle_2] d\xi d\zeta \\ & = 2 \left( \int_0^h \zeta \langle (\partial_x V(\zeta)) \rangle_1 d\zeta - \int_0^h \int_0^\zeta \langle (\partial_x V(\xi)) \rangle_1 d\xi d\zeta \right) \\ & = 2 \left\langle \int_0^h \zeta (\partial_x V(\zeta)) d\zeta - \int_0^h \int_0^\zeta (\partial_x V(\xi)) d\xi d\zeta \right\rangle_1. \end{aligned} \quad (7.20)$$

Similarly, we can solve



$$\begin{aligned}
 & : 4i\varepsilon \left\langle \int_0^h \zeta^2 (\partial_x^2 V(\zeta)) d\zeta - \int_0^h \zeta \int_0^\zeta (\partial_x^2 V(\xi)) d\xi d\zeta \right\rangle_2 \\
 & + 2i\varepsilon^{-1} \left\langle \int_0^h \zeta (\partial_x V(\zeta)) \int_0^\zeta (\partial_x V(\xi)) d\xi d\zeta - \int_0^h \left( \int_0^\zeta (\partial_x V(\xi)) d\xi \right)^2 d\zeta \right\rangle_0 \\
 & - i\varepsilon \left\langle \int_0^h \zeta^2 (\partial_x^4 V(\zeta)) d\zeta - \int_0^h \zeta \int_0^\zeta (\partial_x^4 V(\xi)) d\xi d\zeta \right\rangle_0. \tag{7.21}
 \end{aligned}$$

### 7.3.2 Simplifying integrals over Magnus polytopes

After simplifying terms in the Magnus expansion (2.54) we arrive at expressions such as (7.20) and (7.21), where each integral is of the form

$$I_{\mathcal{S}, \mathbf{P}}(h) = \int_{\mathcal{S}} P(\boldsymbol{\xi}) d\boldsymbol{\xi},$$

where  $\mathbf{P}(\boldsymbol{\xi}) = \prod_{j=1}^s P_j(\xi_j)$  for some  $P_j$ , and  $\mathcal{S}$  is an  $s$ -dimensional polytope of the special form,

$$\mathcal{S} = \{\boldsymbol{\xi} \in \mathbb{R}^s : \xi_1 \in [0, h], \quad \xi_l \in [0, \xi_{m_l}], \quad l = 2, 3, \dots, s\},$$

with  $m_l \in \{1, 2, \dots, l-1\}$ ,  $l = 2, 3, \dots, s$  (Iserles et al. 2000).

Integration over the polytope  $\mathcal{S}$  in the temporal domain has remained an afterthought up to this stage. The special forms of these polytopes, and of certain integrands obtained after expanding the commutators, allows us to simplify the terms of the Magnus expansion further. Integration by parts leads us to the following identities:

$$\int_0^h P_1(\xi_1) \left( \int_0^{\xi_1} P_2(\xi_2) d\xi_2 \right) d\xi_1 = \int_0^h P_2(\xi_1) \left( \int_{\xi_1}^h P_1(\xi_2) d\xi_2 \right) d\xi_1, \tag{7.22}$$

$$\begin{aligned}
 & \int_0^h P_1(\xi_1) d\xi_1 \left( \int_0^{\xi_1} P_2(\xi_2) d\xi_2 \right) \left( \int_0^{\xi_1} P_3(\xi_3) d\xi_3 \right) d\xi_1 \\
 & = \int_0^h \left( \int_{\xi_3}^h P_1(\xi_1) d\xi_1 \right) \left( P_2(\xi_3) \int_0^{\xi_3} P_3(\xi_2) d\xi_2 + P_3(\xi_3) \int_0^{\xi_3} P_2(\xi_2) d\xi_2 \right) d\xi_3.
 \end{aligned} \tag{7.23}$$

In our simplifications, (7.20) and (7.21), we have already encountered integrals over a triangle such as  $\int_0^h \int_0^\zeta (\partial_x V(\xi)) d\xi d\zeta$  and  $\int_0^h \zeta \int_0^\zeta (\partial_x^2 V(\xi)) d\xi d\zeta$ . We can reduce these to integrations over a line by applying the first identity with  $P_1(\xi_1) = 1$ ,  $P_2(\xi_2) = \partial_x V(\xi_2)$  and  $P_1(\xi_1) = \xi_1$ ,  $P_2(\xi_2) = \partial_x V(\xi_2)$ , respectively. Integration over the pyramid in  $\int_0^h \left( \int_0^\zeta (\partial_x V(\xi)) d\xi \right)^2 d\zeta$  is similarly reduced using the second identity with  $P_1(\xi_1) = 1$ ,  $P_2(\xi_2) = \partial_x V(\xi_2)$ ,  $P_3(\xi_3) = \partial_x V(\xi_3)$ .

Although it might be possible to develop a general formalism for extending these observations to higher dimensional polytopes appearing in the Magnus expansion, the two identities presented here suffice for all results presented in our work. Deducing similarly useful identities for reduction of nested integrals in any specific high dimensional polytope case should also be possible.

Applying (7.22) with  $P_1(\zeta) = 1$  and  $P_2(\xi) = \partial_x V(\xi)$ , the second integral in (7.20) can be reduced to a line integral,

$$\int_0^h \int_0^\zeta (\partial_x V(\xi)) d\xi d\zeta = \int_0^h (h - \zeta) (\partial_x V(\zeta)) d\zeta.$$

After similar simplifications in  $\mathfrak{H}$  and applications of the integration identities (7.22) and (7.23), the trees in  $\mathbb{F}_1, \mathbb{F}_3$  and  $\mathbb{F}_4$ , which appear in the truncated Magnus expansion

$$\Theta_4(h) = \text{tree}_1 - \frac{1}{2} \text{tree}_2 + \frac{1}{12} \text{tree}_3 + \frac{1}{4} \text{tree}_4, \quad (7.18)$$

simplify to the forms

$$\text{tree}_1 : ih\varepsilon \partial_x^2 - i\varepsilon^{-1} \int_0^h V(\zeta) d\zeta, \quad (7.24)$$

$$\text{tree}_2 : 4 \left\langle \int_0^h \left( \zeta - \frac{h}{2} \right) (\partial_x V(\zeta)) d\zeta \right\rangle_1, \quad (7.25)$$

$$\begin{aligned} \text{tree}_3 : & -2i\varepsilon^{-1} \int_0^h \int_0^\zeta (2h - 3\zeta) (\partial_x V(\zeta)) (\partial_x V(\xi)) d\xi d\zeta \\ & - 2i\varepsilon \left\langle \int_0^h (h^2 - 3\zeta^2) (\partial_x^2 V(\zeta)) d\zeta \right\rangle_2, \end{aligned} \quad (7.26)$$

$$\begin{aligned} \text{tree}_4 : & 2i\varepsilon^{-1} \int_0^h \int_0^\zeta (\zeta - 2\xi) (\partial_x V(\zeta)) (\partial_x V(\xi)) d\xi d\zeta \\ & + 2i\varepsilon \left\langle \int_0^h (h^2 - 4h\zeta + 3\zeta^2) (\partial_x^2 V(\zeta)) d\zeta \right\rangle_2. \end{aligned} \quad (7.27)$$

The term  $\left\langle \int_0^h \left(\zeta - \frac{h}{2}\right) (\partial_x V(\zeta)) d\zeta \right\rangle_1$ , which occurs in the simplification of the tree

$$\begin{array}{c} \bullet \\ | \\ \bullet \text{---} \bullet \\ | \\ \bullet \end{array} \in \mathbb{F}_3,$$

might seem to be  $\mathcal{O}(h^2)$  at first sight, contradicting our expectations from a tree in  $\mathbb{F}_3$ . A closer look at the special form of the integrand, however, shows that the term is indeed  $\mathcal{O}(h^3)$ . To observe this, consider  $V(\zeta)$  expanded about  $h = 0$  so that  $V(h) = V(0) + \sum_{k=1}^{\infty} h^k V^{(k)}(0)/k!$  and note that the  $h^2$  term  $\int_0^h \left(\zeta - \frac{h}{2}\right) (\partial_x V(0)) d\zeta$  vanishes. Similar care has to be exercised throughout the simplifications in combining the appropriate terms before analysing size.

### 7.3.3 Time symmetry and gain of order

The Magnus expansion  $\tilde{\Theta}_4^{\varepsilon[I]}$  (where the superscript  $[I]$  indicates the integral-preserving Magnus expansion and distinguishes this Magnus expansion from the ones derived in Section 7.2 after discretising integrals), obtained by discarding all  $\mathcal{O}(\varepsilon^{5\sigma-1})$  terms from  $\Theta_4$  (7.18–7.27),

$$\begin{aligned} \tilde{\Theta}_4^{\varepsilon[I]}(h) = & \overbrace{ih\varepsilon\partial_x^2 - i\varepsilon^{-1} \int_0^h V(\zeta) d\zeta}^{\mathcal{O}(\varepsilon^{\sigma-1})} - 2 \overbrace{\left\langle \int_0^h \left(\zeta - \frac{h}{2}\right) (\partial_x V(\zeta)) d\zeta \right\rangle_1}^{\mathcal{O}(\varepsilon^{3\sigma-1})} \\ & + \overbrace{i\varepsilon^{-1} \int_0^h \int_0^\zeta \left(\zeta - \xi - \frac{h}{3}\right) [\partial_x V(\zeta)] [\partial_x V(\xi)] d\xi d\zeta}^{\mathcal{O}(\varepsilon^{4\sigma-1})} \\ & + 2i\varepsilon \overbrace{\left\langle \int_0^h \left(\zeta^2 - h\zeta + \frac{h^2}{6}\right) (\partial_x^2 V(\zeta)) d\zeta \right\rangle_2}^{\mathcal{O}(\varepsilon^{4\sigma-1})}, \end{aligned} \quad (7.28)$$

shares an  $\mathcal{O}(\varepsilon^{5\sigma-1})$  error with the simpler expansion

$$\tilde{\Theta}_3^{\varepsilon[I]}(h) = \overbrace{ih\varepsilon\partial_x^2 - i\varepsilon^{-1} \int_0^h V(\zeta) d\zeta}^{\mathcal{O}(\varepsilon^{\sigma-1})} - 2 \overbrace{\left\langle \int_0^h \left(\zeta - \frac{h}{2}\right) (\partial_x V(\zeta)) d\zeta \right\rangle_1}^{\mathcal{O}(\varepsilon^{3\sigma-1})}. \quad (7.29)$$

This gain of order in case of  $\tilde{\Theta}_3^{\varepsilon[I]}(h)$  is explained by (7.4) and, as we had remarked earlier, comes about due to the time symmetry of the Magnus expansion when truncated by  $\mathbb{E}_k$  (or equivalently, by  $\mathbb{F}_k$ ).

Before rushing on to make a conjecture that all terms of the form  $\mathcal{O}(\varepsilon^{2k\sigma-1})$ , such as the  $\mathcal{O}(\varepsilon^{4\sigma-1})$  term in  $\tilde{\Theta}_4^{\varepsilon[I]}$ , may be discarded at no cost in general, it is worth remembering



that the increased order (7.4) due to time symmetry of an odd-indexed method,  $\tilde{\Theta}_{2p-1}^{\varepsilon[I]}$ , only allows us to discard the ‘penultimate’  $\mathcal{O}(\varepsilon^{2p\sigma-1})$  trees. Thus  $\tilde{\Theta}_5^{\varepsilon[I]}$  will once again feature the discarded  $\mathcal{O}(\varepsilon^{4\sigma-1})$  trees while being free to discard the  $\mathcal{O}(\varepsilon^{6\sigma-1})$  trees, and  $\tilde{\Theta}_7^{\varepsilon[I]}$  will feature trees of sizes  $\mathcal{O}(\varepsilon^{k\sigma-1})$  for  $k = 1, 3, 4, 5, 6, 7$  while being free to discard the  $\mathcal{O}(\varepsilon^{8\sigma-1})$  trees. Terms (to be contrasted with trees) of size  $\mathcal{O}(\varepsilon^{8\sigma-1})$  could appear from the simplification of included trees in such situations, however, and we are prohibited to discard them.

A further exploitation of time symmetry is explored in Section 8.4 where we develop expansions that discard all  $\mathcal{O}(\varepsilon^{2k\sigma-1})$  terms.

### 7.3.4 A simplifying notation

The algebraic workings become increasingly convoluted once we start dealing with larger nested commutators and integrals. It becomes helpful to introduce a notation for the integrals on the line,

$$\mu_{j,k}(h) = \int_0^h \tilde{\mathcal{B}}_j^k(h, \zeta) V(\zeta) d\zeta, \quad (7.30)$$

and integrals over the triangle,

$$\Lambda[f]_{a,b}(h) = \int_0^h \int_0^\zeta f(h, \zeta, \xi) \partial_x^a V(\zeta) \partial_x^b V(\xi) d\xi d\zeta, \quad (7.31)$$

where  $\tilde{\mathcal{B}}$  is a rescaling of the Bernoulli polynomials (Abramowitz & Stegun 1964, Lehmer 1988),

$$\tilde{\mathcal{B}}_j(h, \zeta) = h^j \mathcal{B}_j(\zeta/h).$$

The Bernoulli polynomials and their scaled versions can be described in terms of the Bernoulli numbers,

$$\begin{aligned} \mathcal{B}_n(\zeta) &= \sum_{k=0}^n \binom{n}{k} B_{n-k} \zeta^k, \\ \tilde{\mathcal{B}}_n(h, \zeta) &= \sum_{k=0}^n \binom{n}{k} B_{n-k} \zeta^k h^{n-k}. \end{aligned}$$

The first few Bernoulli polynomials are

$$\begin{aligned} \mathcal{B}_0(\zeta) &= 1, \\ \mathcal{B}_1(\zeta) &= \zeta - \frac{1}{2}, \\ \mathcal{B}_2(\zeta) &= \zeta^2 - \zeta + \frac{1}{6}, \\ \mathcal{B}_3(\zeta) &= \zeta^3 - \frac{3}{2}\zeta^2 + \frac{1}{2}\zeta, \end{aligned}$$

and the scaled versions are

$$\begin{aligned}\tilde{\mathcal{B}}_0(h, \zeta) &= 1, \\ \tilde{\mathcal{B}}_1(h, \zeta) &= \zeta - \tfrac{1}{2}h, \\ \tilde{\mathcal{B}}_2(h, \zeta) &= \zeta^2 - \zeta h + \tfrac{1}{6}h^2, \\ \tilde{\mathcal{B}}_3(h, \zeta) &= \zeta^3 - \tfrac{3}{2}\zeta^2 h + \tfrac{1}{2}\zeta h^2.\end{aligned}$$

Since the  $j$ th scaled Bernoulli polynomial,  $\tilde{\mathcal{B}}_j(h, \zeta)$ , scales as  $\mathcal{O}(h^j)$  for  $\zeta = \mathcal{O}(h)$ , we expect  $\mu_{j,k}(h) = \mathcal{O}(h^{j+k+1})$  but, since integrals of the Bernoulli polynomials vanish,

$$\int_0^h \mathcal{B}_j(h, \zeta) d\zeta = 0, \quad (7.32)$$

the term  $\mu_{j,1}(h)$  gains an extra power of  $h$  (following Taylor analysis) and is  $\mathcal{O}(h^{j+2})$ . With this new notation in place, the Magnus expansions  $\tilde{\Theta}_3^{\varepsilon[I]}$  and  $\tilde{\Theta}_4^{\varepsilon[I]}$  can be presented more concisely,

$$\tilde{\Theta}_3^{\varepsilon[I]}(h) = \overbrace{i h \varepsilon \partial_x^2 - i \varepsilon^{-1} \mu_{0,0}(h)}^{\mathcal{O}(\varepsilon^{\sigma-1})} - 2 \overbrace{\langle \partial_x \mu_{1,1}(h) \rangle_1}_{\mathcal{O}(\varepsilon^{3\sigma-1})}, \quad (7.33)$$

$$\tilde{\Theta}_4^{\varepsilon[I]}(h) = \tilde{\Theta}_3^{\varepsilon[I]}(h) + \overbrace{i \varepsilon^{-1} \Lambda[\phi]_{1,1}(h) + 2i \varepsilon \langle \partial_x^2 \mu_{2,1}(h) \rangle_2}_{\mathcal{O}(\varepsilon^{4\sigma-1})}, \quad (7.34)$$

with

$$\phi(h, \zeta, \xi) = \zeta - \xi - \tfrac{h}{3}.$$

In general, for a polynomial  $p_n(h, \zeta, \xi)$  featuring *only* degree- $n$  terms in  $h, \zeta$  and  $\xi$ , the linear (integral) functional  $\Lambda[p_n]_{a,b}(h)$  is  $\mathcal{O}(h^{n+2})$ . However, the integral of  $\phi$  over the triangle vanishes,

$$\int_0^h \int_0^\zeta \phi(h, \zeta, \xi) d\xi d\zeta = 0, \quad (7.35)$$

lending an extra power of  $h$  to terms featuring  $\Lambda[\phi]_{a,b}(h)$ .

## 7.4 Exponentiation of Magnus expansions

Once a suitably truncated Magnus expansion  $\tilde{\Theta}_p^\varepsilon(h)$  has been derived— $\tilde{\Theta}_5^{\varepsilon[M]}$  or  $\tilde{\Theta}_3^{\varepsilon[I]}$ , for instance—in order to approximate the exact solution,

$$u(h) = \exp(\Theta(h))u(0),$$

by a numerical solution,

$$\tilde{u}(h) = \exp(\tilde{\Theta}_p^\varepsilon(h))u(0),$$

we still need to approximate effectively the exponential  $\exp(\tilde{\Theta}_p^\varepsilon(h))u(0)$ . This is certainly no easier than the challenge of exponentiating  $ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V$  that we face in the case of the time-independent potential: for  $p \geq 2$ ,  $\tilde{\Theta}_p^\varepsilon(h)$  always features a term of the form  $ih\varepsilon\partial_x^2 - i\varepsilon^{-1}f(h)$  that possesses similar structure and size, with  $f(h) = \mathcal{O}(h)$  being some linear functional of the potential. The matter is complicated further by the inclusion of increasingly complicated terms once we consider  $p \geq 3$ , such as the term  $\frac{1}{6}h^3 \langle \partial_x V_1 \rangle_1$  in  $\tilde{\Theta}_5^{\varepsilon[M]}$  (7.17) or  $2 \langle \partial_x \mu_{1,1}(h) \rangle_1$  in  $\tilde{\Theta}_3^{\varepsilon[I]}$  (7.33).

Direct exponentiation via diagonal Padé methods is an even worse approach for a Magnus expansion than it is for the solution of a time-independent Hamiltonian: each exponential here requires  $\mathcal{O}(M^3) = \mathcal{O}(\varepsilon^{-3})$  operations, but now we need to take  $\mathcal{O}(\varepsilon^{-\sigma})$  time steps and the overall cost becomes  $\mathcal{O}(\varepsilon^{-3-\sigma})$ . This holds irrespective of the Magnus expansion under consideration. Higher order Magnus expansions along the lines of Section 7.2 and Section 7.3 involve more terms and become more expensive but are still less expensive than Magnus expansions expressed directly in commutators of matrices (such as those discussed in Section 3.4.1).

Alternatively, one can approximate the exponential of the Magnus expansion using Lanczos iterations, resulting in Magnus–Lanczos schemes. The Magnus expansion, taken as a whole, is  $\mathcal{O}(\varepsilon^{\sigma-1})$  in size. Consequently, at the very least  $m = \mathcal{O}(\varepsilon^{\sigma-1})$  Lanczos iterations are required (for  $\sigma \leq 1$ ) to exponentiate it to reasonable accuracy. The cost per step of the Magnus–Lanczos scheme is, therefore,  $\mathcal{O}(\varepsilon^{\sigma-1}\varepsilon^{-1}\log\varepsilon^{-1})$ . Since we require  $\mathcal{O}(\varepsilon^{-\sigma})$  steps, the global cost comes to  $\mathcal{O}(\varepsilon^{-2}\log\varepsilon^{-1})$ .

The cost is also affected by the specific Magnus expansion that we combine with the Lanczos iterations. The  $\tilde{\Theta}_{2n+1}^\varepsilon(h)$  Magnus expansion with  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  accuracy, simplified in Jordan polynomials of the differential operator, involves solving nested commutators of  $\langle 1 \rangle_2$  and  $\langle V(\zeta) \rangle_0$  of grade  $2n+1$  and lower. Were we to start with such a Magnus expansion,

$$\tilde{\Theta}_{2n+1}^\varepsilon(h) = ih\varepsilon\partial_x^2 + \sum_{k=1}^{2n+1} i^{k+1} \langle f_k \rangle_k,$$

we would require  $2(2n+1) + 2 = 4(n+1)$  FFTs for the computation of  $\tilde{\Theta}_{2n+1}^\varepsilon(h)\mathbf{v}$  in each of the Lanczos iterations.

For a Magnus expansion written in terms of nested commutators of matrices such as the methods discussed in Section 3.4.1, the asymptotic cost would be the same (i.e.  $\mathcal{O}(\varepsilon^{-2}\log\varepsilon^{-1})$ ). However, the number of FFTs would grow much more rapidly than the  $4(n+1)$  FFTs required for a Magnus expansion simplified in  $\mathfrak{H}$  since computing with larger nested commutators becomes progressively more expensive. The most naïve

approach would be to open each commutator  $[A, B] = AB - BA$ , whereby

$$[\mathcal{K}^2, [\mathcal{K}^2, \mathcal{D}_V]]\mathbf{v} = \mathcal{K}^4 \mathcal{D}_V \mathbf{v} - 2\mathcal{K}^2 \mathcal{D}_V \mathcal{K}^2 \mathbf{v} + \mathcal{D}_V \mathcal{K}^4 \mathbf{v}$$

can be computed in  $\mathcal{O}(M \log M)$  operations. However, the number of such terms and therefore the number of FFTs grows much more rapidly in  $n$ .

Yoshida splittings also fare more poorly in the case of time-dependent Hamiltonians. Take the lowest order Yoshida splitting—the Strang splitting. There could be multiple ways to split  $\tilde{\Theta}_{2n+1}^\varepsilon(h)$ . One approach could be to separate the  $\langle 1 \rangle_2$  and the  $\langle \cdot \rangle_0$  terms which are easy to exponentiate,

$$\exp\left(\tilde{\Theta}_{2n+1}^\varepsilon(h)\right) = e^{\frac{1}{2}ih\varepsilon\partial_x^2} e^{\frac{1}{2}if_0} \exp\left(-\langle f_1 \rangle_1 + \sum_{k=3}^{2n+1} i^{k+1} \langle f_k \rangle_k\right) e^{\frac{1}{2}if_0} e^{\frac{1}{2}ih\varepsilon\partial_x^2} + \mathcal{O}(\varepsilon^{3\sigma-1}).$$

However, we still need to repeatedly compose the Strang splitting in order to arrive at the order  $2n + 2$  Yoshida splitting featuring an  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  accuracy. In such a Yoshida splitting, we encounter  $3^n$  exponentials of both  $\frac{1}{2}ih\varepsilon\partial_x^2$  and the central exponent.

Each exponential of  $\partial_x^2$  requires two FFTs. The central exponent can be exponentiated via Lanczos iterations. This time, however, the exponent is  $\mathcal{O}(\varepsilon^{3\sigma-1})$ . For  $\sigma \leq \frac{1}{3}$ , we require at least  $m = \mathcal{O}(\varepsilon^{3\sigma-1})$  Lanczos iterations for reasonable approximation, each of which requires  $4n + 2$  FFTs and costs  $\mathcal{O}((4n + 2)\varepsilon^{-1} \log \varepsilon^{-1})$ . This translates to  $\mathcal{O}((4n + 2) \times 3^n \varepsilon^{3\sigma-1} \varepsilon^{-1} \log \varepsilon^{-1})$  cost for the order  $2n + 2$  Yoshida splitting of the Magnus expansion. However, this merely accounts for one step of the Magnus scheme and propagating over  $[0, T]$  requires  $\mathcal{O}(\varepsilon^{-\sigma})$  time steps, bringing the overall cost to  $\mathcal{O}((4n + 2) \times 3^n \varepsilon^{2\sigma-2} \log \varepsilon^{-1})$  at the very least (remember that the number of Lanczos iterations also grow with  $n$ , something we haven't accounted for here). For  $\sigma = \frac{1}{4}$ , for instance, this is  $\mathcal{O}((4n + 2) \times 3^n \varepsilon^{-3/2} \log \varepsilon^{-1})$ .

By comparison, for  $\sigma > \frac{1}{3}$ , we need  $m = \left\lceil \frac{(2n+3)\sigma-1}{3\sigma-1} \right\rceil$  iterations. For  $\sigma = \frac{1}{2}$ , for instance,  $m = 2n + 1$  iterations are required. The cost of exponentiating the central exponent via Lanczos iterations in this case is  $\mathcal{O}((2n + 1)(4n + 2) \times 3^n \varepsilon^{-1} \log \varepsilon^{-1})$ . Additionally, we have the  $\mathcal{O}(2 \times 3^n \varepsilon^{-1} \log \varepsilon^{-1})$  cost for exponentiating the  $\partial_x^2$  term. The overall cost for propagating to  $T$  is  $\mathcal{O}((8n^2 + 8n + 4) \times 3^n \varepsilon^{-3/2} \log \varepsilon^{-1})$ , which is comparable to the cost for  $\sigma = \frac{1}{4}$ . The accuracy of the method for  $\sigma = \frac{1}{2}$  is much higher, on the other hand. Just like we saw in Section 6.8 for Zassenhaus splittings of Chapter 6, there seems little point in going for  $\sigma \leq \frac{1}{3}$ .

Yet another approach is to separate terms by size to arrive at a Strang splitting

$$\exp\left(\tilde{\Theta}_{2n+1}^\varepsilon(h)\right) = e^{\frac{1}{2}W^{[0]}} \dots e^{\frac{1}{2}W^{[n]}} e^{W^{[n+1]}} e^{\frac{1}{2}W^{[n]}} \dots e^{\frac{1}{2}W^{[0]}} + \mathcal{O}(\varepsilon^{3\sigma-1}),$$

such that  $W^{[0]}$  is the term featuring  $\partial_x^2$ , while  $W^{[k]} = \mathcal{O}(\varepsilon^{(2k-1)\sigma-1})$ . Separating terms by

their spectral size would result in fewer Lanczos iterations for subsequent terms. However, this looks very much like a Zassenhaus splitting and is not much cheaper than one. Moreover, this splitting only features an  $\mathcal{O}(\varepsilon^{3\sigma-1})$  accuracy—the correct Yoshida splitting will be significantly more expensive.

We have seen that the cost of Magnus–Lanczos schemes scales as  $\mathcal{O}(\varepsilon^{-2} \log \varepsilon^{-1})$ . This can be improved upon by performing a Yoshida splitting before exponentiating the central exponent. Such schemes can achieve a  $\mathcal{O}(\varepsilon^{-3/2} \log \varepsilon^{-1})$  cost, which is asymptotically superior. In Chapter 8, we will develop the Magnus–Zassenhaus schemes whose asymptotic costs are similar, but which require fewer FFTs as we go towards higher order methods. This observation will be similar to that made in Section 6.7. However, unlike the  $2 \times 3^n$  FFTs required per step of the Yoshida splitting for time-independent Hamiltonians, we have found that  $(8n^2 + 8n + 4) \times 3^n$  FFTs are required when we combine Magnus expansions with Yoshida splittings. On the other hand, the increase in cost of Zassenhaus splittings for Magnus expansions is seen to be milder. This will result in the Magnus–Zassenhaus schemes being superior not only for very high orders of accuracy, but also for low and moderately high orders.



## Chapter 8

# Magnus–Zassenhaus schemes

As discussed in Section 7.4, the large size and inconvenient structure of commutators occurring in the matrix based Magnus expansions introduced in Section 3.4.1 makes the exponentiation of high-order Magnus expansions extremely expensive. However, even after simplification of the Magnus expansion in  $\mathfrak{H}$ , following Chapter 7, the Magnus–Lanczos schemes are not the most optimal approach for exponentiation.

Asymptotically speaking, the  $\mathcal{O}(\varepsilon^{-3/2} \log \varepsilon^{-1})$  cost of Yoshida splittings of the Magnus expansion (where the central exponent in the Strang splitting is exponentiated via Lanczos iterations) is more favourable. However, these Yoshida methods require  $\tilde{Y}(n) = (8n^2 + 8n + 4) \times 3^n$  FFTs which is significantly higher than the  $Y(n) = 2 \times 3^n$  FFTs required in the case of time-independent Hamiltonians.

Simplifying the Magnus expansion in the Lie algebra  $\mathfrak{H}$  without resorting to spatial discretisation gives us another advantage—it is now possible to resort to a commutator-free symmetric Zassenhaus splitting for  $\exp(\tilde{\Theta}_p^\varepsilon(h))$  along the lines of Chapter 6. The Zassenhaus algorithm provides a neat separation of terms with differing structures and sizes, each of which is easy to exponentiate separately.

In Sections 8.1 and 8.2, two versions of the Magnus expansion— $\tilde{\Theta}_5^{\varepsilon[M]}$  (7.17) and  $\tilde{\Theta}_3^{\varepsilon[I]}$  (7.33), respectively—are successfully combined with Zassenhaus splittings to yield efficient methods— $\mathcal{Z}_{2,\sigma}^{\Theta[M]}$  and  $\mathcal{Z}_{1,\sigma}^{\Theta[I]}$ , respectively—for solving the Schrödinger equation with time-dependent potentials. We find that, in principle, an effective splitting scheme for the Magnus expansion can be developed for time steps of size  $h = \mathcal{O}(\varepsilon^\sigma)$  for any  $\sigma > \frac{1}{3}$ , which is the same restriction we encountered in the case of Zassenhaus splittings for Schrödinger equations with time-independent potentials.

Much like the postponement of spatial discretisation considerations till the last stage, Zassenhaus splittings that commence from integral-preserving Magnus expansions, such as  $\tilde{\Theta}_3^{\varepsilon[I]}$ , eventually require the evaluation of integrals. Some of the options available to us for evaluation of these integrals are considered briefly in Section 8.2.1, following which some numerical results are presented in Section 8.3.

A technical difficulty encountered while attempting to preserve integrals, when starting from an expansion like  $\tilde{\Theta}_p^{\varepsilon[I]}$ , is the occurrence of  $\mathcal{O}(\varepsilon^{4\sigma-1})$ ,  $\mathcal{O}(\varepsilon^{6\sigma-1})$  or, in general,  $\mathcal{O}(\varepsilon^{2k\sigma-1})$  terms that do not occur when splitting the integral-free Magnus expansion  $\tilde{\Theta}_p^{\varepsilon[M]}$ . In Section 8.4 we exploit time symmetry of the Magnus expansion to derive splittings where the exponents  $W^{[k]}$  are  $\mathcal{O}(\varepsilon^{(2k-1)\sigma-1})$  for  $k \geq 2$ . In these new splittings terms of size  $\mathcal{O}(\varepsilon^{2k\sigma-1})$  do not feature, reducing the number of exponents in the splitting (and the cost) to half, thereby making high-order integral-preserving splittings simpler and more feasible.

In Section 8.5 we note that the cost of the Magnus–Zassenhaus schemes grows quadratically in the order desired and is only marginally higher than the cost of Zassenhaus splittings for the time-independent case. Unlike the case of time-independent Hamiltonians, the cost of these Zassenhaus splittings is superior to the corresponding Yoshida splittings even for low and moderately high-order methods as a consequence of the latter approach becoming significantly more expensive.

## 8.1 Approach 1: Discretised integrals

For the sake of simplicity, we demonstrate an example of a Zassenhaus splitting performed on (7.17), the truncated integral and commutator-free Magnus expansion  $\tilde{\Theta}_5^{\varepsilon[M]}(h)$ . The procedure clearly generalises to any  $\tilde{\Theta}_p^{\varepsilon[M]}(h)$ .

In line with our previous examples in Chapter 6, the splitting presented here involves extracting the largest terms (analysed in powers of  $\varepsilon$ ) first. Variants of Zassenhaus splittings based on different choices are possible but not explored here. We commence the Zassenhaus splitting algorithm (summarised in Table 6.1) with  $W^{[0]} = i h \varepsilon \partial_x^2$ . As we had seen in Section 6.2, we could equally well have chosen  $W^{[0]} = -i h \varepsilon^{-1} V_0$ , for instance, to arrive at a variant of the splitting presented here.

The exponent to be split here is the Magnus expansion (7.17),  $\mathcal{W}^{[0]} = \tilde{\Theta}_5^{\varepsilon[M]}(h)$ ,

$$\begin{aligned} \tilde{\Theta}_5^{\varepsilon[M]}(h) = & \overbrace{i h \varepsilon \partial_x^2 - i h \varepsilon^{-1} V_0}^{\mathcal{O}(\varepsilon^{\sigma-1})} - \overbrace{\frac{1}{12} i h^3 \varepsilon^{-1} V_2 - \frac{1}{6} h^3 \langle \partial_x V_1 \rangle_1}_{\mathcal{O}(\varepsilon^{3\sigma-1})} \\ & + \overbrace{\frac{1}{360} h^5 \left[ i \varepsilon^{-1} \left( 2(\partial_x V_2)(\partial_x V_0) - 3(\partial_x V_1)^2 \right) - 2 \langle (\partial_x V_1)(\partial_x^2 V_0) + 3(\partial_x V_0)(\partial_x^2 V_1) \rangle_1 \right]}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\ & + \overbrace{\frac{1}{90} i h^5 \varepsilon \langle \partial_x^2 V_2 \rangle_2 - \frac{1}{90} h^5 \varepsilon^2 \langle \partial_x^3 V_1 \rangle_3}_{\mathcal{O}(\varepsilon^{5\sigma-1})} = \Theta + \mathcal{O}(\varepsilon^{7\sigma-1}). \end{aligned}$$

The first steps of Zassenhaus splitting procedure involve computations of commutators in the sBCH formula. Here, once again, we will need to use the identities (6.3) of the Lie



algebra  $\mathfrak{H}$  for arriving at a commutator-free expression,

$$\begin{aligned}
 [\langle f \rangle_4, \langle g \rangle_0] &= 4 \langle f(\partial_x g) \rangle_3 - 2 \langle 3(\partial_x f)(\partial_x^2 g) + f(\partial_x^3 g) \rangle_1, \\
 [\langle f \rangle_3, \langle g \rangle_0] &= 3 \langle f(\partial_x g) \rangle_2 - \frac{1}{2} \langle 3(\partial_x f)(\partial_x^2 g) + f(\partial_x^3 g) \rangle_0, \\
 [\langle f \rangle_2, \langle g \rangle_2] &= 2 \langle f(\partial_x g) - (\partial_x f)g \rangle_3 + \langle 2(\partial_x^2 f)(\partial_x g) - 2(\partial_x f)(\partial_x^2 g) + (\partial_x^3 f)g - f(\partial_x^3 g) \rangle_1, \\
 [\langle f \rangle_2, \langle g \rangle_1] &= \langle 2f(\partial_x g) - (\partial_x f)g \rangle_2 - \frac{1}{2} \langle 2(\partial_x f)(\partial_x^2 g) + f(\partial_x^3 g) \rangle_0, \\
 [\langle f \rangle_2, \langle g \rangle_0] &= 2 \langle f(\partial_x g) \rangle_1, \\
 [\langle f \rangle_1, \langle g \rangle_1] &= \langle f(\partial_x g) - (\partial_x f)g \rangle_1, \\
 [\langle f \rangle_1, \langle g \rangle_0] &= \langle f(\partial_x g) \rangle_0.
 \end{aligned}$$

Using these simplification rules, we compute the sBCH for the first stage of the Zassenhaus algorithm with  $W^{[0]} = i\hbar\varepsilon\partial_x^2$  and  $\mathcal{W}^{[0]} = \tilde{\Theta}_5^{\varepsilon[M]}(h)$ ,

$$\begin{aligned}
 \mathcal{W}^{[1]} &= \text{sBCH}(-W^{[0]}, \mathcal{W}^{[0]}) \\
 &= -\overbrace{i\hbar\varepsilon^{-1}V_0}^{\mathcal{O}(\varepsilon^{\sigma-1})} + \overbrace{\frac{1}{12}i\hbar^3\varepsilon^{-1}\left(2(\partial_x V_0)^2 - V_2\right) - \frac{1}{6}h^3\langle\partial_x V_1\rangle_1 + \frac{1}{6}i\hbar^3\varepsilon\langle\partial_x^2 V_0\rangle_2}^{\mathcal{O}(\varepsilon^{3\sigma-1})} \\
 &\quad - \overbrace{\frac{1}{24}i\hbar^3\varepsilon\langle\partial_x^4 V_0\rangle}^{\mathcal{O}(\varepsilon^{3\sigma+1})} - \overbrace{\frac{1}{360}i\hbar^5\varepsilon^{-1}\left(8(\partial_x V_0)^2(\partial_x^2 V_0) + 3(\partial_x V_1)^2 - 12(\partial_x V_2)(\partial_x V_0)\right)}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\
 &\quad + \overbrace{\frac{1}{30}h^5\langle 2(\partial_x V_0)(\partial_x^2 V_1) - (\partial_x V_1)(\partial_x^2 V_0) \rangle_1}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\
 &\quad - \overbrace{\frac{1}{720}i\hbar^5\varepsilon\langle 127(\partial_x V_0)(\partial_x^3 V_0) + 130(\partial_x^2 V_0)^2 - 18(\partial_x^2 V_2) \rangle_2}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\
 &\quad + \overbrace{\frac{1}{60}h^5\varepsilon^2\langle\partial_x^3 V_1\rangle_3 - \frac{13}{90}i\hbar^5\varepsilon^3\langle\partial_x^4 V_0\rangle_4}^{\mathcal{O}(\varepsilon^{5\sigma-1})} + \mathcal{O}(\varepsilon^{7\sigma-1}).
 \end{aligned}$$

At the second stage we select the largest remaining element  $W^{[1]} = -i\hbar\varepsilon^{-1}V_0$ , whereby

$$\begin{aligned}
\mathcal{W}^{[2]} &= \text{sBCH}(-W^{[1]}, \mathcal{W}^{[1]}) \\
&= \overbrace{\frac{1}{12}i\hbar^3\varepsilon^{-1}\left(2(\partial_x V_0)^2 - V_2\right) - \frac{1}{6}h^3\langle\partial_x V_1\rangle_1 + \frac{1}{6}i\hbar^3\varepsilon\langle\partial_x^2 V_0\rangle_2}^{\mathcal{O}(\varepsilon^{3\sigma-1})} \\
&\quad - \overbrace{\frac{1}{24}i\hbar^3\varepsilon(\partial_x^4 V_0)}^{\mathcal{O}(\varepsilon^{3\sigma+1})} - \overbrace{\frac{1}{360}i\hbar^5\varepsilon^{-1}\left(13(\partial_x V_0)^2(\partial_x^2 V_0) + 3(\partial_x V_1)^2 - 12(\partial_x V_2)(\partial_x V_0)\right)}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\
&\quad + \overbrace{\frac{1}{30}h^5\langle 2(\partial_x V_0)(\partial_x^2 V_1) - (\partial_x V_1)(\partial_x^2 V_0)\rangle_1}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\
&\quad - \overbrace{\frac{1}{720}i\hbar^5\varepsilon\langle 127(\partial_x V_0)(\partial_x^3 V_0) + 130(\partial_x^2 V_0)^2 - 18(\partial_x^2 V_2)\rangle_2}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\
&\quad + \overbrace{\frac{1}{60}h^5\varepsilon^2\langle\partial_x^3 V_1\rangle_3 - \frac{13}{90}i\hbar^5\varepsilon^3\langle\partial_x^4 V_0\rangle_4}^{\mathcal{O}(\varepsilon^{5\sigma-1})} + \mathcal{O}(\varepsilon^{7\sigma-1}).
\end{aligned}$$

We terminate the procedure by letting  $W^{[2]}$  consist of the  $\mathcal{O}(\varepsilon^{3\sigma-1})$  terms in  $\mathcal{W}^{[2]}$  and are left with  $\mathcal{O}(\varepsilon^{3\sigma+1})$  and  $\mathcal{O}(\varepsilon^{5\sigma-1})$  terms in  $\mathcal{W}^{[3]} = \mathcal{W}^{[2]} - W^{[2]}$ , once we ignore  $\mathcal{O}(\varepsilon^{7\sigma-1})$  terms. Since  $\varepsilon^{3\sigma+1} = \mathcal{O}(\varepsilon^{5\sigma-1})$  under  $\sigma \leq 1$ , it is not problematic to combine these terms. The outcome is the splitting,

$$\mathcal{Z}_{2,\sigma}^{\Theta[M]} = e^{\frac{1}{2}W^{[0]}} e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[2]}} e^{\mathcal{W}^{[3]}} e^{\frac{1}{2}W^{[2]}} e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[0]}} = \exp\left(\tilde{\Theta}_5^{\varepsilon[M]}(h)\right) + \mathcal{O}(\varepsilon^{7\sigma-1}), \quad (8.1)$$

where

$$\begin{aligned}
W^{[0]} &= i\varepsilon h \partial_x^2 = \mathcal{O}(\varepsilon^{\sigma-1}), \\
W^{[1]} &= -i\varepsilon^{-1}hV_0 = \mathcal{O}(\varepsilon^{\sigma-1}), \\
W^{[2]} &= \frac{1}{12}i\varepsilon^{-1}h^3\left(2(\partial_x V_0)^2 - V_2\right) - \frac{1}{6}h^3\langle\partial_x V_1\rangle_1 + \frac{1}{6}i\varepsilon h^3\langle\partial_x^2 V_0\rangle_2 = \mathcal{O}(\varepsilon^{3\sigma-1}), \\
\mathcal{W}^{[3]} &= -\frac{1}{24}i\varepsilon h^3(\partial_x^4 V_0) - \frac{1}{360}i\varepsilon^{-1}h^5\left(13(\partial_x V_0)^2(\partial_x^2 V_0) + 3(\partial_x V_1)^2 - 12(\partial_x V_2)(\partial_x V_0)\right) \\
&\quad + \frac{1}{30}h^5\langle 2(\partial_x V_0)(\partial_x^2 V_1) - (\partial_x V_1)(\partial_x^2 V_0)\rangle_1 \\
&\quad - \frac{1}{720}i\varepsilon h^5\langle 127(\partial_x V_0)(\partial_x^3 V_0) + 130(\partial_x^2 V_0)^2 - 18(\partial_x^2 V_2)\rangle_2 \\
&\quad + \frac{1}{60}\varepsilon^2 h^5\langle\partial_x^3 V_1\rangle_3 - \frac{13}{90}i\varepsilon^3 h^5\langle\partial_x^4 V_0\rangle_4 = \mathcal{O}(\varepsilon^{5\sigma-1}).
\end{aligned}$$

This splitting looks very similar to the splitting  $\mathcal{Z}_{2,\sigma}^{[2]}$  (6.10), derived in Chapter 6 for the time-independent Hamiltonian and the notation  $\mathcal{Z}_{2,\sigma}^{\Theta[M]}$  is very similar too, with  $\Theta[M]$  indicating that we have started from an appropriately high-order integral-free Magnus expansion of the type  $\tilde{\Theta}_p^{\varepsilon[M]}$ . Structurally speaking,  $\mathcal{Z}_{2,\sigma}^{[2]}$  and  $\mathcal{Z}_{2,\sigma}^{\Theta[M]}$  differ in the appearance of odd-indexed terms, such as  $-\frac{1}{6}h^3\langle\partial_x V_1\rangle_1$  which appears in  $W^{[2]}$  here. Apart from this,

the splittings are very similar and the computation of these exponentials is carried out exactly on the same lines as the approach pursued for  $\mathcal{Z}_{2,\sigma}^{[2]}$  in Section 6.4.

The outer exponents are structurally identical to those in  $\mathcal{Z}_{2,\sigma}^{[2]}$  and they are exponentiated in precisely the same manner— $W^{[0]}$  is discretised to  $\tilde{W}^{[0]} = i\varepsilon h \mathcal{K}^2$ , which is exponentiated in  $\mathcal{O}(M \log M)$  operations (where  $M = 2N + 1$  is the number of grid points in the spatial discretisation), while  $W^{[1]} \rightsquigarrow \tilde{W}^{[1]} = -i\varepsilon^{-1} h \mathcal{D}_V$ , which is exponentiated directly in  $\mathcal{O}(M)$  cost.

The remaining exponents,  $W^{[2]}$  and  $\mathcal{W}^{[3]}$ , do not possess a structure amenable to exact exponentiation. However, they are very small— $\mathcal{O}(\varepsilon^{3\sigma-1})$  and  $\mathcal{O}(\varepsilon^{5\sigma-1})$ , respectively—precisely like the corresponding exponents in splitting  $\mathcal{Z}_{2,\sigma}^{[2]}$ . Following the analysis of Lanczos iterations in Section 6.4, the exponentials of these terms can be evaluated to  $\mathcal{O}(\varepsilon^{7\sigma-1})$  accuracy using merely three and two Lanczos iterations, respectively, under the scaling  $\sigma = 1$ .

The number of iterations required do change with the choice of  $\sigma$ . As we saw in Section 6.8, for  $\sigma = \frac{1}{2}$  the term  $W^{[2]}$  is of size  $\mathcal{O}(\varepsilon^{1/2})$  and we would require  $m = 5$  for achieving the  $\mathcal{O}(\varepsilon^{5/2})$  accuracy of the corresponding splitting. A critical stage is reached at  $\sigma = \frac{1}{3}$ , where the term  $W^{[2]}$  becomes  $\mathcal{O}(\varepsilon^0)$  and no longer decreases in size with  $\varepsilon$ . Beyond this, the exponent grows in size with decreasing  $\varepsilon$  and, following the analysis of Section 6.8, smaller  $\sigma$  are no longer cost effective. Thus we are forced to place a limit on the time step scaling and restrict ourselves to  $\frac{1}{3} < \sigma \leq 1$ . This restriction arises out of the Zassenhaus splitting (not the Magnus expansion).

All exponents for which we resort to Lanczos iterations feature Jordan polynomials in  $\partial_x$  of the form  $\langle f \rangle_k = (f \circ \partial_x^k + \partial_x^k \circ f)/2$ . As we saw earlier, these are discretised in a straightforward way,

$$\langle f \rangle_k \rightsquigarrow \frac{1}{2}(\mathcal{D}_f \mathcal{K}^k + \mathcal{K}^k \mathcal{D}_f) = \frac{1}{2} \left( \mathcal{D}_f \mathcal{F}^{-1} \mathcal{D}_{(in\pi)^k} \mathcal{F} + \mathcal{F}^{-1} \mathcal{D}_{(in\pi)^k} \mathcal{F} \mathcal{D}_f \right),$$

and  $\langle f \rangle_k u \rightsquigarrow \frac{1}{2} \left( \mathcal{D}_f \mathcal{F}^{-1} \mathcal{D}_{(in\pi)^k} \mathcal{F} u + \mathcal{F}^{-1} \mathcal{D}_{(in\pi)^k} \mathcal{F} \mathcal{D}_f u \right)$  is computed in  $\mathcal{O}(M \log M)$  operations using four FFTs (more efficient ways of combining FFTs to reduce costs are discussed in Section 6.4.1). Altogether,  $\tilde{W}^{[2]} \mathbf{v}$  can be computed using  $\mathcal{O}(M \log M)$  operations, resulting in an overall cost of  $\mathcal{O}(mM \log M)$  operations for the approximation of  $e^{\tilde{W}^{[2]}} \mathbf{u}$ , where  $m$  is the number of Lanczos iterations required.

The exponents  $W^{[2]}$  and  $\mathcal{W}^{[3]}$  feature spatial derivatives of the potential ( $\partial_x V_0$ ,  $\partial_x V_1$ , and  $\partial_x^2 V_0$ , for instance), which might be available in closed form. However, in general these can be evaluated through numerical differentiation.  $\partial_x^2 V_0$ , for instance, can be approximated by the vector  $\mathcal{K}^2 \mathbf{V}_0$ . We note that the degree of accuracy required in the differentiation here is considerably lower than the spectral accuracy of spectral collocation differentiation matrices. We may, instead, use centred finite-difference differentiation matrices of varying accuracy (depending on the accuracy requirement per term) to achieve

the same feat at an  $\mathcal{O}(M)$  cost. For instance, the term

$$-\frac{1}{6}h^3 \langle \partial_x V_1 \rangle_1$$

appearing in  $W^{[2]}$  needs to be approximated to  $\mathcal{O}(\varepsilon^{7\sigma-1})$  accuracy (the accuracy of  $\mathcal{Z}_{2,\sigma}^{\Theta[M]}$ ). Considering the  $h^3 = \mathcal{O}(\varepsilon^{3\sigma})$  factor and the fact that  $\langle \cdot \rangle_1 = \mathcal{O}(\varepsilon^{-1})$ , we need  $\partial_x V_1$  to be accurate up to  $\mathcal{O}(\varepsilon^{4\sigma})$ . In other words, an order four differentiation method would suffice for this instance of  $\partial_x V_1$  under the scaling  $\sigma = 1$  and an order two method would do for  $\sigma = 1/2$ .

We note that all exponents in our splitting  $\mathcal{Z}_{2,\sigma}^{\Theta[M]}$  are of the form  $i^{k+1} \langle f \rangle_k$ . This is a more general scenario than the Zassenhaus splittings of Chapter 6 where all terms are  $i \langle f \rangle_{2k}$ . Nevertheless, the terms  $i^{k+1} \langle f \rangle_k$  are in  $\mathfrak{H}$  and are discretised as skew-Hermitian matrices once we resort to spectral collocation,

$$i^{k+1} \langle f \rangle_k \rightsquigarrow \frac{1}{2} i^{k+1} (\mathcal{D}_f \mathcal{K}^k + \mathcal{K}^k \mathcal{D}_f).$$

Since the exponential of a skew-Hermitian matrix is unitary, unitary evolution and (consequently) unconditional stability of the method are guaranteed, once again (see Section 6.9).

For a more detailed discussion of the exponentiation methods and stability analysis, we refer the reader to Sections 6.4, 6.5, 6.8 and 6.9 of Chapter 6.

## 8.2 Approach 2: Undiscretised integrals

Integral-preserving Magnus–Zassenhaus schemes can be derived along similar lines as Section 8.1. Once the integrals have been simplified in the Magnus expansion  $\tilde{\Theta}_3^{\varepsilon[I]}(h)$ , for instance, the Zassenhaus splitting procedure can work while being agnostic of the presence of integrals, treating terms such as  $\mu_{0,0}(h)$  and  $\mu_{1,1}(h)$  as any other potential function.

As usual, we discard terms smaller than our error tolerance—for  $\tilde{\Theta}_{2n+1}^{\varepsilon[I]}(h)$  this means discarding terms of size  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$ . The Magnus expansion that we wish to split is,

$$\tilde{\Theta}_3^{\varepsilon[I]}(h) = \overbrace{i h \varepsilon \partial_x^2 - i \varepsilon^{-1} \mu_{0,0}(h)}^{\mathcal{O}(\varepsilon^{\sigma-1})} - \overbrace{2 \langle \partial_x \mu_{1,1}(h) \rangle_1}^{\mathcal{O}(\varepsilon^{3\sigma-1})}. \quad (7.33)$$

We get a five-stage splitting by starting with  $W^{[0]} = i h \varepsilon \partial_x^2$  and then extracting the largest term at each iteration. For a splitting of  $\tilde{\Theta}_3^{\varepsilon[I]}(h)$ , we only need to compute the sBCH at the first stage of the Zassenhaus splitting,

$$\begin{aligned} \mathcal{W}^{[1]} &= \text{sBCH}(-W^{[0]}, \mathcal{W}^{[0]}) \\ &= -\overbrace{i \varepsilon^{-1} \mu_{0,0}(h)}^{\mathcal{O}(\varepsilon^{\sigma-1})} + \overbrace{\frac{1}{6} i h \varepsilon^{-1} (\partial_x \mu_{0,0}(h))^2 - 2 \langle \partial_x \mu_{1,1}(h) \rangle_1 + \frac{1}{6} i h^2 \varepsilon \langle \partial_x^2 \mu_{0,0}(h) \rangle_2}^{\mathcal{O}(\varepsilon^{3\sigma-1})}. \end{aligned}$$

At the next stage we let  $W^{[1]} = -i\varepsilon^{-1}\mu_{0,0}(h)$  and terminate with  $\mathcal{W}^{[2]} = \mathcal{W}^{[1]} - W^{[1]}$  (the leading term in the sBCH of  $\mathcal{W}^{[1]}$  and  $-W^{[1]}$ ), arriving at the splitting

$$\mathcal{Z}_{1,\sigma}^{\Theta[I]} = e^{\frac{1}{2}W^{[0]}} e^{\frac{1}{2}W^{[1]}} e^{\mathcal{W}^{[2]}} e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[0]}} = \exp\left(\tilde{\Theta}_3^{\varepsilon[I]}(h)\right) + \mathcal{O}\left(\varepsilon^{5\sigma-1}\right), \quad (8.2)$$

where

$$\begin{aligned} W^{[0]} &= ih\varepsilon\partial_x^2 = \mathcal{O}\left(\varepsilon^{\sigma-1}\right), \\ W^{[1]} &= -i\varepsilon^{-1}\mu_{0,0}(h) = \mathcal{O}\left(\varepsilon^{\sigma-1}\right), \\ \mathcal{W}^{[2]} &= \frac{1}{6}ih\varepsilon^{-1}\left(\partial_x\mu_{0,0}(h)\right)^2 - 2\langle\partial_x\mu_{1,1}(h)\rangle_1 + \frac{1}{6}ih^2\varepsilon\langle\partial_x^2\mu_{0,0}(h)\rangle_2 = \mathcal{O}\left(\varepsilon^{3\sigma-1}\right). \end{aligned}$$

As a quick sanity check, for time-independent potentials,  $V(x, t) = V(x)$ ,  $\mathcal{Z}_{1,\sigma}^{\Theta[I]}$  reduces to the standard symmetric Zassenhaus splitting  $\mathcal{Z}_{1,\sigma}^{[2]}$  of Chapter 6, which is obtained from (6.10) by ignoring the  $\mathcal{W}^{[3]}$  exponent.

It is only by this stage—having arrived at an asymptotic splitting expressed in operatorial terms—that we start considering discretisation issues. Considerations of spatial and temporal discretisation (in the form of approximation through quadrature) are entirely independent here and one may proceed to address them in any order. As usual, we resort to spectral collocation for spatial discretisation. The considerations of stability and the approximation of exponentials follow along the same line as discussed in Section 8.1 (and previously in Chapter 6). Having preserved integrals, we now consider how to evaluate these for an effective numerical scheme.

### 8.2.1 Evaluation of integrals

In this section we consider the problem of evaluating the integrals appearing in  $\mathcal{Z}_{1,\sigma}^{\Theta[I]}$ :  $\mu_{0,0}(h)$  and  $\mu_{1,1}(h)$ . The integral

$$\mu_{0,0}(h) = \int_0^h V(\zeta) d\zeta$$

appears once in  $W^{[1]}$  and twice in  $\mathcal{W}^{[2]}$ . Since we have already committed an  $\mathcal{O}\left(\varepsilon^{5\sigma-1}\right)$  error in the splitting, we need not approximate the integrals to a greater degree of accuracy. In the term  $\frac{1}{6}ih^2\varepsilon\langle\partial_x^2\mu_{0,0}(h)\rangle_2$ , we need an  $\mathcal{O}\left(\varepsilon^{3\sigma}\right)$  or  $\mathcal{O}\left(h^3\right)$  approximation of  $\mu_{0,0}(h)$  due to the presence of the  $h^2\varepsilon$  scalar factor and since  $\langle f \rangle_2$  is  $\mathcal{O}\left(\varepsilon^{-2}\right)$  for  $f = \mathcal{O}\left(\varepsilon^0\right)$ . However,  $\mathcal{O}\left(h^4\right)$  and  $\mathcal{O}\left(h^5\right)$  approximations are needed for its other occurrences in  $\mathcal{W}^{[2]}$  and  $W^{[1]}$ , respectively. The only occurrence of  $\mu_{1,1}(h)$ ,

$$\mu_{1,1}(h) = \int_0^h \left(\zeta - \frac{h}{2}\right) V(\zeta) d\zeta,$$

demands an  $\mathcal{O}(h^5)$  approximation. Thus, it suffices to evaluate  $\mathcal{O}(h^5)$  accuracy approximations for the integrals  $\mu_{0,0}(h)$  and  $\mu_{1,1}(h)$  once per time step.

In the standard setting, where  $V(h)$  is sufficiently smooth in the temporal domain, we may effectively approximate the integrals using two Gauss–Legendre knots  $t_k := \frac{h}{2}(1 + k/\sqrt{3})$ ,  $k = -1, 1$ , and weights  $w_k = \frac{h}{2}$  (Davis & Rabinowitz 1984),

$$\mu_{0,0}(h) = \left( \frac{V(t_1) + V(t_{-1})}{2} \right) h + \mathcal{O}(h^5),$$

and

$$\mu_{1,1}(h) = \left( \frac{V(t_1) - V(t_{-1})}{4\sqrt{3}} \right) h^2 + \mathcal{O}(h^5).$$

These turn out to be centred finite difference approximations of  $\partial_t^j V(t)$  at  $t = \frac{h}{2}$  for  $j = 0, 1$  on the grid  $\frac{h}{2}(1 + k/\sqrt{3})$ ,  $k = -1, 0, 1$ .

**Note:** In Section 2.3.5 we noted that, in the context of power-truncated Magnus expansions, a quadrature method which usually has an accuracy of  $\mathcal{O}(h^{2n})$  gains an order of accuracy, becoming  $\mathcal{O}(h^{2n+1})$ . This was due to time-symmetry and odd nature of such Magnus expansions, where all even terms in the Taylor series, including error terms, vanish. Thus two Gauss–Legendre knots result in a  $\mathcal{O}(h^5)$  method in our case.

While the approach presented here requires us to carry all integrals along in an undiscretised form, and is thus considerably more involved, it is also more flexible. Little has been assumed about the regularity of  $V$  in the temporal domain so far. It is generally accepted that integrals are to be preferred over derivatives—they are particularly helpful in the case of functions with low regularity and in the case of highly oscillatory functions. Here our decision of not replacing the integrals with Taylor expansion at the outset affords us the ability to effectively handle such cases.

It might be possible to evaluate the exact integral for  $\mu_{0,0}(h)$  and  $\mu_{1,1}(h)$ , for instance, if an analytic expression for  $V$  is available. This is also true for potentials of the form  $V(x, t) = f(t)V(x)$  where an analytic expression for  $f(t)$  is available. If  $f(t)$  is a highly oscillatory envelope, oscillating around the origin, there can be a further decrease in size of integrals involved which can benefit us greatly.

Where an exact integral is not available, integrals featuring highly oscillatory integrands of the form

$$\int_0^1 f(t) e^{i\omega g(t)} dt$$

can be approximated effectively using the Filon method (Iserles & Nørsett 2005).

Unless the the potential  $V$  depends on  $\varepsilon$ , the approximation of the integrals will require  $\mathcal{O}(M)$  operations. Thus, even though the evaluation of integrals needs to be performed at

each step of the Magnus–Zassenhaus splitting scheme, it does not significantly contribute to the costs, which are dominated by the  $\mathcal{O}(M \log M)$  costs of FFTs.

### 8.3 Numerical results

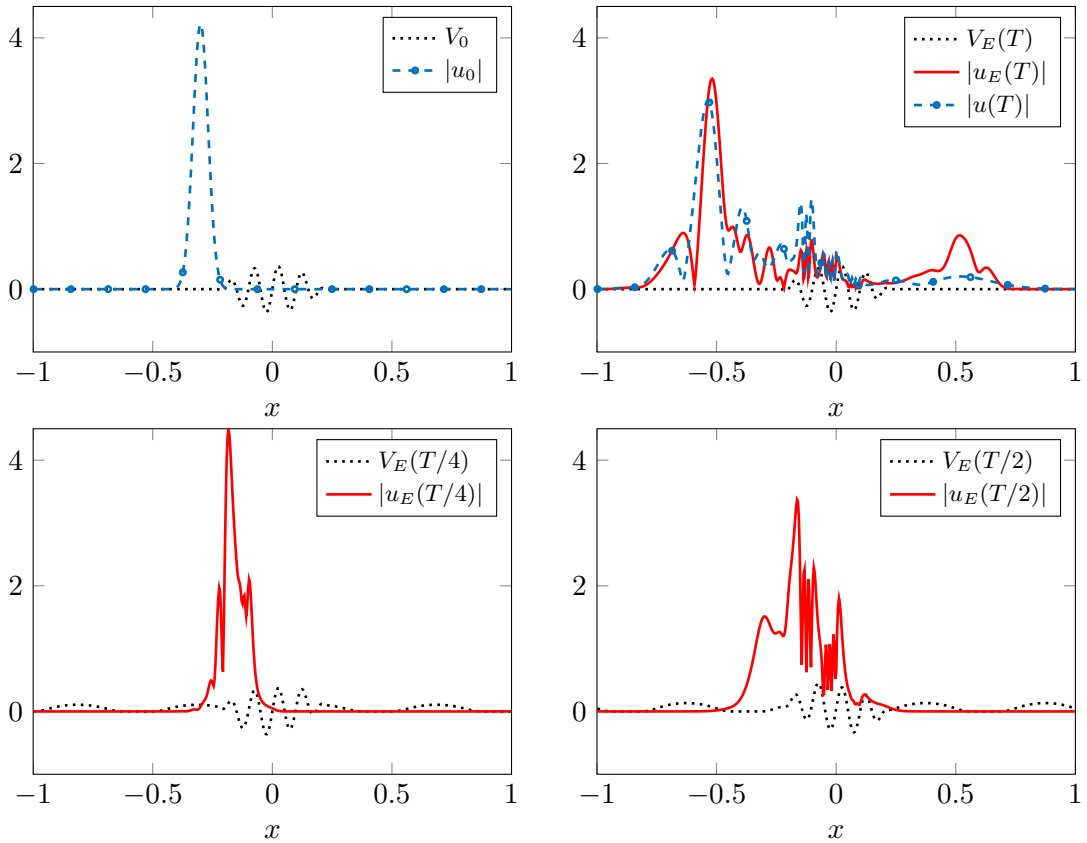
**Example 8.3.1.** Consider the evolution of the wave-packet

$$u_0(x) = (\delta\pi)^{-1/4} \exp\left(ik_0 \frac{(x-x_0)}{\delta} - \frac{(x-x_0)^2}{2\delta}\right), \quad x_0 = -0.3, \quad k_0 = 0.1, \quad \delta = 10^{-3}, \quad (8.3)$$

heading towards the lattice potential

$$V_0 = \hat{\rho}(4x) \sin(20\pi x), \quad (8.4)$$

where  $\hat{\rho}(x) = \text{expbump}(x)$  is the bump function (2.18).



**Figure 8.1:** Initial wave-packet  $u_0$  (top left); final wave-packets at time  $T = 0.75$  (top right):  $u(T)$  under the influence of  $V_0$  and  $u_E(T)$  under the influence of  $V_E(x, t) = V_0 + E(x, t)$  for Example 8.3.1. Intermediate stages of  $u_E$  are shown in the bottom row.

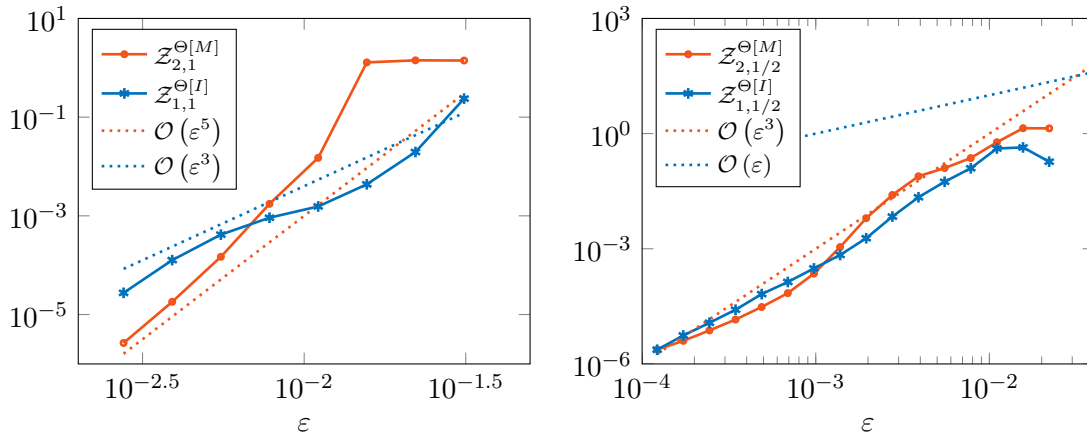
When the semiclassical parameter is  $\varepsilon = 2^{-8}$ , the wave-packet evolves to  $u(T)$  at  $T = 0.75$  in Figure 8.3.1 under the influence of the time-independent potential  $V_0$  alone.

When we excite it using an additional time-dependent potential,

$$E(x, t) = \hat{\rho}(3t - 1)\hat{\rho}(\sin(2\pi(x - t))),$$

so that the wave packet evolves under  $V_E(x, t) = V_0(x) + E(x, t)$ , a significantly larger part of the wave packet is able to make it across the lattice to the right hand side (see  $u_E(T)$  in Figure 8.1).

The excitation pulse is not active for the entire duration since  $\hat{\rho}(3t - 1)$  acts as a smooth envelope simulating the switching on and off of the time-dependent component of the potential. The excited potential is evident at  $t = T/2$  in Figure 8.1.



**Figure 8.2:** Global  $\ell^2$  errors for Example 8.3.1 at  $T = 0.75$  using schemes  $\mathcal{Z}_{2,\sigma}^{\Theta[M]}$  (8.1) and  $\mathcal{Z}_{1,\sigma}^{\Theta[I]}$  (8.2) under  $\sigma = 1$  (left) and  $\sigma = \frac{1}{2}$  (right).

In the splitting scheme  $\mathcal{Z}_{2,1}^{\Theta[M]}$  (8.1), we commit a local  $\ell^2$  error of  $\mathcal{O}(\varepsilon^6)$  per time step. Since the number of time steps is  $\mathcal{O}(\varepsilon^{-1})$ , the global  $\ell^2$  error is  $\mathcal{O}(\varepsilon^5)$ . The global error for  $\mathcal{Z}_{1,1}^{\Theta[I]}$  (8.2) is  $\mathcal{O}(\varepsilon^3)$ .

In Figure 8.2 we present the global  $\ell^2$  errors in the propagation of  $u_0$  to  $u_E(T = 0.75)$  under the influence of  $V_E$  using the schemes  $\mathcal{Z}_{2,\sigma}^{\Theta[M]}$  and  $\mathcal{Z}_{1,\sigma}^{\Theta[I]}$ . The precise scaling used in our experiments for  $\sigma = 1$  is

$$M \sim 5\varepsilon^{-1}, \quad h \sim \varepsilon/2.$$

We note that the expense of evaluating a reference solution by brute force is what limits us to moderately small values of  $\varepsilon$  (see Section 8.3.1). Moreover, for reasonable error decrease in practice, our spatial and temporal resolutions need to be sufficient to resolve the potential correctly, at the very least. Thus, at large values of  $\varepsilon$ , these schemes cannot be expected to be accurate under the scaling prescribed here. However, the scaling chosen here allows us to resolve the potential reasonably for moderate values of  $\varepsilon$ . For very small  $\varepsilon$ , less stringent scalings such as  $M \sim \varepsilon^{-1}$  do work, as we see in the  $\sigma = \frac{1}{2}$  case.



The global  $\ell^2$  error in  $\mathcal{Z}_{2,\frac{1}{2}}^{\Theta[M]}$  is  $\mathcal{O}(\varepsilon^{6\sigma-1}) = \mathcal{O}(\varepsilon^2)$  (since  $\sigma = \frac{1}{2}$ ) and that in  $\mathcal{Z}_{1,\frac{1}{2}}^{\Theta[I]}$  is  $\mathcal{O}(\varepsilon)$ . We resort to the high-order method  $\mathcal{Z}_{2,1}^{\Theta[I]}$  (8.16) (which we encounter in Section 8.4, and whose accuracy is established in Figure 8.3) for generating reference solutions for  $\mathcal{Z}_{2,\frac{1}{2}}^{\Theta[M]}$  and  $\mathcal{Z}_{1,\frac{1}{2}}^{\Theta[I]}$ . This allows us to generate high accuracy reference solutions for very small  $\varepsilon$ . At these scales, reasonable spatial discretisation of the potential is easily achieved with  $M \sim \varepsilon^{-1}$ . However, anticipating that the temporal resolution of the potential would suffer for moderately small  $\varepsilon$  (since we chose  $h = \mathcal{O}(\sqrt{\varepsilon})$ ), we let  $h = \sqrt{\varepsilon}/4$  so that the scaling used for  $\sigma = \frac{1}{2}$  is

$$M \sim \varepsilon^{-1}, \quad h \sim \sqrt{\varepsilon}/4.$$

This has the added advantage of making the exponents smaller, causing the errors to conform to the asymptotic estimates at an earlier stage. The splitting  $\mathcal{Z}_{1,\frac{1}{2}}^{\Theta[I]}$  performs surprisingly good, as we can see in Figure 8.2, reaching  $\mathcal{O}(\varepsilon^3)$  accuracy instead of  $\mathcal{O}(\varepsilon)$  and even beating the higher order scheme  $\mathcal{Z}_{2,\frac{1}{2}}^{\Theta[M]}$  at times (to be fair, to some extent even  $\mathcal{Z}_{2,\frac{1}{2}}^{\Theta[M]}$  outperforms the  $\mathcal{O}(\varepsilon^2)$  error expected out of it)!

### 8.3.1 Finding a reference solution

Since no analytic solution of (3.20), the semiclassical Schrödinger equation with time-dependent potentials, is available, reference solutions must also be obtained via a numerical approach. Unlike Section 6.10.1, we obtain the reference solutions for our numerical experiments by resorting to the exponential midpoint rule (3.25) where we freeze the potential in the middle of the interval followed by a Strang splitting,

$$u_R^1 = e^{\frac{1}{2}i\varepsilon h_R \mathcal{K}^2} e^{-i\varepsilon^{-1} h_R \mathcal{D}_{V(h_R/2)}} e^{\frac{1}{2}i\varepsilon h_R \mathcal{K}^2} u_R^0.$$

We make this choice because directly exponentiating a Hamiltonian (via MATLAB's `expm`) with potential frozen at the middle of the interval is more expensive but no more accurate than the above Strang splitting—freezing the potential already leads to an error which is of the same order as the error of the Strang splitting.

Since such a splitting is also the lowest order in the Magnus–Zassenhaus family of schemes (following our notation, this splitting is  $\mathcal{Z}_{0,\sigma}^{\Theta[I]}$  or, equivalently,  $\mathcal{Z}_{0,\sigma}^{\Theta[M]}$ ), we require very small time steps for convergence—certainly  $h_R \ll h$  is required for the reference solution to possess an error smaller than the splittings  $\mathcal{Z}_{1,1}^{\Theta[I]}$  and  $\mathcal{Z}_{2,1}^{\Theta[M]}$  whose error behaviour we are attempting to verify.

Another factor we must take into account is the growth of oscillations with decreasing  $\varepsilon$ . To capture this, we take a larger constant of proportionality while selecting  $M_R = \mathcal{O}(\varepsilon^{-1})$  degrees of freedom for spatial discretisation while computing the reference solution. The error in the reference solution is estimated by generating another reference solution with finer spatio-temporal resolution, to which the candidate reference solution is compared. A

reference solution is only accepted when this difference is much smaller than the error of the schemes to be compared.

Using such a low order method for generating reference solutions to a high degree of accuracy means generating reference solutions is orders of magnitude slower than the splittings  $\mathcal{Z}_{1,1}^{\Theta[I]}$  and  $\mathcal{Z}_{2,1}^{\Theta[M]}$  requiring validation. The exorbitant cost of reference solutions is what restricts experimental study of numerical errors to moderate values of  $\varepsilon$  for  $\sigma = 1$ . A similar method is utilised for establishing the accuracy of the scheme  $\mathcal{Z}_{2,1}^{\Theta[I]}$ , which is encountered in the following section.

Having established the accuracy of the Magnus–Zassenhaus schemes under  $\sigma = 1$  using this brute force method, we resort to  $\mathcal{Z}_{2,1}^{\Theta[I]}$  for generating reference solutions for the  $\sigma = \frac{1}{2}$  case. There are a couple of reasons for this—firstly, the midpoint rule is the lowest order Magnus–Zassenhaus method  $\mathcal{Z}_{0,\sigma}^{\Theta[I]}$  with a local accuracy of  $\mathcal{O}(\varepsilon^{3\sigma-1})$  and a global accuracy of  $\mathcal{O}(\varepsilon^{2\sigma-1})$ , which is  $\mathcal{O}(1)$  for  $\sigma = \frac{1}{2}$  and  $\mathcal{O}(\varepsilon)$  for  $\sigma = 1$ , forcing us to choose  $\sigma = 1$  for reference solutions; secondly, the high accuracy of  $\mathcal{Z}_{2,1}^{\Theta[I]}$  allows us to approach much smaller values of  $\varepsilon$  than would be practical via the midpoint rule, which has a very low accuracy even for  $\sigma = 1$ .

## 8.4 High order schemes with undiscretised integrals

Arbitrarily high-order Magnus–Zassenhaus splittings for the Schrödinger equation with time-dependent potential (3.21) while keeping the integrals intact can be obtained by starting from a correspondingly high-order integral-preserving Magnus expansion. For a Zassenhaus splitting featuring an error of  $\mathcal{O}(\varepsilon^{7\sigma-1})$ , for instance, we need to consider the Magnus expansion  $\tilde{\Theta}_5^{\varepsilon[I]}$  which is derived by simplifying trees in  $\Theta_5$  and discarding all  $\mathcal{O}(\varepsilon^{7\sigma-1})$  terms,

$$\begin{aligned}
 \tilde{\Theta}_5^{\varepsilon[I]}(h) = & \overbrace{\mathrm{i}h\varepsilon\partial_x^2 - \mathrm{i}\varepsilon^{-1}\mu_{0,0}(h)}^{\mathcal{O}(\varepsilon^{\sigma-1})} - \overbrace{2\langle\partial_x\mu_{1,1}(h)\rangle_1}^{\mathcal{O}(\varepsilon^{3\sigma-1})} + \overbrace{\mathrm{i}\varepsilon^{-1}\Lambda[\phi]_{1,1}(h) + 2\mathrm{i}\varepsilon\langle\partial_x^2\mu_{2,1}(h)\rangle_2}^{\mathcal{O}(\varepsilon^{4\sigma-1})} \\
 & + \overbrace{\frac{1}{6}\langle\Lambda[\psi_1]_{1,2}(h) + \Lambda[\psi_2]_{2,1}(h)\rangle_1}^{\mathcal{O}(\varepsilon^{4\sigma-1})} + \overbrace{\frac{1}{6}\langle\Lambda[\theta_1]_{1,2}(h) + \Lambda[\theta_2]_{2,1}(h)\rangle_1}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\
 & + \overbrace{\frac{4}{3}\varepsilon^2\langle\partial_x^3\mu_{3,1}(h)\rangle_3}^{\mathcal{O}(\varepsilon^{5\sigma-1})} + \overbrace{\frac{1}{4}\mathrm{i}\varepsilon\partial_x^4\mu_{2,1}(h)}^{\mathcal{O}(\varepsilon^{4\sigma+1})} = \Theta(h) + \mathcal{O}(\varepsilon^{7\sigma-1}), \tag{8.5}
 \end{aligned}$$

where

$$\begin{aligned}
 \psi_1(h, \zeta, \xi) &= h^2 - 4h\xi + 2\zeta\xi, \\
 \psi_2(h, \zeta, \xi) &= (h - 2\zeta)^2 - 2\zeta\xi, \\
 \theta_1(h, \zeta, \xi) &= h^2 - 6h\zeta + 6h\xi + 6\zeta\xi + 3\zeta^2 - 12\xi^2, \\
 \theta_2(h, \zeta, \xi) &= h^2 - 6h\zeta + 6h\xi - 6\zeta\xi + 5\zeta^2.
 \end{aligned} \tag{8.6}$$

Integrals of  $\theta_j$  vanish over the triangle,

$$\int_0^h \int_0^\zeta \theta_j(h, \zeta, \xi) \, d\xi \, d\zeta = 0, \quad j = 1, 2, \tag{8.7}$$

lending an extra power of  $h$  to the functionals where  $\theta_j$ s appear. No similar observation about  $\psi_j$ s can be made and, although similar in many regards, functionals featuring them ought not be combined with the corresponding ones featuring  $\theta_j$ s by this stage.

When we commence the Zassenhaus splitting procedure from the truncated Magnus expansion  $\tilde{\Theta}_5^{\varepsilon[I]}$  featuring terms of sizes  $\mathcal{O}(\varepsilon^{k\sigma-1})$ ,  $k = 1, 3, 4, 5$ , the resulting exponential splitting will continue to feature terms of all these sizes. The expansion  $\tilde{\Theta}_7^{\varepsilon[I]}$  will feature  $\mathcal{O}(\varepsilon^{k\sigma-1})$  terms with  $k = 1, 3, 4, 5, 6, 7$ , and its Zassenhaus expansion retains such terms. This is suboptimal—the time-symmetric nature of the power-truncated Magnus expansion  $\Theta_p$  implies that it should be possible to expand  $\Theta_p^{\varepsilon[I]}(h)$  and, therefore,  $\tilde{\Theta}_p^{\varepsilon[I]}(h)$  solely in odd powers of  $h$  for any choice of the potential,  $V$ .

A Zassenhaus splitting commencing from an odd-powered expansion will never introduce even powers of  $h$  since underlying the procedure is a recursive application of the symmetric BCH which features only odd-grade commutators. The objective of this section is to develop such a Magnus–Zassenhaus scheme while preserving integrals in their undiscretised forms. This splitting will not feature  $\mathcal{O}(\varepsilon^{2k\sigma-1})$  terms, reducing the number of exponentials that need to be evaluated in the splitting.

### 8.4.1 Revisiting time symmetry

To examine the time symmetry of the Magnus expansion, we revisit (2.48),

$$u(t+h) = e^{\Theta(t+h,t)} u(t), \tag{2.48}$$

where the exponential,  $\exp(\Theta(t+h, t))$ , is the evolution operator from  $t$  to  $t+h$ . As we had remarked earlier,  $\Theta(t+h, t)$  can be easily recovered from  $\Theta(h, 0)$  (shortened to  $\Theta(h)$ ) by substituting all occurrences of  $A(\zeta)$  with  $A(t+\zeta)$ .

We define  $t_{1/2} = t + h/2$  as the midpoint of the interval  $[t, t+h]$  and rewrite (2.48) as

$$u(t_{1/2} + \frac{h}{2}) = e^{\Theta(t_{1/2} + \frac{h}{2}, t_{1/2} - \frac{h}{2})} u(t_{1/2} - \frac{h}{2}). \tag{8.8}$$

Since  $\Theta(t_{1/2} - \frac{h}{2}, t_{1/2} + \frac{h}{2})$  is the evolution operator from  $t_{1/2} + \frac{h}{2}$  to  $t_{1/2} - \frac{h}{2}$  (going backward in time by length  $h$ ), we also have

$$u(t_{1/2} - \frac{h}{2}) = e^{\Theta(t_{1/2} - \frac{h}{2}, t_{1/2} + \frac{h}{2})} u(t_{1/2} + \frac{h}{2}). \quad (8.9)$$

Combining the two, we find that  $\Theta(t_{1/2} - \frac{h}{2}, t_{1/2} + \frac{h}{2}) = -\Theta(t_{1/2} + \frac{h}{2}, t_{1/2} - \frac{h}{2})$ , so that  $\Theta$  is odd in  $h$  around the midpoint  $t_{1/2}$ . Similarly, the power-truncated Magnus expansions  $\Theta_p^{\varepsilon[I]}(t_{1/2} + \frac{h}{2}, t_{1/2} - \frac{h}{2})$  are also odd in  $h$  around  $t_{1/2}$  due to their time symmetry (Iserles et al. 2001). If we expand  $\Theta_p^{\varepsilon[I]}$  in powers of  $h$  around  $t_{1/2}$ , therefore, we should only get odd powers of  $h$ .

The expansion  $\Theta_p^{\varepsilon[I]}(t_{1/2} + \frac{h}{2}, t_{1/2} - \frac{h}{2}) = \Theta_p^{\varepsilon[I]}(t+h, t)$  can be obtained from  $\Theta_p^{\varepsilon[I]}(h) = \Theta_p^{\varepsilon[I]}(h, 0)$  by substituting all occurrences of  $V(\zeta)$  by  $V(t_{1/2} - \frac{h}{2} + \zeta)$ . Keeping the odd nature of the expansion about  $t_{1/2}$  in mind, we shift the origin to  $t_{1/2}$  by defining

$$U(\zeta) = V(t_{1/2} + \zeta),$$

whereby we need to substitute  $V(\zeta)$  with  $U(\zeta - \frac{h}{2})$ . To arrive at the desired expansion of  $\Theta_p^{\varepsilon[I]}(t+h, t)$ , therefore, we only need to substitute occurrences of  $\mu_{j,k}(h)$  and  $\Lambda[f]_{a,b}(h)$  with the new definitions,

$$\mu_{j,k}(h) = \int_0^h \tilde{B}_j^k(h, \zeta) U\left(\zeta - \frac{h}{2}\right) d\zeta, \quad (8.10)$$

$$\Lambda[f]_{a,b}(h) = \int_0^h \int_0^\zeta f(h, \zeta, \xi) \left[ \partial_x^a U\left(\zeta - \frac{h}{2}\right) \partial_x^b U\left(\xi - \frac{h}{2}\right) \right] d\xi d\zeta. \quad (8.11)$$

Since we have shifted the origin to  $t_{1/2}$ , all odd and even components are to be understood with respect to 0 from this point onwards. This makes identification of the odd components of the Magnus expansion simpler, assuming that the odd and even components of  $U$  can be found.

#### 8.4.2 Identifying odd and even components

A multivariate function  $F$  is said to be odd if  $F(-\zeta) = -F(\zeta)$  and even if  $F(-\zeta) = F(\zeta)$ . The odd and even components,  $F^o$  and  $F^e$ , of a multivariate function  $F$  are defined as

$$F^o(\zeta) = \frac{1}{2} (F(\zeta) - F(-\zeta))$$

and

$$F^e(\zeta) = \frac{1}{2} (F(\zeta) + F(-\zeta)),$$

respectively. It follows that the odd and even components of a product of two multivariate functions are

$$(F(\zeta)G(\zeta))^o = F^e(\zeta)G^o(\zeta) + F^o(\zeta)G^e(\zeta)$$

and

$$(F(\zeta)G(\zeta))^e = F^e(\zeta)G^e(\zeta) + F^o(\zeta)G^o(\zeta),$$

respectively. We stress that the odd-ness or even-ness of a multivariate function is considered with respect to all its parameters and therefore changes with context. In the integration identities,

$$\left( \int_0^s F(r, \zeta) dr \right)^o = \int_0^s F^e(r, \zeta) dr, \quad (8.12)$$

and

$$\left( \int_0^s F(r, \zeta) dr \right)^e = \int_0^s F^o(r, \zeta) dr, \quad (8.13)$$

for instance, these are with respect to  $(s, \zeta)$  on the left hand side but with respect to  $(r, \zeta)$  on the right hand side.

Extending the new definitions of  $\mu$  and  $\Lambda$  (8.10, 8.11), we define

$$\mu_{\star, j, k}(h) = \int_0^h \tilde{B}_j^k(h, \zeta) U^\star \left( \zeta - \frac{h}{2} \right) d\zeta, \quad \star \in \{e, o\}, \quad (8.14)$$

and

$$\Lambda[f]_{\star, a, b}(h) = \int_0^h \int_0^\zeta f(h, \zeta, \xi) \left[ \partial_x^a U \left( \zeta - \frac{h}{2} \right) \partial_x^b U \left( \xi - \frac{h}{2} \right) \right]^\star d\xi d\zeta, \quad \star \in \{e, o\}. \quad (8.15)$$

The even and odd components of previously defined terms  $\mu_{j, k}(h)$  and  $\Lambda[f]_{a, b}(h)$  can be expressed in terms of these new definitions. It should be obvious that  $\tilde{B}_j(h, \zeta)$  is odd in  $(h, \zeta)$  for odd values of  $j$  and even for even values. Consequently,

$$\begin{aligned} (\mu_{j, k}(h))^o &= \int_0^h \left( \tilde{B}_j^k(h, \zeta) U \left( \zeta - \frac{h}{2} \right) \right)^e d\zeta \\ &= \int_0^h \left( \tilde{B}_j^k(h, \zeta) \right)^e U^e \left( \zeta - \frac{h}{2} \right) + \left( \tilde{B}_j^k(h, \zeta) \right)^o U^o \left( \zeta - \frac{h}{2} \right) d\zeta \\ &= \begin{cases} \mu_{e, j, k}(h) & \text{if } jk \text{ is even,} \\ \mu_{o, j, k}(h) & \text{if } jk \text{ is odd.} \end{cases} \end{aligned}$$

In particular,  $\mu_{0, 0}(h)^o = \mu_{e, 0, 0}(h) = \int_0^h U^e \left( \zeta - \frac{h}{2} \right) d\zeta$ . Note carefully the difference of

usage in subscript and superscript. The odd part of  $\Lambda[f]_{a,b}(h)$  is

$$\begin{aligned}
 \left(\Lambda[f]_{a,b}(h)\right)^o &= \int_0^h \int_0^\zeta \left\{ f(h, \zeta, \xi) \left[ \partial_x^a U\left(\zeta - \frac{h}{2}\right) \partial_x^b U\left(\xi - \frac{h}{2}\right) \right] \right\}^o d\xi d\zeta \\
 &= \int_0^h \int_0^\zeta f^o(h, \zeta, \xi) \left[ \partial_x^a U\left(\zeta - \frac{h}{2}\right) \partial_x^b U\left(\xi - \frac{h}{2}\right) \right]^e d\xi d\zeta \\
 &\quad + \int_0^h \int_0^\zeta f^e(h, \zeta, \xi) \left[ \partial_x^a U\left(\zeta - \frac{h}{2}\right) \partial_x^b U\left(\xi - \frac{h}{2}\right) \right]^o d\xi d\zeta \\
 &= \Lambda[f^o]_{e,a,b}(h) + \Lambda[f^e]_{o,a,b}(h).
 \end{aligned}$$

For an odd function such as  $\phi(h, \zeta, \xi) = \zeta - \xi - \frac{h}{3}$ , this reduces to

$$\left(\Lambda[\phi]_{a,b}(h)\right)^o = \Lambda[\phi]_{e,a,b}(h).$$

The even parts  $\mu_{j,k}(h)^e$  and  $\Lambda[f]_{a,b}(h)^e$  can be deduced analogously.

### 8.4.3 Time-symmetrised rewriting

The time symmetry of power truncated Magnus expansions ensures that  $\Theta_p^\varepsilon(h)$  and  $\tilde{\Theta}_p^\varepsilon(h)$  are odd around the origin so that the even part  $\tilde{\Theta}_p^\varepsilon(h)^e$  vanishes, leaving  $\tilde{\Theta}_p^\varepsilon(h) = \tilde{\Theta}_p^\varepsilon(h)^o$ . Even terms in the Magnus expansion may, therefore, be discarded without further analysis. For instance, we conclude that

$$\begin{aligned}
 \tilde{\Theta}_4^{\varepsilon[I]}(h) &= i h \varepsilon \partial_x^2 - i \varepsilon^{-1} (\mu_{0,0}(h))^o - 2 \langle \partial_x (\mu_{1,1}(h))^o \rangle_1 \\
 &\quad + i \varepsilon^{-1} \left( \Lambda[\phi]_{1,1}(h) \right)^o + 2 i \varepsilon \langle \partial_x^2 (\mu_{2,1}(h))^o \rangle_2 \\
 &= i h \varepsilon \partial_x^2 - i \varepsilon^{-1} \mu_{e,0,0}(h) - 2 \langle \partial_x \mu_{o,1,1}(h) \rangle_1 \\
 &\quad + i \varepsilon^{-1} \Lambda[\phi]_{e,1,1}(h) + 2 i \varepsilon \langle \partial_x^2 \mu_{e,2,1}(h) \rangle_2.
 \end{aligned}$$

For an analytic potential,

$$U\left(\zeta - \frac{h}{2}\right) = \sum_{k=0}^{\infty} \frac{U^{(k)}(0)}{k!} \left(\zeta - \frac{h}{2}\right)^k,$$

we expect a gain of a single power of  $h$  in  $\mu_{j,1}(h)$  which, instead of being  $\mathcal{O}(h^{j+1})$ , as might be expected otherwise, turns out to be  $\mathcal{O}(h^{j+2})$  since  $\int_0^h B_j(h, \zeta) U(0) d\zeta$  vanishes. In the case of  $\mu_{e,j,1}(h)$  where we take the even part of the potential  $U$  at  $\zeta - \frac{h}{2}$ , this translates into a gain of two powers of  $h$  since the smallest non-vanishing term is

$$\frac{1}{2} \int_0^h B_j(h, \zeta) U^{(2)}(0) \left(\zeta - \frac{h}{2}\right)^2 d\zeta.$$

Consequently, the term  $2i\varepsilon \langle \partial_x^2 \mu_{e,2,1}(h) \rangle_2$  is  $\mathcal{O}(\varepsilon^{5\sigma-1})$  in contrast to the  $\mathcal{O}(\varepsilon^{4\sigma-1})$  term  $2i\varepsilon \langle \partial_x^2 \mu_{2,1}(h) \rangle_2$  which appears in (7.34). This gain of power arises from discarding the even part  $2i\varepsilon \langle \partial_x^2 \mu_{o,2,1}(h) \rangle_2$  from the full term. Similar analysis shows that  $i\varepsilon^{-1} \Lambda[\phi]_{e,1,1}(h)$  is also  $\mathcal{O}(\varepsilon^{5\sigma-1})$  since the integral of  $\phi$  over the triangle vanishes. We may, therefore, discard this term from the expansion  $\tilde{\Theta}_4^{\varepsilon[I]}$ .

**Note:** After discarding the even terms and the  $\mathcal{O}(\varepsilon^{5\sigma-1})$  odd terms,  $\tilde{\Theta}_4^{\varepsilon[I]}$  and  $\tilde{\Theta}_3^{\varepsilon[I]}$  become identical. This is not entirely surprising since both feature an  $\mathcal{O}(\varepsilon^{5\sigma-1})$  error.

The full power of this approach becomes evident when working on higher-order expansions such as  $\tilde{\Theta}_5^{\varepsilon[I]}(h)$ ,

$$\begin{aligned} \tilde{\Theta}_5^{\varepsilon[I]}(h)^o &= \overbrace{i h \varepsilon \partial_x^2 - i \varepsilon^{-1} \mu_{e,0,0}(h)}^{\mathcal{O}(\varepsilon^{\sigma-1})} - \overbrace{2 \langle \partial_x \mu_{o,1,1}(h) \rangle_1}^{\mathcal{O}(\varepsilon^{3\sigma-1})} \\ &\quad + \overbrace{i \varepsilon^{-1} \Lambda[\phi]_{e,1,1}(h) + 2i\varepsilon \langle \partial_x^2 \mu_{e,2,1}(h) \rangle_2}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\ &\quad + \overbrace{\frac{1}{6} \left\langle \Lambda[\psi_1 + \theta_1]_{o,1,2}(h) + \Lambda[\psi_2 + \theta_2]_{o,2,1}(h) \right\rangle_1}^{\mathcal{O}(\varepsilon^{5\sigma-1})} \\ &\quad + \overbrace{\frac{4}{3} \varepsilon^2 \langle \partial_x^3 \mu_{o,3,1}(h) \rangle_3}^{\mathcal{O}(\varepsilon^{5\sigma-1})} + \overbrace{\frac{1}{4} i \varepsilon \partial_x^4 \mu_{e,2,1}(h)}^{\mathcal{O}(\varepsilon^{5\sigma+1})} = \Theta(h) + \mathcal{O}(\varepsilon^{7\sigma-1}). \end{aligned}$$

In this way, we are able to eradicate all  $\mathcal{O}(\varepsilon^{4\sigma-1})$  terms, which correspond to even powers of  $h$ . Since  $\varphi_j = \psi_j + \theta_j$ ,  $j = 1, 2$ , are even,  $\left( \Lambda[\varphi_j]_{a,b}(h) \right)^o = \Lambda[\varphi_j]_{o,a,b}(h)$ . We also note that  $\Lambda[f]_{o,a,b}(h)$  and  $\mu_{o,j,k}(h)$  gain an extra power of  $h$  since the odd part of the function  $V$  does not feature a constant term,  $V^o(0) = 0$ . The term  $i\varepsilon \partial_x^4 \mu_{e,2,1}(h)$ , which is  $\mathcal{O}(\varepsilon^{5\sigma+1})$ , can now be discarded since  $\sigma \leq 1$ . This process allows us to use half the number of terms for the same degree of accuracy.

#### 8.4.4 Zassenhaus splitting after time-symmetrisation

A Zassenhaus splitting commencing from this Magnus expansion naturally yields an exponential splitting featuring only  $\mathcal{O}(\varepsilon^{\sigma-1})$ ,  $\mathcal{O}(\varepsilon^{3\sigma-1})$  and  $\mathcal{O}(\varepsilon^{5\sigma-1})$  terms,

$$\mathcal{Z}_{2,\sigma}^{\Theta[I]} = e^{\frac{1}{2}W^{[0]}} e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[2]}} e^{W^{[3]}} e^{\frac{1}{2}W^{[2]}} e^{\frac{1}{2}W^{[1]}} e^{\frac{1}{2}W^{[0]}} = \exp\left(\tilde{\Theta}_5^{\varepsilon[I]}(h)\right) + \mathcal{O}(\varepsilon^{7\sigma-1}), \quad (8.16)$$

where

$$\begin{aligned}
W^{[0]} &= i h \varepsilon \partial_x^2 = \mathcal{O}(\varepsilon^{\sigma-1}), \\
W^{[1]} &= -i \varepsilon^{-1} \mu_{e,0,0}(h) = \mathcal{O}(\varepsilon^{\sigma-1}), \\
W^{[2]} &= \frac{1}{6} i h \varepsilon^{-1} (\partial_x \mu_{e,0,0}(h))^2 - 2 \langle \partial_x \mu_{o,1,1}(h) \rangle_1 + \frac{1}{6} i h^2 \varepsilon \langle \partial_x^2 \mu_{e,0,0}(h) \rangle_2 = \mathcal{O}(\varepsilon^{3\sigma-1}), \\
\mathcal{W}^{[3]} &= i \varepsilon^{-1} \Lambda[\phi]_{e,1,1}(h) - \frac{1}{24} i h^2 \varepsilon (\partial_x^4 \mu_{e,0,0}(h)) - \frac{1}{6} i \varepsilon^{-1} (\partial_x \mu_{e,0,0}(h))^2 (\partial_x^2 \mu_{e,2,1}(h)) \\
&\quad - \frac{2}{45} i h^2 \varepsilon^{-1} (\partial_x \mu_{e,0,0}(h))^2 (\partial_x^2 \mu_{e,0,0}(h)) + \frac{1}{6} \langle \Lambda[\varphi_1]_{o,1,2}(h) + \Lambda[\varphi_2]_{o,2,1}(h) \rangle_1 \\
&\quad - h \langle (\partial_x \mu_{e,0,0}(h)) (\partial_x^2 \mu_{o,1,1}(h)) - \frac{1}{3} (\partial_x^2 \mu_{e,0,0}(h)) (\partial_x \mu_{o,1,1}(h)) \rangle_1 \\
&\quad + \frac{1}{30} i h^3 \varepsilon \langle (\partial_x^2 \mu_{e,0,0}(h))^2 - 2 (\partial_x \mu_{e,0,0}(h)) (\partial_x^3 \mu_{e,0,0}(h)) \rangle_2 \\
&\quad + 2i \varepsilon \langle \partial_x^2 \mu_{e,2,1}(h) \rangle_2 + \frac{4}{3} \varepsilon^2 \langle \partial_x^3 \mu_{o,3,1}(h) \rangle_3 \\
&\quad + \frac{1}{3} h^2 \varepsilon^2 \langle \partial_x^3 \mu_{o,1,1}(h) \rangle_3 - \frac{1}{120} i h^4 \varepsilon^3 \langle \partial_x^4 \mu_{e,0,0}(h) \rangle_4 = \mathcal{O}(\varepsilon^{5\sigma-1}),
\end{aligned}$$

where  $\varphi_j = \psi_j + \theta_j$ ,  $j = 1, 2$ . The term  $\frac{1}{24} i h^2 \varepsilon (\partial_x^4 \mu_{e,0,0}(h))$  in  $\mathcal{W}^{[3]}$  is  $\mathcal{O}(\varepsilon^{3\sigma+1})$ . For  $1/2 < \sigma \leq 1$ , this term can be bundled with the  $\mathcal{O}(\varepsilon^{5\sigma-1})$  terms in  $\mathcal{W}^{[3]}$ , whereas for  $\sigma \leq 1/2$  it can be ignored since it is smaller than  $\mathcal{O}(\varepsilon^{7\sigma-1})$ . We remind the reader that the splitting is computationally effective for  $1/3 < \sigma \leq 1$ .

This scheme requires the evaluation of five line integrals,

$$\begin{aligned}
\mu_{e,0,0}(h) &= \int_0^h U^e \left( \zeta - \frac{h}{2} \right) d\zeta, \\
\mu_{o,1,1}(h) &= \int_0^h \left( \zeta - \frac{h}{2} \right) U^o \left( \zeta - \frac{h}{2} \right) d\zeta, \\
\mu_{e,1,2}(h) &= \int_0^h \left( \zeta - \frac{h}{2} \right)^2 U^e \left( \zeta - \frac{h}{2} \right) d\zeta, \\
\mu_{e,2,1}(h) &= \int_0^h \left( \zeta^2 - h\zeta + \frac{1}{6} h^2 \right) U^e \left( \zeta - \frac{h}{2} \right) d\zeta, \\
\mu_{o,3,1}(h) &= \int_0^h \left( \zeta^3 - \frac{3}{2} h \zeta^2 + \frac{1}{2} h^2 \zeta \right) U^o \left( \zeta - \frac{h}{2} \right) d\zeta,
\end{aligned}$$

and three integrals over the triangle,

$$\begin{aligned}
&\int_0^h \int_0^\zeta \phi(h, \zeta, \xi) \left[ \partial_x U \left( \zeta - \frac{h}{2} \right) \partial_x U \left( \xi - \frac{h}{2} \right) \right]^e d\xi d\zeta, \\
&\int_0^h \int_0^\zeta \varphi_1(h, \zeta, \xi) \left[ \partial_x U \left( \zeta - \frac{h}{2} \right) \partial_x^2 U \left( \xi - \frac{h}{2} \right) \right]^o d\xi d\zeta, \\
&\int_0^h \int_0^\zeta \varphi_2(h, \zeta, \xi) \left[ \partial_x^2 U \left( \zeta - \frac{h}{2} \right) \partial_x U \left( \xi - \frac{h}{2} \right) \right]^o d\xi d\zeta,
\end{aligned}$$

once per time step. If analytic expressions are not available, it is possible to approximate these through Gauss–Legendre quadrature. For the  $\mathcal{O}(\varepsilon^{7\sigma-1})$  splitting here, we require

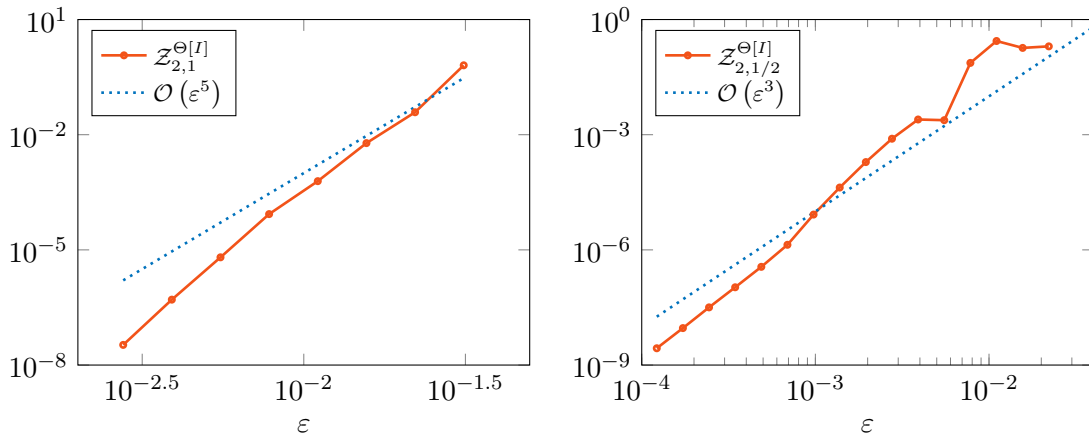


merely three Gauss–Legendre knots  $t_k = h(1 + k\sqrt{3/5})/2, k = -1, 0, 1$ , with weights  $w_k = \frac{5}{18}h, \frac{4}{9}h, \frac{5}{18}h$  (Davis & Rabinowitz 1984), respectively, for the  $\mathcal{O}(h^7)$  accuracy which is required of the quadrature.

**Note:** In Section 2.3.5 we noted that, in the context of power-truncated Magnus expansions, a quadrature method which usually has an accuracy of  $\mathcal{O}(h^{2n})$  gains an order of accuracy, becoming  $\mathcal{O}(h^{2n+1})$ . This was due to time-symmetry and odd nature of such Magnus expansions, where all even terms in the Taylor series, including error terms, vanish. Thus three Gauss–Legendre knots result in a  $\mathcal{O}(h^7)$  method in our case.

Evaluation for the integrals over the line,  $\mu_{\star,j,k}(h)$ , follows directly using these quadrature knots. Integrals over the triangle are also easily evaluated using the same three knots by substituting  $U$  with its interpolating polynomial (see Section 2.3.5).

#### 8.4.5 Numerical results



**Figure 8.3:** Global  $\ell^2$  errors for Example 8.3.1 at  $T = 0.75$  using  $\mathcal{Z}_{2,\sigma}^{\Theta[I]}$  (8.16) under for  $\sigma = 1$  (left) and  $\sigma = \frac{1}{2}$  (right).

Under the scaling  $\sigma = 1$ , the accuracy of the  $\mathcal{Z}_{2,\sigma}^{\Theta[I]}$  scheme is  $\mathcal{O}(\varepsilon^6)$  and we require three and two Lanczos iterations, respectively, for exponentiating  $W^{[2]}$  and  $\mathcal{W}^{[3]}$ . The global  $\ell^2$  error is seen to be better than  $\mathcal{O}(\varepsilon^5)$  through numerical experiments—in Figure 8.3 we use the same setup and scaling choices as elaborated in Section 8.3 for Example 8.3.1. Once again, for  $\sigma = \frac{1}{2}$ , where the error is expected to be  $\mathcal{O}(\varepsilon^2)$ , we use  $\mathcal{Z}_{2,1}^{\Theta[I]}$  for generating reference solutions. This allows us to generate reference solutions for very small  $\varepsilon$ . In Figure 8.3, we find that  $\mathcal{Z}_{2,1/2}^{\Theta[I]}$  outperforms the  $\mathcal{O}(\varepsilon^2)$  error expected out of it, and is closer to  $\mathcal{O}(\varepsilon^3)$  error.

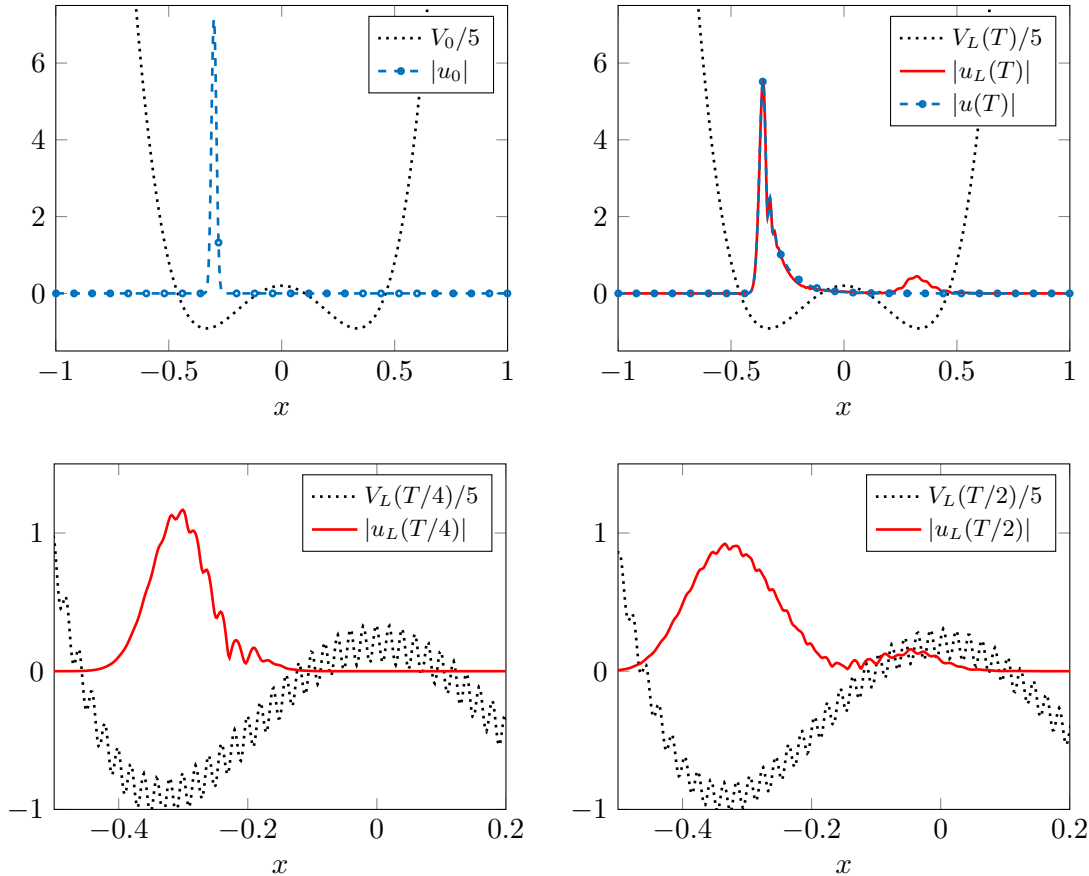
**Example 8.4.1.** Consider the evolution of the wave-packet

$$u_0(x) = \sqrt{10}(\delta\pi)^{-1/4} \exp\left(-\frac{(x-x_0)^2}{2\delta}\right), \quad x_0 = -0.3, \quad \delta = 10^{-3}, \quad (8.17)$$

sitting in the double-well potential

$$V_0 = 450x^4 - 100x^2 + 1. \quad (8.18)$$

Figure 8.4 shows that the wave-packet  $u(T)$  at  $T = 0.1$ , evolving under the influence of the time-independent double-well potential  $V_0$  alone, does not leave the left well (note that the Magnus–Zassenhaus schemes do not restrict us to  $T = 0.1$  in any way—rather, this is a choice that suffices to demonstrate the control in this particular experimental setup and experimental results on numerical accuracy in Figure 8.5 are obtained for  $T = 1$ ).



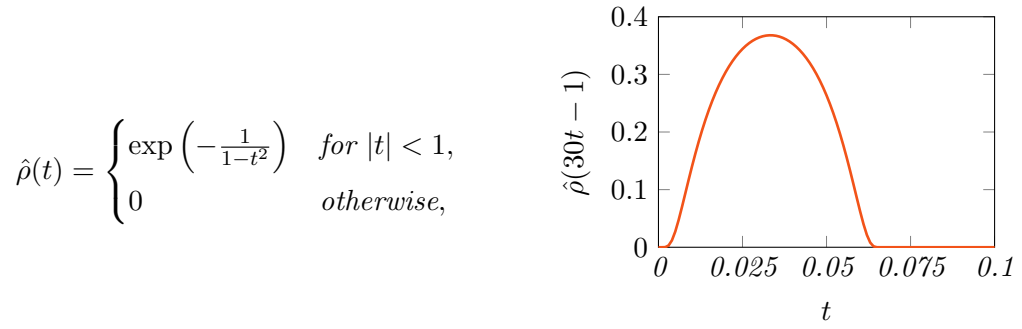
**Figure 8.4:** Initial wave-packet  $u(0) = u_0$  (left), final wave-packets at time  $T = 0.1$  (right):  $u(T)$  under the influence of  $V_0$  alone and  $u_L(T)$  under the influence of  $V_L(x, t) = V_0(x) + L_{100}(x, t)$  in Experiment 8.4.1. For ease of illustration we depict the potential scaled down by a factor of five. The intermediate stages of  $u_L$  at  $T/4$  and  $T/2$  are shown in the bottom row (zoomed in for clarity).

When we excite the wave-packet with the time-dependent potential,

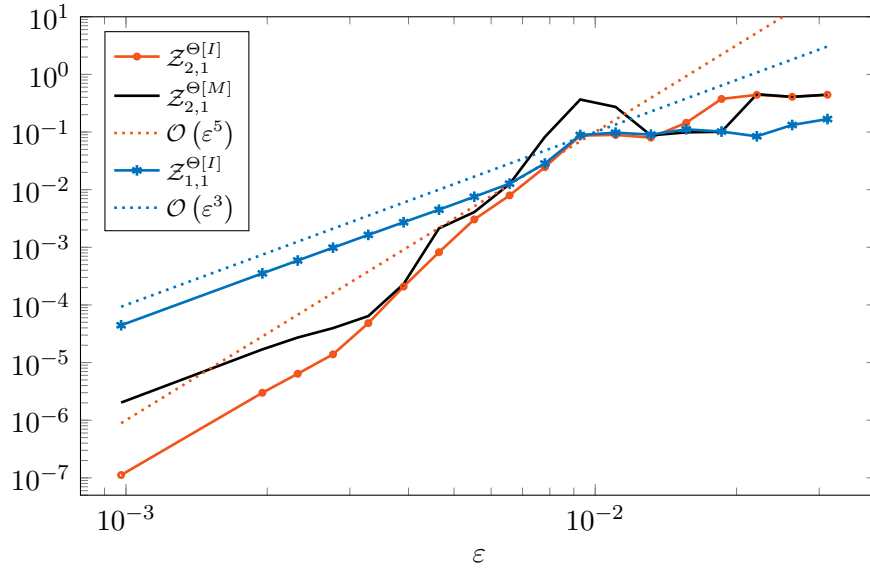
$$L_\omega(x, t) = -2\hat{\rho}(30t - 1) \sin((x - 5t)\pi\omega), \quad (8.19)$$

we find that we are able to induce part of the wave-packet to move to the second well. Figure 8.4 shows the final wave-packet  $u_L(T)$  under the influence of the potential  $V_L(x, t) = V_0 + L_{100}(x, t)$ .

Here the semiclassical parameter is taken to be  $\varepsilon = 0.01$  and  $\hat{\rho}(t) = \text{expbump}(t)$  is the bump function (2.18),



which acts as a smooth envelope simulating the switching on and off of the time-dependent component of the potential.



**Figure 8.5:** Global  $\ell^2$  errors for Example 8.4.1 at  $T = 1$  using  $\mathcal{Z}_{2,1}^{\Theta[I]}$  (8.16),  $\mathcal{Z}_{1,1}^{\Theta[M]}$  (8.1) and  $\mathcal{Z}_{1,1}^{\Theta[I]}$  (8.2).

Although neither the wave-function nor the potential is periodic, we impose periodic boundary conditions at  $x = \pm 1$  in order to resolve spatial oscillations with spectral accu-

racy and restrict the domain to  $[-1, 1]$ . Reasonable results are obtained even though the initial conditions are not periodic since the potential is a confining potential.

The errors for Example 8.4.1 are presented in Figure 8.5. These are the global errors in  $u(x, T)$  where  $T = 1$  is the final time. The precise scaling used in our numerical experiments is

$$M \sim 2\varepsilon^{-1}, \quad h \sim \varepsilon/4.$$

Note that the potential function used out here is particularly tricky on two counts—firstly,  $V_0$  is very large; secondly, and more worryingly, the time varying component  $L_{100}$  has high oscillations in both space and time. In space,  $L_{100}$  has 100 oscillations, whereas in time it has 500. Thus, we require  $h$  to be sufficiently small in order to resolve the temporal behaviour of  $L_{100}$  correctly. This becomes particularly troublesome in the case of  $\sigma = \frac{1}{2}$ —even though it is likely that for  $\varepsilon \rightarrow 0$  the asymptotic accuracy does show up,  $h = \sqrt{\varepsilon}$  would mean that we would need to wait for  $\varepsilon \sim 10^{-6}$  before the temporal behaviour of the potential is correctly resolved, which is the very least we require.

As we have seen repeatedly—in Chapter 6 and here again—the asymptotic estimates only tell half the story and for moderately small values of  $\varepsilon$ , as well as moderately small values of  $h$  caused by small  $\sigma$ , correct spatio-temporal discretisation of the potential as well as the magnitude of the potential and its derivatives needs to be taken into account since these can play a role in making exponents of the splitting as well as errors large in practice.

## 8.5 Costs of Magnus–Zassenhaus schemes

In Section 8.4 we demonstrate that, even while preserving integrals, we can derive Magnus–Zassenhaus schemes where  $\mathcal{O}(\varepsilon^{2k\sigma-1})$  terms do not appear. Consequently, these schemes look structurally identical to the Zassenhaus splitting for time-independent Hamiltonians. The  $\mathcal{O}(\varepsilon^{(2n+3)\sigma-1})$  Magnus–Zassenhaus scheme is

$$\mathcal{Z}_{n,\sigma}^{\Theta[\star]} = e^{\frac{1}{2}W^{[0]}} \cdots e^{\frac{1}{2}W^{[n]}} e^{W^{[n+1]}} e^{\frac{1}{2}W^{[n]}} \cdots e^{\frac{1}{2}W^{[0]}},$$

for  $\star \in M, I$ , where each  $W^{[k]} = \mathcal{O}(\varepsilon^{(2k-1)\sigma-1})$ ,  $k \geq 2$ , is of the form

$$W^{[k]} = \sum_{j=0}^{2k-2} i^{j+1} \langle f_j \rangle_j,$$

following the narrative of Theorem 6.7.1. Evaluating  $W^{[k]}u$  requires  $2 \times (2k-2) + 2 = 4k-2$  FFTs. Unlike the case of Zassenhaus splittings, here we find odd-indexed terms such as  $\langle f_1 \rangle_1$  appearing and the cost is slightly higher, as a consequence. Consequently,  $\mathcal{Z}_{n,\sigma}^{\Theta[\star]}$

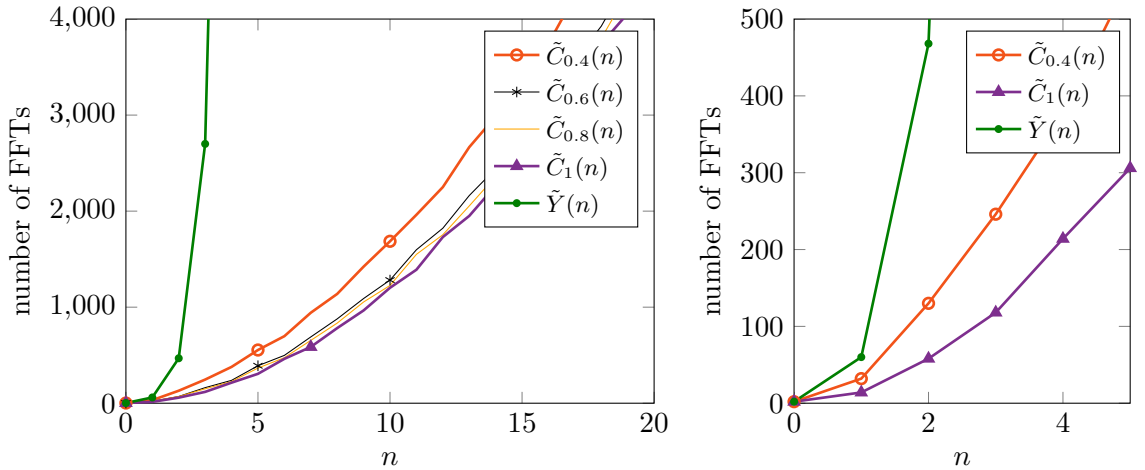
requires

$$\tilde{C}_\sigma(n) = 2 + (4(n+1) - 2) \left\lceil \frac{(2n+3)\sigma - 1}{(2n+1)\sigma - 1} \right\rceil + 2 \sum_{k=2}^n (4k-2) \left\lceil \frac{(2n+3)\sigma - 1}{(2k-1)\sigma - 1} \right\rceil \quad (8.20)$$

FFTs.

We recall from Section 7.4 that a comparable Yoshida splitting requires  $\tilde{Y}(n) = (8n^2 + 8n + 4) \times 3^n$  FFTs for the Magnus expansions simplified in  $\mathfrak{H}$ . The cost of Yoshida splittings for Magnus expansions of Section 3.4.1 which feature commutators of matrices is even higher since high grade commutators are extremely expensive to exponentiate (even via Lanczos iterations).

As we see in Figure 8.6, the costs of Yoshida splittings are found to be universally higher here than the cost of Magnus–Zassenhaus schemes—even for low and moderately high-order accuracies.



**Figure 8.6:** Number of FFTs in the Magnus–Zassenhaus splitting  $\mathcal{Z}_{n,\sigma}^{\Theta[\star]}$  for  $\sigma = 2/5, 3/5, 4/5, 1$ , compared to the  $\tilde{Y}(n) = (8n^2 + 8n + 4) \times 3^n$  FFTs required for a comparable Yoshida splitting. Here, even for small  $n \leq 3$ , the number of FFTs for Yoshida splittings is higher than for  $\mathcal{Z}_{n,\sigma}^{\Theta[\star]}$ .

**Note:** We noted in Section 6.7 that Yoshida splittings are not the most efficient and optimised of high order splittings in existence. On the other hand, no splittings for time-dependent potentials along the lines of the highly optimised splittings of (Blanes et al. 2008) were known to the author at the time of writing. Thus, the choice of Yoshida splittings as a reference seems reasonable.



## Chapter 9

# Formal error analysis

In the entire discourse preceding this chapter, we have designed symmetric Zassenhaus splittings and truncated Magnus expansions of undiscretised operators while assuming  $\partial_x^n = \mathcal{O}(\varepsilon^{-n})$  and  $\langle f \rangle_n = \mathcal{O}(\varepsilon^{-n})$ . This has allowed us to separate terms by powers of  $\varepsilon$  and to estimate the error of our schemes as well as the cost.

The scaling  $\partial_x^n = \mathcal{O}(\varepsilon^{-n})$  is argued by noting that the  $\mathcal{O}(\varepsilon)$  wavelength oscillations in the wave function force us to take  $M = \mathcal{O}(\varepsilon^{-1})$  grid points, whereby  $\partial_x^n$  is replaced by  $\mathcal{K}^n$  upon discretisation, which scales as  $\mathcal{O}(M^n) = \mathcal{O}(\varepsilon^{-n})$ . However, this approach indicates that the analysis of size of terms as well as the error is only valid upon discretisation. This is a cause for concern since it is crucial for the design of our methods, such as the Magnus–Zassenhaus schemes, that we work with undiscretised operators till the very last stage. The jump to the analysis of the undiscretised case is far from trivial, however, and will form the gist of this chapter.

In this chapter we attempt to put the analysis employed in the rest of this thesis on firm theoretical grounds. A common ingredient is the highly oscillatory behaviour of the wave function, proven in Section 9.1 and Section 9.2 via conservation of energy. A reader comfortable with these assumptions might prefer to jump directly to Section 9.3 after consulting Section 9.1.4 and Section 9.1.5 which give a more formal sense to our shorthand  $\partial_x^n = \mathcal{O}(\varepsilon^{-n})$  and  $\langle f \rangle_n = \mathcal{O}(\varepsilon^{-n})$ .

The observations of high oscillations noted in Section 9.1 and Section 9.2 are employed in Section 9.3.2 to derive an error bound for truncated but undiscretised Magnus expansions, following largely the approach of (Hochbruck & Lubich 2003). Error bounds for the spatially discretised Magnus expansions then follow in a straightforward way in Section 9.3.3 by following the analysis of (Bao et al. 2002) and (Pasciak 1980). In doing so, we find that the degree of spatial discretisation required is somewhat higher. However, due to  $\partial_x^n = \mathcal{O}(\varepsilon^{-n})$  holding nevertheless as a consequence of the results in Section 9.1, the finer spatial discretisation affects only the cost of FFTs and not the analysis of error, which is made precise in Section 9.3.

In addition to keeping the error analysis largely independent of spatial discretisation considerations, the analysis presented in this chapter is carried out with the time step  $h$  kept independent of  $\varepsilon$  throughout, except where noted explicitly. This is in contrast to the development of methods presented earlier in this thesis and, in fact, suggests that it should be possible to design methods along the same lines without the power laws governing the scaling of  $h$  and  $M$  in terms of  $\varepsilon$ . However, we suspect that this might make the analysis of cost more involved.

The error analysis for Zassenhaus splittings is presented in Section 9.4. Having been developed earlier in this thesis and being applicable to the simpler case of time-independent potential, it is natural to expect that the analysis for Zassenhaus splittings should be presented first. The reason for this anomaly is that the error analysis for Zassenhaus splittings proceeds by studying the error in truncating the sBCH series (which is used recursively in the symmetric Zassenhaus procedure), which, in turn, can be expressed as a Magnus expansion for the purpose of analysis. Therefore the techniques used in the error analysis of Magnus expansions extend to the analysis of the symmetric Zassenhaus splittings.

## 9.1 Highly oscillatory solutions in the semiclassical regime

Recall the highly oscillatory behaviour of solutions of the Schrödinger equation under the semiclassical regime from (3.10),

$$\left\| \partial_x^k u(t) \right\|_{L^2} \leq C_k \varepsilon^{-k}, \quad k \in \mathbb{Z}^+, \quad t \geq 0, \quad (9.1)$$

where  $C_k$ s are independent of  $\varepsilon$  and  $t$ , along the lines of (Bao et al. 2002, Jin et al. 2011). This growth of spatial derivatives of the wave function is evident from the numerical example presented in Figure 3.2.

In this section we will prove that these bounds hold for the solution of the Schrödinger equations (3.9) and (3.19) for all  $t$  if they do for the initial conditions. The proof strategy involves noting that once the initial condition satisfies (9.1), this implies finite initial energy under the semiclassical limit and, therefore, also the finiteness of energy for all times due to conservation of energy. Finiteness of energy at all times, in turn, implies that derivatives of the wave function must stay bounded by (9.1) at all times.

We will also need to assume that the potential function has bounded derivatives

$$\left\| \partial_x^k V(t) \right\| \leq \hat{C}_k, \quad k \in \mathbb{Z}^+, \quad t \geq 0, \quad (9.2)$$

which also have bounded variation in  $t$ ,

$$\int_0^t \left\| \partial_t \partial_x^k V(\tau) \right\| d\tau \leq \hat{C}_k, \quad k \in \mathbb{Z}^+, \quad t \geq 0, \quad (9.3)$$



where  $\hat{C}_k$ s are once again independent of  $\varepsilon$  and  $t$ . Here the norm on derivatives of  $V$  might be  $L^2$  or  $L^\infty$ . These observations will allow us to formalise the statement  $\partial_x^k = \mathcal{O}(\varepsilon^{-k})$ .

As for the initial conditions, it should be simple to verify whether they satisfy (9.1). For  $k > 1$ , we need to consider conservation of higher moments of the Hamiltonian. In Section 9.2, in the case of truncated but undiscretised Magnus expansions, we will consider the Magnus expansion as a perturbed Hamiltonian whose moments are also conserved.

Note that the proof of (9.1) in Section 9.1.1 and Section 9.1.3 does not rely on an appeal to the non-physical nature of infinite energy under the semiclassical limits. Moreover, we avoid the extremely complex issue of semiclassical limits of observable densities.

The author gratefully acknowledges the input of Weizhu Bao, whose insight on the emergence of this oscillatory behaviour as a consequence of semiclassical limit of energy forms the basis of Lemma 9.2.1. This insight has also inspired proofs of other oscillatory behaviour considered in this chapter.

### 9.1.1 The first derivative in the semiclassical limit

The energy of the quantum system at time  $t$  is the expected value of the Hamiltonian,

$$E(t) = \langle u(t), Hu(t) \rangle.$$

In the case of the standard Hamiltonian with time-independent potential under semiclassical scaling,

$$H = -\varepsilon^2 \partial_x^2 + V(x),$$

the energy can be written as

$$E(t) = \|\varepsilon \partial_x u(t)\|_{L^2}^2 + \langle u(t), Vu(t) \rangle.$$

Naturally, a fundamental physical property that one expects to be conserved in a closed system is the energy. This is indeed the case. Noting that the evolution of the wavefunction is given by

$$u(t) = e^{-itH/\varepsilon} u(0),$$

we can show that the energy is conserved,

$$\begin{aligned} E(t) &= \langle u(t), Hu(t) \rangle = \langle e^{-itH/\varepsilon} u(0), H e^{-itH/\varepsilon} u(0) \rangle \\ &= \langle e^{-itH/\varepsilon} u(0), e^{-itH/\varepsilon} H u(0) \rangle = \langle u(0), H u(0) \rangle = E(0). \end{aligned}$$

The above proof works out because  $H$  commutes with the evolution operator  $e^{-itH/\varepsilon}$ .

Lemma 9.2.1 forms the basis of our assumption that the wave function develops oscillations of wavelength  $\mathcal{O}(\varepsilon)$  in space and time under the semiclassical regime. When considering the semiclassical limit, we will explicitly note the dependence of  $u$ ,  $H$  and  $E$

on  $\varepsilon$  by writing  $u^\varepsilon$ ,  $H^\varepsilon$  and  $E^\varepsilon$ , instead.

**Lemma 9.1.1.** *In the semiclassical limit,  $\varepsilon \rightarrow 0$ , the spatial derivative of the wavefunction remains bounded by inverse of the semiclassical parameter,*

$$\|\partial_x u^\varepsilon(t)\|_{L^2} \leq C\varepsilon^{-1}, \quad t \geq 0, \quad (9.4)$$

if it is bounded at  $t = 0$ ,

$$\|\partial_x u^\varepsilon(0)\|_{L^2} \leq C_0\varepsilon^{-1}.$$

Here  $0 < C < \infty$  is independent of  $\varepsilon$  and  $t$ .

*Proof.* Noting that

$$\|\varepsilon \partial_x u^\varepsilon(t)\|_{L^2}^2 - \|V\| \leq E^\varepsilon(t) \leq \|\varepsilon \partial_x u^\varepsilon(t)\|_{L^2}^2 + \|V\|, \quad t \geq 0,$$

and the fact that  $E^\varepsilon(t) = E^\varepsilon(0)$ , we conclude

$$\|\varepsilon \partial_x u^\varepsilon(t)\|_{L^2}^2 \leq \|\varepsilon \partial_x u^\varepsilon(0)\|_{L^2}^2 + 2\|V\|.$$

Thus

$$\lim_{\varepsilon \rightarrow 0} \|\varepsilon \partial_x u(t)\|_{L^2}^2 \leq C_0^2 + 2\|V\|,$$

so that  $C = \sqrt{C_0^2 + 2\|V\|}$  suffices for (9.4).  $\square$

**Note:** Here, and later, the specific norm is often not written for functions such as  $V$  and it can be taken as  $L^2$  or  $L^\infty$  as convenient.

**Note:**  $C^2$  bounds from above the kinetic energy at time  $t$ . Not only does the kinetic energy needs to be finite, but it cannot be 0 for all cases—certainly not when interesting dynamics are involved. Thus, although the extent of oscillations will depend from case to case, we can safely assume that  $\mathcal{O}(\varepsilon)$  oscillations generally appear in the solution at finite times.

**Note:** (9.4) is a more informal way of writing

$$\lim_{\varepsilon \rightarrow 0} \|\varepsilon \partial_x u(t)\|_{L^2} \leq C.$$

### 9.1.2 Time-dependent potentials

In case of the exact solution for time-dependent potentials, where the Hamiltonian is

$$H(t) = -\varepsilon^2 \partial_x^2 + V(t, x),$$

the energy

$$E(t) = \langle u(t), H(t)u(t) \rangle = \|\varepsilon \partial_x u(t)\|_{L^2}^2 + \langle u(t), V(t)u(t) \rangle,$$

is no longer conserved. However, this is completely attributable to the time-varying part of the potential function, which is an external influence. The energy should, nevertheless, stay finite for the exact solution and one expects the exact solution to satisfy (9.1).

**Lemma 9.1.2.** *In the semiclassical limit,  $\varepsilon \rightarrow 0$ , the spatial derivative of the wavefunction remains bounded by inverse of the semiclassical parameter,*

$$\|\partial_x u^\varepsilon(t)\|_{L^2} \leq C\varepsilon^{-1}, \quad t \geq 0, \quad (9.5)$$

if it is bounded at  $t = 0$ ,

$$\|\partial_x u^\varepsilon(0)\|_{L^2} \leq C_0\varepsilon^{-1},$$

so long as the variations of the time-dependent potential are bounded independently of  $t$  (9.3). Here  $0 < C < \infty$  is independent of  $\varepsilon$  and  $t$ .

*Proof.* The exact solution is obtained as a limit

$$u^\varepsilon(t) = \lim_{n \rightarrow \infty} \prod_{k=1}^n \exp(-ih\varepsilon^{-1}H^\varepsilon(t_k))u^\varepsilon(0),$$

such that  $h = t_{n+1} - t_n$  and  $t_n \rightarrow t$ .

Considering finite  $n$  (and finite  $h$ ), the energy

$$E_k^\varepsilon(t) = \langle u^\varepsilon(t), H^\varepsilon(t_k)u^\varepsilon(t) \rangle$$

is preserved from  $t_k$  to  $t_{k+1}$ ,

$$E_k^\varepsilon(t_{k+1}) = E_k^\varepsilon(t_k),$$

since  $H(t_k)$  commutes with the evolution operator. Thus

$$\begin{aligned} E_k^\varepsilon(t_k) - E_{k-1}^\varepsilon(t_{k-1}) &= E_k^\varepsilon(t_k) - E_{k-1}^\varepsilon(t_k) \\ &= \langle u^\varepsilon(t_k), (H^\varepsilon(t_k) - H^\varepsilon(t_{k-1}))u^\varepsilon(t_k) \rangle \\ &= \langle u^\varepsilon(t_k), (V(t_k) - V(t_{k-1}))u^\varepsilon(t_k) \rangle \\ &\leq \|V(t_k) - V(t_{k-1})\|. \end{aligned}$$

Telescoping this series, we conclude

$$\begin{aligned} E_n^\varepsilon(t_n) - E_0^\varepsilon(t_0) &\leq \sum_{k=1}^n \|V(t_k) - V(t_{k-1})\| \\ &\leq \int_0^{t_n} \|\partial_t V(\tau)\| \, d\tau, \end{aligned}$$

so that under  $n \rightarrow \infty$  and  $t_n \rightarrow t$ , while exploiting the bounded variation of  $V$  due to assumption (9.3),

$$E^\varepsilon(t) \leq E^\varepsilon(0) + \int_0^t \|\partial_t V(\tau)\| \, d\tau \leq E^\varepsilon(0) + \hat{C}_0.$$

Now consider the limit  $\varepsilon \rightarrow 0$ ,

$$\lim_{\varepsilon \rightarrow 0} \|\varepsilon \partial_x u^\varepsilon(t)\|_{L^2}^2 \leq \lim_{\varepsilon \rightarrow 0} \|\varepsilon \partial_x u^\varepsilon(0)\|_{L^2}^2 + \|V(0)\| + \|V(t)\| + \hat{C}_0 \leq C_0^2 + 3\hat{C}_0,$$

so that (9.5) is satisfied with  $C = \sqrt{C_0^2 + 3\hat{C}_0}$ .  $\square$

### 9.1.3 Higher derivatives in the semiclassical limit

In this section we will restrict our attention to time-independent potentials. Corresponding results for time-dependent potentials can be obtained along the lines of this section by combining the observations of Lemma 9.1.2.

**Lemma 9.1.3.** *In the semiclassical limit,  $\varepsilon \rightarrow 0$ , higher spatial derivatives of the wavefunction remain bounded by inverse powers of the semiclassical parameter,*

$$\left\| \partial_x^k u^\varepsilon(t) \right\|_{L^2} \leq C_k \varepsilon^{-k}, \quad k \in \mathbb{Z}^+, \, t \geq 0, \quad (9.1)$$

if they are bounded at  $t = 0$  and  $V$  satisfies (9.2)

$$\left\| \partial_x^k V \right\| \leq \hat{C}_k, \quad k \in \mathbb{Z}^+,$$

where  $0 < \hat{C}_k < \infty$  and  $0 < C_k < \infty$  are independent of  $\varepsilon$  and  $t$ .

*Proof.* For deriving (9.1) for higher derivatives, we note that

$$\begin{aligned} E_m^\varepsilon(t) &= \langle u(t), (H^\varepsilon)^m u(t) \rangle = \langle e^{-itH^\varepsilon/\varepsilon} u(0), (H^\varepsilon)^m e^{-itH^\varepsilon/\varepsilon} u(0) \rangle \\ &= \langle e^{-itH^\varepsilon/\varepsilon} u(0), e^{-itH^\varepsilon/\varepsilon} (H^\varepsilon)^m u(0) \rangle = \langle u(0), (H^\varepsilon)^m u(0) \rangle = E_m^\varepsilon(0), \end{aligned}$$

is also conserved since arbitrary powers of  $H$  also commute with the evolution operator.

We will prove (9.1) for  $k \in \mathbb{Z}^+$  by inducting over  $m$ . In Lemma 9.2.1, we have already proven the  $m = 1$  case. We first consider the case where  $m = 2n$ , assuming that (9.1) holds for all  $k < 2n$ .

**Case 1,  $m = 2n$ :** The  $2n$ th moment of the Hamiltonian can also be written as

$$E_{2n}^\varepsilon(t) = \langle (H^\varepsilon)^n u(t), (H^\varepsilon)^n u(t) \rangle.$$

Note that  $(H^\varepsilon)^n$  is a degree  $2n$  differential operator which can be written as

$$(H^\varepsilon)^n = \sum_{k=0}^{2n} \varepsilon^k f_k \partial_x^k$$

by using the chain rule, where  $f_k = \mathcal{O}(\varepsilon^0)$  are composed of the potential and its derivatives. Thus we may write

$$\begin{aligned} E_{2n}^\varepsilon(t) &= \sum_{k=0}^{2n} \left\langle \varepsilon^k f_k \partial_x^k u^\varepsilon(t), \varepsilon^k f_k \partial_x^k u^\varepsilon(t) \right\rangle + \sum_{\substack{j,k=0 \\ j \neq k}}^{2n} i^{j+k} \left\langle \varepsilon^k f_k \partial_x^k u^\varepsilon(t), \varepsilon^j f_j \partial_x^j u^\varepsilon(t) \right\rangle \\ &= \sum_{k=0}^{2n} \left\| \varepsilon^k f_k \partial_x^k u^\varepsilon(t) \right\|_{L^2}^2 + \sum_{\substack{j,k=0 \\ j \neq k}}^{2n} i^{j+k} \left\langle \varepsilon^k f_k \partial_x^k u^\varepsilon(t), \varepsilon^j f_j \partial_x^j u^\varepsilon(t) \right\rangle. \end{aligned} \quad (9.6)$$

The bounds we need are

$$\begin{aligned} &\sum_{k=0}^{2n} \|f_k\|^2 \left\| \varepsilon^k \partial_x^k u^\varepsilon(t) \right\|_{L^2}^2 - \sum_{\substack{j,k=0 \\ j \neq k}}^{2n} \|f_k\| \|f_j\| \left\| \varepsilon^k \partial_x^k u^\varepsilon(t) \right\|_{L^2} \left\| \varepsilon^j \partial_x^j u^\varepsilon(t) \right\|_{L^2} \leq E_{2n}^\varepsilon(t) \\ &\leq \sum_{k=0}^{2n} \|f_k\|^2 \left\| \varepsilon^k \partial_x^k u^\varepsilon(t) \right\|_{L^2}^2 + \sum_{\substack{j,k=0 \\ j \neq k}}^{2n} \|f_k\| \|f_j\| \left\| \varepsilon^k \partial_x^k u^\varepsilon(t) \right\|_{L^2} \left\| \varepsilon^j \partial_x^j u^\varepsilon(t) \right\|_{L^2}. \end{aligned} \quad (9.7)$$

Our assumptions regarding  $u^\varepsilon(0)$  directly prove the boundedness of  $E_{2n}^\varepsilon(0)$  under  $\varepsilon \rightarrow 0$ , and thus also of  $E_{2n}^\varepsilon(t)$  for all  $t$ . Here (9.2) is crucial for concluding that  $\|f_k\|$  are independent of  $\varepsilon$ .

If we were to assume that  $u^\varepsilon(t)$  satisfies (9.1) for  $k < 2n$  but doesn't for  $k = 2n$ ,

$$\lim_{\varepsilon \rightarrow \infty} \left\| \varepsilon^{2n} \partial_x^{2n} u^\varepsilon(t) \right\|_{L^2} = \infty,$$

we immediately arrive upon a contradiction since we can show that  $\lim_{\varepsilon \rightarrow \infty} E_{2n}^\varepsilon(t) \geq \infty$  under such an assumption. Thus,

$$\lim_{\varepsilon \rightarrow \infty} \left\| \varepsilon^{2n} \partial_x^{2n} u^\varepsilon(t) \right\|_{L^2} = \tilde{C}_{2n}(t)$$

for some  $0 < \tilde{C}_{2n}(t) < \infty$ .

Having established finiteness, we consider a time-independent bound. Consider the limit of (9.7) under  $\varepsilon \rightarrow 0$ , rewritten in the form

$$\begin{aligned} & \|f_{2n}\|^2 \left\| \varepsilon^{2n} \partial_x^{2n} u^\varepsilon(t) \right\|_{L^2}^2 - 2 \|f_{2n}\| \left\| \varepsilon^{2n} \partial_x^{2n} u^\varepsilon(t) \right\|_{L^2} \sum_{k=0}^{2n-1} \|f_k\| \left\| \varepsilon^k \partial_x^k u^\varepsilon(t) \right\|_{L^2} \\ & \leq E_{2n}^\varepsilon(0) - \sum_{k=0}^{2n-1} \|f_k\|^2 \left\| \varepsilon^k \partial_x^k u^\varepsilon(t) \right\|_{L^2}^2 + \sum_{\substack{j,k=0 \\ j \neq k}}^{2n-1} \|f_k\| \|f_j\| \left\| \varepsilon^k \partial_x^k u^\varepsilon(t) \right\|_{L^2} \left\| \varepsilon^j \partial_x^j u^\varepsilon(t) \right\|_{L^2}. \end{aligned}$$

This can be written in the quadratic form

$$ay(\varepsilon)^2 - by(\varepsilon) - c \leq 0,$$

where

$$y(\varepsilon) = \left\| \varepsilon^{2n} \partial_x^{2n} u^\varepsilon(t) \right\|_{L^2},$$

and the coefficients are

$$\begin{aligned} a &= \|f_{2n}\|^2, \\ b &= 2 \|f_{2n}\| \sum_{k=0}^{2n-1} \|f_k\| \left\| \varepsilon^k \partial_x^k u^\varepsilon(t) \right\|_{L^2}, \\ c &= E_{2n}^\varepsilon(0) - \sum_{k=0}^{2n-1} \|f_k\|^2 \left\| \varepsilon^k \partial_x^k u^\varepsilon(t) \right\|_{L^2}^2 + \sum_{\substack{j,k=0 \\ j \neq k}}^{2n-1} \|f_k\| \|f_j\| \left\| \varepsilon^k \partial_x^k u^\varepsilon(t) \right\|_{L^2} \left\| \varepsilon^j \partial_x^j u^\varepsilon(t) \right\|_{L^2}. \end{aligned}$$

The bounds on  $y(\varepsilon)$  are

$$\frac{b - \sqrt{b^2 + 4ac}}{2a} \leq y(\varepsilon) \leq \frac{b + \sqrt{b^2 + 4ac}}{2a}.$$

We know that in the limit  $\varepsilon \rightarrow 0$ ,

$$\begin{aligned} a &= \|f_{2n}\|^2, \\ b &\leq 2 \|f_{2n}\| \sum_{k=0}^{2n-1} \|f_k\| C_k =: b_0, \\ c &\leq E_{2n}^\varepsilon(0) + \sum_{\substack{j,k=0 \\ j \neq k}}^{2n-1} \|f_k\| \|f_j\| C_j C_k =: c_0, \end{aligned}$$

so that  $y(\varepsilon) = \left\| \varepsilon^{2n} \partial_x^{2n} u^\varepsilon(t) \right\|_{L^2} = C_{2n}(t)$  can be bounded from above by the constant  $\frac{b_0 + \sqrt{b_0^2 + 4ac_0}}{2a}$  which is independent of  $t$ .

**Case 2,  $m = 2n+1$ :** The case of  $m = 2n + 1$  follows analogously. As before, we assume that (9.1) holds for all  $k < m$ . The  $m$ th moment of the Hamiltonian can be written as

$$E_{2n+1}^\varepsilon(t) = \langle (H^\varepsilon)^{n+1} u(t), (H^\varepsilon)^n u(t) \rangle.$$

Note that  $(H^\varepsilon)^{n+1}$  is a degree  $2n + 2$  differential operator which can be written as

$$(H^\varepsilon)^{n+1} = \sum_{k=0}^{2n+2} \varepsilon^k g_k \partial_x^k$$

by using the chain rule, where  $g_k = \mathcal{O}(\varepsilon^0)$  are composed of the potential and its derivatives and will be different from  $f_k$ s. In this case we again feature three terms. The first of these is

$$\sum_{k=0}^{2n} \sum_{j=0}^{2n} \left\langle \varepsilon^k g_k \partial_x^k u^\varepsilon(t), \varepsilon^j f_j \partial_x^j u^\varepsilon(t) \right\rangle,$$

which can be bounded in the straightforward way and behaves like the constant term in the quadratic form. The second term is of the form

$$\sum_{j=0}^{2n} \left\langle \varepsilon^{2n+1} g_{2n+1} \partial_x^{2n+1} u^\varepsilon(t), \varepsilon^j f_j \partial_x^j u^\varepsilon(t) \right\rangle,$$

which is a ‘linear’ term (in terms of the  $2n + 1^{\text{th}}$  derivative) and the third is

$$\sum_{j=0}^{2n} \left\langle \varepsilon^{2n+2} g_{2n+2} \partial_x^{2n+2} u^\varepsilon(t), \varepsilon^j f_j \partial_x^j u^\varepsilon(t) \right\rangle,$$

which can be rewritten as

$$\begin{aligned} \left\langle \varepsilon^{2n+2} g_{2n+2} \partial_x^{2n+2} u^\varepsilon(t), \varepsilon^j f_j \partial_x^j u^\varepsilon(t) \right\rangle &= \left\langle \varepsilon^{2n+1} \partial_x^{2n+1} u^\varepsilon(t), \varepsilon \partial_x (g_{2n+2} \varepsilon^j f_j \partial_x^j u^\varepsilon(t)) \right\rangle \\ &= \left\langle \varepsilon^{2n+1} \partial_x^{2n+1} u^\varepsilon(t), g_{2n+2} \varepsilon^{j+1} f_j \partial_x^{j+1} u^\varepsilon(t) \right\rangle \\ &\quad + \left\langle \varepsilon^{2n+1} \partial_x^{2n+1} u^\varepsilon(t), (\partial_x g_{2n+2}) \varepsilon^{j+1} f_j \partial_x^j u^\varepsilon(t) \right\rangle. \end{aligned}$$

The case of  $j = 2n$  provides us the ‘quadratic’ growth (in terms of the  $2n + 1^{\text{th}}$  derivative), while the other cases give us ‘linear’ growth. The rest of the proof for the  $m = 2n + 1$  case follows along the same lines as the  $m = 2n$  case.  $\square$

**Note:** Here, and later, the specific norm is often not written for functions such as  $f_k$  and it can be taken as  $L^2$  or  $L^\infty$  as convenient.

### 9.1.4 Size of symmetrised differential operators

Recall that we use the *big-O* notation to write

$$f(\varepsilon) = \mathcal{O}(g(\varepsilon))$$

if there exists a  $C > 0$  and  $\varepsilon_0 > 0$  such that

$$f(\varepsilon) \leq Cg(\varepsilon), \quad \forall \varepsilon < \varepsilon_0.$$

Thus, another way of writing (9.1) is

$$\partial_x^k u^\varepsilon(t) = \mathcal{O}\left(\varepsilon^{-k}\right),$$

where the size is considered in context of  $\|\cdot\|_{L^2}$ . This also gives rise to our shorthand,

$$\partial_x^k = \mathcal{O}\left(\varepsilon^{-k}\right),$$

which should be understood with the additional caveat that  $\partial_x^k$  acts solely on solutions of the Schrödinger equation in the semiclassical regime that satisfy (9.1).

Since  $\langle f \rangle_k = \frac{1}{2}(f \circ \partial_x^k + \partial_x^k \circ f)$ , applying the chain rule brings us to the form

$$\langle f \rangle_k = f \partial_x^k + \frac{1}{2} \sum_{j=0}^{k-1} \binom{k}{j} (\partial_x^{k-j} f) \partial_x^j.$$

**Lemma 9.1.4.** *The bounds*

$$\lim_{\varepsilon \rightarrow 0} \left\| \varepsilon^k \partial_x^k u^\varepsilon \right\|_{L^2} \leq C_k, \quad k = 0, \dots, n,$$

where  $0 < C_k < \infty$  are independent of  $\varepsilon$ , are equivalent to

$$\lim_{\varepsilon \rightarrow 0} \left\| \varepsilon^k \langle f \rangle_k u^\varepsilon \right\|_{L^2} \leq C_k \|f\|, \quad k = 0, \dots, n, \quad f \in C^n(\mathbb{R}), \quad (9.8)$$

so long as  $f$  obeys (9.2)

$$\left\| \partial_x^k f \right\| \leq \hat{C}_k, \quad k \in \mathbb{Z}^+,$$

where  $0 < \hat{C}_k < \infty$  are independent of  $\varepsilon$ .



*Proof.* This is due to the bound,

$$\begin{aligned} \|f\| \left\| \varepsilon^k \partial_x^k u \right\|_{L^2} - \frac{1}{2} \sum_{j=0}^{k-1} \varepsilon^{k-j} \binom{k}{j} \left\| \partial_x^{k-j} f \right\| \left\| \varepsilon^j \partial_x^j u \right\|_{L^2} &\leq \left\| \varepsilon^k \langle f \rangle_k u \right\|_{L^2} \\ &\leq \|f\| \left\| \varepsilon^k \partial_x^k u \right\|_{L^2} + \frac{1}{2} \sum_{j=0}^{k-1} \varepsilon^{k-j} \binom{k}{j} \left\| \partial_x^{k-j} f \right\| \left\| \varepsilon^j \partial_x^j u \right\|_{L^2}, \end{aligned}$$

where the terms in the sum are  $\mathcal{O}(\varepsilon)$  or smaller (since  $j \leq k-1$  and derivatives of  $f$  are independent of  $\varepsilon$  due to the assumptions) and become inconsequential under the limit  $\varepsilon \rightarrow 0$ , proving (9.8). In the other direction, noting  $\langle 1 \rangle_k = \partial_x^k$ , we conclude that (9.1) is a special case of (9.8).  $\square$

**Note:** This justifies our shorthand  $\langle f \rangle_k = \mathcal{O}(\varepsilon^{-k})$ —again with the caveat that this differential operator acts on solutions of the Schrödinger equation in the semiclassical regime which satisfy (9.1) and  $f$  satisfies (9.2).

### 9.1.5 Size of commutators

In this section we formally restate Lemma 6.6.1 and its corollary

$$\|[\mathcal{A}(\tau_k), [\dots, [\mathcal{A}(\tau_1), \mathcal{A}(\tau_0)] \dots]]_2 = \mathcal{O}(\varepsilon^{-1}), \quad \tau_i \in [0, T], \quad i = 0, \dots, k, \quad k \in \mathbb{Z}^+, \quad (3.26)$$

where  $\mathcal{A}(t) = i\varepsilon \partial_x^2 - i\varepsilon^{-1}V(t)$ .

**Lemma 9.1.5.** *If  $u^\varepsilon$  satisfies (9.1) and  $V(t)$  satisfies (9.2), then for every grade  $n$  commutator  $L_n$  with letters  $B_j \in \{\varepsilon \partial_x^2, \varepsilon^{-1}V(t)\}$ ,  $j = 1, \dots, n$ ,*

$$\lim_{\varepsilon \rightarrow 0} \|\varepsilon L_n(B_1, \dots, B_n) u^\varepsilon\|_{L^2} \leq C, \quad n \in \mathbb{Z}^+, \quad (9.9)$$

where  $0 \leq C < \infty$  is independent of  $\varepsilon$  and  $t$ .

*Proof.* Consider a grade  $n$  commutator  $\tilde{L}_n$  of  $\langle 1 \rangle_2$  and  $\langle V \rangle_0$ , featuring  $k$  occurrences of the letter  $\langle 1 \rangle_2$  and  $n - k$  occurrences of  $\langle V \rangle_0$ . Since  $\langle 1 \rangle_2 \in \mathfrak{G}_2$  and  $\langle V \rangle_0 \in \mathfrak{G}_0$ , using Corollary 5.3.5,

$$\text{ht}(\tilde{L}_n) \leq 2k - n + 1.$$

That is, the highest symmetrised differential operator is of the form  $\langle g \rangle_{2k-n-1}$  (at most), i.e.  $L_n$  can be written as

$$\tilde{L}_n = \sum_{j=0}^{2k-n-1} \langle f_j \rangle_j,$$

where, for simplicity, we have disregarded the fact that the indices can be either odd or even but not both.

The corresponding commutator,  $L_n$ , of  $\varepsilon \langle 1 \rangle_2$  and  $\varepsilon^{-1} \langle V \rangle_0$  is scaled by  $k$  occurrences of  $\varepsilon$  and  $n - k$  occurrences of  $\varepsilon^{-1}$ ,

$$L_n = \varepsilon^{n-2k} \tilde{L}_n = \varepsilon^{n-2k} \sum_{j=0}^{2k-n-1} \langle f_j \rangle_j.$$

The bound (9.9), therefore, follows straight away from Lemma 9.1.4.  $\square$

**Note:** Recall that due to (9.2), the potential and its derivatives are assumed to be bounded independently of  $t$ . This is important in order for  $C$  to be independent of  $t$ .

**Note:** Note that  $C$  can be 0 since a commutator can vanish.

Since commutators of  $\mathcal{A}(t)$  result in commutators of  $\varepsilon \partial_x^2$  and  $\varepsilon^{-1} V(t)$  via linearity, the following Corollary is an immediate consequence.

**Corollary 9.1.6.** *If  $u^\varepsilon$  satisfies (9.1) and  $V(t)$  satisfies (9.2), then*

$$\lim_{\varepsilon \rightarrow 0} \|\varepsilon [\mathcal{A}(\tau_n), [\dots, [\mathcal{A}(\tau_2), \mathcal{A}(\tau_1)] \dots] u^\varepsilon\|_{L^2} \leq C, \quad \tau_j \in [0, t], \quad j = 0, \dots, n, \quad n \in \mathbb{Z}^+, \quad (9.10)$$

where  $0 \leq C < \infty$  is independent of  $\varepsilon$  and  $t$ .

## 9.2 Oscillatory solutions for truncated Magnus expansions

Of interest to us in this section is the case of finite truncations of the Magnus expansion, such as (7.33), (7.34) and (8.5), prior to spatial discretisation. Since the propagation via the truncated Magnus expansion

$$\tilde{u}(t) = e^{\tilde{\Theta}(t)} u_0, \quad (9.11)$$

solves a perturbed equation

$$\tilde{u}'(t) = \tilde{A}(t) u(t), \quad \tilde{u}(0) = u_0, \quad (9.12)$$

we cannot immediately say whether the solution of this equation also develops oscillations of wavelength  $\mathcal{O}(\varepsilon)$ . However, we require this property in the proof of Theorem 9.3.2.

In this section we will see that conservation of a perturbed energy over each time allows us to draw similar conclusions, justifying the application of bounds (9.1) for  $\tilde{u}$ . The proofs follow along the same lines as Section 9.1. The difference is that instead of relying on the conservation of moments of Hamiltonian, the truncated Magnus expansion will be considered as a perturbed Hamiltonian whose moments are conserved at discrete times,  $t_n$ .

### 9.2.1 Conservation of the perturbed Hamiltonian

Writing  $\tilde{\Theta}(t_{n+1}, t_n)$  for an undiscretised truncated Magnus expansions over the time interval  $[t_n, t_{n+1}]$ , we note that the time-stepping procedure

$$\tilde{u}(t_{n+1}) = e^{\tilde{\Theta}(t_{n+1}, t_n)} \tilde{u}(t_n)$$

conserves a modified energy,

$$\tilde{E}_n(t) = \langle \tilde{u}(t), \tilde{H}_n \tilde{u}(t) \rangle,$$

in the step  $t_n$  to  $t_{n+1}$ , where the *perturbed Hamiltonian* in the  $n$ th step is defined as

$$\tilde{H}_n = \tilde{H}(t_{n+1}, t_n) = i_h^\varepsilon \tilde{\Theta}(t_{n+1}, t_n).$$

This is to say

$$\tilde{E}_n(t_{n+1}) = \langle \tilde{u}(t_{n+1}), \tilde{H}_n \tilde{u}(t_{n+1}) \rangle = \langle \tilde{u}(t_n), \tilde{H}_n \tilde{u}(t_n) \rangle = \tilde{E}_n(t_n).$$

This occurs since  $\tilde{H}(t_{n+1}, t_n) = i_h^\varepsilon \tilde{\Theta}(t_{n+1}, t_n)$  commutes with the evolution operator  $e^{\tilde{\Theta}(t_{n+1}, t_n)}$ . Note that unitarity of  $e^{\tilde{\Theta}(t_{n+1}, t_n)}$  and, therefore, the skew-Hermiticity of  $\tilde{\Theta}(t_{n+1}, t_n)$  are crucial. By the same argument, we conclude that

$$\tilde{E}_n^2(t) = \langle \tilde{u}(t), \tilde{H}_n^2 \tilde{u}(t) \rangle = \langle \tilde{H}_n \tilde{u}(t), \tilde{H}_n \tilde{u}(t) \rangle,$$

is also conserved.

Once again, we explicitly note the dependence of  $\tilde{u}, \tilde{H}, \tilde{\Theta}, \tilde{E}_n$  and  $\tilde{E}_n^2$  on  $\varepsilon$  by writing  $\tilde{u}^\varepsilon, \tilde{H}^\varepsilon, \tilde{\Theta}^\varepsilon, \tilde{E}_n^\varepsilon$  and  $\tilde{E}_n^{2,\varepsilon}$ . Considering the semiclassical limit, we conclude that the solution of the perturbed equation solved by the truncated Magnus expansion  $\tilde{\Theta}$  also develops  $\mathcal{O}(\varepsilon)$  wavelengths oscillations.

We consider three cases – where the highest differential operator apart from the Laplacian,  $\partial_x^2$ , is of (i) degree one, (ii) degree two and (iii) degree greater than two. The first two cases will be applicable for (7.33) and (7.34), respectively, while the third will be applicable for (8.5). In some cases there will be restrictions on time-steps. However, these depend solely on the potential function and are independent of the semiclassical parameter,  $\varepsilon$ .

The proof for case (iii) is the most general and can also be used for cases (i) and (ii). However, we believe that it is helpful to first consider the simpler and more concrete cases (i) and (ii) before considering a general procedure. The first of the three cases is dealt with in Lemma 9.2.1.

### 9.2.2 The first non-trivial Magnus expansion

**Lemma 9.2.1.** *Let  $\tilde{u}^\varepsilon(t_n)$  be the perturbed solution obtained under propagation via  $\tilde{\Theta}_3^\varepsilon$  (7.33). In the semiclassical limit,  $\varepsilon \rightarrow 0$ , the spatial derivative of the wavefunction  $\tilde{u}^\varepsilon$  remains bounded by inverse of the semiclassical parameter,*

$$\|\partial_x \tilde{u}^\varepsilon(t_n)\|_{L^2} \leq \tilde{C}\varepsilon^{-1}, \quad n \in \mathbb{Z}^+, \quad (9.13)$$

if it is bounded at  $t_0$  (i.e. for  $u_0^\varepsilon$ ) and  $V$  satisfies (9.2)

$$\|\partial_x^k V(t)\| \leq \hat{C}_k, \quad k \in \{0, 1\}, \quad t \geq 0,$$

and (9.3),

$$\int_0^t \|\partial_t \partial_x^k V(\tau)\| \, d\tau \leq \hat{C}_k, \quad k \in \{0, 1\}, \quad t \geq 0,$$

where  $0 < \hat{C}_k < \infty$  and  $0 < \tilde{C}(t_n) < \infty$  are independent of  $\varepsilon$  and  $t$ .

*Proof.* The proof will proceed by induction on the time step. Firstly, we note that  $\tilde{u}^\varepsilon(t_0) = u_0^\varepsilon$ , which is assumed to satisfy (9.13).

Assuming that (9.13) holds for  $t_n$ , let us assume the contrary for  $t_{n+1}$ ,

$$\lim_{\varepsilon \rightarrow 0} \|\varepsilon \partial_x \tilde{u}(t_{n+1})\|_{L^2} = \infty.$$

The perturbed Hamiltonian in this case is

$$\tilde{H}_n^\varepsilon = i \frac{\varepsilon}{h} \tilde{\Theta}_3^\varepsilon(t_{n+1}, t_n) = -\varepsilon^2 \partial_x^2 + \frac{1}{h} \mu_{0,0}(t_n, t_{n+1}) + 2i \frac{1}{h} \langle \partial_x \mu_{1,1}(t_n, t_{n+1}) \rangle_1,$$

where

$$\mu_{0,0}(t_n, t_n + h) = \int_0^h V(t_n + \zeta) \, d\zeta$$

and

$$\mu_{1,1}(t_n, t_n + h) = \int_0^h \left(\zeta - \frac{h}{2}\right) V(t_n + \zeta) \, d\zeta.$$

The perturbed energy  $\tilde{E}_n^\varepsilon$  at  $t$  is

$$\begin{aligned} \tilde{E}_n^\varepsilon(t) &= \langle \tilde{u}^\varepsilon(t), \tilde{H}_n^\varepsilon \tilde{u}^\varepsilon(t) \rangle \\ &= \|\varepsilon \tilde{u}^\varepsilon(t)\|_{L^2}^2 + \frac{1}{h} \langle \tilde{u}^\varepsilon(t), \mu_{0,0}(t_n, t_{n+1}) \tilde{u}^\varepsilon(t) \rangle + 2i \frac{1}{h} \langle \tilde{u}^\varepsilon(t), \langle \partial_x \mu_{1,1}(t_n, t_{n+1}) \rangle_1 \tilde{u}^\varepsilon(t) \rangle. \end{aligned}$$

Although it might be hard to say much about the inner products appearing here, it suffices

to bound  $\tilde{E}^\varepsilon(t)$  by noting  $\|\tilde{u}^\varepsilon(t)\|_{L^2} = 1$  and

$$\begin{aligned} & \langle \tilde{u}^\varepsilon(t), 2 \langle \partial_x \mu_{1,1}(t_n, t_{n+1}) \rangle_1 \tilde{u}^\varepsilon(t) \rangle \\ &= \langle \tilde{u}^\varepsilon(t), (\partial_x \circ (\partial_x \mu_{1,1}(t_n, t_{n+1})) + (\partial_x \mu_{1,1}(t_n, t_{n+1})) \partial_x) \tilde{u}^\varepsilon(t) \rangle \\ &= - \langle \partial_x \tilde{u}^\varepsilon(t), (\partial_x \mu_{1,1}(t_n, t_{n+1})) \tilde{u}^\varepsilon(t) \rangle + \langle (\partial_x \mu_{1,1}(t_n, t_{n+1})) \tilde{u}^\varepsilon(t), \partial_x \tilde{u}^\varepsilon(t) \rangle \\ &= 2i \operatorname{Im} \langle \partial_x \tilde{u}^\varepsilon(t), (\partial_x \mu_{1,1}(t_n, t_{n+1})) \tilde{u}^\varepsilon(t) \rangle, \end{aligned}$$

which brings us to

$$\begin{aligned} & \|\varepsilon \partial_x \tilde{u}^\varepsilon(t_n)\|_{L^2}^2 - \frac{1}{h} \|\mu_{0,0}(t_n, t_{n+1})\|_{L^2} - \frac{2}{h} \|\partial_x \mu_{1,1}(t_n, t_{n+1})\|_{L^2} \|\varepsilon \partial_x \tilde{u}^\varepsilon(t_n)\|_{L^2} \leq \tilde{E}_n^\varepsilon(t_n) \\ & \leq \|\varepsilon \partial_x \tilde{u}^\varepsilon(t_n)\|_{L^2}^2 + \frac{1}{h} \|\mu_{0,0}(t_n, t_{n+1})\|_{L^2} + \frac{2}{h} \|\partial_x \mu_{1,1}(t_n, t_{n+1})\|_{L^2} \|\varepsilon \partial_x \tilde{u}^\varepsilon(t_n)\|_{L^2}. \end{aligned}$$

Clearly, assuming that  $\tilde{u}^\varepsilon(t_n)$  satisfies (9.13) and  $\partial_x \mu_{1,1}(t_n, t_{n+1})$  is bounded independently of  $\varepsilon$  immediately implies that  $\tilde{E}_n^\varepsilon(t_n)$  stays bounded under the limit  $\varepsilon \rightarrow 0$ .

Working similarly, we conclude that the perturbed energy  $\tilde{E}_n$  at time  $t_{n+1}$  is bounded from below by

$$\tilde{E}_n^\varepsilon(t_{n+1}) \geq \|\varepsilon \partial_x \tilde{u}^\varepsilon(t_{n+1})\|_{L^2}^2 - \frac{1}{h} \|\mu_{0,0}(t_n, t_{n+1})\|_{L^2} - \frac{2}{h} \|\partial_x \mu_{1,1}(t_n, t_{n+1})\|_{L^2} \|\varepsilon \partial_x \tilde{u}^\varepsilon(t_{n+1})\|_{L^2},$$

where the only difference in the bound is in the appearance of the perturbed wave function at time  $t_{n+1}$ . If we assume

$$\lim_{\varepsilon \rightarrow 0} \|\varepsilon \partial_x \tilde{u}(t_{n+1})\|_{L^2} = \infty,$$

this immediately brings us to the conclusion

$$\lim_{\varepsilon \rightarrow 0} \tilde{E}_n^\varepsilon(t_{n+1}) \geq +\infty$$

since the term  $\|\partial_x \mu_{1,1}(t_n, t_{n+1})\|_{L^2} \|\varepsilon \partial_x \tilde{u}^\varepsilon(t_{n+1})\|_{L^2}$  grows at a lower rate than  $\|\varepsilon \partial_x \tilde{u}(t_{n+1})\|_{L^2}^2$ .

However, by conservation of  $\tilde{E}_n^\varepsilon$  over the time step from  $t_n$  to  $t_{n+1}$ ,  $\tilde{E}_n^\varepsilon(t_{n+1}) = \tilde{E}_n^\varepsilon(t_n)$  and  $\tilde{E}_n^\varepsilon(t_{n+1})$  must also stay bounded in the semiclassical limit, contradicting our conclusion. Thus  $\|\varepsilon \partial_x \tilde{u}(t_{n+1})\|_{L^2}$  must stay bounded.

To show that the bound can be derived independently of  $t$ , we recall that  $\tilde{E}_{n-1}(t_n) =$

$\tilde{E}_{n-1}(t_{n-1})$  due to conservation of the perturbed energy over one time step. Thus,

$$\begin{aligned}
 \tilde{E}_n^\varepsilon(t_n) - \tilde{E}_{n-1}^\varepsilon(t_{n-1}) &= \tilde{E}_n^\varepsilon(t_n) - \tilde{E}_{n-1}^\varepsilon(t_n) \\
 &= \langle \tilde{u}^\varepsilon(t_n), \frac{1}{h}(\mu_{0,0}(t_n, t_{n+1}) - \mu_{0,0}(t_{n-1}, t_n)) \tilde{u}^\varepsilon(t_n) \rangle \\
 &\quad + 2i\varepsilon \frac{1}{h} \langle \tilde{u}^\varepsilon(t_n), \langle \partial_x(\mu_{1,1}(t_n, t_{n+1}) - \mu_{1,1}(t_{n-1}, t_n)) \rangle_1 \tilde{u}^\varepsilon(t_n) \rangle \\
 &\leq \frac{1}{h} \|\mu_{0,0}(t_n, t_{n+1}) - \mu_{0,0}(t_{n-1}, t_n)\| \\
 &\quad + \frac{2}{h} \|\partial_x(\mu_{1,1}(t_n, t_{n+1}) - \mu_{1,1}(t_{n-1}, t_n))\| \|\varepsilon \partial_x \tilde{u}^\varepsilon(t_n)\|_{L^2}.
 \end{aligned}$$

Telescoping this series, assuming that  $\tilde{C}$  in (9.13) is independent of  $t_k$  for  $k = 0, \dots, n-1$ , and recalling that derivatives of  $V$  have uniformly bounded variations in  $t$  due to (9.3),

$$\begin{aligned}
 \tilde{E}_n^\varepsilon(t_n) - \tilde{E}_0^\varepsilon(t_0) &\leq \sum_{k=1}^n \frac{1}{h} \|\mu_{0,0}(t_k, t_{k+1}) - \mu_{0,0}(t_{k-1}, t_k)\| \\
 &\quad + \frac{2}{h} \sum_{k=1}^n \|\partial_x(\mu_{1,1}(t_k, t_{k+1}) - \mu_{1,1}(t_{k-1}, t_k))\| \|\varepsilon \partial_x \tilde{u}^\varepsilon(t_k)\|_{L^2} \\
 &\leq \int_0^{t_n} \|\partial_t V(\tau)\| \, d\tau + 2\tilde{C} \int_0^{t_n} \|\partial_t \partial_x V(\tau)\| \, d\tau \\
 &\quad + \frac{2}{h} \|\partial_x(\mu_{1,1}(t_n, t_{n+1}) - \mu_{1,1}(t_{n-1}, t_n))\| \|\varepsilon \partial_x \tilde{u}^\varepsilon(t_n)\|_{L^2} \\
 &\leq \overline{C}_1 + \overline{C}_2 \|\varepsilon \partial_x \tilde{u}^\varepsilon(t_n)\|_{L^2},
 \end{aligned}$$

Using the above inequality, we can show that a bound independent of  $t_n$  can be found along the lines of Lemma 9.1.3.  $\square$

**Note:** The above bound only requires  $\tilde{u}(t_n), V(t) \in C_p^1([-1, 1])$ .

**Note:** It is crucial for the above proof that the  $\|\varepsilon \partial_x \tilde{u}(t_{n+1})\|_{L^2}^2$  term grows faster than other terms. This is guaranteed for the first of the three cases, where the highest differential operator apart from the Laplacian is of degree one. Corollary 9.2.2 deals with the case where there is another degree two differential operator in the perturbed Hamiltonian.

### 9.2.3 The second non-trivial Magnus expansion

**Corollary 9.2.2.** *Let  $\tilde{u}^\varepsilon(t_n)$  be the perturbed solution obtained under propagation via  $\tilde{\Theta}_4^\varepsilon$  (7.34), with an added assumption on the time-step to ensure*

$$\|\partial_x^2 \mu_{2,1}(t_n, t_{n+1})\| < h/2, \quad n \in \mathbb{Z}_0^+.$$

In the semiclassical limit,  $\varepsilon \rightarrow 0$ , the spatial derivative of the wavefunction  $\tilde{u}^\varepsilon$  remains bounded by inverse powers of the semiclassical parameter,

$$\|\partial_x \tilde{u}^\varepsilon(t_n)\|_{L^2} \leq \tilde{C} \varepsilon^{-1}, \quad n \in \mathbb{Z}^+, \quad (9.14)$$

if it is bounded at  $t_0$  (i.e. for  $u_0^\varepsilon$ ) and  $V$  satisfies (9.2),

$$\left\| \partial_x^k V(t) \right\| \leq \hat{C}_k, \quad k \in \{0, 1, 2, 3\}, \quad t \geq 0,$$

and (9.3),

$$\int_0^t \left\| \partial_t \partial_x^k V(\tau) \right\| d\tau \leq \hat{C}_k, \quad k \in \{0, 1, 2, 3\}, \quad t \geq 0,$$

where  $0 < \hat{C}_k < \infty$  and  $0 < \tilde{C} < \infty$  are independent of  $\varepsilon$  and  $t$ .

*Proof.* For the case of  $\tilde{\Theta}_4^\varepsilon(h)$ , we have higher order differential operators such as

$$-\frac{2}{h} \varepsilon^2 \langle \partial_x^2 \mu_{2,1}(t_n, t_{n+1}) \rangle_2$$

appearing in the perturbed Hamiltonian. In this case, letting  $f = \partial_x^2 \mu_{2,1}(t_n, t_{n+1})$  for convenience,

$$\begin{aligned} & |\varepsilon^2 \langle \tilde{u}^\varepsilon(t), 2 \langle f \rangle_2 \tilde{u}^\varepsilon(t) \rangle| \\ &= |\varepsilon^2 \langle \tilde{u}^\varepsilon(t), (\partial_x^2 \circ f + f \partial_x^2) \tilde{u}^\varepsilon(t) \rangle| \\ &= |-\langle \varepsilon \partial_x \tilde{u}^\varepsilon(t), \varepsilon \partial_x (f \tilde{u}^\varepsilon(t)) \rangle - \langle \varepsilon \partial_x (f \tilde{u}^\varepsilon(t)), \varepsilon \partial_x \tilde{u}^\varepsilon(t) \rangle| \\ &\leq 2 \|\varepsilon \partial_x \tilde{u}^\varepsilon(t)\|_{L^2} \|\varepsilon \partial_x (f \tilde{u}^\varepsilon(t))\|_{L^2} \\ &\leq 2 \|\varepsilon \partial_x \tilde{u}^\varepsilon(t)\|_{L^2} (\varepsilon \|\partial_x f\|_{L^2} + \|f\|_{L^2} \|\varepsilon \partial_x \tilde{u}^\varepsilon(t)\|_{L^2}). \end{aligned}$$

In this case our assumption  $\lim_{\varepsilon \rightarrow 0} \|\varepsilon \partial_x \tilde{u}(t_{n+1})\|_{L^2} = \infty$  does not directly allow us to conclude that  $\lim_{\varepsilon \rightarrow 0} \tilde{E}_n^\varepsilon(t_{n+1}) \geq +\infty$  since the above term, in general, grows as fast as the leading term. However, if the time step is chosen to be small enough so that

$$\frac{2}{h} \|f\| = \frac{2}{h} \|\partial_x^2 \mu_{2,1}(t_n, t_{n+1})\| < 1,$$

then the leading term dominates once again. Note carefully that the time step is still not constrained by  $\varepsilon$  – instead it is constrained by the nature of the potential. The reader will recall that  $\mu_{2,1}(t_n, t_{n+1})$  is  $\mathcal{O}(h^4)$ , so that the restriction imposed here is not severe.  $\square$

**Note:** The above bound requires  $\tilde{u}(t_n) \in C_p^1([-1, 1])$  and  $V(t) \in C_p^3([-1, 1])$ .

### 9.2.4 Higher order Magnus expansions

A slightly different approach needs to be followed in the case of higher order Magnus expansions such as  $\tilde{\Theta}_5^\varepsilon(h)$ , developed in Section 8.4. These feature terms such as  $i\frac{4}{3}\varepsilon^3\frac{1}{h}\langle\partial_x^3\mu_{1,3}(t_n, t_{n+1})\rangle_3$  in the perturbed Hamiltonian which are higher degree differential operators than the Laplacian. The strategy pursued here will be along the lines of Lemma 9.1.3.

A general truncated Magnus expansion with  $\mathcal{O}(\varepsilon^{(2p+3)\sigma-1})$  accuracy can be written in the form

$$\tilde{\Theta}_{2p+1}^\varepsilon = \sum_{k=0}^{2p-1} i^{k+1} \varepsilon^{k-1} \langle f_k \rangle_k, \quad (9.15)$$

where our focus is now on the degree of the differential operator and not the size of each term in powers of  $\varepsilon$ . Consequently,  $f_k$  might combine terms of all sizes in terms of  $h$ , as well as higher (positive) powers of  $\varepsilon$ . Thus, although  $f_k$ s might depend on  $\varepsilon$ , they still satisfy (9.2).

**Lemma 9.2.3.** *Let  $\tilde{u}^\varepsilon(t_n)$  be the perturbed solution obtained under propagation via  $\tilde{\Theta}_{2p+1}^\varepsilon$  for  $p \geq 1$ . In the semiclassical limit,  $\varepsilon \rightarrow 0$ , the spatial derivatives of the wavefunction  $\tilde{u}^\varepsilon$  remain bounded by inverse powers of the semiclassical parameter,*

$$\left\| \partial_x^k \tilde{u}^\varepsilon(t_n) \right\|_{L^2} \leq \tilde{C}_k \varepsilon^{-k}, \quad k = 0, \dots, 2p-1, \quad n \in \mathbb{Z}^+, \quad (9.16)$$

if they are bounded at  $t_0$  (i.e. for  $u_0^\varepsilon$ ) and  $V$  satisfies (9.2),

$$\left\| \partial_x^k V(t) \right\| \leq \hat{C}_k, \quad k \in \mathbb{Z}^+, \quad t \geq 0,$$

and (9.3),

$$\int_0^t \left\| \partial_t \partial_x^k V(\tau) \right\| d\tau \leq \hat{C}_k, \quad k \in \mathbb{Z}^+, \quad t \geq 0,$$

where  $0 < \hat{C}_k < \infty$  and  $0 < \tilde{C}_k < \infty$  are independent of  $\varepsilon$  and  $t$ .

*Proof.* The initial condition  $\tilde{u}^\varepsilon(t_0) = u_0^\varepsilon$  is assumed to satisfy (9.16). We now assume that (9.16) is satisfied by  $\tilde{u}^\varepsilon(t_n)$  but not by  $\tilde{u}^\varepsilon(t_{n+1})$ ,

$$\lim_{\varepsilon \rightarrow 0} \left\| \varepsilon^k \partial_x^k \tilde{u}^\varepsilon(t_{n+1}) \right\|_{L^2} = \infty.$$

Let  $k_0$  be the derivative for which this limit goes to infinity the fastest.

Consider the conservation of

$$\tilde{E}_n^{2,\varepsilon}(t) = \left\langle \tilde{u}^\varepsilon(t), \left( \tilde{H}_n^\varepsilon \right)^2 \tilde{u}^\varepsilon(t) \right\rangle = \left\langle \tilde{H}_n^\varepsilon \tilde{u}^\varepsilon(t), \tilde{H}_n^\varepsilon \tilde{u}^\varepsilon(t) \right\rangle$$

when evolving from  $t_n$  to  $t_{n+1}$  under  $\tilde{\Theta}_{2p+1}^\varepsilon(t_n, t_{n+1})$ . The perturbed Hamiltonian is of



the form

$$\tilde{H}_n^\varepsilon = \sum_{k=0}^{2p-1} i^k \varepsilon^k \langle f_k \rangle_k,$$

so that

$$\begin{aligned} \tilde{E}_n^{2,\varepsilon}(t) &= \sum_{k=0}^{2p-1} \left\langle \varepsilon^k \langle f_k \rangle_k \tilde{u}^\varepsilon(t), \varepsilon^k \langle f_k \rangle_k \tilde{u}^\varepsilon(t) \right\rangle + \sum_{\substack{j,k=0 \\ j \neq k}}^{2p-1} i^{j+k} \left\langle \varepsilon^k \langle f_k \rangle_k \tilde{u}^\varepsilon(t), \varepsilon^j \langle f_j \rangle_j \tilde{u}^\varepsilon(t) \right\rangle \\ &= \sum_{k=0}^{2p-1} \left\| \varepsilon^k \langle f_k \rangle_k \tilde{u}^\varepsilon(t) \right\|_{L^2}^2 + \sum_{\substack{j,k=0 \\ j \neq k}}^{2p-1} i^{j+k} \left\langle \varepsilon^k \langle f_k \rangle_k \tilde{u}^\varepsilon(t), \varepsilon^j \langle f_j \rangle_j \tilde{u}^\varepsilon(t) \right\rangle. \end{aligned}$$

The bounds we need are

$$\begin{aligned} \sum_{k=0}^{2p-1} \left\| \varepsilon^k \langle f_k \rangle_k \tilde{u}^\varepsilon(t) \right\|_{L^2}^2 - \sum_{\substack{j,k=0 \\ j \neq k}}^{2p-1} \left\| \varepsilon^k \langle f_k \rangle_k \tilde{u}^\varepsilon(t) \right\|_{L^2} \left\| \varepsilon^j \langle f_j \rangle_j \tilde{u}^\varepsilon(t) \right\|_{L^2} &\leq \tilde{E}_n^{2,\varepsilon}(t) \\ &\leq \sum_{k=0}^{2p-1} \left\| \varepsilon^k \langle f_k \rangle_k \tilde{u}^\varepsilon(t) \right\|_{L^2}^2 + \sum_{\substack{j,k=0 \\ j \neq k}}^{2p-1} \left\| \varepsilon^k \langle f_k \rangle_k \tilde{u}^\varepsilon(t) \right\|_{L^2} \left\| \varepsilon^j \langle f_j \rangle_j \tilde{u}^\varepsilon(t) \right\|_{L^2}. \end{aligned}$$

Our assumptions regarding  $\tilde{u}^\varepsilon(t_n)$  directly prove the boundedness of  $\tilde{E}_n^{2,\varepsilon}(t_n)$  under  $\varepsilon \rightarrow 0$ , and thus also of  $\tilde{E}_n^{2,\varepsilon}(t_{n+1})$ . Here we use the fact that the  $\left\| \varepsilon^k \langle f_k \rangle_k \tilde{u}^\varepsilon(t) \right\|_{L^2}$  also grows at the same asymptotic rate as  $\left\| \varepsilon^k \partial_x^k \tilde{u}^\varepsilon(t) \right\|_{L^2}$  due to (9.8) and (9.2).

While the inner product terms in the second sum are indefinite and lead to bounds that might otherwise become meaningless, the crucial observation in the case of  $\tilde{u}^\varepsilon(t_{n+1})$  is that these terms are always dominated by the  $k_0$  term in the first sum. Thus  $\tilde{u}^\varepsilon(t_{n+1})$  not satisfying (9.16) necessarily means  $\tilde{E}_n^{2,\varepsilon}(t_{n+1}) \geq +\infty$ , bringing us to a contradiction.  $\square$

**Note:** Using the strategy pursued in Lemma 9.1.3, we can prove (9.16) for arbitrarily high  $k \in \mathbb{Z}^+$  by considering conservation of higher moments of  $\tilde{H}_n^\varepsilon$ . Moreover, the bounds can be shown to be independent of  $t$  by following along the lines of Lemma 9.2.1.

### 9.3 Error Analysis for Magnus Expansions

As noted in Section 3.4.1, the Magnus expansion for the Schrödinger equation with time-dependent potential,  $\Theta(h)$ , need not converge since the Hamiltonian is an unbounded operator (note that all results regarding the convergence of the Magnus expansion are restricted to the case of bounded operators). Nevertheless, Hochbruck & Lubich (2003) note that this does not restrict us from designing or analysing methods that are formally

based on the Magnus expansion but in practice work with finite truncations and spatially discretised versions of the expansion.

The error bound obtained by Hochbruck & Lubich (2003) for standard Magnus expansions (under  $\varepsilon = 1$ ) suggests that the fourth order Magnus expansion results in an error,

$$\|u^n - u(t_n)\|_2 \leq Ch^5 t_n \max_{0 \leq t \leq t_n} \|\partial_x^4 u(t)\|_2,$$

which was noted as a cause for concern by the authors due to the presence of  $\partial_x^4 u(t)$  and the typically oscillatory nature of the wavefunction. When this oscillatory behaviour arises from the semiclassical scaling, where  $\hbar \ll \varepsilon \ll 1$ , the bounds might seem to be fairly pessimistic since  $\|\partial_x^4 u\| \leq C\varepsilon^{-4}$ . Moreover, these bounds are derived under the time step scaling  $h \|\mathcal{K}\| \leq c$  for some constant  $c$ , which could be restrictive<sup>1</sup>.

In the semiclassical regime, where the oscillatory behaviour of the wave function is primarily attributable to small values of  $\varepsilon$ , however, we can improve upon these bounds by correctly accounting for powers of  $\varepsilon$ . Additionally, we are able to bypass restrictions of the type  $h \|\mathcal{K}\| \leq c$ . This is achieved by using the results of Section 9.2, which allow us to conclude that the wavefunction propagated under the truncated Magnus expansion also obeys the bounds (9.1).

The resulting error bounds for the discretised version of  $\tilde{\Theta}_\varepsilon^\varepsilon$  are summarised in Theorem 9.3.6,

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \|u^\varepsilon(t_n) - \mathbf{u}^\varepsilon(t_n)_I\|_{L^2} \leq Ct_n h^6, \quad n \in \mathbb{Z}^+,$$

where  $C > 0$  is independent of  $\varepsilon$ ,  $h$  and  $t_n$  but depends on the potential. With the choice of  $h = \mathcal{O}(\varepsilon^\sigma)$  we arrive upon the error bound

$$\|u^\varepsilon(t_n) - \mathbf{u}^\varepsilon(t_n)_I\|_{L^2} = \mathcal{O}(t_n \varepsilon^{6\sigma-1}),$$

which is more in line with the notation used in the previous chapters of this thesis.

**Note:** The error analysis carried out here also applies to standard truncated Magnus expansions (featuring commutators) in the semiclassical regime and generalises to higher dimensions. However, this aspect will not be explored in this thesis.

### 9.3.1 Proof outline

Recall that the Schrödinger equation featuring time-dependent potential,

$$\partial_t u(x, t) = \mathbf{i} (\varepsilon \partial_x^2 - \varepsilon^{-1} V(x, t)) u(x, t), \quad x \in [-1, 1], \quad t \geq 0, \quad (3.20)$$

---

<sup>1</sup>The authors note that the restriction  $h \|\mathcal{K}\| \leq c$  can be bypassed, but in doing so the bounds end up featuring progressively higher derivatives of the wavefunction.

is also written in the form

$$\partial_t u(t) = \mathcal{A}(t) u(t), \quad u(0) = u_0, \quad (3.22)$$

where  $\mathcal{A}(t) = i\varepsilon \partial_x^2 - i\varepsilon^{-1} V(t)$ . The exact solution of this equation will be denoted as  $u(t)$ , as always.

Formally, the solution of this equation can be written via the exponential of the Magnus expansion,

$$u(t) = e^{\Theta(t,0)} u_0.$$

However, as noted earlier, in the case of unbounded operators such as the Hamiltonian (which features the Laplacian), the Magnus expansion as a series may not necessarily converge for any  $t$ . Nevertheless, an approximate solution can be found by time-stepping using the truncated Magnus expansion  $\tilde{\Theta}$ ,

$$\tilde{u}(t_{n+1}) = e^{\tilde{\Theta}(t_{n+1}, t_n)} \tilde{u}(t_n), \quad n \in \mathbb{Z}^+, \quad \tilde{u}(t_0) = u_0. \quad (9.17)$$

In practice, the numerical solution is found by exponentiating the truncated and discretised Magnus expansion,  $\Theta$ ,

$$\mathbf{u}(t_{n+1}) = e^{\Theta(t_{n+1}, t_n)} \mathbf{u}(t_n), \quad n \in \mathbb{Z}^+, \quad \mathbf{u}(t_0) = \mathbf{u}_0, \quad (9.18)$$

where  $\mathbf{u}_0$  is obtained via trigonometric interpolation of the initial conditions  $u_0$ .

The error is analysed in two steps: in the first step we analyse the error due to the truncation of the Magnus, and in the second step we analyse the error due to discretisation of the truncated Magnus expansion, whereby the error in the concrete (fully discretised) schemes is obtained as

$$\|u(t_n) - \mathbf{u}_I(t_n)\|_{L^2} \leq \|u(t_n) - \tilde{u}(t_n)\|_{L^2} + \|\tilde{u}(t_n) - \mathbf{u}_I(t_n)\|_{L^2}, \quad (9.19)$$

where, following the notation of (Bao et al. 2002),  $\mathbf{u}_I$  will stand for the trigonometric interpolant of the data  $\mathbf{u}$  on the grid, i.e interpolating  $\{(x_j, \mathbf{u}_j)\}_{j=-N}^N$ . Thus  $\|\mathbf{u}_I\|_{L^2} = \|\mathbf{u}\|_{\ell^2}$ . Additionally, we use  $f_I$  to denote the trigonometric interpolant of the function  $f$  on our grid  $\{x_{-N}, \dots, x_N\}$ ,  $x_j = 2j/M$ ,  $M = 2N + 1$ .

For the first part, i.e. the error due to truncation of the Magnus expansion, we follow the approach of Hochbruck & Lubich (2003). Propagation via a truncated Magnus expansion is equivalent to the exact solution of a perturbed equation. The error is analysed by studying this perturbation. The design of the truncated Magnus expansions ensures that the error terms are composed of (i) high order nested integrals of nested commutators and (ii) a remainder term. Bounds for the nested integrals of nested commutators follows straightaway from Corollary 9.1.6. For the remainder term we rely on results of

(Hochbruck & Lubich 2003) and the observations of Section 9.2.

The second part of proof, dealing with errors due to discretisation, follows along the lines of (Bao et al. 2002) and (Pasciak 1980), once we consider the Magnus expansion as being frozen over a time interval. The error can then be analysed as stemming solely from the discretisation. In particular, this analysis gives us a handle on the degree of spatial discretisation required for the required accuracy.

### 9.3.2 Error bounds for the truncated Magnus expansion

The perturbed solution  $\tilde{u}$  in (9.17) can be expressed as,

$$\tilde{u}(t) = e^{\tilde{\Theta}(t, t_n)} \tilde{u}(t_n), \quad t \in [t_n, t_{n+1}], \quad (9.20)$$

where the purpose of time stepping is to keep truncation errors low. This, in fact, is how the truncated Magnus expansion is derived: in each time step we find the formal Magnus expansion at an arbitrary  $t$  and then truncate this series, following which we decide to propagate the solution with steps of size  $h = t_{n+1} - t_n$ .

By differentiating (9.20), we find that the perturbed solution satisfies a modified equation

$$\tilde{u}'(t) = \tilde{A}(t) \tilde{u}(t), \quad \tilde{A} = \text{dexp}_{\tilde{\Theta}(t)} \left( \tilde{\Theta}'(t) \right), \quad (9.21)$$

with initial value  $\tilde{u}(t_n)$ , where we have written  $\tilde{\Theta}(t)$  as a shorthand for  $\tilde{\Theta}(t, t_n)$ . In this section we will try to quantify the error,

$$e(t) = \tilde{u}(t) - u(t), \quad (9.22)$$

for the truncated Magnus expansion  $\tilde{\Theta}$ . The following lemma, essentially restating the results of Lemma 4.1 in (Hochbruck & Lubich 2003), allows us to study the error in terms of the perturbation.

**Lemma 9.3.1.** *The error in the time step  $[t_n, t_{n+1}]$  is bounded by*

$$\|e(t_{n+1})\|_{L^2} \leq \|e(t_n)\|_{L^2} + \int_{t_n}^{t_{n+1}} \left\| \left( \tilde{A}(t) - A(t) \right) u(t) \right\|_{L^2} dt. \quad (9.23)$$

*Proof.* Differentiation the error (9.22) and writing  $\tilde{u}(t) = u(t) + e(t)$ ,

$$\begin{aligned} e'(t) &= \tilde{u}'(t) - u'(t) \\ &= \tilde{A}(t) \tilde{u}(t) - A(t) u(t) \\ &= \tilde{A}(t) u(t) + \tilde{A}(t) e(t) - A(t) u(t). \end{aligned}$$

Following the observation in Lemma 4.1 in (Hochbruck & Lubich 2003), we note that

$$\begin{aligned}
 \frac{1}{2} \frac{d}{dt} \|e(t)\|_{L^2([-1,1], \mathbb{C})}^2 &= \frac{1}{2} \langle e', e \rangle_{L^2([-1,1], \mathbb{C})} + \frac{1}{2} \langle e, e' \rangle_{L^2([-1,1], \mathbb{C})} \\
 &= \operatorname{Re} \langle e', e \rangle_{L^2([-1,1], \mathbb{C})} \\
 &= \operatorname{Re} \langle \tilde{A}u - Au + \tilde{A}e, e \rangle \\
 &= \operatorname{Re} \langle \tilde{A}u - Au, e \rangle \\
 &\leq \left\| \tilde{A}u - Au \right\|_{L^2} \|e\|_{L^2}.
 \end{aligned}$$

Here we have used the fact that  $\tilde{A}$  is skew-Hermitian in the case of the Schrödinger equation. This means  $\langle \tilde{A}e, e \rangle = -\langle e, \tilde{A}e \rangle = -\overline{\langle \tilde{A}e, e \rangle}$  and, consequently,  $\langle \tilde{A}e, e \rangle$  is purely imaginary. Noting that  $\frac{d}{dt} \|e\|^2 = 2 \|e\| \frac{d}{dt} \|e\|$ , this inequality immediately gives us

$$\frac{d}{dt} \|e(t)\| \leq \left\| \tilde{A}(t)u(t) - A(t)u(t) \right\|_{L^2}.$$

Integrating this inequality over  $[t_n, t_{n+1}]$  proves the lemma.  $\square$

To estimate the integral in Lemma 9.3.1, we need to study how much

$$\tilde{A}(t) = \operatorname{dexp}_{\tilde{\Theta}(t)} \left( \tilde{\Theta}'(t) \right)$$

deviates from  $A(t)$ . Following the strategy of (Hochbruck & Lubich 2003), we write the  $\operatorname{dexp}$  function as

$$\operatorname{dexp}_{\tilde{\Theta}}(\tilde{\Theta}') = \tilde{\Theta}' + \frac{1}{2}[\tilde{\Theta}, \tilde{\Theta}'] + \dots + \frac{1}{(2p+1)!} \operatorname{ad}_{\tilde{\Theta}}^{2p}(\tilde{\Theta}') + \frac{1}{(2p+2)!} r_{2p+2}(\operatorname{ad}_{\tilde{\Theta}}) \left( \operatorname{ad}_{\tilde{\Theta}}^{2p+1}(\tilde{\Theta}') \right),$$

where  $\tilde{\Theta} = \tilde{\Theta}_{2p+1}^\varepsilon(t, t_n)$  and the remainder  $r_p$  is defined by the series,

$$\frac{e^z - 1}{z} = 1 + \frac{1}{2}z + \dots + \frac{1}{(p-1)!} z^{p-2} + \frac{1}{p!} z^{p-1} r_p(z).$$

Thus, to estimate the integral in (9.23), we must analyse the action of

$$\hat{\mathcal{E}}(t) = \tilde{\Theta}' - A(t) + \frac{1}{2}[\tilde{\Theta}, \tilde{\Theta}'] + \dots + \frac{1}{(2p+1)!} \operatorname{ad}_{\tilde{\Theta}}^{2p}(\tilde{\Theta}') + \frac{1}{(2p+2)!} r_{2p+2}(\operatorname{ad}_{\tilde{\Theta}}) \left( \operatorname{ad}_{\tilde{\Theta}}^{2p+1}(\tilde{\Theta}') \right)$$

on  $u(t)$ . The remainder term is unlike the other terms appearing here, which can be explicitly written down in terms of commutators and bounded by Corollary 9.1.6. After discussing a procedure by which the bounds for the remainder term can also be translated to the common form,  $\left\| \operatorname{ad}_{\tilde{\Theta}}^k(\tilde{\Theta}')v \right\|$ , we can proceed to derive bounds for the specific Magnus expansions  $\tilde{\Theta}_3$  and  $\tilde{\Theta}_5$ .

### The remainder term

From Lemma 5.1 of (Hochbruck & Lubich 2003), we know that the remainder term  $r_p$  has a Fourier transform  $\widehat{r}_p \in L^1(\mathbb{R})$ ,

$$r_p(ix) = \int_{\mathbb{R}} \widehat{r}_p(\xi) e^{i\xi x} d\xi.$$

Consequently, for skew-Hermitian operator  $X$  (where  $r_p(\text{ad}_{\tilde{\Theta}})(X)$  also ends up being skew-Hermitian),

$$\begin{aligned} r_p(\text{ad}_{\tilde{\Theta}})(X) &= \int_{\mathbb{R}} \widehat{r}_p e^{\text{ad}_{\xi\tilde{\Theta}}}(X) d\xi = \int_{\mathbb{R}} \widehat{r}_p \text{Ad}_{\xi\tilde{\Theta}}(X) d\xi \\ &= \int_{\mathbb{R}} \widehat{r}_p e^{\xi\tilde{\Theta}} X e^{-\xi\tilde{\Theta}} d\xi, \end{aligned}$$

from which it follows that

$$\|r_p(\text{ad}_{\tilde{\Theta}})(X)u\|_{L^2} \leq \|\widehat{r}_p\|_{L^1(\mathbb{R})} \sup_{\xi \in \mathbb{R}} \|e^{\xi\tilde{\Theta}} X e^{-\xi\tilde{\Theta}} u\|_{L^2}.$$

Noting that the outermost exponential can be absorbed due to its unitarity, we arrive at the following bound for  $X = \text{ad}_{\tilde{\Theta}}^{2p+1}(B)$ ,

$$\left\| r_{2p+2}(\text{ad}_{\tilde{\Theta}}) \left( \text{ad}_{\tilde{\Theta}}^{2p+1}(B) \right) u(t) \right\|_{L^2} \leq \|\widehat{r}_{2p+2}\|_{L^1(\mathbb{R})} \sup_{\xi \in \mathbb{R}} \left\| \text{ad}_{\tilde{\Theta}}^{2p+1}(B) \exp(-\xi\tilde{\Theta}) u(t) \right\|_{L^2}, \quad (9.24)$$

where  $B = \tilde{\Theta}'_{2p+1}(t, t_n)$ .

We first note that the exact solution,  $u(t)$ , is assumed to satisfy (9.1), i.e. its spatial derivatives are bounded by inverse powers of  $\varepsilon$ ,

$$\lim_{\varepsilon \rightarrow 0} \left\| \varepsilon^k \partial_x^k u^\varepsilon(t) \right\|_{L^2} \leq C_k < \infty, \quad k \in \mathbb{Z}^+.$$

Using  $u(t)$  as the initial condition and following the procedure of Lemma 9.2.3, we can show that

$$w^\varepsilon = \exp(-\xi\tilde{\Theta}(t, t_n)) u(t)$$

also satisfies the oscillatory assumptions

$$\lim_{\varepsilon \rightarrow 0} \left\| \varepsilon^k \partial_x^k w^\varepsilon \right\|_{L^2} \leq C_k < \infty, \quad k \in \mathbb{Z}^+. \quad (9.1)$$

This is since  $\tilde{\Theta}(t, t_n)$  commutes not only with the evolution operator  $\exp(\tilde{\Theta}(t, t_n))$ , but also with  $\exp(-\xi\tilde{\Theta}(t, t_n))$  for any  $\xi \in \mathbb{R}$ , and thus the perturbed Hamiltonian and its moments are conserved. Therefore, the arguments used in Lemma 9.2.3 can be employed here to conclude that  $w$  satisfies (9.1).

### Error bounds for $\tilde{\Theta}_3$

For this part we closely follow the approach in Section 6 of (Hochbruck & Lubich 2003). For the sake of simplicity we consider the time shifted to  $t_n$ , so that our first non-trivial Magnus expansion (7.33) is

$$\tilde{\Theta}_3(h) = i\varepsilon\partial_x^2 - i\varepsilon^{-1}\mu_{0,0}(h) - 2\langle\partial_x\mu_{1,1}(h)\rangle_1, \quad (7.33)$$

as usual. Recalling (2.51), used in the derivation of the Magnus expansion, we note that the truncated Magnus expansion is obtained by setting

$$\tilde{\Theta}'_3(t) = A(t) - \frac{1}{2} \int_0^t [A(\zeta), A(t)] d\zeta,$$

integrating which brings us to

$$\tilde{\Theta}_3(t) = \int_0^t A(\zeta) d\zeta - \frac{1}{2} \int_0^t \int_0^\xi [A(\zeta), A(\xi)] d\zeta d\xi.$$

In our case we used  $t = h$  for time stepping and  $A(t) = i\varepsilon\partial_x^2 - i\varepsilon^{-1}V(t)$  before reducing this expansion to (7.33). However, we will disregard this simplification for now and directly use the (algebraic) results of (Hochbruck & Lubich 2003) to conclude that

$$\begin{aligned} \hat{\mathcal{E}}(t) &= \tilde{\Theta}' - A(t) + \frac{1}{2}[\tilde{\Theta}, \tilde{\Theta}'] + \frac{1}{6}[\tilde{\Theta}[\tilde{\Theta}, \tilde{\Theta}']] + \frac{1}{24}r_4(\text{ad}_{\tilde{\Theta}})(\text{ad}_{\tilde{\Theta}}^3(\tilde{\Theta}')) \\ &= -\frac{1}{12} \int_0^t \int_0^t \int_0^t [A(\zeta), [A(\xi), [A(\chi), A(t)]]] d\chi d\xi d\zeta \\ &\quad - \frac{1}{12} \int_0^t \int_0^t \int_0^\xi [A(\zeta), [A(\chi), A(\xi)], A(t)] d\chi d\xi d\zeta \\ &\quad - \frac{1}{24} \int_0^t \int_0^\zeta \int_0^t [[A(\chi), A(\zeta)], [A(\xi), A(t)]] d\xi d\chi d\zeta \\ &\quad + \mathcal{R}(t) + \frac{1}{24}r_4(\text{ad}_{\tilde{\Theta}})(\text{ad}_{\tilde{\Theta}}^3(\tilde{\Theta}')). \end{aligned}$$

The precise form of  $\mathcal{R}(t)$  is not important, except for noting that it is composed of terms with four integrals (five commutators) and five integrals (six commutators) of  $A$ . Nor, for that matter, does the precise form of the integrals appearing here or the form of  $\text{ad}_{\tilde{\Theta}}^3(\tilde{\Theta}')$  matter, except for noting that they all have three or more integrals.

Due to the observation made in Section 2.3.3, all these terms are  $\mathcal{O}(t^4)$  (not  $\mathcal{O}(t^3)$ ) or smaller. Corollary (9.1.6), further tells us that all these commutators satisfy bounds of the form

$$\lim_{\varepsilon \rightarrow 0} \|\varepsilon [[A(\chi), A(\zeta)], [A(\xi), A(t)]] u^\varepsilon(t)\|_{L^2} \leq \tilde{C} < \infty,$$

for instance, and thus the integrals satisfy

$$\lim_{\varepsilon \rightarrow 0} \left\| \varepsilon \int_0^t \int_0^\zeta \int_0^t [[A(\chi), A(\zeta)], [A(\xi), A(t)]] d\xi d\chi d\zeta u^\varepsilon(t) \right\|_{L^2} \leq \tilde{C}t^4.$$

Similarly,

$$\lim_{\varepsilon \rightarrow 0} \left\| \varepsilon \operatorname{ad}_{\tilde{\Theta}}^3(\tilde{\Theta}') u^\varepsilon(t) \right\|_{L^2} \leq \hat{C}t^4.$$

Using (9.24) and the fact that  $w^\varepsilon = \exp(-\xi\tilde{\Theta}(t, t_n))u^\varepsilon(t)$  also manifests the highly oscillatory behaviour (9.1), we conclude that

$$\lim_{\varepsilon \rightarrow 0} \left\| \varepsilon r_4(\operatorname{ad}_{\tilde{\Theta}}) \left( \operatorname{ad}_{\tilde{\Theta}}^3(\tilde{\Theta}') \right) u^\varepsilon(t) \right\|_{L^2} \leq \overline{C}t^4.$$

It follows immediately that

$$\lim_{\varepsilon \rightarrow 0} \left\| \varepsilon \hat{\mathcal{E}}(t) u^\varepsilon(t) \right\|_{L^2} \leq Ct^4,$$

for some  $0 < C < \infty$ , integrating which bring us to the conclusion

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \|e^\varepsilon(t_{n+1})\|_{L^2} \leq \lim_{\varepsilon \rightarrow 0} \varepsilon \|e^\varepsilon(t_n)\|_{L^2} + Ch^5. \quad (9.25)$$

Telescoping this sum provides the error bounds for the truncated but undiscretised Magnus expansion, summarised in Theorem 9.3.2.

**Theorem 9.3.2.** *If  $\tilde{u}^\varepsilon(t_n)$  is the solution found by time stepping the undiscretised Magnus expansion  $\tilde{\Theta}_3^\varepsilon(h)$ , the exact solution  $u^\varepsilon(t)$  satisfies (9.1) and the potential satisfies (9.2) and (9.3), then*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \|u^\varepsilon(t_n) - \tilde{u}^\varepsilon(t_n)\|_{L^2} \leq Ct_n h^4, \quad n \in \mathbb{Z}^+,$$

for some  $C > 0$  which is independent of  $\varepsilon$ ,  $h$  and  $t_n$  but depends on the potential function.

### A comment on higher order Magnus expansions

We note that the difference of the operators  $\hat{\mathcal{E}}(t) = \tilde{A}(t) - A(t)$  in the previous section, where we carried out the analysis for  $\tilde{\Theta}_3$ , is expressed in terms of a remainder term and terms with three (and more) nested integrals of nested commutators. This should be no surprise—the very derivation of the truncated Magnus expansions  $\tilde{\Theta}_3$  is designed to cancel terms with two and fewer integrals.

In the case of a Magnus expansion such as  $\tilde{\Theta}_5$ , the operator difference  $\hat{\mathcal{E}}(t)$  can be written as a sum of  $r_6(\operatorname{ad}_{\tilde{\Theta}}) \left( \operatorname{ad}_{\tilde{\Theta}}^5(\tilde{\Theta}') \right)$  and nested integrals which are  $\mathcal{O}(t^6)$  or smaller. The latter are bounded in a straightforward manner due to Corollary 9.1.6, while the former is bounded, once again, by using (9.24) and the fact that  $w^\varepsilon = \exp(-\xi\tilde{\Theta}(t, t_n))u^\varepsilon(t)$  obeys (9.1). Thus, we find error bounds for  $\tilde{\Theta}_5^\varepsilon(h)$  along, essentially, the same lines as



before.

**Theorem 9.3.3.** *If  $\tilde{u}^\varepsilon(t_n)$  is the solution found by time stepping the undiscretised Magnus expansion  $\tilde{\Theta}_5^\varepsilon(h)$ , the exact solution  $u^\varepsilon(t)$  satisfies (9.1) and the potential satisfies (9.2) and (9.3), then*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \|u^\varepsilon(t_n) - \tilde{u}^\varepsilon(t_n)\|_{L^2} \leq C t_n h^6, \quad n \in \mathbb{Z}^+,$$

for some  $C > 0$  which is independent of  $\varepsilon$ ,  $h$  and  $t_n$  but depends on the potential function.

In deriving these bounds we have also laid out the procedure for deriving error bounds for a general high order Magnus expansion.

### 9.3.3 Error in discretisation

Recall that the time stepping of the truncated Magnus expansion is the procedure

$$\tilde{u}(t_{n+1}) = e^{\tilde{\Theta}(t_{n+1}, t_n)} \tilde{u}(t_n), \quad n \in \mathbb{Z}^+, \quad \tilde{u}(t_0) = u_0. \quad (9.17)$$

However, as noted before, the numerical solution is found by exponentiating the discretised version of  $\tilde{\Theta}$ ,

$$\mathbf{u}(t_{n+1}) = e^{\Theta(t_{n+1}, t_n)} \mathbf{u}(t_n), \quad n \in \mathbb{Z}^+, \quad \mathbf{u}(t_0) = \mathbf{u}_0. \quad (9.26)$$

The error due to discretisation,

$$\epsilon(t_n) = \tilde{u}(t_{n+1}) - \mathbf{u}(t_{n+1})_I,$$

is bounded as

$$\begin{aligned} \|\epsilon(t_{n+1})\|_{L^2} &= \|\tilde{u}(t_{n+1}) - \mathbf{u}(t_{n+1})_I\|_{L^2} \\ &\leq \left\| e^{\tilde{\Theta}(t_{n+1}, t_n)} \tilde{u}(t_n) - e^{\tilde{\Theta}(t_{n+1}, t_n)} \mathbf{u}(t_n)_I \right\|_{L^2} \\ &\quad + \left\| e^{\tilde{\Theta}(t_{n+1}, t_n)} \mathbf{u}(t_n)_I - \left( e^{\Theta(t_{n+1}, t_n)} \mathbf{u}(t_n) \right)_I \right\|_{L^2} \\ &= \|\epsilon(t_n)\|_{L^2} + \left\| e^{\tilde{\Theta}(t_{n+1}, t_n)} \mathbf{u}(t_n)_I - \left( e^{\Theta(t_{n+1}, t_n)} \mathbf{u}(t_n) \right)_I \right\|_{L^2}, \end{aligned}$$

where we have exploited the unitarity of  $\tilde{\Theta}$ . For the second term, note that  $e^{\tilde{\Theta}(t_{n+1}, t_n)} \mathbf{u}(t_n)_I$  is the solution of

$$y'(t) = h^{-1} \tilde{\Theta}(t_{n+1}, t_n) y(t), \quad t \in [t_n, t_{n+1}], \quad y(t_n) = \mathbf{u}(t_n)_I \quad (9.27)$$

at  $t = t_{n+1}$ , while  $e^{\Theta(t_{n+1}, t_n)} \mathbf{u}(t_n)$  is the solution of

$$\mathbf{y}'(t) = h^{-1} \Theta(t_{n+1}, t_n) \mathbf{y}(t), \quad t \in [t_n, t_{n+1}], \quad \mathbf{y}(t_n) = \mathbf{u}(t_n) \quad (9.28)$$

at  $t = t_{n+1}$ . Note here that  $\tilde{\Theta}(t_{n+1}, t_n)$  and  $\Theta(t_{n+1}, t_n)$  are treated in a time-independent way (i.e. frozen over the interval). Thus the only error here is due to the discretisation of  $\tilde{\Theta}$ .

By following the analysis of (Pasciak 1980), the difference in solution<sup>2</sup> of (9.27) and (9.28) at  $t_{n+1}$  can be bounded as

$$\|y(t_{n+1}) - \mathbf{y}(t_{n+1})_I\|_{L^2} \leq Ch^{-1} \sup_{t \in [t_n, t_{n+1}]} \left\| \tilde{\Theta}(y(t) - \mathbf{y}(t)_I) \right\|_{L^2}.$$

Since  $\tilde{\Theta}_{2p+1}$  can be expressed in terms of  $h \sum_{j=0}^{2p-1} \varepsilon^{j-1} \langle f_j \rangle_j$ s, where  $f_j = \mathcal{O}(1)$  or smaller in terms of  $h$  and  $\varepsilon$ , we conclude following the analysis of (Pasciak 1980),

$$\begin{aligned} \|y(t_{n+1}) - \mathbf{y}(t_{n+1})_I\|_{L^2} &\leq C \sup_{t \in [t_n, t_{n+1}]} \sum_{j=0}^{2p-1} \varepsilon^{j-1} \left\| \langle f_j \rangle_j (y(t) - \mathbf{y}(t)_I) \right\|_{L^2} \\ &\leq \tilde{C}_s \sup_{t \in [t_n, t_{n+1}]} \sum_{j=0}^{2p-1} \varepsilon^{j-1} M^{-s} \left\| \partial_x^{s+j} y(t_n) \right\|_{L^2}, \quad s \in \mathbb{Z}^+, \end{aligned}$$

for an arbitrary  $s$ , where we have used the fact that  $f_j$ s obey (9.2). Consider the scaling

$$M = \mathcal{O}(h^{-\gamma} \varepsilon^{-1}), \quad 0 < \gamma \leq 1,$$

whereby

$$\|y(t_{n+1}) - \mathbf{y}(t_{n+1})_I\|_{L^2} \leq \tilde{C}_s h^{\gamma s} \varepsilon^{-1} \sup_{t \in [t_n, t_{n+1}]} \sum_{j=0}^{2p-1} \left\| \varepsilon^{s+j} \partial_x^{s+j} y(t_n) \right\|_{L^2}, \quad s \in \mathbb{Z}^+.$$

Note that  $y(t_n)$  obeys (9.1), so that

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \|y(t_{n+1}) - \mathbf{y}(t_{n+1})_I\|_{L^2} \leq \hat{C}_s h^{\gamma s}.$$

We now choose  $s = \gamma^{-1}(2p+3)$ , whereby

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \|\epsilon(t_{n+1})\|_{L^2} \leq \lim_{\varepsilon \rightarrow 0} \varepsilon \|\epsilon(t_n)\|_{L^2} + \hat{C}_{\gamma^{-1}(2p+3)} h^{2p+3}.$$

Telescoping this (while noting that  $\epsilon(t_0)$  is also  $\mathcal{O}(h^{2p+3} \varepsilon^{-1})$  due to the accuracy of spectral discretisation) brings us to the following theorem.

**Theorem 9.3.4.** *If  $\tilde{u}(0)$  obeys the bounds (9.1), the potential function obeys (9.2) and (9.3), and  $M = \mathcal{O}(h^{-\gamma} \varepsilon^{-1})$  for some  $0 < \gamma \leq 1$ , then the error due to spatial discretisation*

---

<sup>2</sup>Note that this is zero at  $t_n$  by definition.

of  $\Theta_{2p+3}$  is

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \|\tilde{u}(t_n) - \mathbf{u}(t_n)_I\|_{L^2} \leq Ct_n h^{2p+2}, \quad (9.29)$$

for some  $C > 0$  which is independent of  $\varepsilon$ ,  $h$  and  $t_n$  but depends on the potential function.

**Note:** We can take  $\gamma$  to be arbitrarily small, so that  $M$  can be fairly close to  $\mathcal{O}(\varepsilon^{-1})$ . For instance, with  $h = \mathcal{O}(\varepsilon^{1/2})$  and  $\gamma = 1/2$ , the gridding strategy  $M = \mathcal{O}(\varepsilon^{-1-1/4})$  suffices.

**Note:** Despite the fact that  $M$  is not strictly  $\mathcal{O}(\varepsilon^{-1})$ , our error analysis carries through since it is largely carried out without discretisation considerations and since, due to Lemma 9.1.4,  $\langle f \rangle_k = \mathcal{O}(\varepsilon^{-k})$  continues to hold regardless. However, this does affect the exact cost of FFTs, the impact of which will not be considered here fully.

### 9.3.4 A summary of the error bounds

Combining the results of Theorem 9.3.2 and Theorem 9.3.3 with Theorem 9.3.4, we get the error bounds for our truncated and discretised Magnus expansions.

**Theorem 9.3.5.** *If  $\mathbf{u}^\varepsilon(t_n)$  is the numerical solution found by time stepping the discretised Magnus expansion  $\Theta_3^\varepsilon$ , the exact solution  $u^\varepsilon(t)$  satisfies (9.1), the potential function satisfies (9.2) and (9.3), and the degree of discretisation is chosen so that  $M = \mathcal{O}(h^{-\gamma}\varepsilon^{-1})$  for some  $0 < \gamma \leq 1$ , then*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \|u^\varepsilon(t_n) - \mathbf{u}^\varepsilon(t_n)_I\|_{L^2} \leq Ct_n h^4, \quad n \in \mathbb{Z}^+,$$

for some  $C > 0$  which is independent of  $\varepsilon$ ,  $h$  and  $t_n$  but depends on the potential function and the choice of  $\gamma$ .

**Theorem 9.3.6.** *If  $\mathbf{u}^\varepsilon(t_n)$  is the numerical solution found by time stepping the discretised Magnus expansion  $\Theta_5^\varepsilon$ , the exact solution  $u^\varepsilon(t)$  satisfies (9.1), the potential function satisfies (9.2) and (9.3), and the degree of discretisation is chosen so that  $M = \mathcal{O}(h^{-\gamma}\varepsilon^{-1})$  for some  $0 < \gamma \leq 1$ , then*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \|u^\varepsilon(t_n) - \mathbf{u}^\varepsilon(t_n)_I\|_{L^2} \leq Ct_n h^6, \quad n \in \mathbb{Z}^+,$$

for some  $C > 0$  which is independent of  $\varepsilon$ ,  $h$  and  $t_n$  but depends on the potential function and the choice of  $\gamma$ .

**Note:** Since  $C$  is independent of  $t_n$ , the bounds effectively apply for any time  $T$ .

**Note:** The constraint  $h \|\mathcal{K}\| \leq c$  in (Hochbruck & Lubich 2003) arises in the case of higher order truncations of the Magnus expansion while finding an estimate for

the (discretised) term corresponding to

$$\left\| \partial_x^k \exp(-\xi \tilde{\Theta}) u^\varepsilon(t) \right\|_{L^2}.$$

Here, however, we use the fact that, due to Lemma 9.2.3,  $w^\varepsilon = \exp(-\xi \tilde{\Theta}) u^\varepsilon(t)$  essentially features the same degree of oscillations (9.1) as  $u^\varepsilon(t)$ .

**Note:** The time step  $h$  is taken to be independent of  $\varepsilon$  for the entire analysis. It is natural to contrast this with the methods presented in this thesis where  $h$  is assumed to scale as  $\mathcal{O}(\varepsilon^\sigma)$  for some  $0 < \sigma \leq 1$ . This scaling is used for two reasons:

- (i) It allows us to analyse all terms and errors solely in powers of  $\varepsilon$ , making the narrative easier. This is also helpful in deciding the degree of discretisation,  $M$ .
- (ii) When we combine the Magnus expansion with Zassenhaus splittings or with Lanczos iterations, for that matter, the number of Lanczos iterations depend on the spectral size of the exponent. This analysis becomes easier when carried out in a single *currency*. It points out that Lanczos iterations are not the most optimal choice in the semiclassical regime. Moreover, it imposes the time step restriction  $\sigma > 1/3$  on Magnus–Zassenhaus schemes. This restriction is solely due to cost considerations arising from exponentiation via the Zassenhaus splitting.

## 9.4 Error Analysis for Zassenhaus Splittings

The error analysis for Zassenhaus splittings proceeds by using essentially the same tools that have been employed in the previous sections of this chapter. Our main tool will be to link the error in truncating the sBCH series in each step of the symmetric Zassenhaus splitting procedure to the error analysis for the truncated Magnus expansions.

Moan & Niesen (2008) note that the Baker–Campbell–Hausdorff (BCH) formula can be obtained via the Magnus expansion,

$$\text{BCH}(X, Y) = \Theta_{\hat{A}}(2),$$

where  $\hat{A}(t) = X$  for  $t \in [0, 1]$  and  $\hat{A}(t) = Y$  for  $t \in (1, 2]$ , and  $\Theta_{\hat{A}}(2)$  is the formal Magnus expansion at  $t = 2$  under the time-varying vector field  $\hat{A}(t)$ . By a similar logic,

$$\text{sBCH}(X, Y) = \Theta_A(2),$$

where  $A(t) = X$  for  $t \in [0, \frac{1}{2}] \cup [1\frac{1}{2}, 2]$  and  $A(t) = Y$  for  $t \in (\frac{1}{2}, 1\frac{1}{2})$ . This holds since, by definition, the exponential of the sBCH,

$$e^{\text{sBCH}(X,Y)}u = e^{\frac{1}{2}X}e^Y e^{\frac{1}{2}X}u,$$

involves evolving  $u$  under the influence of  $X$  till  $t = \frac{1}{2}$ , then under  $Y$  till  $t = 1\frac{1}{2}$ , and under  $X$  again till  $t = 2$ . This is also the exact effect of propagating under the formal Magnus expansion of  $A$  till  $t = 2$  since, by definition, it involves evolving under  $A(t) = X$  from  $t = 0$  to  $t = \frac{1}{2}$ , followed by evolution under  $A(t) = Y$  from  $t = \frac{1}{2}$  to  $t = 1\frac{1}{2}$ , and then again under  $A(t) = X$  from  $t = 1\frac{1}{2}$  till  $t = 2$ .

The symmetric Zassenhaus procedure involves recursive applications of sBCH in order to extract the largest terms from the exponent. Thus the overall error can be analysed in two parts: (i) the error due to successive applications of suitably truncated sBCH and (ii) the error due to discretisation. The latter of these can be analysed on exactly the same lines as Section 9.3.3 and will not be discussed in detail.

In the following sections we proceed to analyse the error due to the truncation of the sBCH in successive steps of the Zassenhaus procedure. For the sake of simplicity, we will carry out the analysis for the first non-trivial symmetric Zassenhaus splitting,

$$Z_{1,1}^{[2]} = e^{\frac{1}{2}ih\varepsilon\partial_x^2}e^{-\frac{1}{2}ih\varepsilon^{-1}V}e^{\frac{1}{6}ih^3\varepsilon^{-1}(\partial_x V)^2 + \frac{1}{6}ih^3\varepsilon\langle\partial_x^2 V\rangle_2}e^{-\frac{1}{2}ih\varepsilon^{-1}V}e^{\frac{1}{2}ih\varepsilon\partial_x^2}. \quad (9.30)$$

Extension of this analysis to the splittings such as (6.6) should prove to be relatively straightforward.

#### 9.4.1 Error in the first step of the symmetric Zassenhaus algorithm

Let us consider the first step in the splitting of the exponent  $-it\varepsilon^{-1}H = it\varepsilon\partial_x^2 - it\varepsilon^{-1}V$  and analyse the error arising from the truncation of the sBCH.

**Lemma 9.4.1.** *Define  $X = -ih\varepsilon\partial_x^2$  and  $Y = ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V$ . Assuming that  $u^\varepsilon$  satisfies (9.1) and  $V$  satisfies (9.2), the local error due to the first application of the truncated sBCH in the symmetric Zassenhaus procedure is bounded as follows*

$$\left\| e^Y u^\varepsilon - e^{-\frac{1}{2}X} e^{X+Y - (\frac{1}{24}[Y,X],X] + \frac{1}{12}[Y,X],Y]} e^{-\frac{1}{2}X} u^\varepsilon \right\|_{L^2} \leq Ch^5\varepsilon^{-1}, \quad n \in \mathbb{Z}^+,$$

in the limit  $\varepsilon \rightarrow 0$ , where  $C > 0$  is independent of  $\varepsilon$  and  $h$ .

*Proof.* Consider the evolution of  $u^\varepsilon$  under  $ih\varepsilon\partial_x^2 - ih\varepsilon V$ ,

$$\begin{aligned} e^{ih\varepsilon\partial_x^2 - ih\varepsilon V} u^\varepsilon &= e^{\frac{1}{2}ih\varepsilon\partial_x^2} e^{\text{sBCH}(-ih\varepsilon\partial_x^2, ih\varepsilon\partial_x^2 - ih\varepsilon V)} e^{\frac{1}{2}ih\varepsilon\partial_x^2} u^\varepsilon \\ &= e^{\frac{1}{2}ih\varepsilon\partial_x^2} e^{\Theta_A(2)} e^{\frac{1}{2}ih\varepsilon\partial_x^2} u^\varepsilon, \end{aligned}$$

where we have taken  $X = -ih\varepsilon\partial_x^2$  and  $Y = ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V$  while defining  $A$ . Here the Magnus expansion  $\Theta_A(2)$  and the sBCH expansion are used in a formal sense since they might not converge as a series.

The error in truncating the sBCH series to  $\widetilde{\text{sBCH}}$  by excluding all  $h^5$  or smaller terms, for instance, is equivalent to the error in truncating  $\Theta_A$  to  $\tilde{\Theta}_A$  by excluding all  $h^5$  terms,

$$\begin{aligned} & \left\| e^{ih\varepsilon\partial_x^2 - ih\varepsilon V} u^\varepsilon - e^{\frac{1}{2}ih\varepsilon\partial_x^2} e^{\widetilde{\text{sBCH}}(-it\varepsilon\partial_x^2, it\varepsilon\partial_x^2 - ih\varepsilon V)} e^{\frac{1}{2}ih\varepsilon\partial_x^2} u^\varepsilon \right\|_{L^2} \\ &= \left\| e^{\frac{1}{2}ih\varepsilon\partial_x^2} e^{\Theta_A(2)} e^{\frac{1}{2}ih\varepsilon\partial_x^2} u^\varepsilon - e^{\frac{1}{2}ih\varepsilon\partial_x^2} e^{\tilde{\Theta}_A(2)} e^{\frac{1}{2}ih\varepsilon\partial_x^2} u^\varepsilon \right\|_{L^2} \\ &= \left\| e^{\Theta_A(2)} w^\varepsilon - e^{\tilde{\Theta}_A(2)} w^\varepsilon \right\|_{L^2}, \end{aligned}$$

where

$$w^\varepsilon = e^{\frac{1}{2}ih\varepsilon\partial_x^2} u^\varepsilon.$$

By the conservation of  $\varepsilon^2\partial_x^2$  in the above step and the fact that  $u^\varepsilon$  obeys (9.1), we can show that  $w^\varepsilon$  also obeys the bounds (9.1) by following the approach of Lemma 9.1.3.

The error analysis then follows along the lines of Section 9.3.2 since the formal notation  $e^{\Theta_A(2)} w^\varepsilon$  amounts to the evolution of  $w^\varepsilon$  under  $A$  till  $t = 2$ . Thus, the error in this local step can be characterised by Lemma 9.3.1,

$$\|e(2)\|_{L^2} \leq \int_0^2 \left\| \left( \tilde{A}(t) - A(t) \right) u(t) \right\|_{L^2} dt,$$

where error is zero at  $t = 0$  (in the context of this local step).

Recall that, by definition,

$$\Theta'_3(t) = A(t) - \frac{1}{2} \int_0^t [A(\zeta), A(t)] d\zeta.$$

Thus

$$\begin{aligned} \hat{\mathcal{E}}(t) &= \tilde{\Theta}' - A(t) + \frac{1}{2}[\tilde{\Theta}, \tilde{\Theta}'] + \frac{1}{6}[\tilde{\Theta}[\tilde{\Theta}, \tilde{\Theta}']] + \frac{1}{24}r_4(\text{ad}_{\tilde{\Theta}}) \left( \text{ad}_{\tilde{\Theta}}^3(\tilde{\Theta}') \right) \\ &= -\frac{1}{12} \int_0^t \int_0^t \int_0^t [A(\zeta), [A(\xi), [A(\chi), A(t)]]] d\chi d\xi d\zeta \\ &\quad -\frac{1}{12} \int_0^t \int_0^t \int_0^\xi [A(\zeta), [A(\chi), A(\xi)], A(t)] d\chi d\xi d\zeta \\ &\quad -\frac{1}{24} \int_0^t \int_0^\zeta \int_0^t [[A(\chi), A(\zeta)], [A(\xi), A(t)]] d\xi d\chi d\zeta \\ &\quad + \mathcal{R}(t) + \frac{1}{24}r_4(\text{ad}_{\tilde{\Theta}}) \left( \text{ad}_{\tilde{\Theta}}^3(\tilde{\Theta}') \right) \end{aligned}$$

still holds since it is derived algebraically for any  $A$ . There is a minor difference in the analysis—on the one hand,  $t$  goes from 0 to 2, not 0 to  $h$ , and is  $\mathcal{O}(h^0)$ ; on the other

hand,  $A(t) = \mathcal{O}(h)$ . Thus all terms in  $\hat{\mathcal{E}}$  seem to be  $\mathcal{O}(h^4)$  or smaller. However, by analysing the Taylor series of the inner most commutator,  $[A(\xi), A(\zeta)]$ , along the lines of Section 2.3.3, we find that there is a gain of one power of  $h$ . Thus, all terms in  $\hat{\mathcal{E}}$  are, in fact,  $\mathcal{O}(h^5)$  or smaller.

Once again, following the application of Corollary (9.1.6), and the observations of Section 9.3.2, we conclude

$$\lim_{\varepsilon \rightarrow 0} \left\| \varepsilon \hat{\mathcal{E}}(t) w^\varepsilon \right\|_{L^2} \leq Ch^5, \quad t \in [0, 2],$$

where  $0 < C < \infty$ . Integrating this bring us to the result of this Lemma<sup>3</sup>.  $\square$

**Note:** At many stages in the symmetric Zassenhaus splitting procedure, having truncated the sBCH by the grade of commutators (and, therefore, by powers of  $h$ ), we also discard some terms that are too small in powers of  $\varepsilon$ .

The analysis of this can be carried out by following the initial observations of Section 9.3.3—considering as reference the truncated Magnus expansion  $\tilde{\Theta}_A(2)$  frozen over  $[0, 2]$  while comparing it to the Magnus expansion which is obtained by discarding the said terms from  $\tilde{\Theta}_A(2)$ .

The difference here is that both Magnus expansions are undiscretised. However, the error simply amounts to the difference between the two expansions, i.e. the discarded terms, whose action on the wave function is easily bounded by Lemma 9.1.4.

**Note:** By following an analysis similar to Section 9.2, we can prove that the solution under the undiscretised Zassenhaus splitting (in fact, also at each partial stage such as propagation under  $ih\varepsilon\partial_x^2$ ) continues to obey the assumptions (9.1) if the initial condition does. This will not be proven here.

Thus, using the fact that the simplifications in the algebra  $\mathfrak{G}$  are exact, and the results of Lemma 9.4.1, we conclude,

$$\left\| \left( e^{ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V} - e^{\frac{1}{2}ih\varepsilon\partial_x^2} e^{-ih\varepsilon^{-1}V + \frac{1}{6}ih^3\varepsilon^{-1}(\partial_x V)^2 + \frac{1}{6}ih^3\varepsilon\langle\partial_x^2 V\rangle_2} e^{\frac{1}{2}ih\varepsilon\partial_x^2} \right) u^\varepsilon \right\|_{L^2} \leq Ch^5\varepsilon^{-1},$$

under  $h = \mathcal{O}(\varepsilon^\sigma)$ ,  $0 < \sigma \leq 1$ . Here we have also taken the liberty of discarding the term  $-\frac{1}{24}ih^3\varepsilon(\partial_x^4 V)$  from the central exponent since it is deemed to be too small. It is trivial to see that this term scales as  $\mathcal{O}(h^3\varepsilon)$ , which, strictly speaking, isn't smaller than the error when solely analysing in  $h$ . However, when  $\varepsilon \rightarrow 0$  independent of  $h$ , this term vanishes in the limit. Even if we assume that  $h$  scales with  $\varepsilon$ , but is at most as small as  $\varepsilon$  (under  $\sigma = 1$ ) (which is not really much of a restriction) this term is  $\mathcal{O}(h^5\varepsilon^{-1})$  or smaller.

<sup>3</sup>The difference in this proof is that  $\left\| \varepsilon \hat{\mathcal{E}}(t) w^\varepsilon \right\|_{L^2}$  is already bounded in terms of  $h$ , features the fifth power, and is independent of  $t$ . However, this is only a technicality that appears due to the way in which we have scaled  $t$  from  $[0, 2]$  instead of, say,  $[0, 2h]$ .

### 9.4.2 Error in the second step of the symmetric Zassenhaus algorithm

**Lemma 9.4.2.** *Define  $X = ih\varepsilon^{-1}V$  and  $Y = -ih\varepsilon^{-1}V + \frac{1}{6}ih^3\varepsilon^{-1}(\partial_x V)^2 + \frac{1}{6}ih^3\varepsilon \langle \partial_x^2 V \rangle_2$ , i.e. the central exponent in the previous step. Assuming that  $u^\varepsilon$  satisfies (9.1) and  $V$  satisfies (9.2), the local error in the central exponent due to the second application of the truncated sBCH in the symmetric Zassenhaus procedure is bounded as follows*

$$\left\| e^Y u^\varepsilon - e^{-\frac{1}{2}X} e^{X+Y} e^{-\frac{1}{2}X} u^\varepsilon \right\|_{L^2} \leq Ch^5 \varepsilon^{-1}, \quad n \in \mathbb{Z}^+,$$

in the limit  $\varepsilon \rightarrow 0$ , where  $C > 0$  is independent of  $\varepsilon$  and  $h$ .

*Proof.* Once again, we find

$$\begin{aligned} & \left\| e^Y u^\varepsilon - e^{\frac{1}{2}ih\varepsilon^{-1}V} \widetilde{e^{\text{sBCH}(X,Y)}} e^{\frac{1}{2}ih\varepsilon^{-1}V} u^\varepsilon \right\|_{L^2} \\ &= \left\| e^{\frac{1}{2}ih\varepsilon^{-1}V} e^{\Theta_A(2)} e^{\frac{1}{2}ih\varepsilon^{-1}V} u^\varepsilon - e^{\frac{1}{2}ih\varepsilon^{-1}V} e^{\tilde{\Theta}_A(2)} e^{\frac{1}{2}ih\varepsilon^{-1}V} u^\varepsilon \right\|_{L^2} \\ &= \left\| e^{\Theta_A(2)} w^\varepsilon - e^{\tilde{\Theta}_A(2)} w^\varepsilon \right\|_{L^2}, \end{aligned}$$

where

$$w^\varepsilon = e^{\frac{1}{2}ih\varepsilon^{-1}V} u^\varepsilon.$$

Since  $V$  satisfies (9.2) and  $u^\varepsilon$  satisfies (9.1), simple application of the chain rule shows that  $w^\varepsilon$  also satisfies (9.1).

In this step we truncate the sBCH series immediately—dropping all grade three or higher commutators. This corresponds to the Magnus expansion  $\Theta_1$  where

$$\Theta'_1(t) = A(t).$$

Thus the error is due to

$$\hat{\mathcal{E}}(t) = \tilde{\Theta}' - A(t) + \frac{1}{2}r_2(\text{ad}_{\tilde{\Theta}}) \left( \text{ad}_{\tilde{\Theta}}(\tilde{\Theta}') \right) = \frac{1}{2}r_2(\text{ad}_{\tilde{\Theta}}) \left( \text{ad}_{\tilde{\Theta}}(\tilde{\Theta}') \right).$$

Following the analysis in Section 9.3.2, the bound for the remainder term boils down to bounding a term of the form

$$\left\| [\tilde{\Theta}, \tilde{\Theta}'] \tilde{w}^\varepsilon \right\|_{L^2},$$

which features one integral (from  $t = 0$  to  $t = 2$ ) and one commutator of  $A$  at different times and where  $\tilde{w}^\varepsilon$  obeys (9.1).

Since  $A(t) = \mathcal{O}(h)$ , one might expect this term to be  $\mathcal{O}(h^2)$ , naïvely, and  $\mathcal{O}(h^3)$  after accounting for the gain in power via Taylor analysis. However, this term is, in fact  $\mathcal{O}(h^5)$ . This happens because the  $\mathcal{O}(h)$  term in  $A$  is  $ih\varepsilon^{-1}V$  for all times (except for a change in sign) and thus cancels out except when it is commuted with the  $\mathcal{O}(h^3)$  term.

Following the application of Corollary (9.1.6), and the observations of Section 9.3.2,



we conclude

$$\lim_{\varepsilon \rightarrow 0} \left\| \varepsilon \hat{\mathcal{E}}(t) w^\varepsilon \right\|_{L^2} \leq Ch^5, \quad t \in [0, 2],$$

where  $0 < C < \infty$ . Integrating this bring us to the result of this Lemma.  $\square$

### 9.4.3 Local error in a single step of $\mathcal{Z}_{1,1}^{[2]}$

We now combine the results of Lemma 9.4.1 and Lemma 9.4.2 to analyse the local error in a single step of the first non-trivial Zassenhaus splitting,  $\mathcal{Z}_{1,1}^{[2]}$ .

**Lemma 9.4.3.** *Assuming that  $u^\varepsilon$  satisfies (9.1) and  $V$  satisfies (9.2), the local error in a single step of the symmetric Zassenhaus splitting  $\mathcal{Z}_{1,1}^{[2]}$  is bounded as follows*

$$\left\| e^{ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V} u^\varepsilon - \mathcal{Z}_{1,1}^{[2]} u^\varepsilon \right\|_{L^2} \leq Ch^5 \varepsilon^{-1}, \quad n \in \mathbb{Z}^+,$$

in the limit  $\varepsilon \rightarrow 0$ , where  $C > 0$  is independent of  $\varepsilon$  and  $h$ .

*Proof.* Letting

$$\begin{aligned} X_1 &= -ih\varepsilon\partial_x^2, \\ X_2 &= ih\varepsilon^{-1}V, \\ Y_1 &= ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V, \\ Y_2 &= -ih\varepsilon^{-1}V + \frac{1}{6}ih^3\varepsilon^{-1}(\partial_x V)^2 + \frac{1}{6}ih^3\varepsilon \langle \partial_x^2 V \rangle_2, \end{aligned}$$

and noting that

$$\mathcal{Z}_{1,1}^{[2]} = e^{-\frac{1}{2}X_1} e^{Y_2} e^{-\frac{1}{2}X_1} - e^{-\frac{1}{2}X_1} e^{-\frac{1}{2}X_2} e^{Y_2+X_2} e^{-\frac{1}{2}X_2} e^{-\frac{1}{2}X_1},$$

the error in a single step can be written as

$$\begin{aligned} \left\| \left( e^{Y_1} - \mathcal{Z}_{1,1}^{[2]} \right) u^\varepsilon \right\|_{L^2} &\leq \left\| \left( e^{Y_1} - e^{-\frac{1}{2}X_1} e^{Y_2} e^{-\frac{1}{2}X_1} \right) u^\varepsilon \right\|_{L^2} \\ &\quad + \left\| \left( e^{-\frac{1}{2}X_1} e^{Y_2} e^{-\frac{1}{2}X_1} - e^{-\frac{1}{2}X_1} e^{-\frac{1}{2}X_2} e^{Y_2+X_2} e^{-\frac{1}{2}X_2} e^{-\frac{1}{2}X_1} \right) u^\varepsilon \right\|_{L^2}. \end{aligned}$$

Of these terms, the first is

$$\left\| \left( e^{ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V} - e^{\frac{1}{2}ih\varepsilon\partial_x^2} e^{-ih\varepsilon^{-1}V + \frac{1}{6}ih^3\varepsilon^{-1}(\partial_x V)^2 + \frac{1}{6}ih^3\varepsilon \langle \partial_x^2 V \rangle_2} e^{\frac{1}{2}ih\varepsilon\partial_x^2} \right) u^\varepsilon \right\|_{L^2},$$

which was shown to be bounded by  $Ch^5\varepsilon^{-1}$  at the end of Section 9.4.1 as a consequence of Lemma 9.4.1. For the second term, define

$$w^\varepsilon = e^{\frac{1}{2}ih\varepsilon\partial_x^2} u^\varepsilon,$$

so that

$$\begin{aligned} & \left\| \left( e^{-\frac{1}{2}X_1} e^{Y_2} e^{-\frac{1}{2}X_1} - e^{-\frac{1}{2}X_1} e^{-\frac{1}{2}X_2} e^{Y_2+X_2} e^{-\frac{1}{2}X_2} e^{-\frac{1}{2}X_1} \right) u^\varepsilon \right\|_{L^2} \\ &= \left\| \left( e^{Y_2} e^{-\frac{1}{2}X_1} - e^{-\frac{1}{2}X_2} e^{Y_2+X_2} e^{-\frac{1}{2}X_2} e^{-\frac{1}{2}X_1} \right) w^\varepsilon \right\|_{L^2}. \end{aligned}$$

Following the arguments in Section 9.4.1,  $w^\varepsilon$  also satisfies the bounds of (9.1). Thus the above term is also bounded by  $Ch^5\varepsilon^{-1}$  due to Lemma 9.4.2. Combining the two bounds results in a proof of this lemma.  $\square$

#### 9.4.4 Error bounds for the first non trivial Zassenhaus splitting

**Theorem 9.4.4.** *Let  $u^\varepsilon(t)$  be the exact solution of the Schrodinger equation (3.9) featuring a time-independent potential  $V$ , which satisfies (9.2). Assuming that the initial condition  $u^\varepsilon(0)$  satisfies (9.1), the error in the solution  $\tilde{u}^\varepsilon(t_n)$  obtained by propagating via the symmetric Zassenhaus splitting  $\mathcal{Z}_{1,1}^{[2]}$  at time  $t_n$  is bounded as*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \|u^\varepsilon(t_n) - \tilde{u}^\varepsilon(t_n)\|_{L^2} \leq Ct_n h^4, \quad n \in \mathbb{Z}^+,$$

where  $C > 0$  is independent of  $\varepsilon, h$  and  $t_n$ .

*Proof.* Denoting the error at time  $t_n$  as

$$e(t_n) = u^\varepsilon(t_n) - \tilde{u}^\varepsilon(t_n),$$

it follows that

$$\begin{aligned} \|e(t_{n+1})\|_{L^2} &= \|u^\varepsilon(t_{n+1}) - \tilde{u}^\varepsilon(t_{n+1})\|_{L^2} \\ &= \left\| e^{ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V} u^\varepsilon(t_n) - \mathcal{Z}_{1,1}^{[2]} \tilde{u}^\varepsilon(t_n) \right\|_{L^2} \\ &\leq \left\| e^{ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V} u^\varepsilon(t_n) - e^{ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V} \tilde{u}^\varepsilon(t_n) \right\|_{L^2} \\ &\quad + \left\| e^{ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V} \tilde{u}^\varepsilon(t_n) - \mathcal{Z}_{1,1}^{[2]} \tilde{u}^\varepsilon(t_n) \right\|_{L^2} \\ &= \|e(t_n)\|_{L^2} + \left\| e^{ih\varepsilon\partial_x^2 - ih\varepsilon^{-1}V} \tilde{u}^\varepsilon(t_n) - \mathcal{Z}_{1,1}^{[2]} \tilde{u}^\varepsilon(t_n) \right\|_{L^2}. \end{aligned}$$

As noted at the end of Section 9.4.1, the solution under the Zassenhaus splitting  $\tilde{u}^\varepsilon(t_n)$  can also be shown to obey the assumptions (9.1). Thus we can apply the results of Lemma 9.4.3 to conclude

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \|e(t_{n+1})\|_{L^2} \leq \lim_{\varepsilon \rightarrow 0} \varepsilon \|e(t_n)\|_{L^2} + Ch^5.$$

Telescoping the above brings us to the result of this theorem.  $\square$

Having arrived at the error bounds for the symmetric Zassenhaus splitting in the operatorial form, we point the reader to Section 9.3.3 for an analysis of the fully discrete

case. The difference from Section 9.3.3 is that now we need to break the analysis of error in multiple stages due to the splitting of the exponential. The analysis procedure, however, remains essentially the same. In particular, we must impose the same restriction on the degree of discretisation,  $M$ , in terms of  $h$  and  $\varepsilon$ .



## Chapter 10

# Conclusions and future work

### 10.1 Summary of the thesis

In this thesis we have presented a methodology for deriving arbitrarily high-order splitting methods for the computation of semiclassical Schrödinger equations (3.9) and (3.20). The results of our approach are the asymptotic exponential splittings  $\mathcal{Z}_{2,1}^{[1]}$  (6.4),  $\mathcal{Z}_{2,1}^{[2]}$  (6.6),  $\mathcal{Z}_{2,\sigma}^{[2]}$  (6.10),  $\mathcal{Z}_{2,\sigma}^{\Theta[M]}$  (7.17),  $\mathcal{Z}_{1,\sigma}^{\Theta[l]}$  (8.2) and  $\mathcal{Z}_{2,\sigma}^{\Theta[l]}$  (8.16).

These splittings have a few crucial properties in common: *(i)* they do not feature nested commutators, *(ii)* they are asymptotic—each consecutive exponent in these splittings (except perhaps for one) is progressively smaller, leading to quadratic growth of costs, *(iii)* they are stable—the exponents are all skew-Hermitian and exponentials are unitary.

These commutator-free splittings are derived by working in the Lie algebra of Jordan polynomials in  $\partial_x$  with function coefficients,

$$\langle f \rangle_k = f \bullet \partial_x^k = \frac{1}{2}(\partial_x^k \circ f + f \circ \partial_x^k).$$

In this algebra we are able to simplify commutators by using the rule

$$[\langle f \rangle_k, \langle g \rangle_l] = \sum_{n=0}^{\frac{k+l-1}{2}} \sum_{i=0}^{2n+1} \lambda_{n,i}^{k,l} \langle (\partial_x^i f)(\partial_x^{2n+1-i} g) \rangle_{k+l-2n-1}, \quad (6.2)$$

where the coefficients  $\lambda$  can be read off Table 5.2 or computationally generated via (5.11) or (5.12). The height reduction property in this algebra, summarised in Corollary 5.3.5, leads to nested commutators being much smaller than otherwise expected, making the asymptotic splittings effective.

Crucially, the stability of the numerical methods is guaranteed by the  $\mathbb{Z}_2$ -grading and the symmetrised structure of the linear differential operators that form this Lie algebra. Working in this algebra ensures skew-Hermiticity of all exponents, which leads to unitarity of exponentials and guarantees unitary evolution—a crucial physical property in quantum

mechanics.

We have also discussed the choice of semidiscretisation and effective means of approximating matrix exponentials resulting from the discretisation of the operators appearing in the splittings and the consequences upon the costs of the splittings.

The splittings  $\mathcal{Z}_{2,\sigma}^{\Theta[M]}$ ,  $\mathcal{Z}_{1,\sigma}^{\Theta[I]}$  and  $\mathcal{Z}_{2,\sigma}^{\Theta[I]}$  are high-order methods for solving semiclassical Schrödinger equations featuring time-dependent potentials. These splittings are based on commutator-free Magnus expansions  $\tilde{\Theta}_5^{\varepsilon[M]}$ ,  $\tilde{\Theta}_3^{\varepsilon[I]}$ , and  $\tilde{\Theta}_5^{\varepsilon[I]}$ , respectively, which were derived by simplifying truncated Magnus expansions in the Lie algebra  $\mathfrak{G}$ .

Even though the commutator-free Magnus expansions simplified in  $\mathfrak{G}$  are less expensive to exponentiate via Lanczos iterations (or a combination of Yoshida splittings and Lanczos iterations) than the standard Magnus expansions that feature nested commutators of matrices, they remain fairly expensive to exponentiate.

In the case of time-independent potentials moderately high-order Yoshida splittings are less expensive than Zassenhaus splittings, despite the quadratic costs of Zassenhaus splittings being superior to the exponential costs of Yoshida splittings for very high-order methods. However, for time-independent potentials, we find that the Zassenhaus–Magnus splittings are universally more efficient than the corresponding Yoshida splittings—even for low and moderately high orders of accuracy.

We have also analysed the error of these operatorial splittings and expansions without resorting to discretisation. This is achieved by characterising the oscillatory solutions of the Schrödinger equations in the semiclassical regime.

## 10.2 Zassenhaus splittings in practice

- (i) *Variants of the Zassenhaus splittings.* Variant of the Zassenhaus splittings (such as  $\mathcal{Z}_{2,1}^{[1]}$ , which extracts the term  $V$  before the  $\partial_x^2$  term) should enjoy the same asymptotic behaviour as we have seen. However, extracting subsequent terms in a different way (i.e. not simply depending on their size in powers of  $\varepsilon$ ) could potentially lead to splittings that are more accurate or cost effective in some cases.
- (ii) *Scaling choices.* In the case of moderately small values of  $\varepsilon$ , the spatio-temporal discretisation choices could matter a lot. For instance, for  $\varepsilon = 1/50$  the spatial resolution  $M \approx 50$  is hardly likely to be sufficient for discretising the potential function, which is the very least that we require, and  $h \approx 1/7$ , which is the scaling suggested under  $\sigma = 1/2$ , is likely to be completely inadequate for resolving the temporal behaviour of realistic potentials.

Clearly, in these cases we are nowhere near the asymptotic regime  $\varepsilon \rightarrow 0$  and  $M = \mathcal{O}(\varepsilon^{-1})$  is hardly a constraint to worry about. However, the property of height

reduction still applies in terms of powers of  $M$ , once we estimate

$$\langle f \rangle_k \leq C \|f\|_\infty M^k.$$

- (1) It would be worthwhile exploring the effect of different choices of  $M$  and  $h$  on the accuracy and cost of the Zassenhaus splittings.
  - (2) These choices should correctly take into account the spatio-temporal discretisation of the potential and should ideally be based on tighter estimates of the number of oscillations appearing in the wave-function.
  - (3) If  $M$  is not particularly large, the costs may no longer be dominated by the  $\mathcal{O}(M \log M)$  costs of the FFTs and we might have to take into account the evaluation of the potential function, for instance.
  - (4) If  $M \gg \varepsilon^{-1}$ , finite difference methods could once again become effective.
  - (5) It is likely that a Zassenhaus splitting where we extract terms based on their size (not just asymptotically speaking, but estimated for finite  $h$ ,  $M$  and  $\varepsilon$ , and based on the potential under consideration) would look substantially different and found highly effective in practice.
- (iii) *Yoshida–Zassenhaus hybrid schemes.* Hybrid splittings combining Zassenhaus and Yoshida splittings might benefit from the low cost of Yoshida splittings for low and moderate orders of accuracy and yet feature a quadratic growth of costs. In principle this should be possible by expressing, say, the sixth order Yoshida splitting  $\mathcal{S}_6$  as

$$\mathcal{S}_6 = \exp(A + B + h^7 C_7 + h^9 C_9 + \dots),$$

where  $C_n$  are grade  $n$  commutators, and then using sBCH to extract the  $h^7$  terms from the right hand side, writing

$$e^{-\frac{1}{2}h^7 C_7} \mathcal{S}_6 e^{-\frac{1}{2}h^7 C_7} = \exp(\text{sBCH}(-h^7 C_7, A + B + h^7 C_7 + h^9 C_9 + \dots)).$$

### 10.3 Related equations of quantum mechanics

- (i) *A multivariate setting.* An effective numerical discretisation of the semiclassical Schrödinger equation (3.9), evolving in  $d$ -dimensions requires overcoming some significant challenges. Once algebraic foundations have been properly laid down in accordance with Section 10.4.1, the case for small  $d \geq 1$  should be fairly straightforward, albeit a bit expensive.

Matters are more complicated when  $d$  becomes large and the cost of  $\mathcal{O}(M^d \log M)$  becomes unsustainable. It is clear that, for our methodology to be scaleable to large

dimensions, it must be combined with other approaches, e.g. sparse grids (Bungartz & Griebel 2004), Hagedorn wavepackets (Gradinaru & Hagedorn 2014) and quasi monte carlo lattices (Kämmerer, Kunis & Potts 2012).

- (ii) *Stochastic potentials.* Many quantum systems involve continuous measurement and environmental interference, resulting in random noise modelled by stochastic equations such as the Quantum Stochastic Schrödinger Equation (QSSE).

Magnus based methods are known to be effective for solving stochastic differential equations (Lord et al. 2008). It is worth exploring whether the methods discussed in Chapters 7 and 8 can be extended to deal with stochastic quantum equations such as the Stratonovich Schrödinger equation which features a stochastic potential with multiplicative white-noise.

The integral-preserving Magnus–Zassenhaus schemes  $\mathcal{Z}_{n,\sigma}^{\Theta[I]}$  proceed without resorting to a discretisation of integrals and makes few assumptions regarding the regularity of the potential. These will form the ideal starting point for this investigation.

- (iii) *The nonlinear Schrödinger equation.* A major challenge is to apply our methodology in a nonlinear setting, e.g. to the nonlinear Schrödinger equation

$$i\varepsilon \frac{\partial u}{\partial t} = -\frac{\varepsilon^2}{2m} \frac{\partial^2 u}{\partial x^2} - V(x)u + \lambda|u|^2u.$$

Preliminary investigation seems to indicate that a naïve generalisation does not work, because we do not enjoy the reduction of negative powers of  $\varepsilon$  after commutation with Lie-derivatives corresponding to  $|u|^2$ .

- (iv) *Pauli equation.* The Pauli equation, which has the hamiltonian

$$H = (\boldsymbol{\sigma} \cdot (-i\varepsilon \nabla - \mathbf{A}(\mathbf{x})))^2 + V(\mathbf{x}),$$

predicts the dynamics of spin-1/2 particles. Here  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  is the vector of Pauli matrices,  $\mathbf{A}$  is the magnetic potential and  $V$  is the electric potential. The Hamiltonian can be simplified to the form

$$H = (-i\varepsilon \nabla - \mathbf{A}(\mathbf{x}))^2 - \varepsilon \boldsymbol{\sigma} \cdot \mathbf{B} + V(\mathbf{x}),$$

where  $\mathbf{B} = \nabla \times \mathbf{A}$  is the magnetic field.

In principle, a Zassenhaus splitting for this equation can be derived. However, even if we were to consider the kinetic part reduced to one dimension, we encounter a term  $2i\varepsilon \langle A \rangle_1$  which is  $\mathcal{O}(\varepsilon^0)$  in  $H$  and appears as the  $\mathcal{O}(\varepsilon^{\sigma-1})$  term  $2h \langle A \rangle_1$  in  $-i\hbar H/\varepsilon$ . This term is as large as the Laplacian and the potential but due to the



lack of favourable structure it is not possible to exponentiate it cheaply unless the we resort to very small time steps.

## 10.4 Extension of the algebraic framework

### 10.4.1 Commuting elements of the Lie idealiser

In the 3-dimensional Schrödinger equation, the Hamiltonian has the form

$$H = -\varepsilon^2 \Delta + V,$$

where  $\Delta$  is the Laplacian,

$$\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2.$$

Assuming that we are working with smooth enough functions,  $\partial_x, \partial_y, \partial_z$ , and their powers commute. Moreover, all three of these are elements of the Lie idealiser,

$$[\partial_x, f] = \partial_x f, \quad [\partial_y, f] = \partial_y f, \quad [\partial_z, f] = \partial_z f.$$

An extension of the techniques discussed in this thesis to the case of multivariate Schrödinger equation will require the development of corresponding algebraic structures. Symbolic computation with these elements suggests that the symmetrised elements of the form

$$\langle f \rangle_{k,l,j} = \frac{1}{2} \left( \partial_x^k \partial_y^l \partial_z^j \circ f + f \circ \partial_x^k \partial_y^l \partial_z^j \right)$$

enjoy similar properties of height reduction and symmetry. For simplifying matter, we write

$$\langle f \rangle_{\mathbf{k}} = \langle f \rangle_{k_1, k_2, k_3},$$

where  $\mathbf{k}$  is the multi-index  $(k_1, k_2, k_3)$  which suggests obvious higher-dimensional generalisations. The conjecture is that for all  $f, g \in \mathcal{G}$ ,

$$[\langle f \rangle_{\mathbf{k}}, \langle g \rangle_{\mathbf{l}}] = \sum_{\mathbf{n}=\mathbf{0}}^{\frac{\mathbf{k}+\mathbf{l}-\mathbf{1}}{2}} \langle h_{\mathbf{n}} \rangle_{\mathbf{k}+\mathbf{l}-2\mathbf{n}-\mathbf{1}},$$

for some  $h_{\mathbf{n}} \in \mathcal{G}$  and where the sum is restricted to  $0 \leq n_i \leq \frac{k_i+l_i-1}{2}$ ,  $i = 1, 2, 3$ .

Once height reduction in multiple dimensions is established, the results of Chapter 9 should generalise to multiple dimensions in a very straightforward way. In practical applications, the simplification of commutators in higher dimensions has the potential of creating far too many terms, however. Thus the costs of Zassenhaus splittings may no longer grow quadratically. It would be of interest to see if the terms are more favourable once we consider nested commutators of the Laplacian (not just any polynomial in  $\partial_x, \partial_y, \partial_z$ ).

### 10.4.2 Non-commuting elements of the Lie idealiser

It is simple to see that the components of  $\text{div}$  in cylindrical coordinates are elements of the Lie idealiser,

$$\begin{aligned}\left[\frac{1}{r}\partial_r \circ r, f\right] &= \partial_r f, \\ \left[\frac{1}{r}\partial_\phi, f\right] &= \frac{1}{r}\partial_\phi f, \\ [\partial_z, f] &= \partial_z f.\end{aligned}$$

We note that the *derivations induced* by the components of  $\text{div}$  are identical to the components of  $\text{grad}$ ,  $\nabla = (\partial_r, \frac{1}{r}\partial_\phi, \partial_z)$ ,

$$\frac{1}{r}\partial_r \circ r \rightarrow \partial_r, \quad \frac{1}{r}\partial_\phi \rightarrow \frac{1}{r}\partial_\phi, \quad \partial_z \rightarrow \partial_z.$$

The components of  $\text{grad}$  are trivially seen to be elements of the Lie idealiser,

$$\begin{aligned}[\partial_r, f] &= \partial_r f, \\ \left[\frac{1}{r}\partial_\phi, f\right] &= \frac{1}{r}\partial_\phi f, \\ [\partial_z, f] &= \partial_z f,\end{aligned}$$

and the derivations they induce are identical.

Similar observations hold true in spherical coordinates. The components of  $\text{div}$  induce the derivations

$$\frac{1}{r^2}\partial_r \circ r^2 \rightarrow \partial_r, \quad \frac{1}{r \sin \theta}\partial_\theta \circ \sin \theta \rightarrow \frac{1}{r}\partial_\theta, \quad \frac{1}{r \sin \theta}\partial_\phi \rightarrow \frac{1}{r \sin \theta}\partial_\phi,$$

which are identical to the components of  $\text{grad}$  and the derivations they induce.

However, in neither case do these operators commute. In spherical coordinates, none of the components of  $\text{grad}$  commute,

$$\begin{aligned}\left[\partial_r, \frac{1}{r}\partial_\phi\right] &= -\frac{1}{r^2}\partial_\phi, \\ \left[\partial_r, \frac{1}{r \sin \theta}\partial_\phi\right] &= -\frac{1}{r^2 \sin \theta}\partial_\phi, \\ \left[\frac{1}{r}\partial_\theta, \frac{1}{r \sin \theta}\partial_\phi\right] &= -\frac{\cos \theta}{r^2 \sin \theta}\partial_\phi.\end{aligned}$$

In cylindrical coordinates, the non-commuting components of  $\text{grad}$  are the pair  $\partial_r$  and  $\frac{1}{r}\partial_\phi$ , as we have already seen for spherical coordinates.

An interesting question is whether structures such as Jordan polynomials in elements

of the Lie idealiser could be generalised for non-commutating elements that, nevertheless, do obey some properties that are common to components of grad and div in Cartesian, cylindrical and spherical coordinates, for instance. This would have to be done keeping in mind whether such a generalisation is useful from a numerical point of view—for instance, whether it leads to structural properties such as height reduction and preservation of skew-Hermiticity upon discretisation.

### 10.4.3 Wigner equation

The Wigner equation

$$\partial_t w^\varepsilon(x, k, t) = -k \partial_x w^\varepsilon(x, k, t) + \Theta_{\delta^\varepsilon V} w^\varepsilon(x, k, t),$$

obtained from the semiclassical Schrödinger equation via a Wigner transform (Jin et al. 2011), features the pseudo-differential operator  $\Theta_{\delta^\varepsilon V}$ . Here, the pseudo-differential operator  $\Theta_g$ , for  $g \in \mathcal{G} = C_p^\infty([-1, 1]^2, \mathbb{R})$  (for instance), is defined by its action on  $f \in \mathcal{H} = C_p^\infty([-1, 1]^2, \mathbb{C})$ ,

$$\Theta_g(f(x, k)) = \frac{1}{2\pi} \iint g(x, y) f(x, s) e^{iy(k-s)} dy ds,$$

while  $\delta^\varepsilon V(x, y)$  is the function

$$\delta^\varepsilon V(x, y) = \frac{V(x - \frac{\varepsilon}{2}y) - V(x + \frac{\varepsilon}{2}y)}{\varepsilon}.$$

The space  $\mathcal{C} = \{\Theta_g : g \in \mathcal{G}\}$  is a commutative algebra since

$$\begin{aligned} \Theta_g(\Theta_h(f(x, k))) &= \frac{1}{(2\pi)^2} \iiint h(x, z) g(x, y) f(x, s) e^{iy(k-s)} e^{iz(s-r)} dy dz dr ds \\ &= \frac{1}{(2\pi)^2} \iiint h(x, z) g(x, y) f(x, r) e^{is(z-y)} e^{-izr} e^{iyk} dy dz dr ds \\ &= \frac{1}{2\pi} \iiint h(x, z) g(x, y) f(x, r) \delta(z-y) e^{-izr} e^{iyk} dy dz dr \\ &= \frac{1}{2\pi} \iint h(x, z) g(x, z) f(x, r) e^{iz(k-r)} dz dr \\ &= \Theta_{gh}(f(x, k)), \end{aligned} \tag{10.1}$$

where we have used the fact that  $\delta(x-a) = \frac{1}{2\pi} \int e^{ik(x-a)} dk$  is the Dirac delta distribution and that  $\int \delta(x-a) f(x) dx = f(a)$ . Much like the setting of Section 6.1,  $\mathcal{C}$  is isomorphic to  $\mathcal{G}$ .

Moreover, multiplication by  $k$  and the differential operator  $\partial_x$ , are both elements of the Lie idealiser, inducing the derivations  $i\partial_y$  and  $\partial_x$ , respectively. The proof for  $[\partial_x, \Theta_g] =$

$\Theta_{\partial_x g}$  is trivial and goes as follows,

$$\begin{aligned} [\partial_x, \Theta_g](f(x, k)) &= \frac{1}{2\pi} \int \int (\partial_x g(x, y) f(x, s) + g(x, y) \partial_x f(x, s)) e^{iy(k-s)} dy ds \\ &\quad - \frac{1}{2\pi} \int \int g(x, y) \partial_x f(x, s) e^{iy(k-s)} dy ds \\ &= \frac{1}{2\pi} \int \int \partial_x g(x, y) f(x, s) e^{iy(k-s)} dy ds = \Theta_{\partial_x g}(f(x, k)). \end{aligned}$$

The proof for  $[k, \Theta_g] = \Theta_{i\partial_y g}$  is a bit more involved. Opening the commutator, we arrive at

$$k\Theta_g(f(x, k)) - \Theta_g(kf(x, k)) = \frac{1}{2\pi} \int \int g(x, y) f(x, s) (k - s) e^{iy(k-s)} dy ds.$$

We integrate by parts with  $\partial_y u = i(k - s)e^{iy(k-s)}$  and  $v = -ig(x, y)f(x, s)$ . Noting that  $u = e^{iy(k-s)}$  and  $\partial_y v = -i\partial_y g(x, y)f(x, s)$ ,

$$\begin{aligned} [k, \Theta_g](f(x, k)) &= -\frac{1}{2\pi} \int \int (-i\partial_y g(x, y) f(x, s)) e^{iy(k-s)} dy ds \\ &= \frac{1}{2\pi} \int \int (i\partial_y g(x, y)) f(x, s) e^{iy(k-s)} dy ds \\ &= \Theta_{i\partial_y g}(f(x, k)). \end{aligned}$$

We can also show that  $[\partial_x, k] = 0$ , and that these operators commute.

We find that these properties are similar to the natural extension of  $\mathfrak{G}$  to two dimensions, with the role of the isomorphism played by  $\Theta$  instead of left multiplication,  $\mathcal{M}$ . Instead of the Laplacian and the potential, the operators we have to contend with are  $\Theta_{\delta^\varepsilon V}$  and the polynomial  $k\partial_x$ . We need to explore whether the properties of the symmetrised structures conjectured in Section 10.4.1 would be relevant here for developing effective numerical schemes.

#### 10.4.4 Matrix valued potentials

Matrix-valued potentials appear in TDSEs when we need to consider multiple energy levels at once. Magnus expansion based methods such as (Hochbruck & Lubich 2003) are used once these potentials start featuring time-dependent components (Kormann, Holmgren & Karlsson 2008).

Unfortunately, matrix-valued potentials do not directly fall into the framework of  $\mathfrak{G}$  as a Lie algebra. However suitable extensions for these contexts are being actively sought. Initial findings suggest that a modified version of the height reduction rule holds and it might be possible to extend the favourable computational complexity of the Magnus–Zassenhaus schemes discussed in Chapter 8 to these cases.

# Bibliography

- Abramowitz, M. & Stegun, I. A. (1964), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York.
- Apostol, T. M. (1976), *Introduction to analytic number theory*, Undergraduate texts in mathematics, Springer-Verlag, New York.
- Bader, P., Iserles, A., Kropielnicka, K. & Singh, P. (2014), ‘Effective approximation for the semiclassical Schrödinger equation’, *Found. Comput. Math.* **14**, 689–720.
- Bader, P., Iserles, A., Kropielnicka, K. & Singh, P. (2015), Efficient methods for time-dependence in semiclassical Schrödinger equations. Pre-print available.
- Bao, W., Jin, S. & Markowich, P. A. (2002), ‘On time-splitting spectral approximations for the Schrödinger equation in the semiclassical regime’, *J. Comput. Phys.* **175**, 487–524.
- Blanes, S., Casas, F. & Murua, A. (2006), ‘Symplectic splitting operator methods tailored for the time-dependent Schrödinger equation’, *J. Chem. Phys.* **124**, 105–234.
- Blanes, S., Casas, F. & Murua, A. (2008), ‘Splitting and composition methods in the numerical integration of differential equations’, *Bol. Soc. Esp. Mat. Apl.* **45**, 89–145.
- Blanes, S., Casas, F., Oteo, J. A. & Ros, J. (1999), ‘Magnus and Fer expansions for matrix differential equations: the convergence problem’, *J. Phys. A* **31**(1), 259–268.
- Blanes, S., Casas, F., Oteo, J. A. & Ros, J. (2009), ‘The Magnus expansion and some of its applications’, *Phys. Rep.* **470**, 151–238.
- Blanes, S., Casas, F. & Ros, J. (2000), ‘Improved high order integrators based on the Magnus expansion’, *BIT* **40**(3), 434–450.
- Briggs, J. S. & Rost, J. M. (2001), ‘On the derivation of the time-dependent equation of Schrödinger’, *Found. Phys.* **31**(4), 693–712.
- Bungartz, H.-J. & Griebel, M. (2004), ‘Sparse grids’, *Acta Numerica* **13**, 147–269.

- Casas, F. & Murua, A. (2009), ‘An efficient algorithm for computing the Baker–Campbell–Hausdorff series and some of its applications’, *J. Math. Phys.* **50**, 033513–1–033513–23.
- Coutinho, S. C. (1997), ‘The many avatars of a simple algebra’, *Amer. Math. Monthly* **104**(7), 593–604.
- Crouch, P. & Grossman, R. (1993), ‘Numerical integration of ordinary differential equations on manifolds’, *J. Nonlinear Sci.* **3**, 1–33.
- Davis, P. J. & Rabinowitz, P. (1984), *Methods of Numerical Integration*, 2nd edn, Academic Press, Orlando, FL.
- Dixmier, J. (1968), ‘Sur les algèbres de Weyl’, *Bull. Soc. Math. Fr.* **96**, 209–242.
- Dynkin, E. B. (1947), ‘Evaluation of the coefficients of the Campbell–Hausdorff formula’, *Dokl. Akad. Nauk. SSSR* **57**, 323–326.
- Faou, E. (2012), *Geometric Numerical Integration and Schrödinger Equations*, Zurich Lectures in Advanced Mathematics, Europ. Math. Soc., Zürich.
- Fornberg, B. (1998), *A Practical Guide to Pseudospectral Methods*, Cambridge University Press, Cambridge.
- Gaim, W. & Lasser, C. (2014), ‘Corrections to Wigner type phase space methods’, *Nonlinearity* **27**(12), 2951.
- Golub, G. H. & Van Loan, C. F. (1996), *Matrix Computations*, 3rd edn, Johns Hopkins University Press, Baltimore.
- Gradinaru, V. & Hagedorn, G. A. (2014), ‘Convergence of a semiclassical wavepacket based time-splitting for the Schrödinger equation’, *Numer. Math.* **126**(1), 53–73.
- Griffiths, D. J. (2004), *Introduction to Quantum Mechanics*, 2nd edn, Prentice Hall, Upper Saddle River, NJ.
- Hairer, E., Lubich, C. & Wanner, G. (2006), *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd edn, Springer Verlag, Berlin.
- Hall, B. C. (2003), *Lie Groups, Lie Algebras, and Representations*, Graduate Texts in Mathematics, Springer, New York.
- Hesthaven, J. S., Gottlieb, S. & Gottlieb, D. (2007), *Spectral Methods for Time-Dependent Problems*, Cambridge University Press, Cambridge.

- Higham, N. J. & Al-Mohy, A. H. (2010), ‘Computing matrix functions’, *Acta Numerica* **19**, 159–208.
- Hildebrand, F. (1987), *Introduction to Numerical Analysis*, Dover books on advanced mathematics, Dover Publications, New York.
- Hochbruck, M. & Lubich, C. (1997), ‘On Krylov subspace approximations to the matrix exponential operator’, *SIAM J. Numer. Anal.* **34**, 1911–1925.
- Hochbruck, M. & Lubich, C. (2003), ‘On Magnus integrators for time-dependent Schrödinger equations’, *SIAM J. Numer. Anal.* **41**(3), 945–963.
- Iserles, A. (2008), *A First Course in the Numerical Analysis of Differential Equations*, 2nd edn, Cambridge University Press, Cambridge.
- Iserles, A., Kropielnicka, K. & Singh, P. (2015), ‘On the discretisation of the semiclassical Schrödinger equation with time-dependent potential’. Technical Report NA2015/02.
- Iserles, A., Munthe-Kaas, H. Z., Nørsett, S. P. & Zanna, A. (2000), ‘Lie-group methods’, *Acta Numerica* **9**, 215–365.
- Iserles, A. & Nørsett, S. P. (1999), ‘On the solution of linear differential equations in Lie groups’, *Phil. Trans. R. Soc. A* **357**, 983–1019.
- Iserles, A. & Nørsett, S. P. (2005), ‘Efficient quadrature of highly oscillatory integrals using derivatives’, *Proc. R. Soc. A* **461**(2057), 1383–1399.
- Iserles, A., Nørsett, S. P. & Rasmussen, A. (2001), ‘Time symmetry and high-order Magnus methods’, *Appl. Numer. Math.* **39**(3–4), 379–401.
- Jahnke, T. & Lubich, C. (2000), ‘Error bounds for exponential operator splittings’, *BIT Numer. Math.* **40**, 735–744.
- Jin, S., Markowich, P. & Sparber, C. (2011), ‘Mathematical and computational methods for semiclassical Schrödinger equations’, *Acta Numerica* **20**, 121–209.
- Kämmerer, L., Kunis, S. & Potts, D. (2012), ‘Interpolation lattices for hyperbolic cross trigonometric polynomials’, *J. Complexity* **28**(1), 76–92.
- Khesin, B. & Wendt, R. (2009), *The Geometry of Infinite-Dimensional Groups*, A Series of Modern Surveys in Mathematics, Springer-Verlag, Berlin Heidelberg.
- Kormann, K., Holmgren, S. & Karlsson, H. O. (2008), ‘Accurate time propagation for the Schrödinger equation with an explicitly time-dependent Hamiltonian’, *J. Chem. Phys.* **128**(18).

- Kreiss, H.-O. & Oliger, J. (1972), ‘Comparison of accurate methods for the integration of hyperbolic equations’, *Tellus* **24**(3), 199–215.
- Lee, J. M. (2012), *Introduction to Smooth Manifolds*, Graduate Texts in Mathematics, Springer, New York.
- Lehmer, D. H. (1988), ‘A new approach to Bernoulli polynomials’, *Am. Math. Monthly* **95**(10), 905–911.
- Lord, G., Malham, S. J. A. & Wiese, A. (2008), ‘Efficient strong integrators for linear stochastic systems’, *SIAM J. Numer. Anal.* **46**(6), 2892.
- Lubich, C. (2008), *From quantum to classical molecular dynamics: reduced models and numerical analysis*, Zurich Lectures in Advanced Mathematics, Europ. Math. Soc., Zürich.
- Magnus, W. (1954), ‘On the exponential solution of differential equations for a linear operator’, *Comm. Pure Appl. Math.* **7**, 649–673.
- Markowich, P. A., Pietra, P. & Pohl, C. (1999), ‘Numerical approximation of quadratic observables of Schrödinger-type equations in the semi-classical limit’, *Numer. Math.* **81**(4), 595–630.
- McLachlan, R. I. & Quispel, G. R. W. (2002), ‘Splitting methods’, *Acta Numerica* **11**, 341–434.
- Moan, P. C. & Niesen, J. (2008), ‘Convergence of the Magnus series’, *Found. Comput. Math.* **8**(3), 291–301.
- Moler, C. & Loan, C. V. (1978), ‘Nineteen dubious ways to compute the exponential of a matrix’, *SIAM Review* **20**, 801–836.
- Munthe-Kaas, H. (1999), ‘High order Runge-Kutta methods on manifolds’, *Appl. Numer. Math.* **29**, 115–127.
- Munthe-Kaas, H. & Owren, B. (1999), ‘Computations in a free Lie algebra’, *Phil. Trans. R. Soc. A* **357**(1754), 957–981.
- Munthe-Kaas, H. Z., Quispel, G. R. W. & Zanna, A. (2001), ‘Generalized polar decompositions on lie groups with involutive automorphisms’, *Found. Comput. Math.* **1**(3), 297–324.
- Murua, A. (2010), ‘The symmetric Baker–Campbell–Hausdorff formula up to terms of degree 20 written in a Hall basis or a Lyndon basis’, <http://www.ehu.es/ccwmuura/research/sbchHall19.dat>.



- Orszag, S. A. (1969), ‘Numerical methods for the simulation of turbulence’, *Phys. Fluids* **12**(12), II–250.
- Orszag, S. A. (1972), ‘Comparison of pseudospectral and spectral approximation’, *Stud. Appl. Math.* **51**(3), 253–259.
- Oteo, J. A. (1991), ‘The Baker–Campbell–Hausdorff formula and nested commutator identities’, *J. Math. Phys.* **32**, 419–424.
- Park, T. J. & Light, J. C. (1986), ‘Unitary quantum time evolution by iterative Lanczos reduction’, *J. Chem. Phys.* **85**(10), 5870–5876.
- Pasciak, J. E. (1980), ‘Spectral and pseudo spectral methods for advection equations’, *J. Math. Comput.* **35**(152), 1081–1092.
- Reutenauer, C. (1993), *Free Lie Algebras*, London Maths Soc. Monographs **7**, Oxford University Press, Oxford.
- Sheng, Q. (1994), ‘Global error estimates for exponential splitting’, *IMA J. Numer. Anal.* **14**(1), 27–56.
- Singh, P. (2015), Algebraic theory for higher order methods in computational quantum mechanics. arXiv:1510.06896 [math.NA].
- Strang, G. (1968), ‘On the construction and comparison of difference schemes’, *SIAM J. Numer. Anal.* **5**(3), 506–517.
- Tal Ezer, H. & Kosloff, R. (1984), ‘An accurate and efficient scheme for propagating the time dependent Schrödinger equation’, *J. Chem. Phys.* **81**, 3967–3976.
- Tal-Ezer, H., Kosloff, R. & Cerjan, C. (1992), ‘Low-order polynomial approximation of propagators for the time-dependent Schrödinger equation’, *J. Comput. Phys.* **100**(1), 179–187.
- Trefethen, L. N. (2000), *Spectral Methods in MATLAB*, SIAM, Philadelphia.
- Trotter, H. F. (1959), ‘On the product of semi-groups of operators’, *Proc. Amer. Math. Soc.* **10**(4), 545–551.
- Yoshida, H. (1990), ‘Construction of higher order symplectic integrators’, *Phys. Lett.* **150**, 262–268.