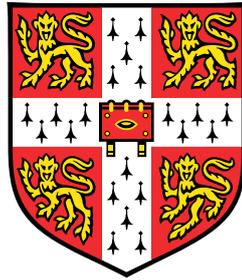


Analytical and numerical techniques for wave scattering



Georg Andreas Maierhofer

Supervisors: Professor Nigel Peake
Professor Arieh Iserles

Department of Applied Mathematics and Theoretical Physics
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or am concurrently submitting, for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or is being concurrently submitted, for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

The contents of Chapter 4 are original joint work with Prof. Daan Huybrechs (KU Leuven). The main part is in press at *Advances in Computational Mathematics* (Maierhofer and Huybrechs, 2022b) and the results of Appendix 4.F are under consideration at a peer-reviewed journal (Maierhofer and Huybrechs, 2022a). More specifically, the mathematical results presented in this chapter are my own work, and Prof. Huybrechs provided technical guidance on planning the research path throughout the project and worked jointly on the Julia code which produces the computational results shown in this chapter. This Julia code builds on several existing packages that were previously written and are maintained by Prof. Huybrechs, most importantly `IntegralEquations2D.jl` (Huybrechs, 2021).

The remainder of the results presented in this thesis, as well as the Python and Matlab code which produces the remainder of the computational results, are my own original work, under the supervision of Prof. Nigel Peake and Prof. Arieh Iserles. The results of Chapters 2 & 3 have been published as two separate articles in the *Journal of Sound and Vibration* (Maierhofer and Peake, 2020, 2022) and the results of Chapter 5 are currently under consideration at a peer-reviewed journal (Maierhofer et al., 2021).

Although this thesis is my own work, I will use the pluralis modestiae throughout.

Georg Andreas Maierhofer
15 June, 2021

Abstract

Analytical and numerical techniques for wave scattering

Georg Andreas Maierhofer

In this thesis, we study the mathematical solution of wave scattering problems which describe the behaviour of waves incident on obstacles and are highly relevant to a raft of applications in the aerospace industry. The techniques considered in the present work can be broadly classed into two categories: analytically based methods which use special transforms and functions to provide a near-complete mathematical description of the scattering process, and numerical techniques which select an approximate solution from a general finite-dimensional space of possible candidates.

The first part of this thesis addresses an analytical approach to the scattering of acoustic and vortical waves on an infinite periodic arrangement of finite-length flat blades in parallel mean flow. This geometry serves as an unwrapped model of the fan components in turbo-machinery. Our contributions include a novel semi-analytical solution based on the Wiener–Hopf technique that extends previous work by lifting the restriction that adjacent blades overlap, and a comprehensive study of the composition of the outgoing energy flux for acoustic wave scattering on this array of blades. These results provide an insight into the importance of energy conversion between the unsteady vorticity shed from the trailing edges of the cascade blades and the acoustic field. Furthermore, we show that the balance of incoming and outgoing energy fluxes of the unsteady field provides a convenient tool for understanding several interesting scattering symmetries on this geometry.

In the second part of the thesis, we focus on numerical techniques based on the boundary integral method which allows us to write the governing equations for zero mean flow in the form of Fredholm integral equations. We study the solution of these integral equations using collocation methods for two-dimensional scatterers with smooth and Lipschitz boundaries. Our contributions are as follows: Firstly, we explore the extent to which least-squares oversampling can improve collocation. We provide rigorous analysis that proves guaranteed convergence for small amounts of oversampling and shows that superlinear oversampling can ensure faster asymptotic convergence rates of the method. Secondly, we examine the computation of the entries in the discrete linear system representing

the continuous integral equation in collocation methods for hybrid numerical-asymptotic basis spaces on simple geometric shapes in the context of high-frequency wave scattering. This requires the computation of singular highly-oscillatory integrals and we develop efficient numerical methods that can compute these integrals at frequency-independent cost. Finally, we provide a general result that allows the construction of recurrences for the efficient computation of quadrature moments in a broad class of Filon quadrature methods, and we show how this framework can also be used to accelerate certain Levin quadrature methods.

*To my parents, Birgitta and Hans-Jörg, and my brother, Thomas,
for their love, patience and unwavering support.*

Acknowledgements

First and foremost, I would like to thank my supervisors Prof. Nigel Peake and Prof. Arieh Iserles. It has been a tremendous privilege to work under their guidance for the past four years.

I would like to thank Nigel for his steady support in my scientific endeavours and his openness to so many directions of research, which has afforded me the wonderful opportunity to pursue my interests and work on projects that I am truly passionate about. Nigel's calm and positive style has been a great source of motivation, especially during the times when projects did not go to plan, and helped me persevere in the face of challenge. Throughout the years I have been grateful for Nigel's willingness to meet and discuss research, regardless of how busy his schedule was. Indeed, working under his supervision not only offered me the chance to learn from him about exciting mathematical ideas, but also to admire his purposefulness and genuine devotion to the students and scientific ideals of the University of Cambridge, displayed by Nigel as a teacher, supervisor, and in various roles of academic leadership. I truly wish that someday I will be able to replicate Nigel's commitment to the advancement of knowledge through research and education in the physical sciences.

At the same time, I would like to thank Arieh for his unrivalled enthusiasm for mathematical research, which has inspired me to see the beauty and find true joy in my work. Throughout the years, I have been amazed by his generosity in spending time helping his research students. Arieh has always been available for a chat, to exchange ideas, and to proofread my writings, regardless of how many papers or other projects he was working on. His commitment to research and supervision has not only supported my work directly but has also instilled a sense of responsibility that has helped me to challenge myself to live up to this standard. During my time at Cambridge, I have not come across a single topic in numerical analysis and, in fact, any mathematical field, that Arieh was not happy to discuss and for which he was not aware of exciting theorems and stories to share. I can only strive and hope that, along my scientific journey, I will be able to acquire a fraction of his mathematical knowledge.

Next, I would like to express my sincere gratitude to Prof. Daan Huybrechs, most importantly for proposing and for working together on such an exciting research project (Chapter 4 of this thesis). In his considerate and friendly style, Daan has taught me a lot

about integral equations, numerical quadrature and wave scattering, and gave me a chance to explore a truly exciting field of research. I am especially grateful for his hospitality during my visit to KU Leuven in 2018 and for his flexibility and commitment to making our collaboration work even when an extended visit was prevented by the ongoing pandemic.

It has been a pleasure to work alongside so many brilliant and kind mathematicians in the Waves group and the Cambridge Centre for Analysis (CCA) over the past years. From the Waves group, I would like to express my gratitude in particular to Dr Lorna Ayton and Dr Anastasia Kisil for many exciting and helpful conversations about research. I am indebted to Anastasia for the opportunity to be part of the steering committee for the Mathematical Analysis in Acoustics special interest group of UKAN, which allowed me to learn about ongoing and future research goals from some of the leading experts of my time. I am also grateful to Dr Mark Spivack for acting as my graduate advisor, and both to him and Dr Orsola Rath Spivack for their enthusiasm and support and for many interesting conversations about scattering in random media. I would also like to thank Dr Stephen Cowley for his unwavering support of students and junior researchers, which has helped me since my first research internship in DAMTP. I am grateful for many helpful and interesting discussions with Mungo Aitken, Dr Peter Baddoo, Dr David Baker, Sam Bradford, Dr Ed Brambley, Alistair Hales, Chris Sear, and Dr Mike Smith. I am especially grateful to Matthew Priddin for a truly exciting collaboration on some topics beyond the scope of this thesis, and many helpful conversations concerning the topics presented here.

Being a member of the CCA has allowed me to learn a great deal about analysis beyond my present research in waves. In this regard, I am grateful to the Cambridge Image Analysis group, and especially to Dr Angelica Aviles-Rivero, Dr Martin Benning, Dr Matthias Ehrhardt and Prof. Carola-Bibiane Schönlieb, for their collaboration on several projects outside the scope of this thesis. I would also like to thank Dr Luca Magri from the Department of Engineering, who taught me a lot about Lyapunov exponents and ergodic systems during my CCA external project. I am similarly grateful to Daniel Heydecker for collaboration and many fun coffee breaks, and to the many bright colleagues in the CCA alongside whom I had the pleasure to work during the past four years. These include Dr Andrew Celsus, Dr Matthew Colbrook, and Karen Luong as fellow numerical analysts, and in particular, my office mate Peter Taylor, whom I thank for providing an enjoyable and collegiate environment to work in. I would like to extend a special thank you to Mungo Aitken, Karen Luong and Dr Shuvrangs Das for proofreading a draft of this thesis. It is impossible to imagine a CCA without the person keeping watch over all CCA cohorts, Tessa Blackman, and I would like to thank her for always keeping an open door, for many nice CCA coffee breaks, and for providing essential support in administrative matters, without which we could not carry out our research.

My work presented here has benefited from numerous scientific discussions with many

experts in the field. In particular, I would like to express my gratitude to Prof. David Abrahams, Prof. Simon Chandler-Wilde, Dr Alfredo Deaño, Dr Andrew Gibbs, Dr Anders Hansen, Dr Dave Hewett, Prof. Stefano Serra-Capizzano, Dr Pranav Singh, and Dr Marcus Webb, for their patience and enthusiasm in explaining and exchanging mathematical concepts and ideas.

I am also grateful for having had the privilege to work and learn as a member of Trinity College since the beginning of my undergraduate studies. The collegiate community has looked after my well-being and academic development, and has allowed me to forge many lifelong friendships. Within the college, I would like to thank especially my tutor Prof. Caterina Ducati and my tutorial administrator Janice Chambers for their support throughout my studies and during the challenges brought about by the pandemic. I am also grateful to Prof. John Hinch for acting as my college mentor.

I would also like to gratefully acknowledge financial support through a research scholarship and travel grants from my college, travel support from the Cambridge Philosophical Society and support from the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/L016516/1 for my studies at the CCA.

There were many close friends without whom my time in Cambridge would have been much less enjoyable. I would like to thank in particular Eero, Diana and Jedrek, Irene and Matthew, Katya and Will, Matthew, Olivér, and Wuyhun, for the many good times together and for keeping me sane during the busier times of my studies.

None of this would have been possible without the unconditional love and support from my family, especially from my parents, Birgitta and Hans-Jörg, and from my brother Thomas. They have always been there to build me up when things did not go to plan, and ready to celebrate with me when things went well. I am particularly grateful for their advice and help in navigating so many of life's challenges during these past years and for their proofreading of this thesis and of many of my previous texts and papers – they are part of a very small number of people that will have read the content of this thesis in its entirety.

Finally, I want to thank Maria, my companion, friend and love, who has stuck with me through thick and thin. Her caring support and kind spirit have helped me navigate the tough times, her mathematical knowledge has been of invaluable help in countless discussions about research, and her cheerfulness has provided an enduring source of joy during my studies and beyond.

Contents

List of Figures	xvii
1 Introduction	1
1.1 Background	2
1.1.1 Cascades of blades in jet engines	3
1.1.2 Computational wave scattering	7
1.2 Mathematical techniques and preliminaries	12
1.2.1 The Wiener–Hopf technique	12
1.2.2 Boundary integral equations	18
1.2.3 Highly oscillatory quadrature and Filon methods	24
1.3 Thesis structure	30
2 Wave scattering by an infinite cascade of non-overlapping blades	35
2.1 Introduction	35
2.2 The equations of motion and mathematical formulation	38
2.2.1 Boundary and edge conditions	39
2.3 Formulation as a Wiener–Hopf problem	42
2.3.1 The scalar Wiener–Hopf kernel	43
2.3.2 Derivation of Wiener–Hopf equations	44
2.4 Solution using the Wiener–Hopf technique	46
2.4.1 Factorisation of κ	46
2.4.2 The uncoupled leading edge problem	46
2.4.3 Asymptotic behaviour of the kernel factors and residues	47
2.4.4 The trailing edge correction	49
2.4.5 The leading edge correction	51
2.5 Reduction to an infinite algebraic system	53
2.5.1 Stable formulation under $\text{Im } \tilde{\Omega} \rightarrow 0^-$	54
2.5.2 Approximate solution of the linear system	55
2.6 Equivalence to previous solutions for overlapping cascades when $d < 1$	56
2.6.1 Review of previous work in the case of overlap	56

2.6.2	Proof of equivalence when $d < 1$	57
2.7	Total unsteady lift and far-field behaviour	59
2.7.1	Total unsteady lift	59
2.7.2	Far-field behaviour	60
2.8	Numerical examples and results	62
2.8.1	Total unsteady lift	62
2.8.2	Far field behaviour	65
2.9	Conclusions	68
	List of symbols	69
2.A	The kernel κ and its factorisation	70
2.B	Wiener–Hopf splitting for the trailing edge correction	71
2.C	Reduction to an infinite algebraic system	74
2.D	The finite section method and convergence	76
2.E	Expressions for the solution with acoustic downstream incidence	78
3	Acoustic and hydrodynamic power of wave scattering by an infinite cascade of blades in mean flow	83
3.1	Introduction	83
3.2	Problem statement	86
3.2.1	Dispersion relation for free space solutions	87
3.2.2	Form of the far-field	89
3.3	An energy balance for the cascade of blades	90
3.3.1	Time-averaged energy balance in terms of amplitudes	91
3.3.2	Interpretations of the interaction terms	94
3.4	Results and discussion	96
3.4.1	Symmetries in the field and zero acoustic reflection	96
3.4.2	The significance of the hydrodynamic and energy conversion terms in balancing the energy	104
3.4.3	Negative acoustic energy absorption and sound power generation in wave-cascade scattering	105
3.5	Conclusions	109
	List of symbols	111
4	Oversampled collocation methods for boundary element methods in two dimensions	113
4.1	Introduction	113
4.2	An oversampled collocation method	118
4.2.1	Mathematical assumptions	120
4.2.2	From least-squares to a discrete Bubnov–Galerkin method	123

4.3	Oversampling in collocation methods and in approximation on Hilbert spaces	125
4.3.1	Stable function approximation on Hilbert spaces	125
4.3.2	Least-squares oversampled collocation method	126
4.4	Convergence analysis of the oversampled collocation method	127
4.4.1	Strang estimate for convergence on energy space	128
4.4.2	Superconvergence and the discrete Aubin–Nitsche lemma	131
4.4.3	Exact expression for the error for equispaced spline bases	133
4.5	Oversampled collocation in specific settings	139
4.5.1	Non-equispaced sampling points	140
4.5.2	Lipschitz domains	141
4.6	Numerical results	143
4.6.1	Smooth domains with equispaced sampling	144
4.6.2	Suboptimal choice of collocation points	147
4.6.3	Polygonal scatterers	150
4.7	Conclusions	152
	List of symbols	155
4.A	Error estimate in the discrete inner product	156
4.B	Error estimate for the discrete inner product non-uniform collocation points	157
4.C	Derivation of exact error expression for equispaced grids	158
4.D	Pseudodifferential form of the single layer operator	161
4.E	Perturbation argument for modified oversampled collocation	164
4.F	Perturbation argument for oversampled collocation (Laplace)	166
4.G	Perturbation argument for oversampled collocation (Single layer Helmholtz)	175
5	Recursive moment computation in Filon methods	187
5.1	Introduction	187
5.2	The extended Filon method	190
5.2.1	The Achilles’ heel of Filon methods: Moment computation	191
5.2.2	Fast interpolation at Filon–Clenshaw–Curtis points	192
5.3	Recursive moment computation in Filon methods	193
5.3.1	Recursive moment computation for Filon–Clenshaw–Curtis methods	198
5.4	Application to integrals with algebraic singularities and stationary points .	200
5.4.1	Stability analysis of the recurrences	202
5.4.2	Stability results for the initial regime	202
5.4.3	Change of behaviour of homogeneous solutions and Oliver’s algorithm	204
5.4.4	Numerical examples and comparison to previous work	205
5.5	Application to high-frequency wave scattering	208
5.5.1	Recursive moment computation	210
5.5.2	Initial conditions	210

5.5.3	Behaviour of homogeneous solutions and initial stability	212
5.5.4	Numerical evidence of stable forward propagation	214
5.5.5	Wave scattering on a screen in two dimensions	215
5.6	Conclusions	219
	List of symbols	221
5.A	Fast interpolation at Clenshaw–Curtis points, mid- and endpoint derivatives	222
5.B	Proof of Theorem 5.4.1	223
5.C	Proof of Theorem 5.4.2	226
5.D	Expression for initial moments in §5.5.2	232
5.E	Proof of Proposition 5.5.2	232
6	An efficient Levin–Clenshaw–Curtis method for a class of highly oscillatory integrals	235
6.1	Introduction	235
6.2	The Levin method	236
6.3	Accelerating the Levin method using banded matrix computations of Chebyshev coefficients	239
6.3.1	Algorithm for the efficient construction of the Levin method	244
6.4	Numerical examples	244
6.4.1	Linear phase function	244
6.4.2	Quadratic phase function	246
6.5	Conclusions	247
	List of symbols	249
7	Concluding remarks	251
7.1	Future work by chapter	252
	References	257

List of Figures

1.1	The predominant design in commercial aircraft engines, the turbofan. . . .	4
1.2	Schematic of a row of rotor blades and a row of stator blades.	5
1.3	The minimum degrees of freedom required by a spline Galerkin boundary element method to solve a wave scattering problem to fixed accuracy as the frequency increases.	8
1.4	Sketch of a domain- and a boundary-based mesh.	10
1.5	Complex half planes R^\pm and strip of overlap $R^+ \cap R^- \subset \mathbb{C}$	13
1.6	The branch cuts of γ	14
1.7	The trapezoidal rule ($\nu = 9$) applied to the real part of $I_\omega[1/(1+x^2)]$	25
2.1	The cascade geometry with mean-flow and blade labels.	38
2.2	The regions of different asymptotic behaviour of $ \kappa $ in the complex α -plane.	48
2.3	Total lift \mathcal{L} for $M \in [0.2, 0.9]$	63
2.4	Total lift $ \mathcal{L} $ as a function of K_3	64
2.5	Total lift $ \mathcal{L} $ as a function of α_0	65
2.6	Relative transmission and reflection amplitudes of the incident wave. . . .	66
2.7	Relative transmission and reflection amplitudes as a function of d	67
2.8	The domain \mathcal{D} of algebraic behaviour of $ \kappa^+ $ in the complex α -plane. . . .	73
3.1	Sketch of cascade of blades with incident acoustic wave.	86
3.2	Dispersion curves for radiation modes away from the cascade.	88
3.3	Control domain for the energy balance.	90
3.4	Relative modal pressures of the first transmitted and reflected mode, with $\alpha_0 = 30^\circ$ and $\Omega M = \pi/4$	97
3.5	Relative modal pressures of the first transmitted and reflected mode, with $\alpha_0 = 30^\circ$ and $\Omega M = 5\pi/4$	98
3.6	Special symmetries in the field for zero mean flow.	99
3.7	The contributions to the outgoing energy flux as percentage of the total incoming power.	105
3.8	The acoustic and hydrodynamic power for varying incident frequency. . . .	106
3.9	Relative acoustic energy absorption by the wake (linear scale).	107

3.10	Relative outgoing sound powers (dB scale).	108
4.1	The improved convergence properties of oversampled collocation on a smooth scatterer in the single layer potential formulation.	116
4.2	Sketch of the mesh refinement for $J = 4$	134
4.3	Error in the numerical method for a smooth circular scatterer, using the single layer potential formulation and linear splines.	145
4.4	The effect of linear oversampling for the same experiment as in Fig. 4.3.	147
4.5	Error in an interior field point for wave scattering on a smooth domain, with equispaced points that are offset from the equispaced spline mesh.	148
4.6	Double layer formulation of the interior Dirichlet problem for the Helmholtz equation. The sampling points are drawn uniformly at random.	150
4.7	Error in the numerical method for a regular pentagonal scatterer, using the single layer potential (order $2\alpha = -1$) and linear splines ($d = 1$), with offset equispaced collocation points.	151
4.8	Single layer formulation of the exterior Dirichlet problem for the Helmholtz equation on a pentagonal scatterer.	152
5.1	The growth of solutions to Eq. (5.18) as measured by $\left\ \prod_{n=1}^N \mathbf{A}_j \right\ $	206
5.2	Comparison of our direct Filon method using recursive moment computation, with the literature.	207
5.3	The growth of solutions to Eq. (5.18) as measured by $\left\ \prod_{n=1}^N \mathbf{B}^{(j)} \right\ $	215
5.4	Scattering of a Gaussian beam on a perfectly conducting finite screen.	216
5.5	Relative error of the direct Filon method for evaluating $\mathcal{S}(\partial_n \psi_i)(s_m)$	219
6.1	The error in the Levin–Clenshaw–Curtis method for $I_\omega^{(1)}[f]$ as a function of ω for fixed $\nu = 4, 64, 128$	245
6.2	The error and timing of the Levin–Clenshaw–Curtis method as a function of the number of collocation points ν for fixed frequency ω	246
6.3	The error in the Levin–Clenshaw–Curtis method for $I_\omega^{(2)}[f]$	247

Chapter 1

Introduction

The generation, propagation and deflection of waves plays an important role in a vast range of modern engineering applications. These include systems where the controlled generation of acoustic and electromagnetic waves is desired for remote and non-invasive measurements such as radar, magnetic resonance imaging, sonar and ultrasound; systems for the exchange of information such as telecommunication networks; structures where controlled deflection enhances entertainment such as the acoustics of opera houses; and even systems for the targeted delivery of energy in medical applications such as high-intensity focused ultrasound and radiation therapy. The UK Acoustics Network estimates that in 2019 the acoustics industry alone contributed £4.6 billion to the United Kingdom's economy annually, employing over 16,000 people in 750 companies nationwide (UK Acoustics Network, 2019). In the same year the electech sector (encompassing applications such as radar and telecommunications) generated a revenue of around £100 billion and employed more than 1 million people in over 45,000 businesses nationwide (Innovate UK, 2019). Together the two sectors make up a significant portion of the UK's GDP each year.

Acoustic waves can also play an antagonistic role: uncontrolled production of sound is an undesired by-product of nearly any movement of mechanical parts at rapid speed, including in aeroplanes, industrial gas turbines, and wind turbines. Such sound waves are typically referred to as noise and form an enormous portion of the negative impact of industrialisation on nature and human life. The European Environment Agency estimates that at least 20% of the EU population lives in areas where environmental noise levels are harmful to health (European Environment Agency, 2020). According to this report, in addition to an estimated 12,000 premature deaths caused by noise pollution on the European territory alone, industrial and traffic noise causes a range of physiological and behavioural responses in wildlife resulting in increased mortality, reduced reproductive success and, hence, to overall lower population densities. Consequently, a large part of national and international acoustics research is devoted to the control and reduction of noise.

1.1 Background

The first step in controlling waves (whether for targeted generation, or for the reduction of unwanted noise) lies in the understanding of the underlying mechanisms and the solution of the ‘forward problem’, i.e. the description of how waves produced by specific sources are deflected by obstacles. Sound waves are manifestations of pressure fluctuations in a fluid, and their propagation in inviscid isentropic flows is described by the acoustic wave equation, which follows from the linearised momentum and continuity equations,

$$\frac{D^2 p}{Dt^2} - c_0^2 \Delta p = 0, \quad (1.1)$$

where $p = p(\mathbf{x}, t)$ is the unsteady pressure field (the quantity describing acoustic fluctuations), c_0 is the speed of sound in the medium, $\Delta = \nabla^2$ is the Laplacian, and D/Dt is the material derivative. In general c_0 may depend on \mathbf{x}, t but for the purpose of this thesis we focus on homogeneous fluids, for which c_0 is constant.

Alternatively, the unsteady field in inviscid isentropic flows may be described by the unsteady velocity field $\mathbf{u}(\mathbf{x}, t)$, which also satisfies the acoustic wave equation $D^2 \mathbf{u}/Dt^2 - c_0^2 \Delta \mathbf{u} = 0$. This velocity field can be decomposed, $\mathbf{u} = \mathbf{u}_v + \mathbf{u}_a$, into a solenoidal part \mathbf{u}_v (with $\nabla \cdot \mathbf{u}_v = 0$) which does not contribute to the acoustics and an irrotational part \mathbf{u}_a (with $\nabla \times \mathbf{u}_a = 0$) which can be expressed in terms of a scalar velocity potential $\phi(\mathbf{x}, t)$ such that $\mathbf{u}_a = \nabla \phi$. The unsteady velocity potential $\phi(\mathbf{x}, t)$ again satisfies the acoustic wave equation, $D^2 \phi/Dt^2 - c_0^2 \Delta \phi = 0$.

Electromagnetic waves are governed by a similar equation, the electromagnetic wave equation (which follows from Maxwell’s equations)

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} - \frac{1}{\mu \varepsilon} \nabla^2 \mathbf{E} = \mathbf{0}, \quad \frac{\partial^2 \mathbf{B}}{\partial t^2} - \frac{1}{\mu \varepsilon} \nabla^2 \mathbf{B} = \mathbf{0},$$

where \mathbf{E}, \mathbf{B} are the electric and magnetic field respectively, μ is the permeability and ε is the permittivity of the medium.

In many problems it is appropriate to assign a harmonic time-dependence to the unsteady field (essentially by considering the modal components of the field) and to write, for instance, $p(t, \mathbf{x}) = \text{Re}(\tilde{p}(\mathbf{x})e^{i\omega t})$. Henceforth, it is understood that physical quantities are equal to the real part of the corresponding quantities in the equations and the real part is not written in the interest of brevity. In this case, the spatial part \tilde{p} satisfies an elliptic partial differential equation, the Helmholtz or the convected Helmholtz equation, which for zero mean flow takes the simple form

$$\Delta \tilde{p} + k^2 \tilde{p} = 0,$$

where $k = \omega/c_0$ is the wave number.

In this thesis we shall focus on the scalar wave equation Eq. (1.1). Nevertheless, we note that many of the techniques for the solution of wave scattering problems, specifically those concerning the computational methods developed in Chapters 4-6, extend to the electromagnetic case, and that the relevant literature often shares ideas between the electromagnetics and the acoustics communities.

We shall be interested in solving Eq. (1.1) subject to suitable boundary conditions which are used to describe the shape and the surface properties of relevant obstacles by which the waves are deflected. Our main goal is to develop mathematical techniques to solve these ‘wave scattering’ problems. Depending on the context ‘to solve’ means to produce numerical values of the scattered field that are computed to a specified precision, to produce results that demonstrate the overall qualitative nature of the scattered waves, or to provide some physical insight into certain properties of the scattered field for a given obstacle.

1.1.1 Cascades of blades in jet engines

One major source of noise pollution is commercial air traffic. As a result of an ever growing amount of air traffic and increasing environmental awareness, governing bodies have imposed progressively more stringent requirements on the permissible noise emission of aircraft over the recent decades. The most up-to-date set of goals in the European aviation sector is established in the EU’s vision Flightpath 2050 (European Commission, 2012) which aims to reduce the perceived noise of flying aircraft by 65% in 2050 relative to the noise level of a typical new aircraft in 2000. This has prompted a number of transnational research efforts including the ‘Silent’ Aircraft Initiative (Dowling et al., 2007) and the ENOVAL (Engine Module Validators) program (ENOVAL Consortium, 2019) seeking to reduce perceived noise and noise emissions through a radical redesign of airframe and engine components respectively.

Currently, the predominant engine design on commercial aircraft is the turbofan engine. In this design a large rotor fan forms the first stage of the compressor with the majority of the mass flow generated by the fan bypassing the engine core. The remainder of the flow enters the engine core where it is further compressed, mixed with fuel and ignited, and, after powering the turbine, exits the core in the form of jet exhaust. The thrust of the jet engine results from both the bypass flow and the exhaust stream.

A schematic of such a turbofan engine is shown in Fig. 1.1a. In Fig. 1.1b we see an artist’s rendering of the latest turbofan engine developed by Rolls-Royce, called the Trent 7000. It is apparent that an integral component of the engine consists of the fan blades, which provide a major proportion of the thrust, and the compressor blades, which are used to compress air entering the combustor. A design change in these turbofan engines that

has proven particularly advantageous over the past decades is an increase in the bypass ratio (BPR) (the ratio between the mass flow rate of the bypass air to the air entering the engine core as shown in Fig. 1.1a) which allowed engine manufacturers to enhance the fuel efficiency and to reduce the jet noise arising from the mixing of the hot jet exhaust with cold ambient air (Peake and Parry, 2012). For example the Trent 7000 has a BPR of 10:1, and ENOVAL Consortium (2019) proposed the next generation of ultra-high bypass engines to work in even larger regimes with $12 < \text{BPR} < 20$. At the same time, this development led to an increase in the relative importance of fan and compressor noise (Peake and Parry, 2012).

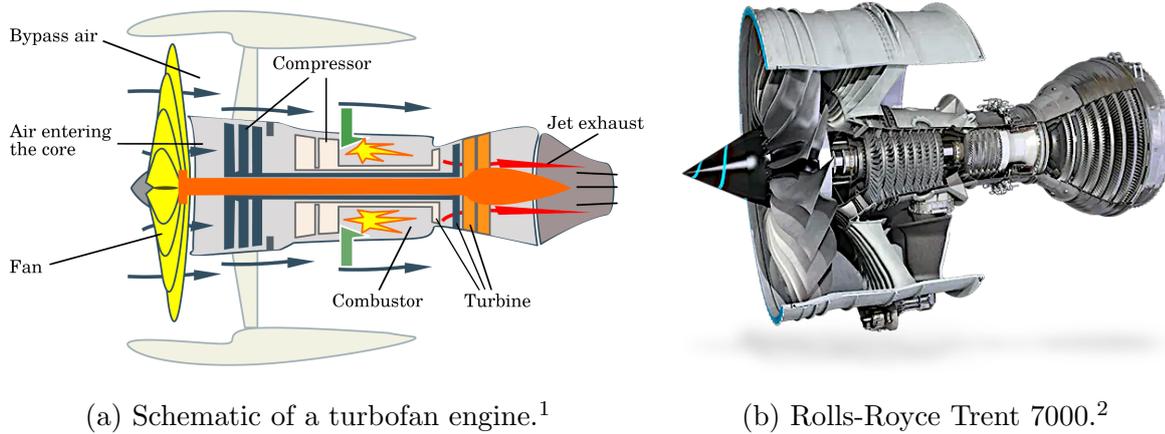


Figure 1.1: The predominant design in commercial aircraft engines, the turbofan.

The fan and compressor components are arrangements of blades with alternating rows of rotating blades (rotors) that compress the flow and provide thrust, and stationary vanes (stators) that direct the flow to avoid excessive swirl. A simplified model of a single rotor and stator row (called a compressor stage) is shown in Fig. 1.2.

In the study of noise emissions of turbomachinery one is therefore interested in both the effects of scattering of acoustic waves on single rows of blades (which can be compounded to provide an understanding of the ‘blockage’ by the fan components of sound generated elsewhere) and the rotor-stator interaction noise. The latter arises from the interaction of vorticity waves, which are generated from the movement of the rotors through the fluid, with the solid stator blades. In both cases it is reasonable to assume that the flow is radially restricted to near cylindrical surfaces (Posson et al., 2010, Baddoo, 2020), which allows us to unwrap and provide a 2D representation of the important features of the geometry as is shown in Fig. 1.2. If, in addition, the hub-tip ratio of the fan is close to unity chord-wise effects can be included, by expanding in terms of span-wise wavenumber (Goldstein, 1976).

¹Image from Adobe Stock, modified and reproduced under the appropriate license.

²Reproduced with permission. Copyright Rolls-Royce plc.

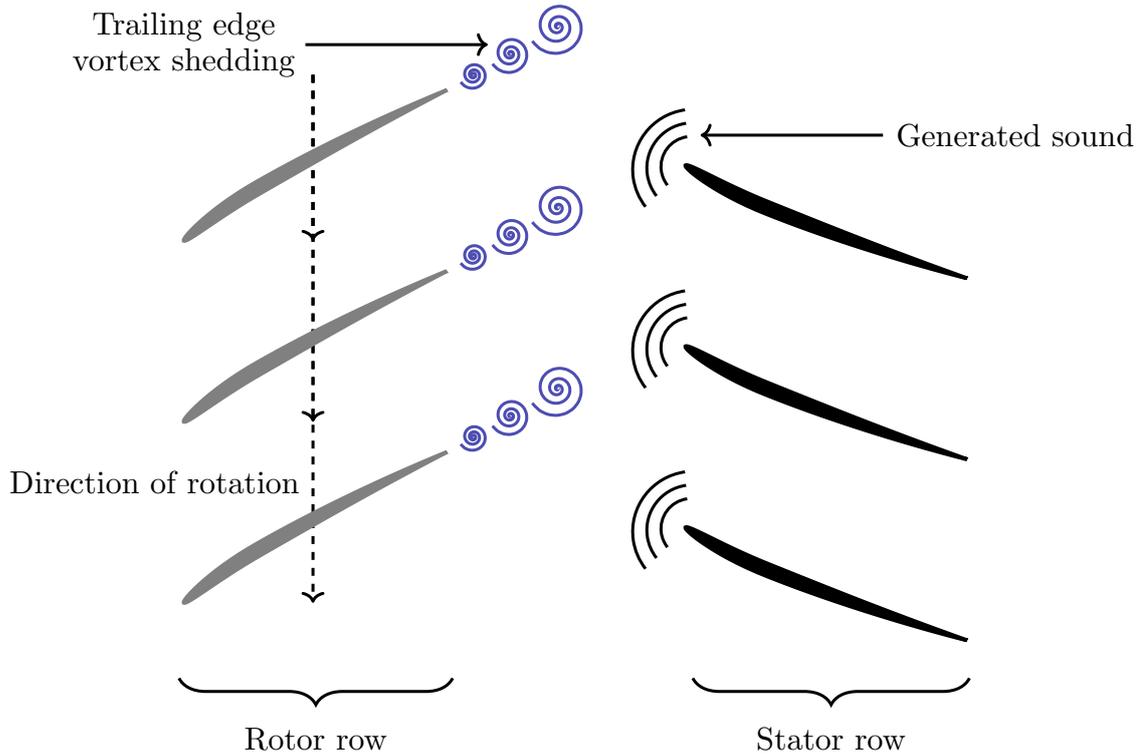


Figure 1.2: The cascade of blades is an unwrapped model of blade rows in turbomachinery.

We shall henceforth refer to this infinitely extending arrangement of blades as a ‘cascade of blades’. Motivated by this modelling application, wave scattering on the cascade of blades has received a large amount of attention over several decades, including studies by Whitehead (1970), Kaji and Okazaki (1970), Koch (1971), Peake (1992, 1993), Glegg (1999), Baddoo and Ayton (2020*b*), and Baddoo (2020). We will provide a more detailed description of the specifics of past contributions in §2.1, but note that several interesting open questions remain. One aspect which we address in this thesis is that all previous solutions based on the Wiener–Hopf technique for this problem imposed a requirement of overlap of consecutive blades in the cascade, and in Chapter 2 we provide a Wiener–Hopf solution that lifts this restriction.

A particularly interesting feature in this model for turbomachinery blade rows is the inevitable presence of nonzero mean flow (for our purposes assumed subsonic), which leads to profound consequences in the properties of the acoustic field. The large Reynolds number in the flow permits the use of a simplified inviscid model for the flow in the body of the fluid, however viscous boundary effects and resulting flow separation at the sharp trailing edge of each cascade blade cannot be ignored. It is understood that these effects can to a large extent be accounted for in the simplified inviscid model by introducing a single edge condition, which is known as the ‘Kutta–Joukowski hypothesis’ and more commonly simply referred to as the ‘Kutta condition’ (Crighton, 1985). The viscous effects in the boundary layer result in the shedding of vorticity from the sharp trailing

edges, which when imposing the Kutta condition is modelled in the inviscid equations using infinitely thin vortex sheets attached at the cascade trailing edges – the wake. The interaction of the unsteady vortical component of the field with the irrotational acoustic component facilitates an exchange of energy between these components. As a result of this energy exchange, and from the outset perhaps somewhat surprisingly, acoustic energy is no longer conserved in the presence of vorticity (in contrast to the total energy in the fluid which, of course, is conserved).

This effect has been subject to a large amount of research over the second half of the twentieth century. While it was well-known since the pioneering work by Lighthill (1952) that vorticity and turbulence can generate noise (i.e. pass energy to the acoustic field) it took several more decades until the converse effect, i.e. the generation of vorticity through sound, was understood. The first time this effect was conclusively described experimentally was in the work by Bechert et al. (1977), who found significant attenuation (at low frequencies) of the transmitted sound power of pure tones exiting a nozzle in the presence of a low Mach number jet flow. The details of past contributions will be discussed in §3.1, but we mention that this observed sound attenuation was first explained analytically in the work of Howe (1979) and Myers (1986). The attenuation of sound by conversion into vorticity has potential for exciting engineering applications, such as the work by Vér (1982), who managed to use this effect to suppress acoustic resonances in exchange heaters by inserting thin perforated plates aligned parallel with the mean flow into the heat exchanger cavity (the generation of vorticity in the perforations of the screen results in damping of the acoustic field).

This exchange of energy is of interest also for understanding the properties of the scattering process on blade rows and we provide a study of the energy balance for this system in Chapter 3. Indeed, it was found by Rienstra (1981) that, for single sharp trailing edges, energy can be exchanged in both ways between these components of the field, and we confirm that also for blade rows this effect can lead both to significant attenuation as well as to production of sound.

Finally, we mention that there are a number of interesting special properties of the field arising from the scattering of incident acoustic waves on the cascade of blades, including symmetries and zero reflection at non-trivial angles of attack. In reference to those properties one may consider blade rows as perfectly transmitting metamaterials at special angles of attack as demonstrated by Porter (2021). In Chapter 3 we show rigorously that zero reflection in the leading acoustic mode at special angles of attack is present for an arbitrary choice of incident frequency and hence extend previous results to the regime beyond the first modal cut-on frequency.

1.1.2 Computational wave scattering

Over the past decades the development of computers with ever greater processing power has opened up many new opportunities for (aero-)acoustical and electromagnetic research. Indeed, in the 1990s Sir James Lighthill foresaw the dawn of a “second golden age of aeroacoustics” facilitated by computer simulations (Lighthill, 1993). While some of the earliest wave scattering simulations could afford only a few dozen degrees of freedom (for instance Banaugh and Goldsmith (1963) were restricted to just 72 degrees of freedom in their numerical solution of the Helmholtz equation on an IBM 709), today aeroacoustical computations involving millions of data points are not uncommon (Lele and Nichols, 2014).

The availability of high-fidelity simulations has tremendous ramifications for manufacturing processes, for instance through the use of aeroacoustical simulations in the design process of modern turbomachinery, and electromagnetic simulations for the design of compact antennae. However, computing power is not limitless and there are hard bounds on what is feasible even with today’s state-of-the-art: in 2019 Rolls-Royce estimated that a full-fidelity multiphysics simulation (including thermo-mechanics, electromagnetics, computational aeroacoustics and computational fluid dynamics) of a complete gas-turbine engine during operation would require about one billion core hours per calculation. At a cost of about 1p per core hour in 2019 this would equate to a total cost of roughly £10 million for just a single calculation, which is clearly outside the remits of practicality (Lapworth, 2019). Thus the construction of efficient numerical algorithms supported by theoretical formulations and reduced models is extremely important even in the modern ‘age of high performance computing’.

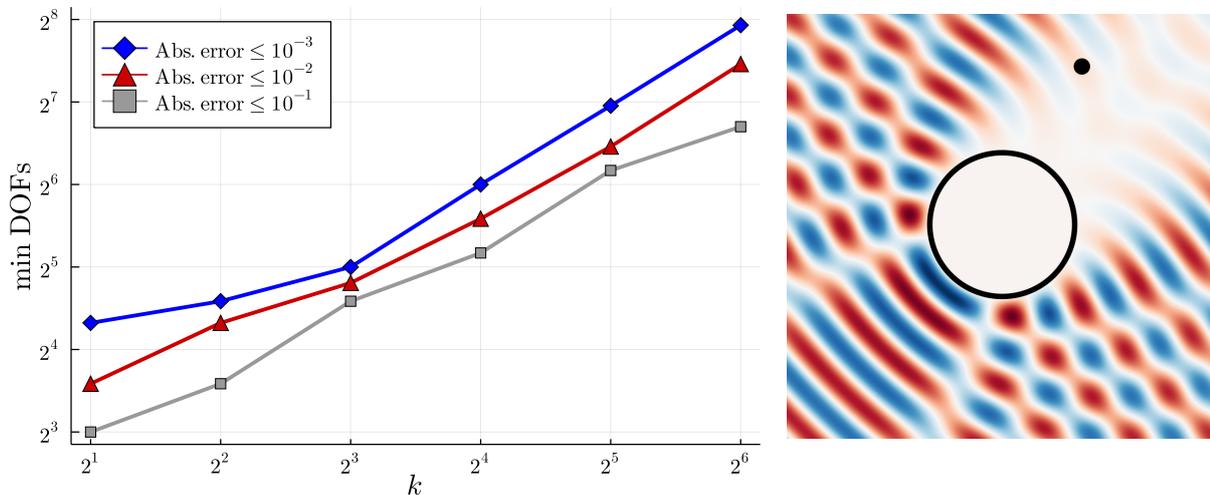
Indeed a particular challenge is posed by accurate wave propagation with minimal numerical artefacts (Lele and Nichols, 2014). This problem becomes especially pronounced at large frequencies: at the heart of modern numerical techniques (such as Galerkin, collocation, and Nystöm methods) lie Taylor expansions, even though they are sometimes disguised by the formalism of order and regularity in Sobolev spaces (Engquist et al., 2009). Thus the numerical error typically scales like a derivative of the approximated solution, which is large for highly-oscillatory functions, meaning oscillatory problems become more expensive to solve with classical techniques as the frequency increases. A practical way of judging whether a wave scattering problem strays in the ‘high-frequency’ regime is the dimensionless quantity kL , where k is the wavenumber as defined above and L is a typical length-scale of the scattering obstacle (Chandler-Wilde, Graham, Langdon and Spence, 2012). Typically for classical methods one requires a few mesh points per wavelength, meaning the number of mesh points scales like $(kL)^d$, where d is the dimension of the domain of the problem. Several important technologies operate in the high-frequency regime in this sense, for example:

- **Sonar** operates at frequencies of roughly 100 kHz – 1 MHz (Andrews, 2003), and

the speed of sound in water is approximately 1500 ms^{-1} , which means $k = \omega/c_0 \approx 66 - 666 \text{ m}^{-1}$. Thus if we wish to detect underwater structures of a length of a few meters the non-dimensional quantity $kL \approx 10^2 - 10^3$.

- **Fifth-generation wireless (5G)** operates at a frequency of $3.4 - 3.8 \text{ GHz}$ (Ofcom and HM Government, 2020) and the speed of light is roughly $3 \times 10^8 \text{ ms}^{-1}$, which means $k \approx 10 \text{ m}^{-1}$ and so if we wish to understand the propagation of these waves in the presence of civil structures which have a length scale of about a dozen meters, we expect $kL \approx 10^2$.

The effect of an increase in computational cost with increasing wavenumber k can be seen on the simple example in Fig. 1.3. Here we consider the minimum number of degrees of freedom (DOFs) that a spline Galerkin boundary element method requires to solve a simple plane wave scattering problem on a disk. Specifically, in this example we solve the exterior Dirichlet problem for the Helmholtz equation on the unit disk using a single layer integral formulation as described in Eq. (1.16). In the Galerkin method the trial and test spaces are chosen as continuous piecewise linear splines defined on an equispaced mesh on the boundary of the disk, $\{x \in \mathbb{R}^2 \mid |x| = 1\}$. The number of DOFs corresponds to the number of mesh points. The computational domain, the boundary of the disk, has dimension $d = 1$, so we expect a linear increase in cost with k which we indeed observe in Fig. 1.3a.



(a) Minimum DOFs required to solve to fixed absolute error. (b) The geometry and field point.

Figure 1.3: The minimum degrees of freedom (DOFs) required by a spline Galerkin boundary element method to solve the wave scattering problem in (b) to a fixed absolute error at a typical field point. We observe a clear linear increase in cost as we increase the wavenumber.³

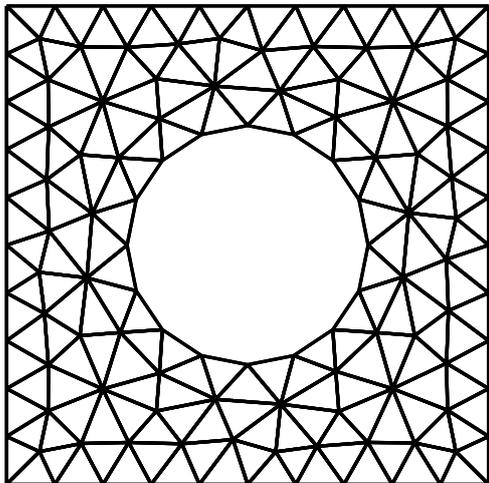
³Figure produced using the Julia package `IntegralEquations2D.jl` (Huybrechs, 2021).

As a consequence further development of efficient computational techniques for wave scattering is of immense importance. The principle of most modern solvers of partial differential equations (PDEs) relies on the selection of an optimal candidate solution from a sufficiently large (but finite dimensional) trial space. The specifics of the method usually determine a set of optimality conditions which serve as a way for the algorithm to judge the accuracy of approximation of the true solution by a given candidate. Thus the speed and usefulness of a numerical method typically depends on the following aspects:

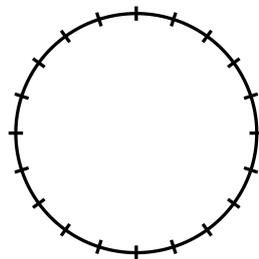
- **The size of the trial space:** A larger trial space (i.e. more degrees of freedom) naturally permits the close approximation of a larger class of functions. However, an increase in degrees of freedom also leads to an inevitable increase in complexity of the PDE solver.
- **The cost of evaluating the optimality conditions:** Of course, the best optimality condition would be an orthogonal projection of the true solution onto the trial space. However, knowledge of the true solution is not available and so typically one encounters a trade-off between simple optimality conditions (such as collocation equations) that are cheap and easy to evaluate and more involved ones (such as orthogonality with respect to bilinear forms in Galerkin methods) that are harder to evaluate but yield provably more accurate results. An example of this provably faster convergence rate for Galerkin methods can be found in the study by Arnold and Wendland (1983) who provided optimal convergence estimates for spline (i.e. piecewise polynomial) collocation methods for integral equations on 1D boundaries, which are then compared to known optimal convergence estimates for Galerkin methods.

For computational wave scattering in zero mean flow (or uniform mean flow after a suitable convective transform) arguably one of the most groundbreaking achievements of the twentieth century is the development of the boundary integral equation method. We provide a slightly more detailed introduction to the relevant mathematical results in §1.2.2 but in essence the method transforms partial differential equations on a domain into integral equations over its boundary. This means the computational domain reduces in dimension and thus the number of degrees of freedom required to approximate the solution to specified accuracy can be significantly smaller than in domain-based methods (as is indicated in Fig. 1.4). Additionally, for wave scattering, boundary integral formulations are attractive as they automatically incorporate a radiation condition at infinity, which needs to be imposed artificially for domain-based methods (for instance using the method of perfectly matched layers introduced by Berenger (1994)). Although these observations do not provide a complete picture of the advantages and disadvantages when compared to domain-based methods, in many settings the boundary integral equation method

can provide a competitive alternative to domain-based computations. Nevertheless, we highlight that boundary integral formulations also bring inherent challenges, for instance the fact that the discretisation matrices are dense (as opposed to sparse matrices arising in domain-based methods with basis functions of compact support) and involve singular integral operators whose numerical evaluation is difficult. There are ways of overcoming these problems, for instance by using specially designed numerical quadrature for these singular integral operators (cf. the work by Schwab and Wendland (1992) and our methods developed in Chapter 5), and fast multipole methods introduced by Rokhlin (1985) that result in linear discretisation systems which, albeit dense, can be solved at significantly reduced cost. At the same time, domain-based methods can perform very well for wave scattering problems, and the method of perfectly matched layers has also been successfully applied in a range of complex settings for example in recent work by Bonnet-Ben Dhia et al. (2016). For this thesis of particular relevance is the success that boundary integral equation methods have seen in the solution of high-frequency wave scattering problems, where so-called hybrid numerical-asymptotic schemes have been developed that can numerically approximate the solution of the relevant integral equations at cost that increases only very slowly or not at all in the frequency of the wave field Chandler-Wilde, Graham, Langdon and Spence (2012). We describe these hybrid numerical-asymptotic methods in some more detail below.



(a) Domain-based mesh.



(b) Boundary-based mesh.

Figure 1.4: Reduction of the dimension of the domain in boundary element methods.

The development of the boundary integral equation method followed on from profound insights by some of the most brilliant mathematicians over the past two centuries, including Green (1828), Fredholm (1903) and Ritz (1909), to name but a few. The history of the method is rich and complex, and it is impossible to point towards a single seminal paper that introduced the boundary integral equation method (also referred to as the boundary element method) as it is researched and known today (see Cheng and Cheng, 2005, for a

comprehensive review). However, we highlight several important contributions by Jaswon (1963), Symm (1963), Kupradze and Aleksidze (1964), and Rizzo (1967) who were amongst the first to directly exploit boundary integral formulations for computations in potential theory and elasticity. Their work helped spark a wider interest and research effort in this approach during the second half of the twentieth century (Cheng and Cheng, 2005). In the numerical analysis part of this thesis (Chapters 4 & 5) we focus on the solution of these boundary integral equations.

Once the PDE is put in integral form one can solve the resulting Fredholm integral equation using a range of techniques, including Galerkin, Nyström and collocation methods. From a theoretical perspective the Galerkin method appears to have many advantages including provable stability and convergence (Hsiao et al., 2017), and, in some cases, faster convergence rates compared to collocation (Arnold and Wendland, 1983). However, the optimality conditions in collocation methods (and even more so for fully discrete Nyström methods) are much easier to evaluate, which makes collocation an attractive alternative. Nevertheless, rigorous convergence analysis of collocation methods for integral equations is relatively scarce in the literature and restricted to isolated instances (Sloan, 1992). In Chapter 4 of this thesis, motivated by recent practical applications by Huybrechs and Olteanu (2019) and Gibbs et al. (2020), we take a non-standard point of view of collocation methods by introducing oversampling, i.e. by taking more collocation points than the dimension of the trial space. In the aforementioned practical applications it was found that oversampling can help improve the convergence properties of collocation methods, specifically by enhancing their robustness towards the choice of collocation points when compared to a non-oversampled version of the same method. As we shall see in Chapter 4 oversampling can, in some cases, even improve the asymptotic convergence rate of the collocation method. We provide an initial, but rigorous analysis that helps explain some of these observations in Chapter 4.

The specific challenge of high-frequency wave scattering has been subject to a large amount of research since the 1970s. One of the first successful approaches seeking to alleviate the problem of increased cost for larger frequencies was put forward by Burnside et al. (1975) in the context of electromagnetic wave scattering, who proposed to incorporate an Ansatz from the geometrical theory of diffraction in the numerical method. A similar approach was taken by Uncles (1976) for an example of acoustic plane wave scattering by a rigid sphere. In essence, the Ansatz ensures the numerical method only has to solve for slowly varying components of the field (i.e. the amplitudes), thus reducing the required size of the trial space and hence the cost at large frequencies enormously. This idea was taken forward and has evolved into so-called hybrid numerical-asymptotic boundary integral methods that incorporate some asymptotic knowledge of the field in the basis space to achieve near frequency-independent cost. A review of recent progress can be found in

Chandler-Wilde, Graham, Langdon and Spence (2012).

However, a choice of oscillatory trial functions also brings new challenges, specifically for the evaluation of the optimality conditions in the method. Amongst those challenges is the assembly of the collocation system for hybrid numerical-asymptotic methods which requires the computation of highly oscillatory integrals. Although highly oscillatory quadrature has seen considerable development over the past two decades (Deaño et al., 2017), there are still many open questions when it comes to applying efficient quadrature to hybrid numerical-asymptotic methods in wave scattering. Of particular interest for this thesis are Filon methods which are described in §1.2.3. These methods are generally advantageous because, if applied correctly, they allow the accurate computation of highly-oscillatory integrals at frequency-independent cost. However, in the context of hybrid numerical-asymptotic (HNA) wave-scattering, their application has not fully matured yet. Until now it has proven infeasible to construct Filon methods that compute the integrals relevant to HNA methods at uniform cost to fixed relative accuracy. This is an issue that we overcome for some individual cases of interest in Chapter 5.

1.2 Mathematical techniques and preliminaries

In this thesis we rely on a number of mathematical techniques, most importantly the Wiener–Hopf method, the boundary integral equation method, and highly oscillatory quadrature. Here we shall introduce the main features and important results available in the literature for each method.

1.2.1 The Wiener–Hopf technique

The technique was first described by Wiener and Hopf (1931) as a way to find exact solutions to integral equations with convolution type kernels in the context of the radiation equilibrium of stars. Comprehensive introductions can be found in the classical book by Noble (1958), in the thesis of Kisil (2015) and in the lecture notes by Abrahams and Aitken (2019). Here we provide an overview of the main methodology for the scalar case that will be used in Chapter 2 of this thesis. Throughout we shall focus on the ‘Jones’s method’ version of the Wiener–Hopf technique as described by Noble (1958).

The Wiener–Hopf technique provides a way to solve explicitly (or approximately) a class of linear partial differential equations (PDEs) with semi-infinite boundary conditions. The basic idea is to transform the linear PDE and work in the Fourier domain, in which case sometimes the original boundary value problem can be reduced to solving a scalar

and complex equation of the form:

$$K(\alpha)\Phi^+(\alpha) + \Psi^-(\alpha) + F(\alpha) = 0, \quad (1.2)$$

which holds in some strip $R^+ \cap R^-$ around the real axis, with

$$R^+ = \{\alpha \in \mathbb{C} \mid \text{Im}(\alpha) > -\epsilon\}, \quad R^- = \{\alpha \in \mathbb{C} \mid \text{Im}(\alpha) < \epsilon\},$$

for some $\epsilon > 0$ (see sketch in Fig. 1.5). In the following we will often refer to R^\pm as the ‘upper/lower half-plane’ respectively.

In Eq. (1.2) Φ^+ , Ψ^- are the unknowns and quantities with superscript ‘ \pm ’ are complex functions that are analytic on R^\pm , and that have at worst algebraic growth as $\alpha \rightarrow \infty$ in R^\pm , respectively. This means, more precisely, that there are constants $C_\pm > 0$ and $n \in \mathbb{N}$ such that

$$|\Phi^+(\alpha)| \leq C_+ |\alpha|^n \wedge |\Psi^-(\alpha)| \leq C_- |\alpha|^n, \text{ for all } \alpha \in R^\pm. \quad (1.3)$$

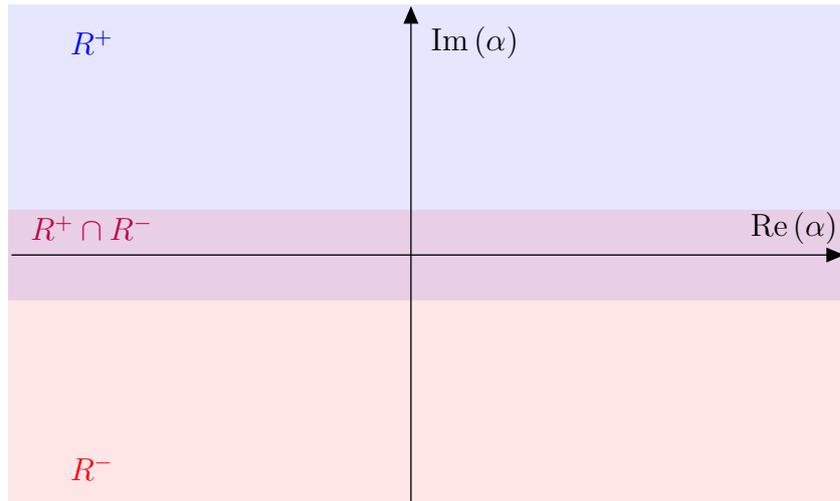


Figure 1.5: Complex half planes R^\pm and strip of overlap $R^+ \cap R^- \subset \mathbb{C}$.

Example 1.2.1. Consider the Helmholtz equation $(\Delta + k^2)u = 0$ on $\{(x, y) \mid y > 0\}$ with semi-infinite boundary conditions

$$\begin{aligned} u(x, 0) &= f(x), & x < 0, \\ \frac{\partial}{\partial y} u(x, 0) &= g(x), & x > 0. \end{aligned}$$

Taking the Fourier transform $U(\alpha, y) = \int_{-\infty}^{\infty} u(x, y) e^{i\alpha x} dx$, reduces the equation to

$$\frac{\partial^2}{\partial y^2} U(\alpha, y) - \gamma(\alpha)^2 U(\alpha, y) = 0,$$

where $\gamma(\alpha)^2 = \alpha^2 - k^2$ and we take the branch cuts of γ as shown in Fig. 1.6a, which is such that the real part of the function takes non-negative values on the real axis. More precisely we choose the following branch of the multivalued complex function γ :

$$\gamma(\alpha) = |\alpha + k|^{\frac{1}{2}} e^{i\frac{\theta_1}{2}} |\alpha - k|^{\frac{1}{2}} e^{i\frac{\theta_2}{2}},$$

where $\theta_1 = \arg(\alpha + k) \in [-3\pi/2, \pi/2)$ and $\theta_2 = \arg(\alpha - k) \in [-\pi/2, 3\pi/2)$. If we enforce $u(x, y)$ to be bounded as $y \rightarrow +\infty$ for any x (a radiation condition) then $U(\alpha, y) = A(\alpha) e^{-\gamma(\alpha)y}$. Restricting to the boundary at $y = 0$ yields the following condition

$$\frac{\partial U}{\partial y}(\alpha, 0) = -\gamma(\alpha) U(\alpha, 0).$$

In order to arrive at a complex equation valid on a strip $R^+ \cap R^-$ in the form of Eq. 1.2, we now introduce a small amount of artificial damping by taking $\text{Im } k = -\epsilon < 0$. This has the effect of moving the branch cuts of γ away from the real axis, thus creating a strip $R^+ \cap R^-$ on which γ is analytic and has non-negative real part, as indicated in Fig. 1.6b.

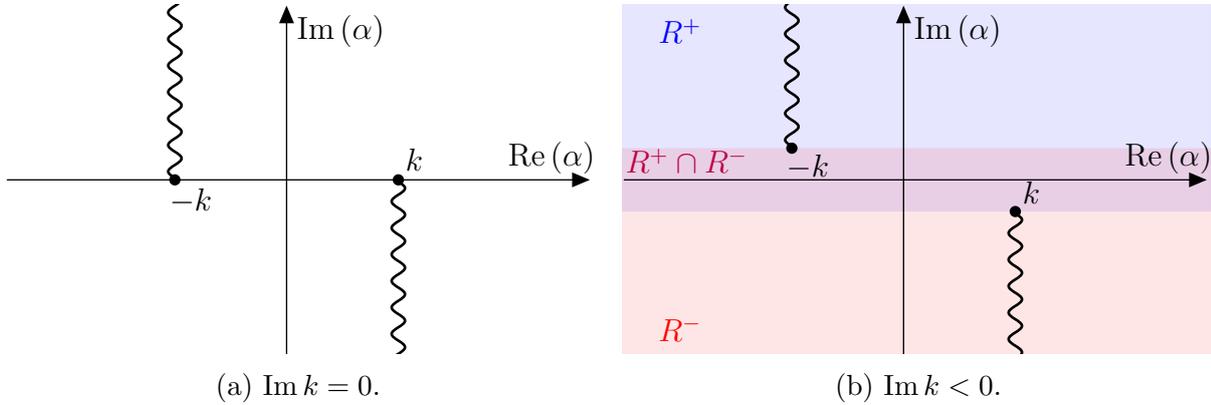


Figure 1.6: The branch cuts of γ .

Finally, imposing the boundary conditions results in a complex analytic equation on $R^+ \cap R^-$ with the desired form Eq. (1.2):

$$\gamma(\alpha) U^+(\alpha, 0) + \frac{\partial U^-}{\partial y}(\alpha, 0) + (\gamma(\alpha) F^-(\alpha) + G^+(\alpha)) = 0,$$

where we have used the following notations for the half-line Fourier transforms:

$$U^-(\alpha, y) := \int_{-\infty}^0 e^{i\alpha x} u(x, y) dx, \quad U^+(\alpha, y) := \int_0^{\infty} e^{i\alpha x} u(x, y) dx, \quad (1.4)$$

and we assumed the unknowns u and $\partial u/\partial y$ decay sufficiently fast as $x \rightarrow +\infty$ and $x \rightarrow -\infty$ respectively, such that their half-line Fourier transforms are well-defined and analytic functions in R^+ and R^- respectively.

The following theorem provides a set of sufficient conditions for the (half-line) Fourier transform of a function to be analytic in a complex strip or half-plane.

Theorem 1.2.2 (Thm. A in Noble (1958, p. 23)). *Let $f(x)$ be a function of the real variable x such that $|f(x)| \leq A \exp(\tau_- x)$ as $x \rightarrow +\infty$ and $|f(x)| \leq B \exp(\tau_+ x)$ as $x \rightarrow -\infty$, with $\tau_-, \tau_+ \in \mathbb{R}$ and $\tau_- < \tau_+$. Suppose the Fourier inversion theorem holds for $f(x)$, and define*

$$F(\alpha) = \int_{-\infty}^{\infty} f(x) e^{i\alpha x} dx. \quad (1.5)$$

Then $F(\alpha)$ is an analytic function of α on the domain $\text{Im } \alpha \in (\tau_-, \tau_+)$ and for any $\tau \in (\tau_-, \tau_+)$ we have

$$f(x) = \frac{1}{2\pi} \int_{i\tau-\infty}^{i\tau+\infty} F(\alpha) e^{-i\alpha x} d\alpha, \quad \forall x \in \mathbb{R}.$$

Note the special case when $\tau_{\pm} = \pm\infty$, implies that the Fourier half transform F^{\pm} as defined in Eq. (1.4) is analytic in $\text{Im } \alpha \gtrless \tau_{\pm}$. By the case $\tau_+ = +\infty$ we mean specifically that if $f(x) \equiv 0$ for any $x < X_0$, some $X_0 < 0$, and we have $|f(x)| \leq A \exp(\tau_- x)$ for $x \geq 0$, and if in addition the Fourier inversion theorem holds for $f(x)$, then $F(\alpha)$ defined as in Eq. (1.5) is an analytic function of α on the domain $\text{Im } \alpha \in (\tau_-, +\infty)$, and for any $\tau \in (\tau_-, +\infty)$ we have

$$f(x) = \frac{1}{2\pi} \int_{i\tau-\infty}^{i\tau+\infty} F(\alpha) e^{-i\alpha x} d\alpha, \quad \forall x \in \mathbb{R}.$$

The analogous meaning is assigned to the case $\tau_- = -\infty$. Moreover, by Noble (1958, Eq. (1.74)), if $f \in C^1((0, \infty))$, i.e. is continuously differentiable on $(0, \infty)$, and satisfies appropriate constraints on its behaviour at $x = +\infty$ (for instance it suffices that $|f(x)| \leq A \exp(\tau x)$ for some $\tau < 0$), then for any $\eta \in (-1, 1)$:

$$f(x) \sim Ax^{\eta} \text{ as } x \rightarrow 0^+ \Rightarrow F^+(\alpha) \sim A\Gamma(\eta+1)e^{\frac{1}{2}\pi i(\eta+1)}\alpha^{-\eta-1} \text{ as } \alpha \rightarrow \infty, \text{Im } \alpha > \tau_-. \quad (1.6)$$

The statement remains true when $\eta \in [1, \infty)$ as long as f has a sufficient number of derivatives on $(0, \infty)$ (with appropriate decay at $x = +\infty$), but for the purposes of this

thesis we shall only use the result in the case $\eta \in (-1, 1)$. This gives us a means to control the growth of F^+ at infinity, thus enforcing the crucial condition that F^+ is at worst of algebraic growth in its domain. The analogous statement is also true for F^- after inverting the signs. The Wiener–Hopf technique then prescribes the following procedure to find an analytic solution to Eq. (1.2): Firstly we factorise the so-called Wiener–Hopf kernel K by expressing it in the form

$$K(\alpha) = K^+(\alpha)K^-(\alpha) \text{ for all } \alpha \in R^+ \cap R^-, \quad (1.7)$$

where $K^+(\alpha), K^-(\alpha)$ are nonzero and analytic in the respective half-plane R^\pm and such that both K^\pm and $1/K^\pm$ are at worst of algebraic growth in the corresponding domain. Although in the scalar case this factorisation is no more challenging than the additive splitting described below, in the case when K is matrix valued (i.e. when the functions Φ^+, Ψ^- are vector valued) a constructive Wiener–Hopf factorisation is, in general, very challenging to perform. All Wiener–Hopf kernels considered in this thesis are scalar, thus we shall not discuss the matrix case in any detail. However, we highlight that the topic of matrix Wiener–Hopf factorisations offers many beautiful results, including classes of matrix kernels that can be factorised exactly (Daniele, 1978, Khrapkov, 1971) and approximate methods for matrix Wiener–Hopf factorisation (Abrahams, 1997, Kisil, 2018, Priddin et al., 2020). Having performed the multiplicative factorisation Eq. (1.7) we rewrite (1.2) as

$$K^+(\alpha)\Phi^+(\alpha) = -\frac{1}{K^-(\alpha)}\Psi^-(\alpha) - \frac{1}{K^-(\alpha)}F(\alpha).$$

The next step is to additively decompose the ‘forcing’ term:

$$-\frac{1}{K^-(\alpha)}F(\alpha) = A^+(\alpha) + B^-(\alpha), \quad \text{for all } \alpha \in R^+ \cap R^-, \quad (1.8)$$

where again A^+ is analytic and at worst of algebraic growth in R^+ and the same holds for B^- in R^- . Here a useful approach is given by an exact ‘Cauchy-type’ splitting.

Theorem 1.2.3 (Thm. B in Noble (1958, p. 13)). *Let $F(\alpha)$ be a complex function that is analytic in an open neighbourhood of the closed strip $\tau_- \leq 0 \leq \tau_+$ such that $|F(\alpha)| < C(1 + |\operatorname{Re} \alpha|)^{-p}$ for some $C, p > 0$, in $\tau_- \leq \operatorname{Im} \alpha \leq \tau_+$. Then, for $\tau_- < c < d < \tau_+$, we may define*

$$F^+(\alpha) = \frac{1}{2\pi i} \int_{-\infty+ic}^{\infty+ic} \frac{F(\xi)}{\xi - \alpha} d\xi, \quad F^-(\alpha) = -\frac{1}{2\pi i} \int_{-\infty+id}^{\infty+id} \frac{F(\xi)}{\xi - \alpha} d\xi,$$

which are such that

$$F(\alpha) = F^+(\alpha) + F^-(\alpha), \quad \text{when } \text{Im } \alpha \in (c, d).$$

Furthermore, F^+ is analytic in $\text{Im } \alpha > c$ and F^- is analytic in $\text{Im } \alpha < d$.

Using Eq. (1.8), we have that Eq. (1.2) is equivalent to

$$K^+(\alpha)\Phi^+(\alpha) - A^+(\alpha) = -\frac{1}{K^-(\alpha)}\Psi^-(\alpha) + B^-(\alpha), \quad \alpha \in R^+ \cap R^-. \quad (1.9)$$

In Eq. (1.9) the left hand side is defined and analytic on R^+ , and the right hand side is defined and analytic on R^- , so we can use the uniqueness of analytic continuation in the following form.

Theorem 1.2.4 (Analytic continuation, see Weisstein (2002, p. 73)). *Let f_1 and f_2 be analytic functions on domains Ω_1 and Ω_2 (i.e. Ω_1, Ω_2 are connected open subsets of \mathbb{C}), respectively, and suppose that the intersection $\Omega_1 \cap \Omega_2$ is not empty and that $f_1 = f_2$ on $\Omega_1 \cap \Omega_2$. Then f_2 is called an analytic continuation of f_1 to Ω_2 , and vice versa. Moreover, if it exists, the analytic continuation of f_1 to Ω_2 is unique.*

This allows us to uniquely define an entire function $E(\alpha)$ on the whole complex plane, by setting

$$E(\alpha) := \begin{cases} K^+(\alpha)\Phi^+(\alpha) - A^+(\alpha), & \text{if } \alpha \in R^+, \\ -\frac{1}{K^-(\alpha)}\Psi^-(\alpha) + B^-(\alpha), & \text{if } \alpha \in R^-. \end{cases}$$

At this point we note that, by the algebraic behaviour of the forcing terms A^+, B^- and of K^\pm , and by our assumptions on the unknowns Φ^+, Ψ^- in Eq. (1.3), that E has algebraic behaviour in the entire complex plane. Thus we can use the extended form of Liouville's theorem:

Theorem 1.2.5 (Liouville's theorem, see Noble (1958, p. 6)). *If $f(\alpha)$ is an entire function such that there exist real constants M, p with $|f(\alpha)| \leq M|\alpha|^p$ for all $\alpha \in \mathbb{C}$, then f is a polynomial of degree at most $\lfloor p \rfloor$.*

Note in particular if we have a bound of this form for E with $p < 1$, then E must be constant, and if $p < 0$ this constant must be zero. Additionally, for any value of $p \in \mathbb{R}$, if there is a particular direction along which $|E|$ decays to zero, then E must be identically zero. Therefore, in this case we can immediately conclude

$$\begin{aligned} \Phi^+(\alpha) &= \frac{A^+(\alpha)}{K^+(\alpha)} \text{ for all } \alpha \in R^+, \\ \Psi^-(\alpha) &= K^-(\alpha)B^-(\alpha) \text{ for all } \alpha \in R^-; \end{aligned}$$

thus we have solved the original problem. Even in the case when $p > 1$, or when the constant is nonzero, this method provides a solution of the original problem up to a finite (usually small) number of constants, which often can be determined by physical conditions or constraints. This is the basic idea of the Wiener–Hopf method for scalar equations which we will apply in Chapter 2.

1.2.2 Boundary integral equations

Comprehensive introductions to the topic of boundary integral equations can be found in the classical book by McLean (2000), and in the recent review papers by Chandler-Wilde, Graham, Langdon and Spence (2012) and by Hsiao et al. (2017). Of particular interest to this thesis is the exterior scattering problem in the two-dimensional case, with Dirichlet boundary conditions (i.e. sound-soft obstacles). Although we note that the Neumann problem, interior scattering and higher-dimensional scattering problems can also be expressed in a similar boundary integral formulation, we shall focus on the following particular case in to this thesis.

We assume that our scattering obstacle is represented by a bounded Lipschitz domain $\Omega \subset \mathbb{R}^2$. This means Ω is a bounded connected open subset of \mathbb{R}^2 whose boundary, $\Gamma = \partial\Omega$, is locally the graph of a Lipschitz function. We assume further that $\mathbb{R}^2 \setminus \Omega$ is connected (so that Ω is simply connected, i.e. does not have any holes), and that the domain Ω is only on one side of its boundary Γ . We wish to solve the following problem on $\Omega_+ := \mathbb{R}^2 \setminus \text{cl}\Omega$:

$$\begin{aligned} \Delta\phi(x) + k^2\phi(x) &= 0, \text{ for } x \in \Omega_+, \\ \phi(x) + \phi_{\text{inc}}(x) &= 0, \text{ for } x \in \partial\Omega_+, \end{aligned} \tag{1.10}$$

where ϕ_{inc} corresponds to the boundary values (the ‘trace’) of some incident field, and $\phi(x)$ is the unknown velocity potential of the scattered field. In the context of boundary integral equations we shall follow the convention of using plain x to denote coordinates in \mathbb{R}^2 (rather than bold-face \mathbf{x} used in the remainder of this thesis). This applies also to our treatment in Chapter 4. An important physical observation is that scattered waves must be outgoing at infinity, which means that for our compact scatterer we need to impose a radiation condition in order to ensure the solution is both physical and unique: The appropriate radiation condition in this case is the Sommerfeld radiation condition (Chandler-Wilde, Graham, Langdon and Spence, 2012, Eq. (2.9)). We say that ϕ satisfies the Sommerfeld radiation condition in two dimensions if

$$\lim_{r:=|x|\rightarrow\infty} \sup_{x/r \in \mathbb{S}^1} r^{1/2} \left(\frac{\partial}{\partial r} + ik \right) \phi(x) = 0, \tag{1.11}$$

where $\mathbb{S}^1 = \{x \in \mathbb{R}^2 \mid \|x\| = 1\}$ is the 1-sphere. In order to make rigorous statements about solutions to the scattering problem, we make use of the formalism of Sobolev spaces, which will specifically allow us to state the existence and uniqueness of solutions to Eq. (1.10).

Sobolev spaces and norms

Let us firstly recall the usual definition of $L^2(\mathbb{R}^2)$,

$$L^2(\mathbb{R}^2) := \left\{ f : \mathbb{R}^2 \rightarrow \mathbb{C} \text{ measurable} \mid \int_{\mathbb{R}^2} |f(x)|^2 dx < \infty \right\}.$$

Using the fact that the Fourier transform $\mathcal{F} : L^2(\mathbb{R}^2) \rightarrow L^2(\mathbb{R}^2)$ is a continuous isomorphism, we then define Sobolev spaces of positive real order, $s \in \mathbb{R}_{\geq 0}$, by

$$H^s(\mathbb{R}^2) := \{f \in L^2(\mathbb{R}^2) \mid \|f\|_{H^s} < \infty\},$$

where $\|f\|_{H^s(\mathbb{R}^2)} := \left(\int_{\mathbb{R}^2} (1 + |\xi|^2)^s |(\mathcal{F}u)(\xi)|^2 d\xi \right)^{\frac{1}{2}}.$

This also allows us to define Sobolev spaces of negative order as the continuous dual spaces, $H^s(\mathbb{R}^2) := (H^{-s}(\mathbb{R}^2))^*$, $s < 0$. From this definition of Sobolev spaces and norms on \mathbb{R}^2 we can define domain-based Sobolev spaces on Ω_+ by restriction. We define for $s \in \mathbb{R}$ and any smooth compactly supported $f \in C_{\text{comp}}^\infty(\text{cl } \Omega_+)$:

$$\|f\|_{H^s(\Omega_+)} := \inf_{\tilde{f} \in C^\infty(\mathbb{R}^2), \tilde{f}|_{\Omega_+} = f} \|\tilde{f}\|_{H^s(\mathbb{R}^2)}.$$

Then $H^s(\Omega_+)$ can be defined as the completion of $C_{\text{comp}}^\infty(\text{cl } \Omega_+)$ (i.e. functions with compact support in $\text{cl } \Omega_+$ that are infinitely differentiable up to the boundary Γ) with respect to the norm $\|f\|_{H^s(\Omega_+)}$. We will not discuss in detail the rich and interesting properties of domain-based Sobolev spaces, since for the purpose of this thesis our interest is focused on Sobolev spaces on the boundary. These spaces can be defined for general Lipschitz boundaries, but the definition is easier to formulate if the boundary can be covered by a single chart, i.e. if $\partial\Omega$ is the graph of a 1-periodic Lipschitz map with Lipschitz inverse

$$z : [0, 1) \rightarrow \Gamma \subset \mathbb{R}^2$$

$$t \mapsto (z_1(t), z_2(t)).$$

If the boundary Γ is not covered by a single chart, we can still find (by compactness) an open cover of finitely many charts and use a partition of unity splitting to extend the definitions from a single chart to the full boundary. This procedure is described in more detail in Chandler-Wilde, Graham, Langdon and Spence (2012, p. 278), and not repeated here in the interest of brevity. We define the boundary space $L^2(\Gamma)$ for measurable

functions on Γ in the usual sense, with the norm

$$\|g\|'_{L^2(\Gamma)} := \left(\int_{\Gamma} |g|^2 ds_{\Gamma} \right)^{\frac{1}{2}} = \left(\int_0^1 |g(z(t))|^2 |z'(t)| dt \right)^{\frac{1}{2}},$$

where ds_{Γ} is the line element on the boundary Γ , i.e. $ds_{\Gamma} = |z'(t)| dt$ in terms of the parametrisation z of the boundary. The integral is well-defined in the Lebesgue sense, since for Lipschitz functions $z'(t)$ exists almost everywhere and is essentially bounded. To begin with, we focus on the case when z is a smooth diffeomorphism meaning that z is bijective, infinitely differentiable as a periodic function on $[0, 1)$ and that $z'(t) \neq 0$ for all $t \in [0, 1)$. In this case the norm $\|\cdot\|'_{L^2(\Gamma)}$ is equivalent to the norm $\|\cdot\|_{L^2(\Gamma)}$ which we define as

$$\|g\|_{L^2(\Gamma)} := \|g \circ z\|_{L^2([0,1])} := \left(\int_0^1 |g(z(t))|^2 dt \right)^{\frac{1}{2}}.$$

Based on this, we could define weak derivatives and integer-order Sobolev norms on the boundary by requiring that the first s weak derivatives of g tangential to Γ lie in $L^2(\Gamma)$ in the above sense, resulting also in a norm $\|g\|'_{H^s(\Gamma)}$. However, since z is a diffeomorphism, Sobolev norms defined in this way are equivalent to the Sobolev norms of the mapped function on $[0, 1)$, i.e. for any $s \in \mathbb{R}$ there is a non-negative constant C_s (which depends on the boundary Γ), such that

$$C_s^{-1} \|g\|'_{H^s(\Gamma)} \leq \|g \circ z\|_{H^s([0,1])} \leq C_s \|g\|'_{H^s(\Gamma)}. \quad (1.12)$$

Note, for $g \in H^s$ with $s < 0$ (i.e. for g an element of the dual space of H^{-s}) we mean by $g \circ z$ the linear map on $H^{-s}([0, 1))$ defined by

$$g(\chi) := g(\chi \circ z^{-1}), \quad \forall \chi \in H^{-s}([0, 1)).$$

Thus we can equivalently define the Sobolev norms and spaces for any $s \geq 0$ as

$$\begin{aligned} H^s(\Gamma) &:= \{g : \Gamma \rightarrow \mathbb{C} \text{ measurable} \mid g \circ z \in H^s([0, 1))\}, \\ \|g\|_{H^s(\Gamma)} &:= \|g \circ z\|_{H^s([0,1])}, \end{aligned} \quad (1.13)$$

where we use the following definitions of the Sobolev norm and space on $[0, 1)$. We define $H^s([0, 1))$ to be the completion of $C^\infty([0, 1))$ under the Sobolev norm $\|\cdot\|_{H^s([0,1])}$, where

$$\|g\|_{H^s([0,1])} := \left(|\hat{g}_0|^2 + \sum_{0 \neq m \in \mathbb{Z}} |m|^{2s} |\hat{g}_m|^2 \right)^{\frac{1}{2}}, \quad (1.14)$$

and $\hat{g}_m := \int_0^1 e^{-2\pi i m t} f(t) dt$, $m \in \mathbb{Z}$, are the Fourier coefficients of the function g . The definition of $\|\cdot\|_{H^s([0,1])}$ makes sense both for positive and negative real values of s and, in particular, provides us with a way to evaluate the Sobolev norms for any $g \in L^1([0,1])$. We define negative order Sobolev spaces on the boundary as the continuous dual spaces, i.e. for any $s < 0$ we define $H^s(\Gamma) := (H^{-s}(\Gamma))^*$ where X^* denotes here the continuous dual of a Banach space. This definition (Eqs. (1.13)-(1.14)) of Sobolev spaces on the boundary Γ is valid also when z is only a Lipschitz function with Lipschitz inverse.

There is a natural correspondence between the spaces $H^s(\Omega_+)$ and $H^s(\Gamma)$ given by the exterior trace operator γ_+ . We define γ_+ for $f \in C_{\text{comp}}^\infty(\text{cl } \Omega_+)$ as

$$\gamma_+ f := f|_\Gamma.$$

As proved by Costabel (1988) (see also Chandler-Wilde, Graham, Langdon and Spence (2012, Eq. (A.17))), this is a bounded linear map in the following sense

$$\|\gamma_+ f\|_{H^{s-1/2}(\Gamma)} \leq C_s \|f\|_{H^s(\Omega_+)}, \quad \frac{1}{2} < s < \frac{3}{2},$$

whenever the parametrisation of the boundary, z , is Lipschitz. When z is a smooth diffeomorphism, the bound holds for all $s > \frac{1}{2}$. Thus we can extend γ_+ to a bounded linear map

$$\gamma_+ : H^s(\Omega_+) \rightarrow H^{s-1/2}(\Gamma),$$

for the aforementioned values of s , i.e. $1/2 < s < 3/2$ when z is Lipschitz and $s > 1/2$ when z is a smooth diffeomorphism. This is the final ingredient required to make rigorous statements about the solutions of the exterior scattering problem from Eq. (1.10).

Theorem 1.2.6 (Existence and uniqueness in the exterior scattering problem, see Chandler-Wilde, Graham, Langdon and Spence (2012, Thm. 2.10)). *Given $\phi_{\text{inc}} \in H^{\frac{1}{2}}(\Gamma)$, there is a unique $\phi \in C^2(\Omega_+)$ such that Eq. (1.10) holds, $\zeta \phi \in H^1(\Omega_+)$ for all $\zeta \in C_{\text{comp}}^\infty(\mathbb{R}^2)$, $\gamma_+ \phi = -\phi_{\text{inc}}$, and ϕ satisfies the radiation condition Eq. (1.11).*

Boundary integral formulation

Thus the (Dirichlet) boundary values of a solution to the Helmholtz equation on Ω_+ uniquely determine the value of the field everywhere in the domain. This abstract statement can be made more explicit in terms of so-called boundary layer potentials. Assume for the time being that Γ is smooth and that $\phi \in C^2(\text{cl } \Omega_+)$ solves $\Delta \phi + k^2 \phi = 0$ in Ω_+ . Then, by Green's second identity, we have (see for instance Thm. 2.6.2 in Huybrechs (2006)) for all

$x \in \Omega_+$

$$\phi(x) = \int_{\Gamma} \left(\frac{\partial \phi}{\partial n_y}(y) G(x, y) - \phi(y) \frac{\partial G}{\partial n_y}(x, y) \right) ds_y, \quad (1.15)$$

where $G(x, y) = \frac{i}{4} H_0^{(1)}(k|x - y|)$ is the Green's function of the Helmholtz equation with wavenumber k , and $\partial/\partial n_y$ is the normal derivative on the boundary Γ pointing outwards from Ω_+ , i.e. into the scatterer Ω . This direction is well-defined since we assumed that Ω is on one side of $\partial\Omega$ only. In the above $H_0^{(1)}$ denotes the Hankel function of the first kind of order zero (Abramowitz and Stegun, 1965, Eq. (9.1.3)). This motivates the introduction of the following layer potentials,

$$\begin{aligned} Su(x) &= \int_{\Gamma} G(x, y) u(y) ds_y, & x \in \Omega_+, \\ Du(x) &= \int_{\Gamma} \frac{\partial G}{\partial n_y}(x, y) u(y) ds_y, & x \in \Omega_+, \end{aligned}$$

for $x \in \Omega_+$. S and D are linear maps taking functions defined on the boundary Γ to functions defined on the full domain Ω_+ . They are continuous in the following sense: for any $\zeta \in C_{\text{comp}}^{\infty}(\mathbb{R}^2)$, denote by ζS the composition of S with pointwise multiplication by ζ , then the layer potentials have continuous extensions to bounded linear maps

$$\begin{aligned} \zeta S &: H^{s-\frac{1}{2}}(\Gamma) \rightarrow H^{s+1}(\Omega_+), \\ \zeta D &: H^{s+\frac{1}{2}}(\Gamma) \rightarrow H^{s+1}(\Omega_+), \end{aligned}$$

for $-1/2 \leq s \leq 1/2$ (Chandler-Wilde, Graham, Langdon and Spence, 2012, Thm. 2.15). Moreover, if $k > 0$, for any $u \in H^{-1}(\Gamma)$ the functions $Su, Du \in C^2(\mathbb{R}^2 \setminus \Gamma)$ and they satisfy Eq. (1.10) and Eq. (1.11) pointwise (Chandler-Wilde, Graham, Langdon and Spence, 2012, Thm. 2.14). Intuitively speaking, one might therefore expect to be able to express the solution $\phi(x)$ of Eq. (1.10) in the form $\phi(x) = (Su)(x)$ or $\phi(x) = (Du)(x)$ for some unknown function $u : \Gamma \rightarrow \mathbb{C}$, and then take an appropriate limit as x approaches the boundary in Eq. (1.15) to find an integral equation for u on the boundary. This is called the ‘indirect approach’ for a boundary integral formulation of the Helmholtz exterior problem and can be formalised as follows.

Firstly, we define the single- and double-layer operators for $u \in C(\Gamma)$ as

$$\begin{aligned} \mathcal{S}u(x) &= \int_{\Gamma} G(x, y) u(y) ds_y, & x \in \Gamma, \\ \mathcal{D}u(x) &= \int_{\Gamma} \frac{\partial G}{\partial n_y}(x, y) u(y) ds_y, & x \in \Gamma. \end{aligned}$$

These are continuous linear maps in that

$$\begin{aligned}\mathcal{S} &: H^{s-\frac{1}{2}}(\Gamma) \rightarrow H^{s+\frac{1}{2}}(\Gamma), \\ \mathcal{D} &: H^{s+\frac{1}{2}}(\Gamma) \rightarrow H^{s+\frac{1}{2}}(\Gamma),\end{aligned}$$

for all $|s| \leq 1/2$ (Chandler-Wilde, Graham, Langdon and Spence, 2012, Thm. 2.17). These boundary operators and potentials allow us to formulate the following Fredholm integral equation versions of the exterior scattering problem Eqs. (1.10) & (1.11).

Theorem 1.2.7 (Eq. (2.63) in Chandler-Wilde, Graham, Langdon and Spence (2012)). *Given $q \in H^{1/2}(\Gamma)$, we have $\phi = Su : \Omega_+ \rightarrow \mathbb{C}$ satisfies $\gamma_+\phi = q$ and Eqs. (1.10) & (1.11) pointwise if and only if*

$$Su = q \text{ in } H^{\frac{1}{2}}(\Gamma). \quad (1.16)$$

Theorem 1.2.8 (Eq. (2.64) in Chandler-Wilde, Graham, Langdon and Spence (2012)). *Given $q \in H^{1/2}(\Gamma)$, we have $\phi = Du : \Omega_+ \rightarrow \mathbb{C}$ satisfies $\gamma_+\phi = q$ and Eqs. (1.10) & (1.11) pointwise if and only if*

$$\left(\frac{1}{2}\mathcal{I} + \mathcal{D}\right)u = q \text{ in } H^{\frac{1}{2}}(\Gamma), \quad (1.17)$$

where $\mathcal{I} : u \mapsto u$ is the identity map.

Thus, in order to solve the exterior Dirichlet problem Eq. (1.10) we can instead solve the integral equations Eqs. (1.16) and (1.17). The solubility of these boundary integral equations can be summarised in the following way.

Theorem 1.2.9 (Thm. 2.25 in Chandler-Wilde, Graham, Langdon and Spence (2012)). *For $k > 0$ and Γ Lipschitz, we have for any $|s| \leq 1/2$ that the mappings*

$$\mathcal{S} : H^{s-\frac{1}{2}}(\Gamma) \rightarrow H^{s+\frac{1}{2}}(\Gamma) \quad (1.18)$$

$$\frac{1}{2}\mathcal{I} + \mathcal{D} : H^{s+\frac{1}{2}}(\Gamma) \rightarrow H^{s+\frac{1}{2}}(\Gamma) \quad (1.19)$$

are Fredholm of index zero (an introduction to Fredholm theory of operators on Banach spaces can be found in (Evans, 2010, Appendix D.5) and (Edmunds and Evans, 2018, §1)). Furthermore, for $|s| \leq 1/2$, \mathcal{S} is invertible between the spaces indicated above if and only if the interior homogeneous Dirichlet problem,

$$\begin{aligned}\Delta\phi(x) + k^2\phi(x) &= 0, \text{ for } x \in \Omega, \\ \phi(x) &= 0, \text{ for } x \in \Gamma,\end{aligned}$$

only has the trivial solution, and $\frac{1}{2}\mathcal{I} + \mathcal{D}$ is invertible between the spaces indicated above if and only if the interior homogeneous Neumann problem,

$$\begin{aligned}\Delta\phi(x) + k^2\phi(x) &= 0, \text{ for } x \in \Omega, \\ \frac{\partial}{\partial n}\phi(x) &= 0, \text{ for } x \in \Gamma,\end{aligned}$$

only has the trivial solution.

In summary, this Fredholm property means that the operators $\mathcal{S}, \frac{1}{2}\mathcal{I} + \mathcal{D}$ are continuous isomorphisms on the spaces indicated in Eqs. (1.18)-(1.19), away from a countable set of resonance frequencies $0 < k_1 < k_2 < \dots$ (which of course depend on Ω). If the boundary Γ is smooth, this result extends to all $s \in \mathbb{R}$. Therefore, away from resonances, the integral equations Eqs. (1.16) & (1.17) provide a suitable formulation to solve the exterior scattering problem. We note that the problem with resonances can be overcome by combined formulations, essentially taking a linear combination of the equations above, which can be constructed to have unique solutions for all values of $k > 0$. For the purpose of this thesis, the above integral formulations suffice, and we refer the interested reader to Chandler-Wilde, Graham, Langdon and Spence (2012, §2.6).

1.2.3 Highly oscillatory quadrature and Filon methods

The efficient approximation of highly oscillatory integrals is a notoriously difficult problem. In this section we provide a brief introduction to recent developments concerning efficient methods for the computation of such integrals. Of particular relevance to this thesis are the so-called Filon methods, and our attention shall therefore be focused on the relevant background of these methods. A comprehensive introduction to highly oscillatory quadrature in its many forms can be found in the recent monograph by Deaño et al. (2017).

Let us begin by outlining the difficulty that classical methods face when trying to approximate a highly oscillatory integral. Consider a finite range Fourier transform

$$I_\omega[f] := \int_{-1}^1 f(x)e^{i\omega x} dx,$$

where we take, for simplicity, f to be an infinitely differentiable real valued function $f : [-1, 1] \rightarrow \mathbb{R}$. Suppose we apply the trapezoidal rule with $\nu + 2$ points, denoted by

$$\mathcal{Q}_T^{[\nu]}[f(x)\exp(i\omega x)] = \frac{1}{\nu+1} \sum_{n=0}^{\nu+1}{}'' f\left(\frac{2n}{\nu+1} - 1\right) \exp\left(i\omega\left(\frac{2n}{\nu+1} - 1\right)\right), \quad (1.20)$$

where \sum'' means that the first and final terms in the sum are halved. This approximation arises from interpolating the integrand $f(x)\exp(i\omega x)$ with a piecewise linear function as

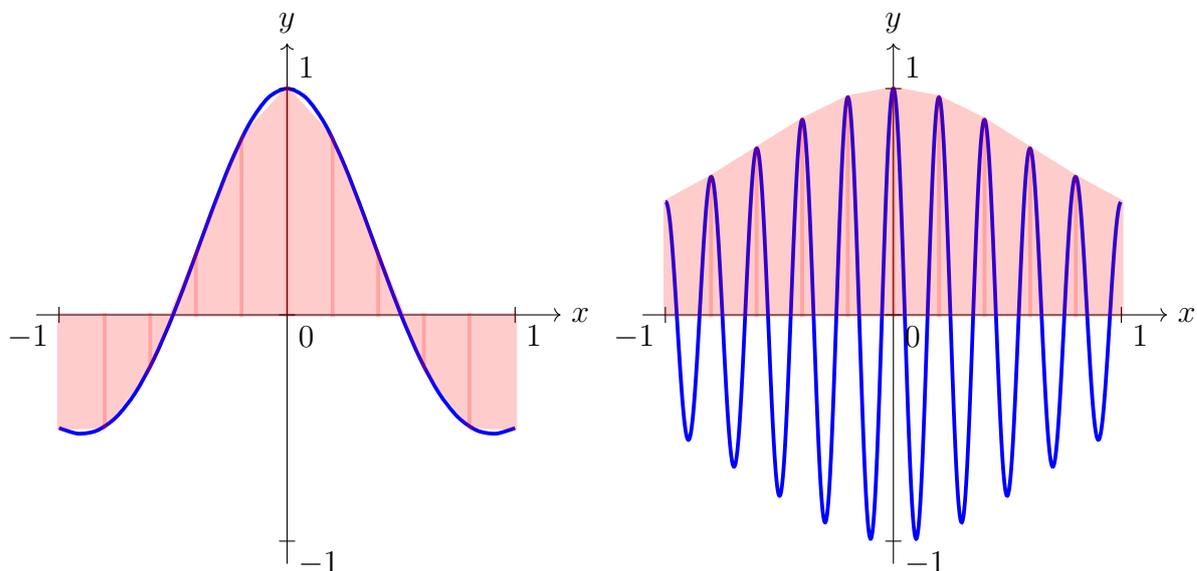
sketched for the real part of the integrand in Fig. 1.7a, and then integrating this piecewise linear function exactly. In the example of Fig. 1.7 we took $f(x) = 1/(1+x^2)$ and fixed $\nu = 9$. As can be seen in Fig. 1.7b, when the oscillations of the integrand increase for fixed ν , the piecewise linear interpolation will fail to resolve the oscillations accurately. More precisely we note for fixed ν we may consider the discrete set of frequencies $\omega = 2(\nu+1)\pi\omega_0$, $\omega_0 \in \mathbb{N}$ for which the approximation Eq. (1.20) becomes

$$\mathcal{Q}_T^{[\nu]}[f(x) \exp(i2\pi(\nu+1)\omega_0 x)] = \frac{1}{\nu+1} \sum_{n=0}^{\nu+1} f\left(\frac{2n}{\nu+1} - 1\right),$$

and is hence independent of ω_0 . On the other hand we know by the Riemann-Lebesgue lemma that

$$I_{2\pi(\nu+1)\omega_0}[f] \rightarrow 0 \quad \text{as } \omega_0 \rightarrow \infty.$$

Thus there is a clear mismatch between the asymptotic behaviour of the integral $I_\omega[f]$ and the trapezoidal rule $\mathcal{Q}_T^{[\nu]}[f(x) \exp(i\omega x)]$ as $\omega \rightarrow \infty$. The only way to overcome this problem in the trapezoidal rule is to increase ν as we increase ω (because our only free parameter is the number of quadrature nodes). A similar problem appears in classical Gaussian quadrature, where the full oscillatory integrand is interpolated by a polynomial (instead of a piecewise linear function) which is then integrated exactly: in order to resolve the oscillations accurately the degree of the polynomial (and hence of the quadrature nodes and the cost) has to increase with ω .



(a) For $\omega = \pi$ the approximation is accurate. (b) For $\omega = 10\pi$ the oscillations are not resolved.

Figure 1.7: The trapezoidal rule ($\nu = 9$) applied to the real part of $I_\omega[1/(1+x^2)]$.

To overcome this increase in cost we need to construct quadrature rules that match the asymptotic behaviour of $I_\omega[f]$: By the method of stationary phase we have (Deaño et al., 2017, Eq. (2.1))

$$I_\omega[f] \sim - \sum_{k=0}^{\infty} \frac{1}{(-i\omega)^{k+1}} [f^{(k)}(1)e^{i\omega} - f^{(k)}(-1)e^{-i\omega}], \quad \omega \rightarrow \infty, \quad (1.21)$$

where we used the symbol \sim to denote that the right hand side is a Poincaré expansion of $I_\omega[f]$ as $\omega \rightarrow \infty$, i.e. for all $N \in \mathbb{N}$

$$\left| I_\omega[f] + \sum_{k=0}^{N-1} \frac{1}{(-i\omega)^{k+1}} [f^{(k)}(1)e^{i\omega} - f^{(k)}(-1)e^{-i\omega}] \right| = \mathcal{O}(\omega^{-(N+1)}), \quad \omega \rightarrow \infty.$$

Thus one notices that the asymptotic behaviour of $I_\omega[f]$ is uniquely determined by the function- and derivative-values of f at the endpoints of the domain of integration, $\{f^{(k)}(\pm 1)\}_{k=0}^{\infty}$. Whilst a construction of the following form was first described by Filon (1930) it was realised much later by Iserles (2004, 2005) and Iserles and Nørsett (2004) that this observation can be used to construct a quadrature for $I_\omega[f]$ that is uniformly accurate as the frequency increases: For this we firstly construct p , an interpolating polynomial of degree $\nu + 1$ satisfying the following interpolation conditions

$$p(\pm 1) = f(\pm 1), \quad \text{and} \quad p(c_l) = f(c_l), \quad l = 1, \dots, \nu, \quad (1.22)$$

for some specified interior interpolation points $-1 < c_1 < \dots < c_\nu < 1$. Then the Filon quadrature rule is defined as

$$\mathcal{Q}_\omega^{[\nu]}[f] := I_\omega[p] = \int_{-1}^1 p(x)e^{i\omega x} dx.$$

Suppose for a moment that we can compute $I_\omega[p]$ exactly. Then, by Eq. (1.21) it is easy to see that the first term in the Poincaré series of the quadrature rule matches the one from the integral $I_\omega[f]$, and therefore

$$|I_\omega[f] - \mathcal{Q}_\omega^{[\nu]}[f]| = \mathcal{O}(\omega^{-2}), \quad \omega \rightarrow \infty.$$

This means for a fixed number of Filon quadrature nodes ν the error in the Filon method actually decays as the frequency increases. While this behaviour could also be achieved by approximating $I_\omega[f]$ with the first term in the asymptotic series in Eq. (1.21), the interior interpolation points c_1, \dots, c_ν ensure that there is a uniform convergence even for moderate frequencies since integration by parts allows us to show estimates of the following form (essentially equivalent to Lemma 1.3 in (Melenk, 2010)).

Proposition 1.2.10 (Filon paradigm for I_ω). *Let $f \in C^2([0, 1])$ and be p the unique polynomial of degree $\nu+1$ satisfying Eq. (1.22). Then there is a constant $C > 0$, independent of $f, \{c_j\}_{j=1}^\nu$, such that:*

$$|I_\omega[f] - \mathcal{Q}_\omega^{[\nu]}[f]| = |I_\omega[f] - I_\omega[p]| \leq C\omega^{-2} (\|f' - p'\|_\infty + \|f'' - p''\|_\infty),$$

for all $\omega > 0$.

This reduces the error in the Filon method to the uniform error (of derivatives) of interpolation at the points $-1 = c_0 < c_1 < \dots < c_\nu < c_{\nu+1} = 1$. There are a number of ways to estimate $\|f^{(j)} - p^{(j)}\|_\infty$. One possibility is via the Hermite interpolation formula as was used by Melenk (2010). Another way is to relate the error to the regularity of f in periodic Sobolev norms on $[0, 2\pi]$ via the change of variable $x = \cos \theta$, this approach was taken by Domínguez et al. (2011). Finally, in our opinion, a very elegant way is via optimal error bounds using the Peano kernel theorem; in particular we can use the following result due to Shadrin (1995): Define the nodal polynomial for the interpolation problem Eq. (1.22) as $\tilde{r}(x) = (x^2 - 1) \prod_{j=1}^\nu (x - c_j)$, then we have the following bounds (where the constants in the inequalities are optimal over $f \in C^{\nu+1}([-1, 1])$):

$$\|f^{(j)} - p^{(j)}\|_\infty \leq \frac{\|\tilde{r}^{(j)}\|_\infty}{(\nu + 1)!} \|f^{(\nu+1)}\|_\infty. \quad (1.23)$$

We can combine Eq. (1.23) with Prop. 1.2.10 and the trivial estimate

$$|I_\omega[f] - \mathcal{Q}_\omega^{[\nu]}[f]| \leq \int_{-1}^1 |f(x) - p(x)| dx \leq 2\|f - p\|_\infty \quad (1.24)$$

to find:

Corollary 1.2.11. *For any $f \in C^\infty([-1, 1])$, $\nu \in \mathbb{N}$, $\omega > 0$:*

$$|I_\omega[f] - \mathcal{Q}_\omega^{[\nu]}[f]| \leq \min \{2\|\tilde{r}\|_\infty, C\omega^{-2} (\|\tilde{r}'\|_\infty + \|\tilde{r}''\|_\infty)\} \frac{\|f^{(\nu+1)}\|_\infty}{(\nu + 1)!}, \quad (1.25)$$

where $C > 0$ is a constant independent of $f, \{c_j\}_{j=1}^\nu, \nu$, and ω .

Proof. From Eq. (1.23) we have

$$\|f - p\|_\infty \leq \|\tilde{r}\|_\infty \frac{\|f^{(\nu+1)}\|_\infty}{(\nu + 1)!}.$$

Thus, we can combine this with Eq. (1.24) to find the first bound on the right hand side

of Eq. (1.25):

$$|I_\omega[f] - \mathcal{Q}_\omega^{[\nu]}[f]| \leq 2\|f - p\|_\infty \leq 2\|\tilde{r}\|_\infty \frac{\|f^{(\nu+1)}\|_\infty}{(\nu+1)!}. \quad (1.26)$$

To find the second estimate on the right hand side of Eq. (1.25) we proceed analogously by combining Eq. (1.23) with Prop. 1.2.10. \square

In essence the estimate in Prop. 1.2.10 and Corollary 5.5.3 guarantees, as long as the points $\{c_j\}_{j=1}^\nu$ are suitably spaced, good convergence properties of the method uniformly in frequency. For example if f is analytic in a neighbourhood of the complex unit ball $\{|z| \leq 1\}$ then Cauchy's integral formula yields the following bound on the derivatives of f for any $x \in [-1, 1]$:

$$|f^{(\nu+1)}(x)| = \left| \frac{(\nu+1)!}{2\pi i} \oint_{|z|=1} \frac{f(z)}{(z-x)^{\nu+2}} dz \right| \leq (\nu+1)! \sup_{|z|=1} |f(z)|,$$

where the integral is taken counter-clockwise over the unit circle. Suppose we take $\{c_j\}_{j=0}^{\nu+1}$ to be Legendre points, i.e. the zeros of the Legendre polynomial $P_{\nu+2}$, then $\tilde{r}(x)$ is a normalised Legendre polynomial with the normalisation (cf. Abramowitz and Stegun, 1965, Eq. 22.3.8)

$$\tilde{r}(x) = \frac{(\nu+2)!}{2^{\nu+2} \left(\frac{1}{2}\right)_{\nu+2}} P_{\nu+2}(x), \quad (1.27)$$

where $(1/2)_{\nu+2}$ denotes the Pochhammer symbol. It is well known (cf. Abramowitz and Stegun, 1965, Eq. 22.14.1) that $\|P_n\|_\infty \leq 1$ for all $n \in \mathbb{N}$, and moreover we have (as a direct consequence of Rodrigues' formula (cf. Abramowitz and Stegun, 1965, Eq. 22.11.5))

$$\begin{aligned} \frac{d}{dx} P_{n+1} &= (n+1)P_n + \frac{d}{dx} P_n, \quad \text{for } n \geq 0, \\ \frac{d^2}{dx^2} P_{n+1} &= (n+2) \frac{d}{dx} P_n + \frac{d^2}{dx^2} P_n, \quad \text{for } n \geq 1. \end{aligned}$$

These two identities allow us to show the following bounds by induction, which are valid for all $\nu \in \mathbb{N}$:

$$\|P'_{\nu+2}\|_\infty \leq \frac{(\nu+2)(\nu+3)}{2}, \quad \|P''_{\nu+2}\|_\infty \leq \frac{(\nu+2)(\nu+3)(\nu+4)(\nu+5)}{8}.$$

Combining these estimates with Eq. (1.27) and the simple estimate

$$\frac{(\nu+2)!}{\left(\frac{1}{2}\right)_{\nu+2}} = \frac{(\nu+2)(\nu+1) \cdots 1}{\left(\nu + \frac{3}{2}\right) \left(\nu + \frac{1}{2}\right) \cdots \frac{1}{2}} \leq 2(\nu+1)$$

we find for all $\nu \in \mathbb{N}$

$$\begin{aligned}\|\tilde{r}\|_\infty &\leq 2^{-(\nu+1)}(\nu+2), \\ \|\tilde{r}'\|_\infty &\leq 2^{-\nu}(\nu+2)^2(\nu+3), \\ \|\tilde{r}''\|_\infty &\leq 2^{-\nu+2}(\nu+2)^2(\nu+3)(\nu+4)(\nu+5).\end{aligned}$$

Therefore, for functions f analytic in a neighbourhood of the complex unit disk, Corollary 5.5.3 guarantees that there is a constant $C > 0$ independent of ν, ω such that for all $\omega \geq 0$ and $\nu \in \mathbb{N}$:

$$|I_\omega[f] - \mathcal{Q}_\omega^{[\nu]}[f]| \leq C2^{-\nu} \min\{\nu, \omega^{-2}\nu^5\} \sup_{|z|=1} |f(z)|,$$

i.e. such that we have exponential convergence in ν uniformly in the frequency ω . Moreover, for fixed ν the error decreases in ω , i.e. is smaller for larger frequencies. Note this type of analysis can be carried out also for other interpolation points or classes of functions. For instance Melenk (2010) provided results concerning the behaviour of the Filon error for amplitudes f that are analytic in just a small complex neighbourhood of the interval $[-1, 1]$ by applying the Hermite interpolation formula directly to the right hand side of Prop. 1.2.10.

This idea of constructing the Filon method by matching asymptotic behaviour of exact integral also extends to more complicated oscillators. For instance integrals of the form

$$I_\omega[f; g] = \int_{-1}^1 e^{i\omega g(x)} f(x) dx$$

can also be approximated efficiently by taking p as an interpolating polynomial that matches the values of f at ± 1 and at any point $x \in (-1, 1)$ with $g'(x) = 0$ (i.e. the stationary points of the integrand). The corresponding Filon method is defined by

$$\mathcal{Q}_{\omega, g}^{[\nu]}[f] := I_\omega[p; g] = \int_{-1}^1 e^{i\omega g(x)} p(x) dx.$$

By matching the values of f at the endpoints and stationary points, the Filon method is again seen to match the asymptotic behaviour of $I_\omega[f, g]$, and hence results in an error that decays in ω and intermediate interpolation points for p can again be added to ensure uniform convergence. An interesting further development of this methodology is the extended Filon method first described by Iserles and Nørsett (2005), which achieves an even better match of the asymptotic behaviour of the integral by interpolating derivative values of f . For instance an extended Filon method for $I_\omega[f]$ can be constructed by taking q to be the unique polynomial of degree $2s + \nu + 1$ satisfying the Hermite-type interpolation

conditions

$$q^{(j)}(\pm 1) = f^{(j)}(\pm 1), \quad j = 0, \dots, s, \quad \text{and} \quad q(c_l) = f(c_l), \quad l = 1, \dots, \nu, \quad (1.28)$$

where as above $-1 < c_1 < \dots < c_\nu < 1$. By Eq. (1.21) this leads to an asymptotic error of the form

$$|I_\omega[f] - \mathcal{Q}_\omega^{[\nu]}[f]| = \mathcal{O}(\omega^{-(s+2)}), \quad \omega \rightarrow \infty.$$

The idea extends also to multivariate integrals (see Deaño et al. (2017, §3.2)) but our interest in this thesis remains focused on the univariate case. A central question we address in Chapter 5 is how one may compute $I_\omega[p]$ efficiently and stably for integrals with complicated oscillatory kernel functions. Finally, we highlight that there are number of interesting and efficient alternatives to Filon methods, including the method of numerical steepest descent developed by Huybrechs and Vandewalle (2006) and Levin methods introduced by Levin (1982, 1996) the latter of which we will briefly discuss in Chapter 6.

1.3 Thesis structure

We now provide an overview of the structure and the main contributions in each chapter of this thesis. The novel scientific content is provided in Chapters 2–6 and for each of those chapters we include a list of symbols (notation may change between chapters due to differing mathematical conventions) which is located after the conclusion of the chapter and before any appendices. A few concluding remarks and an outlook towards future research are provided in the final Chapter 7.

Chapter 2

In this chapter, we consider the scattering of waves by an infinitely extending three-dimensional cascade of finite-length flat blades in subsonic flow at zero angle of attack. As mentioned in §1.1.1 this geometry is of specific relevance as it provides a model for the components in turbofan engines. We study the scattering problem analytically, considering both acoustical and vortical incident fields, spanwise wavenumbers and transverse mean flow.

Our main contributions are a Wiener–Hopf solution of the scattering problem which lifts the restriction that adjacent blades overlap, a condition that had thus far been crucial for the (semi-)analytical study of this problem. Our approach relies on the solution of three coupled scalar Wiener–Hopf problems, corresponding to an uncoupled leading-edge approximation, and a subsequent trailing-edge and leading-edge correction. We provide

closed-form expressions for the far-field behaviour of the scattered velocity potential upstream and downstream of the cascade as well as for the total unsteady lift on each blade in the cascade, depending only on the solution of an infinite matrix equation. Finally, we provide a rigorous convergence analysis guaranteeing that the matrix system can be truncated to find numerical values of this solution.

Chapter 3

In this chapter, we study the balance of outgoing and incoming power for acoustic wave scattering by the cascade of flat blades in uniform subsonic mean flow based on the solution we found in Chapter 2. In nonzero mean flow the Kutta condition at the trailing edge of the cascade plates leads to the production of vorticity and the appearance of unsteady vortex sheets attached at the trailing edges, which results in a hydrodynamic power flux and a mechanism for energy conversion between the unsteady vorticity and the acoustic field. Although the scattering problem itself has been subject to extensive previous research, a comprehensive study of the composition of the outgoing energy flux for this geometry has not previously been conducted.

Our main contribution is a comprehensive study of this composition and we use energy balance arguments to provide analytical proofs of certain symmetries of the field with respect to the angle of incidence relative to the cascade face when there is no mean flow and of the effect of zero acoustic reflection at certain angles of incidence in the case of mean flow. The latter is a new result that holds for all frequencies (i.e. also beyond the first modal cut-on). We provide several numerical experiments covering a range of parameter values that show significant attenuation of sound at low frequencies and demonstrate that the vortex shedding can lead to an increase in sound power output over the incident field. Our numerical results also demonstrate that, for certain parameter values, this can result in a reflected acoustic wave with amplitude larger than the amplitude of the incident wave.

Chapter 4

In this chapter, we explore the extent to which oversampling can improve both robustness and convergence properties of collocation methods when applied as a method of solving Fredholm integral equations. Our interest is motivated by considering boundary integral equation formulations of the exterior Dirichlet problem of the Helmholtz equation.

Our main contributions in this chapter are several results which form the first rigorous convergence analysis for least-squares collocation methods. We demonstrate how oversampling impacts the convergence properties of the method in specific settings. This includes the case of optimal, equispaced collocation points for smooth scatterers, for which we show that already oversampling by a constant factor J (leading to a rectangular

system of size $JN \times N$) leads to significant improvement in the error, and that quadratic oversampling leads to a convergence rate that is faster than in the standard Galerkin formulation. We also provide rigorous results that guarantee convergence in several more general cases including the case of suboptimal irregularly-spaced collocation points and Lipschitz domains. Our results are illustrated using a number of numerical examples which show that the predicted amount of oversampling required to achieve certain convergence rates is close to optimal.

Chapter 5

In this chapter, we study the efficient approximation of highly oscillatory integrals using Filon methods, motivated by the desire to efficiently assemble the discretisation matrix for collocation methods based on a hybrid numerical-asymptotic basis in high-frequency wave scattering. A crucial step in the implementation of these methods is the accurate and fast computation of the Filon quadrature moments.

Our main contributions in this chapter include a general framework that shows how moment recurrences can be constructed for a wide class of oscillatory kernel functions, based on the observation that many physically relevant kernel functions are in the null space of a linear differential operator whose action on the Filon interpolation basis is represented by a banded (infinite) matrix. We discuss in further detail the application to two classes of particular interest: integrals with algebraic singularities and stationary points and integrals involving a Hankel function. We provide rigorous stability results for the moment computation for the first of these classes and demonstrate how the corresponding Filon method results in an accurate approximation at truly frequency-independent cost. For the Hankel kernel, we derive a Filon paradigm describing the convergence behaviour of the method in terms of the frequency and the number of Filon quadrature points. Finally, we show how Filon methods with recursive moment computation can be applied to compute efficiently integrals arising in hybrid numerical-asymptotic collocation methods for high-frequency wave scattering on a screen.

Chapter 6

In this chapter, we study a further application of a result proved in Chapter 5, specifically concerning the recursive computation of Chebyshev coefficients of functions in the null-space of certain differential operators.

Our main contribution in this chapter is the development of an efficient Levin method, which we term the Levin–Clenshaw–Curtis method and which allows us to approximate certain highly oscillatory integrals at frequency independent cost at rapid speed comparable to a fast implementation of Clenshaw–Curtis quadrature. Although the permissible

oscillatory kernels that can be treated with this method are fewer than for the Filon method discussed in Chapter 5, the Levin method removes the need for moment computation and hence presents an attractive and efficient alternative. Our Levin–Clenshaw–Curtis method is able to reduce the cost of a usual Levin method from $\mathcal{O}(\nu^3)$ to just $\mathcal{O}(\nu \log \nu)$ operations, where ν represents the analogue of the number of quadrature points in the Levin method. We provide several numerical examples demonstrating the favourable properties of this approach.

Chapter 2

Wave scattering by an infinite cascade of non-overlapping blades

2.1 Introduction

In this chapter of the thesis we consider a Wiener–Hopf solution to the scattering of vortical and acoustic waves incident on a cascade of blades which, as introduced in §1.1.1, serves as a model for fan components in turbomachinery. As indicated in the introductory chapter, this scattering problem has been subject to a large amount of research over the past decades. Before describing our new contributions let us provide a slightly more detailed overview of past work. The groundwork for the analytical solution of these types of problems was laid in a sequence of works by Carlson and Heins (1947), Heins and Carlson (1947) and Heins (1950), who considered an analogous problem of electromagnetic scattering by a cascade of perfectly conducting semi-infinite blades extending to $x = +\infty$. One of the earliest works considering a cascade of finite-length blades was presented by Kaji and Okazaki (1970), who considered a sound wave incident on the cascade and developed a numerical scheme based on a distribution of doublets to match the pressure jump on the blades. The problem was further studied by Whitehead (1970), who developed the ‘LINSUB’ code, which provides a numerical scheme to solve for a variety of incident fields on a cascade located in subsonic mean-flow. This approximate solution is based on expanding the full pressure field in terms of separable waves in the Prandtl–Glauert plane and using a collocation type approach to match the upwash velocity at the blades. A more analytically-based approach was given by Koch (1971), who used a scalar Wiener–Hopf formulation (of similar nature to Carlson and Heins (1947)) to solve the problem of incident sound waves exactly (up to the truncation of an infinite system of linear equations) in the case of overlapping blades, thus allowing the computation of the scattered radiation in this case.

More recent work has focused on extending these results to incorporate more general

geometries and incident fields, as well as to allow for a more efficient, and hence practically feasible solution. Peake (1992) derived an asymptotic kernel factorisation for the large reduced frequency regime and used a small number of iterates in a Schwarzschild-type approach (Schwarzschild, 1901) to produce an efficient scheme for an incident vorticity wave, which allows for the computation of the unsteady lift distribution. This asymptotic kernel factorisation was later extended by Peake and Kerschen (1995) to be uniformly valid even for close to cut-off conditions of the radiation modes or duct modes. A subsequent work by Peake (1993) solves the complete system exactly using a Wiener–Hopf type approach and presents an expression for the unsteady lift distribution that is valid for arbitrary reduced frequency. In both cases the method of solution relies on an assumption of overlap of adjacent blades in order to arrive at a coupling of the Wiener–Hopf equations that describe the leading-edge and trailing-edge interaction problems. Glegg (1999) further extended the analysis to blades in a three-dimensional setting, including the effects of spanwise wave number and cross-flow (which, as mentioned in the introductory chapter, allow for the modelling of chord-wise effects in a rotor/stator blade row if the hub-tip ratio of the fan is close to unity), and analysed in detail the unsteady loading and far-field behaviour of the scattered field for an incident vortical gust. Glegg’s method of solution can be motivated by a Schwarzschild-type distribution of boundary conditions, and results in a system of four coupled scalar Wiener–Hopf equations, which takes a similar form to the one derived by Peake (1993). These can then be solved, and an exact expression for the solution can be found which relies only on the solution of an infinite linear matrix equation.

In recent years several studies have succeeded in analysing the effect of realistic blade geometry, periodicity and mean flow on the cascade scattering problem. In particular, Peake and Kerschen (1997, 2004) studied the effect of blade mean loading on the generation of noise in the case of mean flow that is aligned at a nonzero incidence angle to the cascade blades. Evers and Peake (2002) included the effect of small but nonzero camber and thickness in a cascade of airfoils. The analytical studies of the cascade also provided a basis for Posson et al. (2010), who were able to use the two-dimensional cascade solution as a source distribution for the scattering problem on an annular blade-row. Ayton and Peake (2013) used the same velocity potential and streamfunction coordinate system as had been applied in Peake and Kerschen (1997, 2004) to analyse realistic airfoil shapes by reduction to flat blades and appropriate boundary conditions. This coordinate system was successfully used by Baddoo and Ayton (2020*b*) to find an exact solution to the scattering problem by an infinite cascade of airfoils again based on the solution presented by Glegg (1999). In a related approach Baddoo and Ayton (2020*a*) were able to include more complex boundary conditions such as porosity of the blades.

However, in all of the aforementioned work which uses the Wiener–Hopf technique

to find either exact or approximate solutions to the scattering problem, the restriction to the case of overlapping blades is required. This restriction was necessary in previous studies because it effectively allows the Wiener–Hopf analysis to be based on the duct modes, which are the only modes present in overlapping parts of the geometry. If the blades do not overlap the modal structure is more complicated and the Wiener–Hopf solution needs to account for that. The main novelty presented in this chapter is the first Wiener–Hopf solution of the problem which does not rely on overlapping blades and is in fact valid for arbitrary blade spacing. This is achieved by a formulation as a system of coupled scalar Wiener–Hopf equations, and by judiciously choosing an appropriate additive Wiener–Hopf splitting (which is described in detail in §2.4.4), that allows us to reduce the scattering problem to an infinite system of linear equations. Our additive splitting is based on pole removal at the radiation modes, which leads to a discrete linear system with decaying coefficients that can be solved by truncation regardless of the choice of overlap. In contrast, the coefficients of the linear systems in previous work using the Wiener–Hopf method would increase exponentially when the blades do not overlap, so do not provide a convergent method in this case. We derive our solution in a very general three-dimensional setting, allowing for effects of spanwise wave number and cross-flow, as well as considering both incident vorticity and acoustic waves. This solution allows us to derive exact expressions for the field away from the cascade structure and the unsteady lift on the blades, which, similarly to the previous work, are in closed form apart from the need to solve the aforementioned linear system numerically. Albeit we focus our attention on the case of flat plates and meanflow that is aligned with the plates, we expect our methodology can be extended to the case of nonzero angles of attack and realistic airfoil shapes of the blades using the aforementioned velocity potential and streamfunction coordinate system exploited by Peake and Kerschen (1997, 2004). We comment in more detail on possible extensions of our work in §7.1. These expressions form the basis for further study of the energy balance in this system as described in Chapter 3.

We begin this chapter by outlining the equations of motion together with the relevant boundary conditions in §2.2. This is followed by a formulation of the scattering problem as a system of coupled Wiener–Hopf equations in §2.3, which is then solved using a Cauchy-type additive splitting in §2.4. The solution can be reduced to an infinite system of linear equations and this reduction is described in §2.5. It is possible to derive expressions for the total unsteady lift and the far-field sound based on our solution to the scattering problem, and these expressions are provided in §2.7. Finally, we provide numerical results using our solution in §2.8 which are used to demonstrate the accuracy of our method based on previous work in the overlapping case and to study the effects of cascade geometry and spanwise wavenumber on the scattered field. Our results are summarised in the concluding remarks in §2.9.

2.2 The equations of motion and mathematical formulation

We consider a cascade of blades of zero thickness and camber, which are parallel to the xz -plane, of finite length in the x -direction, staggered in the y -direction and extend to infinity in the spanwise z -direction as shown in Fig. 2.1.

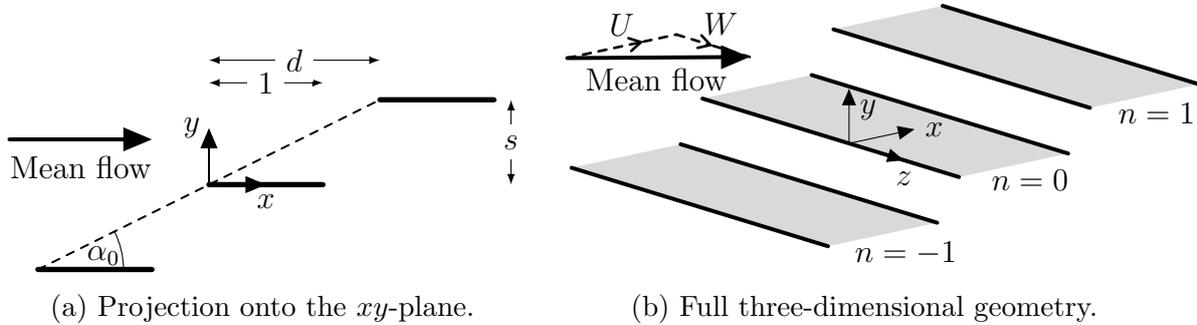


Figure 2.1: The cascade geometry with mean-flow and blade labels.

We assume the blades lie in an inviscid isentropic fluid with nonzero uniform subsonic mean flow $\mathbf{u} = (U, 0, W)$, with corresponding speed $U_\infty := \sqrt{U^2 + W^2}$ and Mach number $M := U_\infty/c_0$, where c_0 is the undisturbed speed of sound. We suppose that the flow is perturbed by an incoming wave (which could correspond to an acoustic wave or a vortical gust) that is incident from $x = -\infty$, and which has the following velocity field

$$\mathbf{u}_{\text{inc}} = \mathbf{A} e^{i\Omega t - i\mathbf{K} \cdot \mathbf{x}},$$

where $\mathbf{K} = (K_1, K_2, K_3)$ is the non-dimensionalised wave vector. The form of the incident field for acoustic waves and harmonic gusts is discussed in more detail in §2.2.1. As described in §1.1, by Goldstein's splitting theorem (Goldstein, 1976, pp. 220–222) the perturbed velocity potential can be decomposed into a solenoidal part (which is zero for incident acoustic waves, and equal to the incident field for vorticity waves) and an irrotational part corresponding to the scattered field (plus the incident field for acoustic waves). The scattered component $\tilde{\phi}(x, y, z, t)$ of this velocity potential satisfies, as described in Chapter 1

$$\frac{1}{c_0^2} \frac{D^2 \tilde{\phi}}{Dt^2} - \Delta \tilde{\phi} = 0, \quad (2.1)$$

where $\Delta = \nabla^2$ is the Laplacian in three dimensions. We non-dimensionalise our equations by rescaling the physical quantities as follows: We rescale lengths by l the blade chord, times by l/U_∞ , density perturbations by the undisturbed fluid density ρ_0 , pressure fluctuations by $\rho_0 U_\infty U$, and we let $M_x := U/c_0$, $M_z := W/c_0$ be the Mach numbers of the x , z -

components of the flow. Note that the rescaling of times by l/U_∞ is chosen in order for our analysis to fit into the context and notation of the relevant literature using a similar non-dimensionalisation (Peake, 1992, 1993, Peake and Kerschen, 2004). However, this choice means that $U_\infty \rightarrow 0$ becomes a singular limit and we need to be careful when taking the limit of zero mean flow in the results of this chapter. Nevertheless, we note that the correct physical limit is recovered if we take $M \rightarrow 0$ with $M\Omega = k$ fixed in the following non-dimensionalised algebraic expressions. Finally, the normalised blade stagger d , and interblade spacing s are defined as in Fig. 2.1a, resulting in the stagger angle $\alpha_0 = \arctan(s/d)$. Crucially, we wish to include the case of non-overlapping blades, so that d may be greater than 1.

We look for time-harmonic solutions of the form:

$$\tilde{\phi}(\mathbf{x}, t) = \phi(x, y)e^{i\Omega t - iK_3 z},$$

where $\Omega = \omega l/U_\infty$ is the reduced frequency, and $K_3 = lk_3$ is the non-dimensional spanwise wavenumber. Then, by Eq. (2.1), ϕ must satisfy the dimensionless equation

$$\beta^2 \partial_x^2 \phi + \partial_y^2 \phi - 2i\tilde{\Omega} M_x M \partial_x \phi + \left(\tilde{\Omega}^2 M^2 - K_3^2 \right) \phi = 0, \quad (2.2)$$

where we denoted the Prandtl–Glauert number by $\beta^2 = 1 - M_x^2$, and we call $\tilde{\Omega} = \Omega - K_3 \frac{M_z}{M}$ the effective reduced frequency, where we require $\tilde{\Omega}^2 M_x^2 \geq K_3^2 \beta^2$ for propagating acoustic modes to exist (as will be seen from the dispersion relation Eq. (2.6)).

2.2.1 Boundary and edge conditions

Let us label the blades by $0, \pm 1, \pm 2, \dots$, then for incident harmonic waves (either vortical or acoustical) the (time-dependent) upwash normal to the n^{th} blade (the y -velocity of the incoming wave on the blades in the upward pointing direction) is given by

$$V \exp(i\Omega t - iK_1 x + in\sigma), \quad (2.3)$$

where $K_1 = lk_1$ is the non-dimensional x -component of the wavevector, and we assume that the incident wave is periodic in the transverse direction, with the change in its phase between adjacent blades (the so-called interblade phase angle) being $\sigma = -dK_1 - sK_2$. We impose a Kutta condition at the trailing edges of the cascade blades which allows us to incorporate a significant amount of the effect of viscosity in the boundary layer of a real airfoil in viscous flow into our inviscid equation of motion Eq. (2.1) (Crighton, 1985). The Kutta condition essentially models the trailing edge stagnation point of airfoils in viscous flow which leads to the shedding of vorticity. This vortex shedding process is commonly modelled (in simplified form) in the inviscid equation by vortex sheets attached at the

trailing edges of the cascade blades, henceforth referred to as the ‘wake’. In summary, ϕ must therefore satisfy the following conditions:

- (i) The total normal velocity (i.e. the sum of upwash and normal component of the scattered potential) must vanish on the rigid stationary blades, i.e.

$$\frac{\partial\phi}{\partial y} = -Ve^{in\sigma - iK_1x} \text{ on } \{0 \leq x - nd \leq 1, y = ns\}.$$

- (ii) The unsteady pressure, which is given in dimensionless form by

$$p = -\frac{D\tilde{\phi}}{Dt} = -\left(i\tilde{\Omega}\frac{M}{M_x} + \frac{\partial}{\partial x}\right)\phi,$$

is continuous away from the blades.

- (iii) The scattered field must satisfy a radiation condition, which, with our introduction of a small amount of fictitious damping $\text{Im}\Omega < 0$ in the formulation of the problem as a Wiener–Hopf system, as described in §2.3.2, is equivalent to requiring ϕ to be bounded at infinity.
- (iv) The incident field has a periodicity as specified in Eq. (2.3). We require the scattered potential to exhibit a similar property, namely for all $-\infty < x < \infty$, $0 \leq y \leq s$, $n \in \mathbb{Z}$:

$$\phi(x, y) = e^{-in\sigma}\phi(x + nd, y + ns). \quad (2.4)$$

This is essentially a Bloch condition as we move from one cell in the cascade to the next. This means, in particular, that we need only determine the velocity potential in the first cascade cell $\{-\infty < x < \infty, 0 \leq y \leq s\}$, and the solution everywhere else is determined from Eq. (2.4).

- (v) The total velocity normal to the blades $\partial\phi/\partial y$ must be continuous everywhere (which is a consequence of the continuity of pressure and the consideration of infinitely thin blades).
- (vi) The scattered field satisfies the Kutta condition at the trailing edge (see for instance Peake (1992) and Crighton (1985)), i.e. $[p]$ is non-singular at the points $(x, y) = (1 + nd, ns)$, and the usual inverse square-root singularity at the leading edge, i.e. ϕ has an inverse square-root singularity at the points $(x, y) = (nd, ns)$.

In (vi) we introduce the notation $[\cdot]$ for the jump in an unsteady quantity across the first cascade cell, so that for instance the jump in unsteady pressure p is given by

$$[p](x) := p(x, 0^+) - e^{-i\sigma} p(x + d, s^-) = p(x, 0^+) - p(x, 0^-). \quad (2.5)$$

In the present work we are particularly interested in incident (harmonic) vortical gusts and acoustic waves as described below.

Incident harmonic gust

A harmonic vortical disturbance, representing for instance a component of the wake shed from an upstream blade row, corresponds to an incident velocity field of the form

$$\mathbf{u}_{\text{inc}} = \mathbf{A} e^{i\Omega t - i\mathbf{K} \cdot \mathbf{x}},$$

which must satisfy mass conservation $\nabla \cdot \mathbf{u}_{\text{inc}} = 0$, i.e. $\mathbf{A} \cdot \mathbf{K} = 0$, and is convected with the flow, i.e. $D\mathbf{u}_{\text{inc}}/Dt = 0$. This results in the dispersion relation $K_1 = \tilde{\Omega} M M_x^{-1}$ and implies $\sigma = -d\tilde{\Omega} M M_x^{-1} - sK_2$.

Incident acoustic waves

Acoustic waves, representing perhaps noise generated elsewhere in the aeroengine, have a velocity potential proportional to $\exp(i\Omega t - iK_1 x - iK_2 y - iK_3 z)$ and which satisfies the convected wave equation Eq. (2.1). Thus we find the following dispersion relation

$$\beta^2 \left(K_1 + \frac{\tilde{\Omega} M_x M}{\beta^2} \right)^2 + K_2^2 + K_3^2 = \frac{\tilde{\Omega}^2 M^2}{\beta^2}. \quad (2.6)$$

This ellipsoid in wave phase space can be parametrised as follows:

$$K_1 = \frac{\tilde{\Omega} M \cos \varphi \sin \theta}{1 + M_x \cos \varphi \sin \theta}, \quad K_2 = \frac{\tilde{\Omega} M \sin \varphi \sin \theta}{1 + M_x \cos \varphi \sin \theta}, \quad K_3 = \frac{\tilde{\Omega} M \cos \theta}{1 + M_x \cos \varphi \sin \theta}. \quad (2.7)$$

From this parametrisation we can immediately extract the incident direction of the wave (in terms of azimuthal angle φ and polar angle θ). Moreover, in terms of φ, θ we have

$$\sigma = \frac{(-\sin \theta) \tilde{\Omega} M}{1 + M_x \cos \varphi \sin \theta} (d \cos \varphi + s \sin \varphi).$$

We focus our attention initially on the case when acoustic waves are incident from upstream, which for us means $s(M_x + \sin \theta \cos \varphi) - d \sin \theta \sin \varphi \geq 0$ and is specified in more detail in §2.3.1. The case of downstream incident acoustic waves can be treated analogously and the relevant expressions of the Wiener–Hopf analysis and far-field are provided for

completeness in Appendix 2.E.

2.3 Formulation as a Wiener–Hopf problem

The periodicity condition (iv) allows us to restrict our attention to the first cascade cell, $\{-\infty < x < \infty, 0 \leq y \leq s\}$. We aim to apply the Fourier transform directly to the differential equation and boundary conditions in order to arrive at a Wiener–Hopf formulation, i.e. we take the approach described in §1.2.1. Let us use the following convention for the x -wise Fourier transform:

$$\Phi(\alpha, y) := \int_{-\infty}^{\infty} \phi(x, y) e^{i\alpha x} dx.$$

Fourier transforming Eq. (2.2) and boundary condition (v) we find

$$\frac{\partial^2 \Phi}{\partial y^2}(\alpha, y) - \gamma^2(\alpha) \Phi(\alpha, y) = 0, \quad (2.8)$$

$$\frac{\partial \Phi}{\partial y}(\alpha, 0^+) - e^{-i\sigma - id\alpha} \frac{\partial \Phi}{\partial y}(\alpha, s^-) = 0, \quad (2.9)$$

where $\gamma^2(\alpha) = \alpha^2 \beta^2 + 2\alpha \tilde{\Omega} M_x M - (\tilde{\Omega}^2 M^2 - K_3^2)$ and the branch cuts of γ are chosen such that the function always has positive real part in the strip $R^+ \cap R^- \subset \mathbb{C}$ which is defined in §2.4.1. We also note that the Fourier transform of the pressure fluctuations is related to the Fourier-transformed velocity potential by

$$P(\alpha, y) = - \left(i\tilde{\Omega} \frac{M}{M_x} - i\alpha \right) \Phi(\alpha, y), \quad \text{and} \quad [P](\alpha) = - \left(i\tilde{\Omega} \frac{M}{M_x} - i\alpha \right) [\Phi](\alpha). \quad (2.10)$$

The solution to Eq. (2.8) takes the form $\Phi(\alpha, y) = A(\alpha) \exp(\gamma(\alpha)y) + B(\alpha) \exp(-\gamma(\alpha)y)$, which together with the periodicity Eq. (2.9) and with Eq. (2.10) implies after a few steps of algebra that

$$[P](\alpha) = \kappa(\alpha) \frac{\partial \Phi}{\partial y}(\alpha, 0), \quad (2.11)$$

where the scalar Wiener–Hopf kernel κ is given by

$$\begin{aligned} \kappa(\alpha) &= \frac{2 \left(\tilde{\Omega} \frac{M}{M_x} - \alpha \right) (\cos(\sigma + d\alpha) - \cosh(\gamma s))}{i\gamma \sinh(\gamma s)} \\ &= \frac{\left(\tilde{\Omega} \frac{M}{M_x} - \alpha \right)}{i\gamma \sinh(\gamma s)} (1 - e^{-i\sigma - \gamma s - id\alpha}) (1 - e^{-i\sigma + \gamma s - id\alpha}) e^{i\sigma + id\alpha}. \end{aligned}$$

We observe that $\kappa(\alpha)$ is single-valued in the complex plane, since its direct dependence on γ is even.

2.3.1 The scalar Wiener–Hopf kernel

One can easily check (as described by Peake (1992)) that κ is meromorphic, with simple poles located at $\alpha = k_n^\pm$, $n \in \mathbb{N}$ where

$$k_n^\pm = \begin{cases} \frac{-M_x M \tilde{\Omega} \mp \sqrt{M_x^2 M^2 \tilde{\Omega}^2 - \beta^2 (n^2 \pi^2 s^{-2} + K_3^2 - M^2 \tilde{\Omega}^2)}}{\beta^2} & \text{if } n \leq p, \\ \frac{-M_x M \tilde{\Omega} \pm i \sqrt{-M_x^2 M^2 \tilde{\Omega}^2 + \beta^2 (n^2 \pi^2 s^{-2} + K_3^2 - M^2 \tilde{\Omega}^2)}}{\beta^2} & \text{if } n > p, \end{cases}$$

and p is the largest integer such that $M_x^2 M^2 (\text{Re } \tilde{\Omega})^2 - \beta^2 (n^2 \pi^2 s^{-2} + K_3^2 - M^2 (\text{Re } \tilde{\Omega})^2) \geq 0$. Furthermore κ has simple zeros at the convected wavenumber $\alpha = \tilde{\Omega} \frac{M}{M_x}$ and at $\alpha = \sigma_m^\pm$, $m \in \mathbb{Z}$, where for $-r \leq m \leq q$:

$$\sigma_m^\pm = \frac{-(s^2 M_x M \tilde{\Omega} + d\sigma + 2d\pi m)}{s^2 \beta^2 + d^2} \mp \frac{\sqrt{\left(s^2 M_x M \tilde{\Omega} + d\sigma + 2d\pi m\right)^2 - (s^2 \beta^2 + d^2) \left((\sigma + 2\pi m)^2 - s^2 \left(\tilde{\Omega}^2 M^2 - K_3^2\right)\right)}}{s^2 \beta^2 + d^2},$$

and for $m > q, m < -r$:

$$\sigma_m^\pm = \frac{-(s^2 M_x M \tilde{\Omega} + d\sigma + 2d\pi m)}{s^2 \beta^2 + d^2} \pm \frac{i \sqrt{-\left(s^2 M_x M \tilde{\Omega} + d\sigma + 2d\pi m\right)^2 + (s^2 \beta^2 + d^2) \left((\sigma + 2\pi m)^2 - s^2 \left(\tilde{\Omega}^2 M^2 - K_3^2\right)\right)}}{s^2 \beta^2 + d^2},$$

with $-r, q$ being the smallest and largest integer respectively such that

$$\left(s^2 M_x M (\text{Re } \tilde{\Omega}) + d\sigma + 2d\pi m\right)^2 - (s^2 \beta^2 + d^2) \left((\sigma + 2\pi m)^2 - s^2 \left((\text{Re } \tilde{\Omega})^2 M^2 - K_3^2\right)\right) \geq 0.$$

We call k_n^\pm the duct modes (which are cut-on if and only if $n \leq p$) and σ_m^\pm the radiation modes (which are cut-on if and only if $-r \leq m \leq q$). These are the coherent acoustic modes that appear in this geometry as a result of the arrangement of the blades in the cascade, and $\alpha = \tilde{\Omega} M / M_x$ is the hydrodynamic mode supporting the wake downstream of the blades. Note only cut-on radiation modes carry energy to $\pm\infty$ and the cut-off modes are evanescent. The radiation modes for cascade scattering are an analogue of Bragg angles for periodic media and are the coherent waves present in the scattered field as a

result of the periodic nature of the geometry. This periodicity means the field is split into a discrete set of modes, as can be seen from the dispersion relation derived in §3.2.1. Their physical interpretation is discussed to some extent by Peake (1993, §5).

Incident acoustic waves

For acoustic waves we can verify using the parametrisation Eq. (2.7) that

$$K_1 = \begin{cases} \sigma_0^- & \text{if } s(M_x + \sin \theta \cos \varphi) - d \sin \theta \sin \varphi \geq 0, \\ \sigma_0^+ & \text{if } s(M_x + \sin \theta \cos \varphi) - d \sin \theta \sin \varphi < 0, \end{cases} \quad (2.12)$$

i.e. $K_1 = \sigma_0^-$ if the incident wave is travelling downstream with respect to the cascade stagger (i.e. is incident from upstream), and $K_1 = \sigma_0^+$ if the wave travels upstream (i.e. is incident from downstream). For simplicity we restrict our analysis initially to the case of incident waves from upstream (i.e. for acoustic waves when $K_1 = \sigma_0^-$), but the scattering problem can be solved in an analogous (reflected) way in the case of upstream travelling waves incident from downstream as well, as is described in Appendix 2.E.

2.3.2 Derivation of Wiener–Hopf equations

A convenient notation for the zeros of the Wiener–Hopf kernel κ is, for all $m \in \mathbb{Z}$,

$$\tilde{\sigma}_m^+ = \sigma_m^+, \quad \tilde{\sigma}_m^- = \begin{cases} \tilde{\Omega} \frac{M}{M_x}, & \text{if } m = 0, \\ \sigma_{m-1}^-, & \text{if } m \geq 1, \\ \sigma_m^-, & \text{if } m \leq -1, \end{cases}$$

which allows us to write the Neumann boundary condition for incident acoustic and vorticity waves as

$$\frac{\partial \phi}{\partial y} = -V e^{in\sigma - i\tilde{\sigma}_\eta^- x} \text{ on } \{0 \leq x - nd \leq 1, y = ns\},$$

where our analysis facilitates a solution for any $-r \leq \eta \leq q$ and in particular for the physically relevant values $\eta = 0, 1$, which correspond to incident gusts ($\eta = 0$) and incident (downstream travelling) acoustic waves ($\eta = 1$).

We follow the standard practice of assuming that $\tilde{\Omega}$ has a small negative imaginary part $\text{Im } \tilde{\Omega} < 0$, which will be taken to zero at the end of the analysis. By construction this introduces a small negative imaginary part for both types of incident fields, and physically amounts to a small amount of damping downstream (for both types of incident fields) and ensures that certain integral transforms in the later analysis are well-defined.

We begin by deriving the system of Wiener–Hopf equations which we shall use to solve the problem. We use an approach similar to Glegg (1999) (although we are able to reduce the number of equations from four to three) and split the boundary value problem into three problems on $\{-\infty < x < \infty, 0 \leq y \leq s\}$, each of which has to satisfy only a combination of two semi-infinite boundary conditions hence facilitating the application of the Wiener–Hopf technique on coupled scalar equations. We define ϕ_1, ϕ_2, ϕ_3 , each to satisfy Eq. (2.2) on the first cascade cell (i.e. on the space between the first and second blade of the cascade, $\{-\infty < x < \infty, 0 \leq y \leq s\}$), the radiation condition (iii) and the continuity of the blade-normal velocity everywhere (v) (which translates to the jump in normal velocity across the first cascade strip, as defined in Eq. (2.5), being zero). Finally, we impose the following pairs of semi-infinite boundary conditions (together with an appropriate distribution of the edge conditions (vi)) on ϕ_j and the corresponding pressures $p_j, j = 1, 2, 3$:

- Leading-edge interaction with incident field:

$$\begin{aligned}\frac{\partial \phi_1}{\partial y}(x, 0) &= -V e^{-i\bar{\sigma}_\eta^- x} \text{ on } x > 0, \\ [p_1](x) &= 0 \text{ on } x < 0,\end{aligned}$$

and ϕ_1 has the conventional inverse square-root singularity at the leading edge $x = 0$.

- Trailing-edge correction:

$$\begin{aligned}\frac{\partial \phi_2}{\partial y}(x, 0) &= 0 \text{ on } x < 1, \\ [p_1 + p_2 + p_3](x) &= 0 \text{ on } x > 1,\end{aligned}$$

and ϕ_2 satisfies the Kutta condition at the trailing edge $x = 1$.

- Leading-edge correction:

$$\begin{aligned}\frac{\partial \phi_3}{\partial y}(x, 0) &= 0 \text{ on } x > 0, \\ [p_2 + p_3](x) &= 0 \text{ on } x < 0,\end{aligned}$$

and ϕ_3 has at worst an inverse square-root singularity at the leading edge $x = 0$.

Thus ϕ_1 corresponds to the scattered potential of a wave incident on a cascade of *semi-infinite* blades, while ϕ_2, ϕ_3 act jointly to correct the pressure jumps downstream of the trailing edge and to ensure that $\phi = \phi_1 + \phi_2 + \phi_3$ is a solution to the original problem. Since the derivation of Eq. (2.11) relied only on the time-reduced convected wave-equation, together with the radiation condition and the continuity of normal velocity, we find that

each of $\phi_j, j = 1, 2, 3$, satisfies Eq. (2.11) as well. If we then transform the additional boundary conditions we arrive at the following system of scalar Wiener–Hopf equations:

$$\frac{[P_1^+](\alpha)}{\kappa(\alpha)} = \frac{\partial \Phi_1^-}{\partial y}(\alpha, 0) - \frac{iV}{\alpha - \tilde{\sigma}_\eta^-}, \quad (2.13)$$

$$\frac{\partial \tilde{\Phi}_2^+}{\partial y}(\alpha, 0) = \frac{1}{\kappa(\alpha)} \left([\tilde{P}_2^-](\alpha) - [\tilde{P}_1^+](\alpha) - [\tilde{P}_3^+](\alpha) \right), \quad (2.14)$$

$$\frac{\partial \Phi_3^-}{\partial y}(\alpha, 0) = \frac{1}{\kappa(\alpha)} \left([P_3^+](\alpha) - [P_2^-](\alpha) \right), \quad (2.15)$$

where we used the following notation of half-line Fourier transforms for a given function ψ :

$$\begin{aligned} \Psi^+(\alpha, y) &:= \int_0^\infty \psi(x, y) e^{i\alpha x} dx, & \Psi^-(\alpha, y) &:= \int_{-\infty}^0 \psi(x, y) e^{i\alpha x} dx, \\ \tilde{\Psi}^+(\alpha, y) &:= e^{-i\alpha} \int_1^\infty \psi(x, y) e^{i\alpha x} dx, & \tilde{\Psi}^-(\alpha, y) &:= e^{-i\alpha} \int_{-\infty}^1 \psi(x, y) e^{i\alpha x} dx. \end{aligned}$$

2.4 Solution using the Wiener–Hopf technique

2.4.1 Factorisation of κ

Given $\text{Im } \tilde{\Omega} < 0$ there is ϵ such that $\pm \text{Im } \tilde{\sigma}_m^\pm > \epsilon$ and $\pm \text{Im } k_n^\pm > \epsilon$. Defining the overlapping half-planes $R^\pm := \left\{ \alpha \in \mathbb{C} \mid \pm \text{Im } \alpha > -\epsilon \right\}$ we observe that each of Eqs. (2.13)-(2.15) is valid in the strip $R^+ \cap R^-$ and we note that the half-line Fourier transforms are analytic in R^\pm according to their superscripts (under the *a priori* assumption that the solutions ϕ_1, ϕ_2, ϕ_3 decay sufficiently fast along the x -direction).

Using the Weierstrass factorisation theorem (see Thm. 5.14 in Conway (1978, p. 170) or the special case given in Noble (1958, p. 40)), and applying the procedure for analysing the asymptotic behaviour of infinite products as outlined by Noble (1958), it is possible to construct functions κ^+, κ^- which are analytic in R^+, R^- respectively, satisfy $\kappa = \kappa^+ \kappa^-$, and have the following algebraic growth behaviour:

$$|\kappa^+(\alpha)| \sim C^+ |\alpha|^{\frac{1}{2}} \text{ as } \alpha \rightarrow \infty \text{ in } R^+, \quad (2.16)$$

$$|\kappa^-(\alpha)| \sim C^- |\alpha|^{-\frac{1}{2}} \text{ as } \alpha \rightarrow \infty \text{ in } R^-, \quad (2.17)$$

for some nonzero constants C^\pm . The details of this splitting are described in §2.A.

2.4.2 The uncoupled leading edge problem

We notice that Eq. (2.13) is uncoupled from the remaining two Wiener–Hopf equations. The equation corresponds essentially to the scattering problem on a cascade of semi-infinite

blades, and explicit solutions have been provided by a number of authors. Amongst the earliest work on this problem is Carlson and Heins (1947), and solutions are also given by Mani and Horvay (1970), Glegg (1999), and Peake (1992). Here we briefly summarise the main steps taken by Peake in the solution of this uncoupled problem (see Peake, 1992, pp. 267-273). Using the aforementioned kernel factorisation we can rewrite Eq. (2.13) in the form

$$\begin{aligned} \frac{[P_1^+](\alpha)}{\kappa^+(\alpha)} + \frac{iV\kappa^-(\tilde{\sigma}_\eta^-)}{\alpha - \tilde{\sigma}_\eta^-} &= \kappa^-(\alpha) \frac{\partial \Phi_1^-}{\partial y}(\alpha, 0) - \kappa^-(\alpha) \frac{iV}{\alpha - \tilde{\sigma}_\eta^-} \left(1 - \frac{\kappa^-(\tilde{\sigma}_\eta^-)}{\kappa^-(\alpha)} \right) \\ &= E_1(\alpha), \end{aligned}$$

which is valid in $R^+ \cap R^-$. Thus analytic continuation allows us to define an entire function $E_1(\alpha)$. As mentioned above we assume that the unsteady field ϕ_1 possesses the conventional inverse square-root singularity (condition (vi) from §2.2.1) at the leading edge $x = 0$, which results in the following asymptotic behaviour (Peake, 1992, p. 269):

$$\begin{aligned} \frac{\partial \Phi_1^-}{\partial y}(\alpha, 0) &= \mathcal{O}\left(\alpha^{-\frac{1}{2}}\right) \text{ as } \alpha \rightarrow \infty \text{ in } R^-, \\ [P_1^+](\alpha) &= \mathcal{O}\left(\alpha^{-\frac{1}{2}}\right) \text{ as } \alpha \rightarrow \infty \text{ in } R^+. \end{aligned}$$

This inverse square-root singularity is a direct consequence of the flow being effectively incompressible and therefore satisfying Laplace's equation nearby the leading edge. This behaviour, together with Liouville's theorem, implies $E_1(\alpha) \equiv 0$ and therefore yields the solution to Eq. (2.13)

$$[P_1^+](\alpha) = -iV\kappa^-(\tilde{\sigma}_\eta^-) \frac{\kappa^+(\alpha)}{\alpha - \tilde{\sigma}_\eta^-}, \quad \frac{\partial \Phi_1^-}{\partial y}(\alpha, 0) = \frac{iV}{\alpha - \tilde{\sigma}_\eta^-} \left(1 - \frac{\kappa^-(\tilde{\sigma}_\eta^-)}{\kappa^-(\alpha)} \right). \quad (2.18)$$

2.4.3 Asymptotic behaviour of the kernel factors and residues

Before we can solve Eq. (2.14) and Eq. (2.15) we first need to look more closely at the asymptotic behaviour of the kernel factors κ^\pm and their residues. We find that $|\kappa|$ has two asymptotic regimes - one regime with constant asymptotic behaviour, and one with exponential growth:

- As $\alpha \rightarrow \infty$ with $\pm \arg \alpha \in \left(\arctan\left(\frac{s\beta}{d}\right), \frac{\pi}{2} \right) \cup \left(\frac{\pi}{2}, \pi - \arctan\left(\frac{s\beta}{d}\right) \right)$,

$$|\kappa(\alpha)| \sim \frac{2e^{d|\operatorname{Im} \alpha| - \beta s |\operatorname{Re} \alpha|}}{\beta}.$$

- And as $\alpha \rightarrow \infty$ with $\arg \alpha \in \left(-\arctan\left(\frac{s\beta}{d}\right), \arctan\left(\frac{s\beta}{d}\right) \right)$ or

$$\arg \alpha \in \left(\pi - \arctan \left(\frac{s\beta}{d} \right), \pi + \arctan \left(\frac{s\beta}{d} \right) \right),$$

$$|\kappa(\alpha)| \sim \frac{2}{\beta}.$$

These regimes are sketched in Fig. 2.2: $|\kappa|$ has exponential growth at infinity in regions ② and ④ and is asymptotically constant in regions ① and ③. Now since κ^+ can be expressed in R^- by $\kappa^+ = \frac{\kappa}{\kappa^-}$, this means that κ^+ is not only of algebraic growth in R^+ but in fact in the following larger part of the complex plane

$$R^+ \cup \left\{ \alpha \mid \arg \alpha \in \left(-\arctan \left(\frac{s\beta}{d} \right), \arctan \left(\frac{s\beta}{d} \right) \right) \cup \left(\pi - \arctan \left(\frac{s\beta}{d} \right), \pi + \arctan \left(\frac{s\beta}{d} \right) \right) \right\},$$

which corresponds to a strip around the real axis together with regions ①, ② and ③. An analogous statement is true for κ^- . Furthermore one finds that $\frac{1}{\kappa^\pm}$ is a meromorphic function, with simple poles and growth of order $\mathcal{O}(\alpha^{-\frac{1}{2}})$ away from its poles (because either it decays algebraically or it decays exponentially).

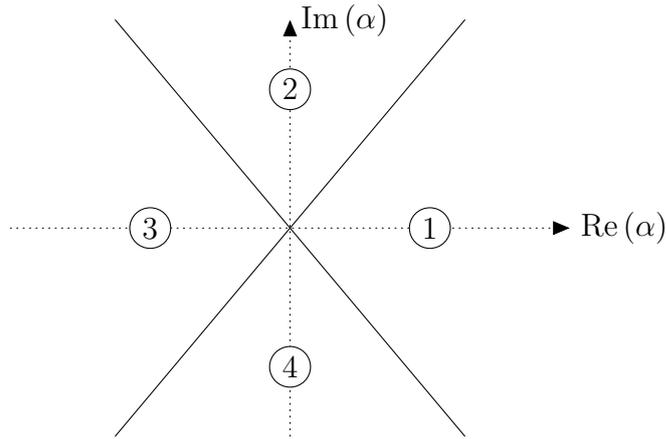


Figure 2.2: The regions of different asymptotic behaviour of $|\kappa|$ in the complex α -plane.

In order to understand the poles of κ^\pm , $\frac{1}{\kappa^\pm}$ it suffices, by way of expressing the functions as above, to understand the residues of κ , $\frac{1}{\kappa}$. We can quickly check that for each $m \in \mathbb{Z}$, $n \in \mathbb{N}$:

$$\text{Res} \left(\frac{1}{\kappa}, \sigma_m^+ \right) = \frac{i\gamma(\sigma_m^+)^2}{(-2)(\tilde{\Omega}_{M_x} - \sigma_m^+) (id\gamma(\sigma_m^+) + s\sigma_m^+)}, \quad (2.19)$$

$$\text{Res} \left(\frac{1}{\kappa}, \sigma_m^- \right) = \frac{i\gamma(\sigma_m^-)^2}{2(\tilde{\Omega}_{M_x} - \sigma_m^-) (id\gamma(\sigma_m^-) - s\sigma_m^-)}, \quad (2.20)$$

$$\text{Res}(\kappa, k_n^+) = \frac{2(\tilde{\Omega}\frac{M}{M_x} - k_n^+)(\cos(\sigma + dk_n^+) - (-1)^n)}{(-1)^n i k_n^+ s}, \quad (2.21)$$

$$\text{Res}(\kappa, k_n^-) = \frac{2(\tilde{\Omega}\frac{M}{M_x} - k_n^-)(\cos(\sigma + dk_n^-) - (-1)^n)}{(-1)^n i k_n^- s}. \quad (2.22)$$

2.4.4 The trailing edge correction

We can rewrite Eq. (2.14) as

$$\kappa^+(\alpha) \frac{\partial \tilde{\Phi}_2^+}{\partial y}(\alpha, 0) = \frac{1}{\kappa^-(\alpha)} \left([\tilde{P}_2^-](\alpha) - [\tilde{P}_1^+](\alpha) - [\tilde{P}_3^+](\alpha) \right).$$

Thus to proceed with the Wiener–Hopf technique we must additively split the term

$$g(\alpha) = \frac{1}{\kappa^-(\alpha)} \left([\tilde{P}_1^+](\alpha) + [\tilde{P}_3^+](\alpha) \right).$$

We now recall from Eq. (2.17) that $\frac{1}{\kappa^-}$ is of order $\mathcal{O}(\alpha^{\frac{1}{2}})$ uniformly in some strip containing the real axis, i.e. in a set of the form $(-\infty, \infty) \times (-\epsilon, \epsilon)$, such that $(-\infty, \infty) \times [-\epsilon, \epsilon] \subset R^+ \cap R^-$. It is also possible to show (as is proved in detail in Appendix 2.B) that

$$[\tilde{P}_1^+](\alpha) = \mathcal{O}(\alpha^{-1})$$

as $\alpha \rightarrow \infty, \alpha \in R^+$, and hence that $[p_1]$ is non-singular at $x = 1$. Imposing the unsteady Kutta condition (condition (vi) in §2.2.1) at the trailing edge to $[p_2]$ and noting that $[p_3](x) = -[p_1](x) - [p_2](x)$, for $x > 1$, we conclude that the pressure jump $[p_3](x)$ must be non-singular at $x = 1$, and it can thus be shown (as we saw in Eq. (1.6)) that $[\tilde{P}_3^+](\alpha) = \mathcal{O}(\alpha^{-1})$ as $\alpha \rightarrow \infty, \alpha \in R^+$. Hence there is a strip \mathcal{S} containing the real axis and a constant C such that

$$\left| \frac{1}{\kappa^-(\alpha)} \left([\tilde{P}_1^+](\alpha) + [\tilde{P}_3^+](\alpha) \right) \right| \leq C (1 + |\text{Re } \alpha|)^{-\frac{1}{2}} \text{ for all } \alpha \in \mathcal{S}.$$

Thus the assumptions of Thm. B in Noble (1958, p. 13) are satisfied and we can choose $c > 0$ and construct two functions such that $g = g_- + g_+$, where

$$g_+(\alpha) = \frac{1}{2\pi i} \int_{-\infty - ic}^{\infty - ic} \frac{g(\zeta)}{\zeta - \alpha} d\zeta \text{ and } g_-(\alpha) = -\frac{1}{2\pi i} \int_{-\infty + ic}^{\infty + ic} \frac{g(\zeta)}{\zeta - \alpha} d\zeta. \quad (2.23)$$

Let us shrink R^\pm such that

$$R^- = \left\{ \alpha \in \mathbb{C} \mid \text{Im } \alpha < \delta \right\} \text{ and } R^+ = \left\{ \alpha \in \mathbb{C} \mid \text{Im } \alpha > -\delta \right\}$$

for some $0 < \delta < c$, then the functions g_{\pm} are analytic in R^{\pm} respectively. Furthermore they are also bounded in their respective half-plane, as is proved in Appendix 2.B. We now observe that $\frac{g(\zeta)}{\zeta - \alpha}$ is meromorphic in an open neighbourhood of $\{\zeta | \text{Im } \zeta \geq -c\}$, with simple poles at σ_m^+ , $m \in \mathbb{Z}$ and, by Eqs. (2.16) & (2.17), decay of $\mathcal{O}\left(\zeta^{-\frac{3}{2}}\right)$. Thus we can use the residue theorem to express $g_-(\alpha)$ in terms of the following infinite series of pole contributions:

$$g_-(\alpha) = \sum_{m \in \mathbb{Z}} \frac{1}{\alpha - \sigma_m^+} \text{Res} \left(\frac{1}{\kappa^-}, \sigma_m^+ \right) \left([\tilde{P}_1^+] (\sigma_m^+) + [\tilde{P}_3^+] (\sigma_m^+) \right). \quad (2.24)$$

We note that the splitting which we are arrived at in Eq. (2.24) is simply the classical method of pole removal. The novelty in our present approach is to use this pole removal for the offending poles, the radiation modes σ_m^+ , which are natural to the geometry irrespective of the value of $d > 0$. This is in contrast to previous work which relied on the removal of duct modes k_n^{\pm} . Indeed the corresponding splitting used in previous studies (for instance Peake (1992) and Glegg (1999)) relies essentially on the following observation: If $d < 1$ then $[p_1](x) + [p_3](x)$, $x > 1$, can be written as an (infinite) linear combination of duct modes k_n^- . This is because in the region $x > 1$ the consecutive semi-infinite boundary conditions form a duct geometry (cf. §2.3.2). Thus, when $d < 1$, their half-line Fourier transform is of the form

$$[\tilde{P}_1^+] (\alpha) + [\tilde{P}_3^+] (\alpha) = \sum_{n \in \mathbb{N}} \frac{L_n}{\alpha - k_n^-} \quad (2.25)$$

for some complex constants L_n . With this knowledge one can close the contour of integration of $g_+(\alpha)$ in the lower half plane and derive a linear system from the discrete pole contributions at the duct modes k_n^- , $n \in \mathbb{N}$. However Eq. (2.25) is no longer valid when $d > 1$, because the modal structure of the pressure $[p_1](x) + [p_3](x)$ becomes more complicated in the region $1 < x < d$, since there is no overlap of consecutive boundary conditions in this part of the domain. As such the additive splitting and linear system from previous work are restricted to the case $d < 1$. We observe, however, that no such assumption was necessary to arrive at Eq. (2.24), indeed collecting the pole contributions at the radiation modes avoids making a distinction between overlapping and non-overlapping geometries altogether and forms a valid splitting for any value of $d > 0$. Now we note by Eq. (2.19), and Eqs. (2.54) & (2.55), that the series in Eq. (2.24) converges locally uniformly in $\alpha \in R^-$ (since the terms in the series decay like $m^{-3/2}$).

This allows us to recast Eq. (2.14) into

$$\begin{aligned} \kappa^+(\alpha) \frac{\partial \tilde{\Phi}_2^+}{\partial y}(\alpha, 0) + g_1^+(\alpha) &= \frac{1}{\kappa^-(\alpha)} \left[\tilde{P}_2^- \right](\alpha) \\ &\quad - \sum_{m \in \mathbb{Z}} \frac{1}{\alpha - \sigma_m^+} \text{Res} \left(\frac{1}{\kappa^-}, \sigma_m^+ \right) \left(\left[\tilde{P}_1^+ \right](\sigma_m^+) + \left[\tilde{P}_3^+ \right](\sigma_m^+) \right) \\ &= E_2(\alpha), \end{aligned}$$

which by analytic extension defines an entire function $E_2(\alpha)$. By the unsteady Kutta condition at the trailing edge, both the normal velocity and the pressure field are non-singular at $x = 1, y = 0$. Thus there is a constant C such that

$$\left| \frac{\partial \tilde{\Phi}_2^+}{\partial y}(\alpha, 0) \right| \leq \frac{C}{1 + |\text{Im } \alpha|} \quad \text{if } \alpha \in R^+, \quad \text{and} \quad \left| \left[\tilde{P}_2^- \right](\alpha) \right| \leq \frac{C}{1 + |\text{Im } \alpha|} \quad \text{if } \alpha \in R^-, \quad (2.26)$$

which implies together with our previous analysis that $|E_2(\alpha)| \leq C|\alpha|^{\frac{1}{2}}$ uniformly in \mathbb{C} , and thus by the extended Liouville theorem $E_2(\alpha)$ must be constant. Moreover, $E_2(iy) \rightarrow 0$ as $y \rightarrow -\infty$ and thus E_2 must be identically zero, which then implies:

$$\frac{1}{\kappa^-(\alpha)} \left[\tilde{P}_2^- \right](\alpha) = \sum_{m \in \mathbb{Z}} \frac{1}{\alpha - \sigma_m^+} \text{Res} \left(\frac{1}{\kappa^-}, \sigma_m^+ \right) \left(\left[\tilde{P}_1^+ \right](\sigma_m^+) + \left[\tilde{P}_3^+ \right](\sigma_m^+) \right). \quad (2.27)$$

2.4.5 The leading edge correction

We proceed similarly to the trailing edge correction: firstly note that Eq. (2.15) is equivalent to

$$\kappa^-(\alpha) \frac{\partial \Phi_3^-}{\partial y}(\alpha, 0) = \frac{1}{\kappa^+(\alpha)} \left(\left[P_3^+ \right](\alpha) - \left[P_2^- \right](\alpha) \right).$$

Thus, we must additively split the term

$$h(\alpha) = \frac{1}{\kappa^+(\alpha)} \left[P_2^- \right](\alpha).$$

We now observe that by construction $[p_2](x) = -[p_3](x)$, for $x < 0$. Additionally we assumed that $[p_3](x)$ exhibits at worst the conventional inverse square-root singularity at the leading edge (cf. §2.3.2). Therefore $[p_2](x)$ must also have at worst an inverse square-root singularity at $x = 0$ and thus $\left[P_2^- \right](\alpha) = \mathcal{O}(\alpha^{-1/2})$ as $\alpha \rightarrow \infty, \alpha \in R^-$. This means that there is a strip \mathcal{S} around the real axis and a constant C such that

$$|h(\alpha)| \leq C (1 + |\text{Re } \alpha|)^{-\frac{1}{2}} \quad \forall \alpha \in \mathcal{S}.$$

Thus we may proceed as before to apply Thm. B from Noble (1958, p. 13) and construct two functions h_+, h_- which are bounded and analytic in R^+, R^- respectively, and are given by

$$h_+(\alpha) = \frac{1}{2\pi i} \int_{-\infty-ic}^{\infty-ic} \frac{h(\zeta)}{\zeta - \alpha} d\zeta, \quad \text{and} \quad h_-(\alpha) = \frac{1}{2\pi i} \int_{-\infty+ic}^{\infty+ic} \frac{h(\zeta)}{\zeta - \alpha} d\zeta.$$

Since $\frac{h(\zeta)}{\zeta - \alpha}$ is a meromorphic function with simple poles in an open neighbourhood containing $\{\zeta | \text{Im } \zeta \leq c\}$, we can use the residue theorem to collect the contributions of poles and express h_+ as the following convergent series:

$$h_+(\alpha) = \sum_{m \in \mathbb{Z}} \frac{1}{\alpha - \tilde{\sigma}_m^-} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) [P_2^-] (\tilde{\sigma}_m^-),$$

which converges locally uniformly due to the decay in the residues. This is again simply a pole removal splitting based on the radiation modes $\tilde{\sigma}_m^-$. We highlight again that as in §2.4.4 this additive Wiener–Hopf splitting relying on radiation modes is the main difference to previous work by Peake (1992) and Glegg (1999), and is valid for all $d > 0$, hence overcomes the restriction to overlapping blades. Thus Eq. (2.15) is equivalent to

$$\begin{aligned} \kappa^-(\alpha) \frac{\partial \Phi_3^-}{\partial y}(\alpha, 0) - h_-(\alpha) &= \frac{1}{\kappa^+(\alpha)} [P_3^+] (\alpha) - \sum_{m \in \mathbb{Z}} \frac{1}{\alpha - \tilde{\sigma}_m^-} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) [P_2^-] (\tilde{\sigma}_m^-) \\ &= E_3(\alpha). \end{aligned}$$

As for the previous two Wiener–Hopf equations this defines, by analytic continuation, an entire function $E_3(\alpha)$. We further assumed that ϕ_3 has at worst an inverse square root type singularity at the leading edge, which implies that the growth of $\frac{\partial \Phi_3^-}{\partial y}(\alpha, 0)$, $[P_3^+] (\alpha)$ in R^-, R^+ is similar to the corresponding growth of $\frac{\partial \Phi_1^-}{\partial y}(\alpha, 0)$, $[P_1^+] (\alpha)$ respectively, which amounts to (cf. §2.4.2)

$$\begin{aligned} \frac{\partial \Phi_3^-}{\partial y}(\alpha, 0) &= \mathcal{O}(\alpha^{-\frac{1}{2}}) \quad \text{as } \alpha \rightarrow \infty \text{ in } R^-, \\ [P_3^+] (\alpha) &= \mathcal{O}(\alpha^{-\frac{1}{2}}) \quad \text{as } \alpha \rightarrow \infty \text{ in } R^+. \end{aligned}$$

Thus the entire function $E_3(\alpha)$ is bounded. Furthermore if we take $\alpha \rightarrow \infty$ with $\arg \alpha = \pi/2$, then the right hand side of the equation decays to zero, thus Liouville’s theorem combined with this observation implies $E_3(\alpha) \equiv 0$. Therefore, in particular

$$\frac{1}{\kappa^+(\alpha)} [P_3^+] (\alpha) = \sum_{m \in \mathbb{Z}} \frac{1}{\alpha - \tilde{\sigma}_m^-} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) [P_2^-] (\tilde{\sigma}_m^-). \quad (2.28)$$

2.5 Reduction to an infinite algebraic system

In order to reduce Eq. (2.18), (2.27) and (2.28) to a discrete matrix equation we define the following coefficients (for $m \in \mathbb{Z}$)

$$A_m := [\tilde{P}_1^+] (\tilde{\sigma}_m^+), \quad B_m := [P_2^-] (\tilde{\sigma}_m^-), \quad C_m := [\tilde{P}_3^+] (\tilde{\sigma}_m^+).$$

Then using Eq. (2.27) we find by Fourier inversion and several changes of order of summation and integration, which are rigorously justified in Appendix 2.C:

$$B_j = [P_2^-] (\tilde{\sigma}_m^-) = \sum_{m \in \mathbb{Z}} \mathbf{G}_{jm} (A_m + C_m),$$

where

$$\mathbf{G}_{jm} = \frac{-1}{2\pi i} \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_m^+ \right) \int_{-\infty+i\epsilon}^{\infty+i\epsilon} \frac{e^{i\alpha} \kappa^-(\alpha)}{(\alpha - \tilde{\sigma}_j^-)(\alpha - \tilde{\sigma}_m^+)} d\alpha. \quad (2.29)$$

We can also use Fourier inversion on Eq. (2.28) (which is given in full detail in Appendix 2.C) to arrive at a second collection of linear equations. This allows us to reduce the scattering problem to the following linear system:

$$B_j = \sum_{m \in \mathbb{Z}} \mathbf{G}_{jm} (A_m + C_m), \quad (2.30)$$

$$C_j = \sum_{m \in \mathbb{Z}} \mathbf{F}_{jm} B_m, \quad (2.31)$$

where \mathbf{G}_{jm} are as in Eq. (2.32) and

$$\mathbf{F}_{jm} = \frac{1}{2\pi i} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) \int_{\Gamma^-} \frac{e^{-i\alpha} \kappa^+(\alpha)}{(\alpha - \tilde{\sigma}_j^+)(\alpha - \tilde{\sigma}_m^-)} d\alpha, \quad (2.32)$$

where the contour of integration Γ^- behaves at its tails like

$$\arg \alpha \sim \begin{cases} -\tilde{\epsilon}, & \text{Re } \alpha \rightarrow +\infty, \\ \pi + \tilde{\epsilon}, & \text{Re } \alpha \rightarrow -\infty. \end{cases}$$

for some $0 < \tilde{\epsilon} \ll 1$. A sketch of Γ^- is shown in Fig. 2.8. In this system B_j, C_j are unknown coefficients which determine the solution to the scattering problem, while A_j are known and using the results from §2.4.2 we can express these as

$$A_j = \frac{-V}{2\pi} \kappa^-(\tilde{\sigma}_\eta^-) \int_{\Gamma^-} \frac{e^{-i\alpha} \kappa^+(\alpha)}{(\alpha - \tilde{\sigma}_j^+)(\alpha - \tilde{\sigma}_\eta^-)} d\alpha,$$

where, as we outlined in §2.3.2, the incidence parameter η , takes values $-r \leq \eta \leq q$, with $\eta = 0$ corresponding to the case of an incident gust, and $\eta = 1$ corresponding to the case of an incident sound wave.

2.5.1 Stable formulation under $\text{Im } \tilde{\Omega} \rightarrow 0^-$

We now wish to reduce the fictitious damping $\text{Im } \tilde{\Omega}$ to zero. Note in the above analysis we chose the contours of integration in the system Eqs. (2.30)-(2.31), i.e. the oriented curves Γ^- and $\text{Im } \alpha = \epsilon$, such that all poles and zeros in R^+ are above the curves and all poles and zeros in R^- are below the curves. Therefore if we were to *fix our current contours of integration* and then considered the limit $\text{Im } \tilde{\Omega} \rightarrow 0^-$, the poles of the cut-on modes would then have to cross the contours of integration in order to move onto the real-axis. In addition, if we allowed the contours of integration to change according to the above restrictions when $\text{Im } \tilde{\Omega} \rightarrow 0^-$, then the contours would coalesce on the real axis (at least in a neighbourhood of the origin) and the poles corresponding to cut-on modes would end up on the contours. This means that our current contours (although appropriate for the analytical treatment) are not suitable for numerical evaluation in the physical case of zero damping (i.e. when $\text{Im } \tilde{\Omega} = 0$). In order to overcome this problem we change contours of integration to a set of contours that allows us to evaluate the linear system numerically even when all the cut-on modes are located on the real axis. Thus we consider the following change of contours: Choose $|\text{Im } \tilde{\Omega}| = \delta$ sufficiently small such that there is a constant $\epsilon_1 > 0$ such that for all $0 \leq \text{Im } \tilde{\Omega} \leq \delta$, we have $\text{Im } \tilde{\Omega} \frac{M}{M_x} > -\epsilon_1$,

$$|\text{Im } \sigma_m^\pm| < \epsilon_1, \text{ if } -r \leq m \leq q+1, \text{ and } |\text{Im } \sigma_m^\pm| > \epsilon_1, \text{ if } m > q+1, m < -r,$$

i.e. such that all cut-on modes are inside a strip of width $2\epsilon_1$ around the real axis, and all cut-off modes are outside. Then we can change the contours as follows, in each case picking up a pole contribution if the corresponding mode is cut-on:

$$\begin{aligned} F_{jm} &= \frac{1}{2\pi i} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) \int_{-\infty+i\epsilon_1}^{\infty+i\epsilon_1} \frac{e^{-i\alpha} \kappa^+(\alpha)}{(\alpha - \tilde{\sigma}_j^+)(\alpha - \tilde{\sigma}_m^-)} d\alpha \\ &\quad + \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) \begin{cases} e^{-i\tilde{\sigma}_j^+} \kappa^+(\tilde{\sigma}_j^+) \frac{1}{\tilde{\sigma}_j^+ - \tilde{\sigma}_m^-}, & \text{if } \tilde{\sigma}_j^+ \text{ is cut-on,} \\ 0, & \text{otherwise,} \end{cases} \\ G_{jm} &= \frac{-1}{2\pi i} \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_m^+ \right) \int_{-\infty-i\epsilon_1}^{\infty-i\epsilon_1} \frac{e^{i\alpha} \kappa^-(\alpha)}{(\alpha - \tilde{\sigma}_j^-)(\alpha - \tilde{\sigma}_m^+)} d\alpha \\ &\quad + \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_m^+ \right) \begin{cases} e^{i\tilde{\sigma}_j^-} \kappa^-(\tilde{\sigma}_j^-) \frac{1}{\tilde{\sigma}_j^- - \tilde{\sigma}_m^+}, & \text{if } \tilde{\sigma}_j^- \text{ is cut-on,} \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

$$A_j = -\frac{V}{2\pi} \kappa^-(\tilde{\sigma}_\eta^-) \int_{-\infty+i\epsilon_1}^{\infty+i\epsilon_1} \frac{e^{-i\alpha} \kappa^+(\alpha)}{(\alpha - \tilde{\sigma}_j^+)(\alpha - \tilde{\sigma}_\eta^-)} d\alpha$$

$$- iV \begin{cases} e^{-i\tilde{\sigma}_j^+} \kappa^+(\tilde{\sigma}_j^+) \kappa^-(\tilde{\sigma}_\eta^-) \frac{1}{\tilde{\sigma}_j^+ - \tilde{\sigma}_\eta^-}, & \text{if } \tilde{\sigma}_j^+ \text{ is cut-on,} \\ 0, & \text{otherwise.} \end{cases}$$

In these new expressions we are able to take the limit $\text{Im } \tilde{\Omega} \rightarrow 0^-$ while keeping the contour of integration fixed, thus allowing us to solve the resulting system without the fictitious damping (by setting $\text{Im } \tilde{\Omega} = 0$).

2.5.2 Approximate solution of the linear system

The system Eqs. (2.30) & (2.31) is equivalent to:

$$(I - GF)B = GA, \quad (2.33)$$

$$C = FB, \quad (2.34)$$

where I is the identity operator. For our numerical results in §2.8 we solve Eq. (2.33) approximately using the finite section method (truncation), together with a numerical approximation to the integral coefficients F, G . The precise version of the finite section method employed in our numerical examples as well as a convergence analysis is described in Appendix 2.D.

In light of Chapters 5 & 6 we mention that the cost of solution of the scattering problem using this present Wiener–Hopf-based approach is frequency-independent and indeed increases with frequency for the following reasons. Firstly, an increase in frequency leads to a larger number of cut-on radiation and duct modes (cf. the explicit expressions given in §2.3.1) which means that the onset of asymptotic decay of the entries of matrices and vectors in Eqs. (2.30) & (2.31) will occur at larger indices. This means more equations must be retained to provide an accurate approximation using the finite section method. Secondly, the larger number of cut-on poles and zeros will lead to a larger number of alternating zeros and poles of κ on the real axis, which effectively results in oscillations of κ^\pm along the contours of integration for the integral expressions for F, G, A given in §2.5.1. Therefore the numerical implementation has to compute rapidly varying integrals which, as explained in §1.2.3, is computationally expensive unless tailored quadrature methods can be developed. On the upside however, we highlight the work by Peake (1992) who provided an asymptotic kernel factorisation of κ that is valid in the regime $\tilde{\Omega} \gg 1$, and indicates the possibility of reducing the cost of the Wiener–Hopf solution using asymptotic insights.

2.6 Equivalence to previous solutions for overlapping cascades when $d < 1$

Before moving on to consider different observables of the scattering problem which can be evaluated numerically using our solution, we briefly demonstrate that in the case of overlapping blades ($d < 1$) our solution based on pole removal at the radiation modes $\tilde{\sigma}_m^\pm$ is equivalent to the method based on pole removal at duct modes k_n^\pm which was described by Peake (1993) and Glegg (1999).

2.6.1 Review of previous work in the case of overlap

In order to demonstrate this relationship, we briefly summarize the main points of the arguments given in these two papers adopted to the notation and context of this thesis. As already mentioned in Eq. (2.25) when $d < 1$ one can show from physical considerations in the region of overlap $d < x < 1$ (which is a duct, hence only duct modes k_n^\pm are present), that there are complex constants D_n, E_n, F_n such that

$$\left[\tilde{P}_1^+ \right] (\alpha) = \sum_{n \in \mathbb{N}} \frac{D_n e^{-ik_n^-}}{\alpha - k_n^-}, \quad \left[P_2^- \right] (\alpha) = \sum_{n \in \mathbb{N}} \frac{E_n}{\alpha - k_n^+}, \quad \left[\tilde{P}_3^+ \right] (\alpha) = \sum_{n \in \mathbb{N}} \frac{F_n e^{-ik_n^-}}{\alpha - k_n^-}. \quad (2.35)$$

This observation, which is valid only in the case of overlap, provides the basis for an alternative additive splitting to the one which we described in §2.4.4 and §2.4.5: For instance one can split the term

$$\begin{aligned} \frac{1}{\kappa^-(\alpha)} \left[\tilde{P}_1^+ \right] (\alpha) &= \frac{1}{\kappa^-(\alpha)} \sum_{n \in \mathbb{N}} \frac{D_n e^{-ik_n^-}}{\alpha - k_n^-} \\ &= \sum_{n \in \mathbb{N}} \frac{D_n e^{-ik_n^-}}{\alpha - k_n^-} \frac{1}{\kappa^-(k_n^-)} + \sum_{n \in \mathbb{N}} \frac{D_n e^{-ik_n^-}}{\alpha - k_n^-} \left(\frac{1}{\kappa^-(\alpha)} - \frac{1}{\kappa^-(k_n^-)} \right). \end{aligned}$$

This means, when $d < 1$, we can perform a pole removal based on duct modes k_n^\pm instead of our pole removal based on radiation modes $\tilde{\sigma}_m^\pm$. This additive splitting can be used to solve the Wiener–Hopf problems Eqs. (2.14) & (2.15), and in this setting reduce the scattering problem to an infinite system of linear equations which takes the form ($l \in \mathbb{N}$)

$$E_l = - \sum_{n \in \mathbb{N}} \frac{e^{i(k_l^+ - k_l^-)} \text{Res}(\kappa^-, k_l^+)}{k_l^+ - k_n^-} \frac{1}{\kappa^-(k_n^-)} (D_n + F_n), \quad (2.36)$$

$$F_l = - \sum_{n \in \mathbb{N}} \frac{\text{Res}(\kappa^+, k_l^-)}{\kappa^+(k_n^+)} \frac{1}{k_l^- - k_n^+} E_n. \quad (2.37)$$

Here the coefficients D_n are known in terms of $V, \tilde{\sigma}_\eta$ (as a direct consequence of the solution to the uncoupled leading edge problem in Eq. (2.18)) and E_n, F_n are to be determined from Eqs. (2.36)-(2.37).

2.6.2 Proof of equivalence when $d < 1$

Our aim is now to express D_n, E_n, F_n in terms of A_m, B_m, C_m and to see if we can recover Eqs. (2.30) & (2.31) from Eqs. (2.36) & (2.37). The two systems are equivalent in the following sense:

Claim 2.6.1. *Let $0 < d < 1$, and be ϕ_1, ϕ_2, ϕ_3 be a solution to the three coupled boundary value problems as described in §2.3.2. Let all Fourier (half-)transforms be defined as before, and let the constants $A_m, B_m, C_m, m \in \mathbb{Z}$, and $D_l, E_l, F_l, l \in \mathbb{N}$, be defined as*

$$\begin{aligned} A_m &= \left[\tilde{P}_1^+ \right] (\tilde{\sigma}_m^+), & D_l &= e^{ik_n^-} \text{Res} \left(\left[\tilde{P}_1^+ \right] (\alpha), k_n^- \right), \\ B_m &= \left[P_2^- \right] (\tilde{\sigma}_m^-), & E_l &= \text{Res} \left(\left[P_2^- \right] (\alpha), k_n^+ \right), \\ C_m &= \left[\tilde{P}_3^+ \right] (\tilde{\sigma}_m^+), & F_l &= e^{ik_n^-} \text{Res} \left(\left[\tilde{P}_3^+ \right] (\alpha), k_n^- \right). \end{aligned}$$

Then the following equivalence is true:

$$\left\{ A_m, B_m, C_m \text{ satisfy (2.30)-(2.31)} \right\} \iff \left\{ D_l, E_l, F_l \text{ satisfy (2.36)-(2.37)} \right\}.$$

Proof. We shall demonstrate here only that D_l, E_l, F_l satisfying Eq. (2.37) implies that A_m, B_m, C_m satisfy Eq. (2.31) and we note that all other implications follow similarly. From Eq. (2.35) we find that

$$A_m = \sum_{n \in \mathbb{N}} \frac{D_n e^{-ik_n^-}}{\tilde{\sigma}_m^+ - k_n^-}, \quad B_m = \sum_{n \in \mathbb{N}} \frac{E_n}{\tilde{\sigma}_m^- - k_n^+}, \quad C_m = \sum_{n \in \mathbb{N}} \frac{F_n e^{-ik_n^-}}{\tilde{\sigma}_m^+ - k_n^-}.$$

This implies, using (2.37), for all $j \in \mathbb{Z}$:

$$C_j = - \sum_{n \in \mathbb{N}} \frac{e^{-ik_n^-}}{\tilde{\sigma}_j^+ - k_n^-} \sum_{l \in \mathbb{N}} \frac{\text{Res}(\kappa^+, k_n^-)}{\kappa^+(k_l^+)} \frac{1}{k_n^- - k_l^+} E_l. \quad (2.38)$$

At this point the following identity (valid when $d < 1$) proves useful:

Lemma 2.6.2.

$$-\frac{1}{\kappa^+(k_l^+)} \frac{1}{k_n^- - k_l^+} = \sum_{m \in \mathbb{Z}} \frac{1}{k_n^- - \tilde{\sigma}_m^-} \frac{1}{\tilde{\sigma}_m^- - k_l^+} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right). \quad (2.39)$$

Proof. Both of the above sides are found to be equal to the following integral

$$- \int_{-\infty}^{\infty} \frac{1}{(k_n^- - \alpha)(\alpha - k_l^+) \kappa^+(\alpha)} d\alpha,$$

by closing the contour in the correct complex half-plane and collecting the appropriate residue contributions (notice that k_n^- is a zero of $\frac{1}{\kappa^+}$ hence does not contribute a residue when the contour is closed in the lower half plane). The contours can be closed on either side since we showed in §2.4.3 that $\frac{1}{\kappa^+}$ has at worst $\mathcal{O}(\alpha^{-\frac{1}{2}})$ growth away from its poles in either half-plane. \blacksquare

Using Eqs. (2.38) & (2.39) we have

$$\begin{aligned} C_j &= \sum_{n \in \mathbb{N}} \frac{e^{-ik_n^-}}{\tilde{\sigma}_j^+ - k_n^-} \sum_{l \in \mathbb{N}} \sum_{m \in \mathbb{Z}} \frac{1}{k_n^- - \tilde{\sigma}_m^-} \frac{1}{\tilde{\sigma}_m^- - k_l^+} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) \text{Res}(\kappa^+, k_n^-) E_l \\ &= \sum_{m \in \mathbb{Z}} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) \sum_{n \in \mathbb{N}} \frac{e^{-ik_n^-}}{\tilde{\sigma}_j^+ - k_n^-} \frac{1}{k_n^- - \tilde{\sigma}_m^-} \text{Res}(\kappa^+, k_n^-) \sum_{l \in \mathbb{N}} \frac{1}{\tilde{\sigma}_m^- - k_l^+} E_l \\ &= \sum_{m \in \mathbb{Z}} F_{jm} B_m, \end{aligned}$$

which is precisely Eq. (2.31). The final line follows by considering the form of F_{jm} , which according to Eq. (2.32) is

$$F_{jm} = \frac{1}{2\pi i} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) \int_{\Gamma^-} \frac{e^{-i\alpha} \kappa^+(\alpha)}{(\alpha - \tilde{\sigma}_j^+)(\alpha - \tilde{\sigma}_m^-)} d\alpha.$$

We recall from the discussion in §2.4.3 that κ^+ grows at worst like $\mathcal{O}(e^{d|\text{Im} \alpha})$ as $\alpha \rightarrow \infty$ in R^- . Therefore, when $d < 1$, we can close the contour of F_{jm} in R^- to obtain only contributions at the simple poles of κ^+ which are given by

$$F_{jm} = - \sum_{n \in \mathbb{N}} \frac{1}{k_n^- - \tilde{\sigma}_j^+} e^{-ik_n^-} \text{Res}(\kappa^+, k_n^-) \frac{1}{k_n^- - \tilde{\sigma}_m^-} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right). \quad (2.40)$$

Thus we have indeed shown that Eq. (2.37) implies Eq. (2.31), and by using analogous arguments we can prove that Eq. (2.36) implies Eq. (2.30). The implication in the other direction follows analogously by considering the expressions Eqs. (2.57) & (2.58) which allow us, when $d < 1$, to express E_l and F_l respectively in terms of A_m, B_m, C_m using a similar expression of the corresponding integrals in terms of pole contributions as above. \square

Remark 2.6.3. *We have thus shown that, although our approach, using a pole removal splitting based on radiation modes $\tilde{\sigma}_m^\pm$, works for general values of $d > 0$, it is equivalent to previous work using pole removal at duct modes k_n^\pm , when $d < 1$. This is mainly because in the overlapping case the coefficients in the linear system Eqs. (2.30) & (2.31) simplify*

as was shown in Eq. (2.40). This trick is not possible when $d > 1$, however the integral expressions for F_{jm}, G_{jm} remain valid in the general case.

2.7 Total unsteady lift and far-field behaviour

We can now use our solution to find expressions for the total unsteady lift and the form of the velocity potential in the far-field given an incident gust, $\eta = 0$, and an incident sound wave, $\eta = 1$.

2.7.1 Total unsteady lift

The total unsteady lift on a single blade is (recalling the boundary condition (ii))

$$\mathcal{L} = \int_0^1 [p](x) dx = \int_{-\infty}^{\infty} [p](x) dx = [P](0).$$

Given our solution we can evaluate this quantity as follows: By the boundary condition (ii) we have

$$[P](\alpha) = [P_1^+](\alpha) + [P_2^+](\alpha) + [P_3^+](\alpha), \quad (2.41)$$

and we showed earlier that

$$[P_1^+](\alpha) = -iV\kappa^-(\tilde{\sigma}_\eta^-) \frac{\kappa^+(\alpha)}{\alpha - \tilde{\sigma}_\eta^-}, \quad (2.42)$$

$$[P_3^+](\alpha) = \kappa^+(\alpha) \sum_{m \in \mathbb{Z}} \frac{B_m}{\alpha - \tilde{\sigma}_m^-} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right). \quad (2.43)$$

Furthermore

$$[P_2^+](\alpha) = e^{i\alpha} [\tilde{P}_2^+](\alpha) + e^{i\alpha} [\tilde{P}_2^-](\alpha) - [P_2^-](\alpha), \quad (2.44)$$

and we can express these in terms of our known quantities A_m, B_m, C_m as follows (for Eq. (2.45), (2.49) & (2.50) we used Fourier inversion and then the appropriate forward half-transform):

$$[P_2^-](\alpha) = \sum_{m \in \mathbb{Z}} (A_m + C_m) \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_m^+ \right) \left(\frac{-1}{2\pi i} \right) \int_{-\infty+i\epsilon}^{\infty+i\epsilon} \frac{e^{i\alpha' \kappa^-(\alpha')}}{(\alpha' - \alpha)(\alpha' - \tilde{\sigma}_m^+)} d\alpha', \quad (2.45)$$

$$[\tilde{P}_2^-](\alpha) = \kappa^-(\alpha) \sum_{m \in \mathbb{Z}} (A_m + C_m) \frac{1}{\alpha - \tilde{\sigma}_m^+} \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_m^+ \right), \quad (2.46)$$

(2.47)

$$\left[\tilde{P}_2^+ \right] (\alpha) = - \left[\tilde{P}_1^+ \right] (\alpha) - \left[\tilde{P}_3^+ \right] (\alpha), \quad (2.48)$$

$$\left[\tilde{P}_1^+ \right] (\alpha) = \frac{-iV}{2i\pi} \int_{-\infty-i\epsilon}^{\infty-i\epsilon} \frac{e^{-i\alpha'} \kappa^+(\alpha')}{(\alpha' - \alpha)(\alpha' - \tilde{\sigma}_\eta^-)} \kappa^-(\tilde{\sigma}_\eta^-) d\alpha', \quad (2.49)$$

$$\left[\tilde{P}_3^+ \right] (\alpha) = \frac{1}{2i\pi} \sum_{m \in \mathbb{Z}} B_m \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) \int_{\Gamma^-} \frac{e^{-i\alpha'} \kappa^+(\alpha')}{(\alpha' - \alpha)(\alpha' - \tilde{\sigma}_m^-)} d\alpha'. \quad (2.50)$$

Similarly to §2.5.1 we can express the contour integrals in a form that is stable under $\text{Im } \tilde{\Omega} \rightarrow 0^-$ and in summary the total unsteady lift has the following form which can be evaluated numerically:

$$\begin{aligned} \mathcal{L} &= iV \kappa^-(\tilde{\sigma}_\eta^-) \frac{1}{2\pi i} \int_{-\infty+i\epsilon_1}^{\infty+i\epsilon_1} \frac{e^{-i\alpha} \kappa^+(\alpha)}{\alpha(\alpha - \tilde{\sigma}_\eta^-)} d\alpha \\ &\quad - \sum_{m \in \mathbb{Z}} \text{Res} B_m \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) \frac{1}{2i\pi} \int_{-\infty+i\epsilon_1}^{\infty+i\epsilon_1} \frac{e^{-i\alpha} \kappa^+(\alpha)}{\alpha(\alpha - \tilde{\sigma}_m^-)} d\alpha \\ &\quad + \sum_{m \in \mathbb{Z}} (A_m + C_m) \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_m^+ \right) \frac{1}{2\pi i} \int_{-\infty-i\epsilon_1}^{\infty-i\epsilon_1} \frac{e^{i\alpha} \kappa^-(\alpha)}{\alpha(\alpha - \tilde{\sigma}_m^+)} d\alpha. \end{aligned}$$

Here $\epsilon_1 > 0$ is such that when $\text{Im } \tilde{\Omega} = 0$ all cut-off modes lie outside the strip $|\text{Im } \alpha| \leq \epsilon_1$.

2.7.2 Far-field behaviour

We aim to understand the behaviour of the scattered potential ϕ far downstream ($x > 0, |x| \gg 1$) and far upstream ($x < 0, |x| \gg 1$) of the blades. To do so we observe the following identity, which follows after a few steps of algebra from Eq. (2.8) together with the periodicity Eq. (2.9) and the expression for the pressure Eq. (2.10):

$$\Phi(\alpha, y) = [P](\alpha) \frac{\cosh(\gamma y) e^{i\sigma + id\alpha} - \cosh(\gamma(y-s))}{\kappa(\alpha) \sinh(\gamma s) \gamma}.$$

Thus we can find the velocity potential ϕ by computing the inverse Fourier transform

$$\phi(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\alpha x} [P](\alpha) \frac{\cosh(\gamma y) e^{i\sigma + id\alpha} - \cosh(\gamma(y-s))}{\kappa(\alpha) \sinh(\gamma s) \gamma} d\alpha. \quad (2.51)$$

We can then express $[P]$ using Eqs. (2.41)-(2.50). The resulting integrand when multiplied by $e^{-i\alpha x}$ decays exponentially in the upper half complex plane if $sx < yd$, and decays exponentially in the lower half complex plane if $s(x-1) > yd$. In both cases we can then close the contour of integration, and collect the appropriate pole contributions. This yields, after a few steps of algebra and taking care when exchanging order of summation and integration, the following exact expressions which are valid outside of the cascade structure

i.e. in the regions indicated. The far-field consists of only the contributions from the cut-on modes, as the cut-off modes will decay exponentially as we move downstream/upstream of the cascade respectively.

Far field downstream

For $s(x-1) > yd$ and $0 \leq y \leq s$:

$$\phi(x, y) = \sum_{m \in \mathbb{Z}} e^{-ix\tilde{\sigma}_m^-} f_m(y) \text{Res} \left(\frac{1}{\kappa}, \tilde{\sigma}_m^- \right) \frac{1}{\gamma(\tilde{\sigma}_m^-)} (-i) [P] (\tilde{\sigma}_m^-), \quad (2.52)$$

where

$$\begin{aligned} (-i) [P] (\tilde{\sigma}_m^-) = & \left[iB_m - i \sum_{l \in \mathbb{Z}} (A_l + C_l) e^{i\tilde{\sigma}_m^-} \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_l^+ \right) \frac{\kappa^-(\tilde{\sigma}_m^-)}{\tilde{\sigma}_m^- - \tilde{\sigma}_l^+} \right. \\ & - iV \kappa^-(\tilde{\sigma}_\eta^-) e^{i\tilde{\sigma}_m^-} \frac{1}{2\pi} \int_{-\infty+i\epsilon_1}^{\infty+i\epsilon_1} \frac{e^{-i\alpha}}{\alpha - \tilde{\sigma}_\eta^-} \frac{\kappa^+(\alpha)}{\alpha - \tilde{\sigma}_m^-} d\alpha \\ & \left. + \sum_{l \in \mathbb{Z}} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_l^- \right) B_l e^{i\tilde{\sigma}_m^-} \frac{1}{2\pi} \int_{-\infty+i\epsilon_1}^{\infty+i\epsilon_1} \frac{e^{-i\alpha}}{\alpha - \tilde{\sigma}_l^-} \frac{\kappa^+(\alpha)}{\alpha - \tilde{\sigma}_m^-} d\alpha \right]. \end{aligned}$$

In the above expression we have introduced

$$f_m(y) = \frac{\cosh(\gamma(\tilde{\sigma}_m^-)y) e^{i\sigma + id\tilde{\sigma}_m^-} - \cosh(\gamma(\tilde{\sigma}_m^-)(y-s))}{\sinh(\gamma(\tilde{\sigma}_m^-)s)},$$

which can be further simplified for the acoustic modes (i.e. when $m \neq 0$):

$$f_m(y) = \begin{cases} e^{\gamma(\tilde{\sigma}_{m-1}^-)y}, & \text{if } m \geq 1, \\ e^{\gamma(\tilde{\sigma}_m^-)y}, & \text{if } m < 0. \end{cases}$$

We observe that the far field downstream of the blades consists only of radiating modes that travel downstream, and a hydrodynamic mode $\tilde{\Omega} \frac{M}{M_x}$ which supports the wake.

Far field upstream

For $sx < yd$ and $0 \leq y \leq s$:

$$\phi(x, y) = \sum_{m \in \mathbb{Z}} e^{-i\tilde{\sigma}_m^+ x} \text{Res} \left(\frac{1}{\kappa}, \tilde{\sigma}_m^+ \right) \frac{e^{-\gamma(\tilde{\sigma}_m^+)y}}{\gamma(\tilde{\sigma}_m^+)} i [P] (\tilde{\sigma}_m^+), \quad (2.53)$$

where

$$\begin{aligned} i[P](\tilde{\sigma}_m^+) &= V\kappa^-(\tilde{\sigma}_\eta^-) \frac{\kappa^+(\tilde{\sigma}_m^+)}{\tilde{\sigma}_m^+ - \tilde{\sigma}_\eta^-} + i\kappa^+(\tilde{\sigma}_m^+) \sum_{l \in \mathbb{Z}} \frac{B_l}{\tilde{\sigma}_m^+ - \tilde{\sigma}_l^-} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_l^- \right) \\ &+ \sum_{l \in \mathbb{Z}} \frac{A_l + C_l}{2\pi} \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_l^+ \right) \int_{-\infty - i\epsilon_1}^{\infty - i\epsilon_1} \frac{e^{i\alpha}}{\alpha - \tilde{\sigma}_l^+} \frac{\kappa^-(\alpha)}{\alpha - \tilde{\sigma}_m^+} d\alpha - ie^{i\tilde{\sigma}_m^+} (A_m + C_m). \end{aligned}$$

Similarly to the far field downstream, the scattered potential upstream consist only of radiating modes which travel upstream. Moreover, there is no contribution from the hydrodynamic mode, since there is no wake upstream of the blades.

2.8 Numerical examples and results

The expressions for total unsteady lift and far-field amplitudes which we derived in §2.7 allow us to provide numerical results on the effects of cascade geometry and incident-wave properties on the scattered potential. The effects of incidence angle, spanwise wave numbers and mean flow are studied in detail in Kaji and Okazaki (1970), Koch (1971), and Glegg (1999), and we therefore do not aim to provide a comprehensive parametric study of these effects, but rather we show typical results that can be achieved using our method, with particular focus on the effect of blade spacing and of non-overlapping cascade geometries. Furthermore, we note that the effect of the cross-flow W was absorbed in our definition of $\tilde{\Omega}$, so that a larger cross-flow simply results in a smaller value of the effective reduced frequency, and therefore we need not study the effect of cross-flow separately. We also provide examples that allow comparison to previous results in order to validate the accuracy of our method. Finally, we note that m , the number of equations retained when solving Eq. (2.33) using the finite section method (see §2.5.2), was chosen to ensure the results shown are close to the limit value as $m \rightarrow \infty$ (i.e. the algorithm has converged). The typical number chosen was $m \approx 70$, but for large mean flow significantly fewer equations were required whilst for small mean flow a larger number was retained ($m \approx 250$ when $M = \sqrt{M_x^2 + M_z^2} \approx 0.2$).

2.8.1 Total unsteady lift

For the total unsteady lift we begin by comparing our solution to results provided by Peake (1993) for overlapping blades in a two-dimensional setting. For this we consider the same parameter settings as in Peake (1993, p. 269): $M_z = 0, K_3 = 0, \eta = 0$ (i.e. an incident gust of reduced frequency Ω), $\tilde{\Omega} = \Omega = 1.0, \sigma = 2\pi/3, s = d = 1/\sqrt{2}, M = M_x \in [0.2, 0.9]$.

The results can be seen in Figs. 2.3a and 2.3c and our solution appears to be in very good agreement with the reference result. The rapid variation observed in Fig. 2.3c

corresponds to the first downstream radiating mode σ_0^- becoming cut-on.

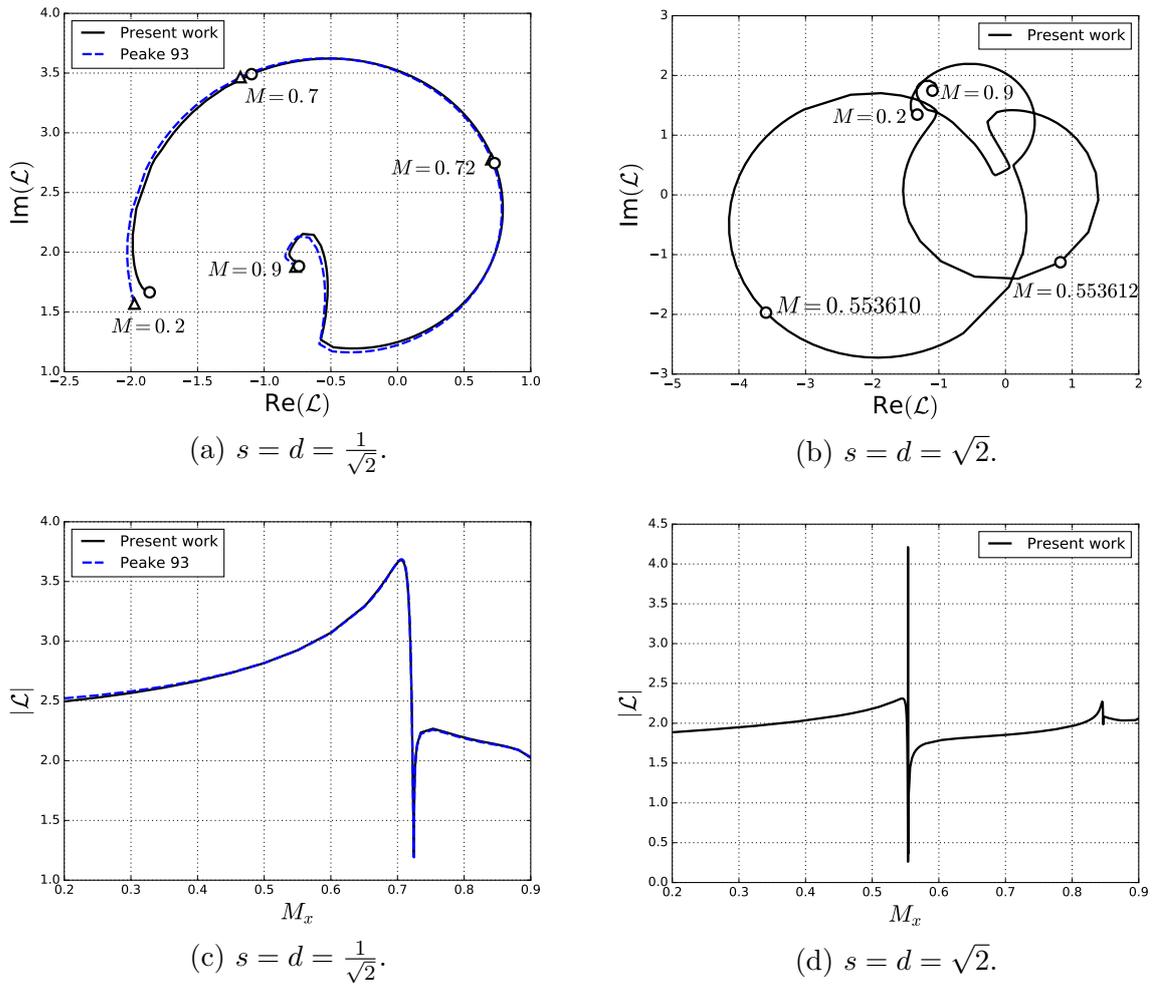


Figure 2.3: Total lift \mathcal{L} for $M \in [0.2, 0.9]$.

However, some discrepancy between our present work and the reference solution (Peake, 1993) can be observed in Fig. 2.3a for small and large values of M . This is due to the algebraic decay in the coefficients of our linear system Eqs. (2.33)-(2.34), which is found to be slower (not in the algebraic order, but in the constant multiplying this decay rate) for small and large values of M , which means in effect more terms must be retained in the linear system. In this particular numerical example the number of equations retained ($m \approx 250$) was at the limit of what could be achieved in reasonable computing time. We expect that more efficient implementation and an appropriate preconditioning of the linear system might lead to improved performance, however we also highlight that for potential practical application a relative accuracy below 5% is often sufficient.

In Figs. 2.3b and 2.3d we plot the total unsteady lift in a case of non-overlapping blades (inaccessible to Peake (1993) and other previous work), with $s = d = \sqrt{2}$, and the remaining parameters as before. We observe overall a similar shape of the total lift. The increased blade spacing however has the effect of the radiating modes cutting on at lower

Mach numbers which results in a rapid variation at $M \approx 0.55$ (corresponding to σ_0^- cutting on) and at $M \approx 0.85$ (corresponding to σ_1^- cutting on). Moreover, comparing Fig. 2.3c and Fig. 2.3d we can see that the amplitude of the unsteady lift is (almost always) reduced in the non-overlapping case, as would be expected for a more widely spaced cascade.

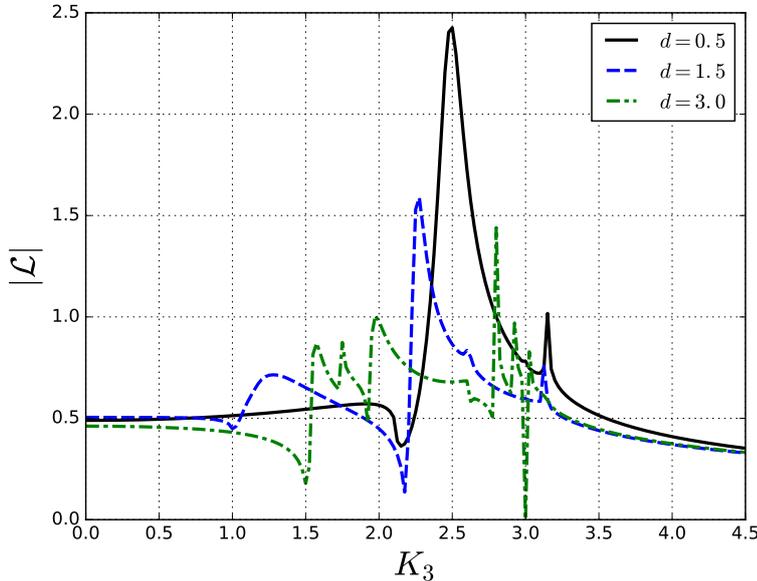


Figure 2.4: Total lift $|\mathcal{L}|$ as a function of K_3 .

Next we study the effect of spanwise wavenumber on the total lift. We consider $M = M_x = 0.3, \tilde{\Omega} = \Omega = 10.0, \eta = 0$ (i.e. an incident gust), $\sigma = 3\pi/4, s \in \{0.6, 1.8, 3.6\}, \alpha = 5\pi/18$ and vary $K_3 \in [0, 4.5]$. The total unsteady lift is shown in Fig. 2.4, where in the black curve we can see the result for a case of overlapping blades and the remaining ones correspond to non-overlapping blades. In the case $d = 0.5$ the effect of K_3 is most apparent: From the formulae in §2.3.1 increasing K_3 sufficiently far has the effect of cutting off the acoustic modes. Indeed, we observe a rapid variation in the lift around $K_3 \approx 2.4$ which is just after the radiating mode σ_0^- has cut off. For increased blade spacing we observe the interplay of this cut-off effect with increasing K_3 to the effect of more modes becoming cut-on as spacing increases. For $d = 3.0, K_3 = 0$ all of $\sigma_m^\pm, -2 \leq m \leq 2$, and $k_n^\pm, -3 \leq n \leq 3$, are cut-on, however as we increase K_3 they all cut off successively resulting in the rapid variations observed in the corresponding green curve. As we reach $K_3 = \frac{\Omega M_x}{\beta} \approx 3.14$ all of the acoustic modes have to become cut-off since by the dispersion relation Eq. (2.6) there are no propagating acoustic modes for $K_3 > \frac{\Omega M_x}{\beta}$. This is indeed observed in Fig. 2.4.

Finally, we consider the effect of the cascade stagger angle α_0 on the total lift. In Fig. 2.5 we see the dependence of the total unsteady lift on α_0 , with $\sqrt{s^2 + d^2} = 1.5$ fixed. The remaining parameters were chosen as $M = M_x = 0.3, M_z = 0, \eta = 0, \sigma = 3\pi/4, \tilde{\Omega} = \Omega = 1.0, K_3 \in \{0, 1.0, 2.0, 3.0\}$. We note that here we have $d > 1$ for $\alpha_0 < 48.18^\circ$.

Most importantly we observe the smooth transition of \mathcal{L} from the non-overlapping to the overlapping regime as α_0 increases from 0. We also observe an apparently singular behaviour as $\alpha_0 \rightarrow 0$, which corresponds to the limit of no vertical separation between the blades where our model is clearly invalid.

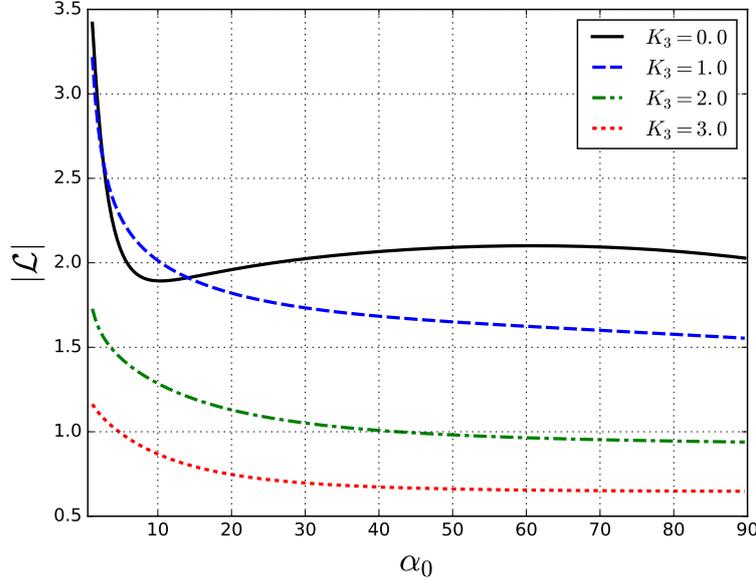


Figure 2.5: Total lift $|\mathcal{L}|$ as a function of α_0 .

2.8.2 Far field behaviour

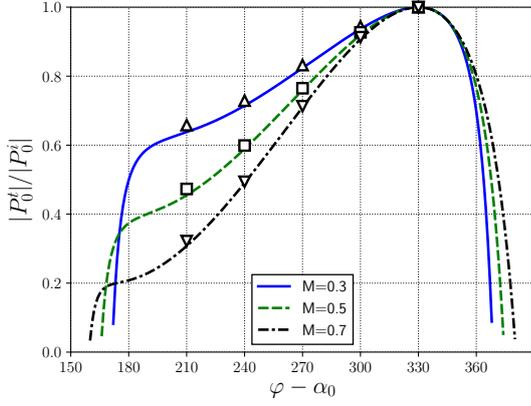
We now consider the far field sound for an incident acoustic wave as derived in §2.7.2. We begin by reproducing results given by Koch (1971, p. 125), who considered the transmission and reflection amplitudes relative to the incident amplitude as a function of angle of incidence. In order to match Koch's settings we restrict ourselves to the case $M_z = 0, \theta = \pi/2$, which means we consider an incident acoustic wave ($\eta = 1$) with $K_1 = \frac{\tilde{\Omega}M \cos \varphi}{1+M_x \cos \varphi}, K_2 = \frac{\tilde{\Omega}M \sin \varphi}{1+M_x \cos \varphi}, K_3 = 0$. We further choose $(s, d) \in \{(\sin \alpha_0, \cos \alpha_0), (2 \sin \alpha_0, 2 \cos \alpha_0), (3 \sin \alpha_0, 3 \cos \alpha_0)\}, M = M_x \in \{0.3, 0.5, 0.7\}, k_0 = 0.25\pi, \alpha_0 = \pi/6$ and $\tilde{\Omega} = \Omega = k_0/M$. In Fig. 2.6 we show the transmission and reflection amplitudes relative to the incident amplitude $I = \gamma(\sigma_1^-)^{-1}V$.

These amplitudes $T_m, R_m, m \in \mathbb{Z}$, are such that

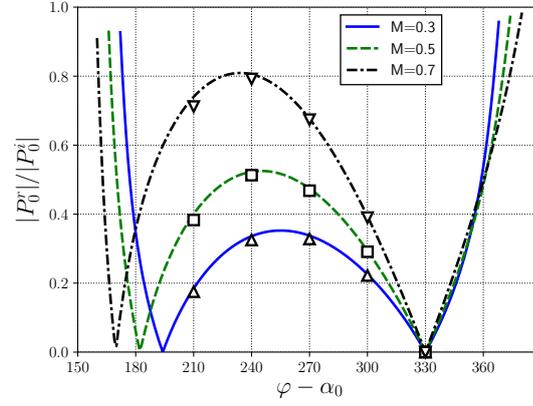
$$\phi(x, y) = \begin{cases} T_{\tilde{\Omega} \frac{M}{M_x}} e^{-ix \tilde{\Omega} \frac{M}{M_x}} f_0(y) + \sum_{m \in \mathbb{Z}} T_m e^{-ix \sigma_m^- + \gamma(\sigma_m^-)y} & x > d + 1, \\ \sum_{m \in \mathbb{Z}} R_m e^{-i\sigma_m^+ x - \gamma(\sigma_m^+)y} & x < 0, \end{cases}$$

and they are explicitly given in Eq. (2.52) and (2.53). In order to match Koch's setting we plot the modal pressure corresponding to the first transmitted and reflected acoustic mode

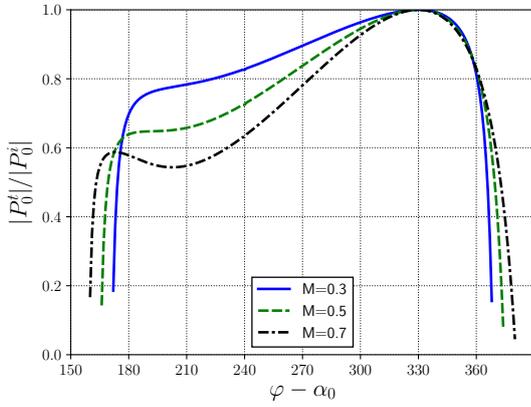
of the total field $\phi + \phi_{\text{inc}}$. In our case these are given by $P_0^t = i(\sigma_0^- - \tilde{\Omega}M/M_x)(T_0 + I)$, $P_0^i = i(\sigma_0^- - \tilde{\Omega}M/M_x)I$, $P_0^r = i(\sigma_0^+ - \tilde{\Omega}M/M_x)R_0$.



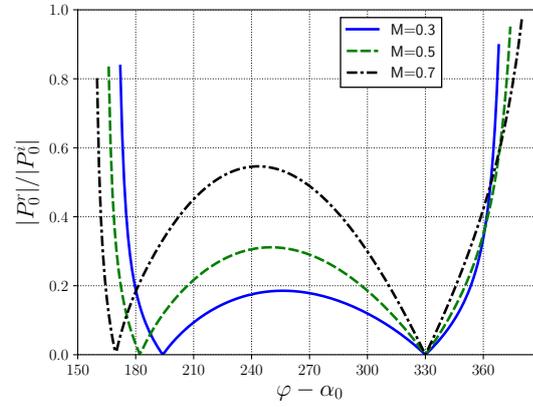
(a) $d = \cos \alpha_0 \approx 0.87$



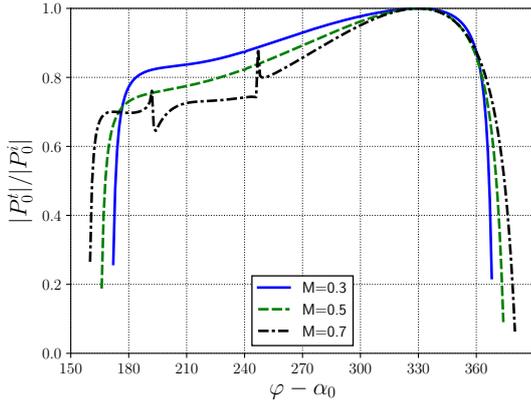
(b) $d = \cos \alpha_0 \approx 0.87$



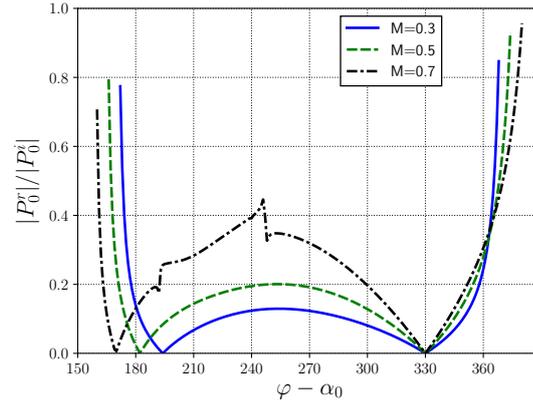
(c) $d = 2 \cos \alpha_0 \approx 1.73$



(d) $d = 2 \cos \alpha_0 \approx 1.73$



(e) $d = 3 \cos \alpha_0 \approx 2.60$



(f) $d = 3 \cos \alpha_0 \approx 2.60$

Figure 2.6: Relative transmission and reflection amplitudes of the incident wave. The points in Figs. 2.6a and 2.6b represent reference values taken from Koch (1971).

The results are shown in comparison to a number of point values taken from Koch

(1971) in Figs. 2.6a and 2.6b. Our solution is found to be in good agreement with the cited results. We are also able to show the corresponding results when the blade spacing is increased. The overall trend to be observed is an increase in $|P_0^t/P_0^i|$ and a decrease in $|P_0^r|/|P_0^i|$ as the blade spacing increases. The increased blade spacing also allows for more interesting effects to occur: We observe rapid variations in Figs. 2.6e and 2.6f when $M = 0.7$ around $\varphi - \alpha_0 \approx 190^\circ, 247^\circ$ which correspond to radiating and duct modes becoming cut-on. In this example we focused our attention to acoustic waves which are incident from upstream, but we note that a similar plot is made available, and studied further also for downstream incidence in §3.4.

The effect of increased transmission and reduced reflection is further observed in our final result, which shows the relative transmission and reflection amplitudes for the modes $\sigma_0^-, \sigma_1^-, \sigma_0^+$ as a function of blade spacing. This result is shown in Fig. 2.7, where we consider a similar setting as in Fig. 2.6 with $M = M_x = 0.7, \varphi = 7\pi/6, \alpha_0 = \pi/6$ and vary $d \in [0.5, 6]$ with s/d fixed.

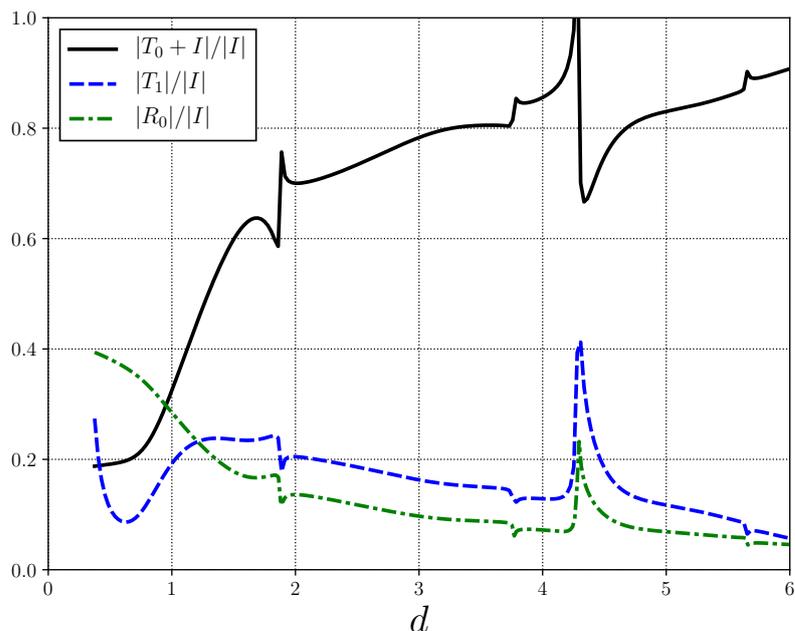


Figure 2.7: Relative transmission and reflection amplitudes as a function of d .

The figure demonstrates two central effects of the increasing blade spacing on the far field: Firstly, the increase in spacing results in an increasing trend of $|T_0 + I|/|I|$ to 1 and a decreasing trend of all other amplitudes to 0. Secondly, as d increases, the radiating modes and duct modes become cut-on successively, resulting in the rapid variations observed in the graph. Both effects can be interpreted to mean that, as an acoustic obstacle, the cascade becomes more permeable as the spacing increases.

2.9 Conclusions

In this chapter we have developed the first Wiener–Hopf solution to the scattering problem of a cascade of finite-length non-overlapping blades. The solution allowed us to express the scattered field only relying on the solution of an infinite algebraic system with decaying coefficients, which can be provably truncated to achieve convergent numerical approximations. In contrast, all previous work resulted in systems that are well-behaved when $d \leq 1$, but whose coefficients are exponentially increasing when $d > 1$, thus cannot yield a valid approximation in that case. Our solution applies in a general regime including effects of spanwise wavenumber and cross-flow, as well as incident acoustic and vortical waves, and we provided explicit expressions for the total unsteady lift on each blade as well as the far-field sound upstream and downstream of the cascade. We found that our solution applies to both the overlapping, $d < 1$, and the non-overlapping regime, $d \geq 1$, which showed that a number of features such as the form of the far-field is shared between the regimes, and that quantities such as the total lift transition smoothly as d increases across $d = 1$. We have also provided extensive numerical results demonstrating both the accuracy of our solution in comparison to previous work, and the type of predictions that we are able to achieve for the first time in the non-overlapping regime. This includes the study of transmission amplitudes as the gap spacing increases as well as the effect of the stagger angle α on the lift for large solidity. Our results confirm the physical expectation that the cascade becomes more permeable with increased blade spacing.

We will use the present Wiener–Hopf solution in Chapter 3 to study the energy balance of acoustic wave scattering on this cascade of blades.

List of symbols

Symbol	Description
M_x	Chordwise Mach number.
M_z	Spanwise Mach number.
M	Total Mach number.
P_j^t, P_j^r	Transmitted and reflected modal pressures (for incident sound field).
R^\pm	Overlapping complex half planes.
T_m, R_m	Modal amplitudes of the transmitted and reflected field respectively.
α_0	Cascade stagger angle.
β	Prandtl–Glauert number of chordwise flow.
V	Amplitude of normal velocity of incident wave on the blade.
η	Mode number of incident wave ($\eta = 1$ is acoustic, $\eta = 0$ is vortical wave).
κ	Scalar Wiener–Hopf kernel.
\mathbf{K}	Non-dimensional wave vector of the incident field.
\mathbf{u}	Unsteady velocity field.
\mathcal{L}	Total unsteady lift on a single blade.
D/Dt	Material derivative.
F, G	Infinite matrix operators in the discrete solution of the Wiener–Hopf system.
ϕ	Spatial component of unsteady velocity potential.
σ_n	Streamwise wavenumber of radiation modes.
σ	Interblade phase angle.
θ	Polar incident angle for acoustic waves.
$\tilde{\phi}$	Time-dependent unsteady velocity potential.
$\tilde{\sigma}_n^\pm$	Zeros of the Wiener–Hopf kernel κ .
φ	Azimuthal incident angle for acoustic waves.
$\tilde{\Omega}$	Effective reduced frequency.
d	Blade stagger.
k_n^\pm	Streamwise wavenumber of duct modes.
p	Unsteady time-harmonic pressure field.
s	Interblade spacing.

2.A The kernel κ and its factorisation

For completeness we provide a brief overview of the Wiener–Hopf factorisation of $\kappa(\alpha)$. This and very similar kernels have already been factorised by Koch (1971), Peake (1992) and Glegg (1999), and the reader is referred to these papers for a detailed treatment. The kernel κ is meromorphic with simple poles at k_n^\pm and zeros at $\tilde{\Omega} \frac{M}{M_x}, \sigma_m^\pm$ which cluster linearly along the following rays as $|m|, n \rightarrow \infty$:

$$\sigma_m^\pm \sim 2\pi \frac{-d \pm is\beta}{s^2\beta^2 + d^2} m + \frac{\mp isdMM_x\tilde{\Omega}/\beta \pm is\beta\sigma - MM_x\tilde{\Omega}s^2 - d\sigma}{s^2\beta^2 + d^2} \text{ as } m \rightarrow +\infty, \quad (2.54)$$

$$\sigma_{-m}^\pm \sim 2\pi \frac{d \pm is\beta}{s^2\beta^2 + d^2} m + \frac{\pm isdMM_x\tilde{\Omega}/\beta \mp is\sigma\beta - MM_x\tilde{\Omega}s^2 - d\sigma}{s^2\beta^2 + d^2} \text{ as } m \rightarrow +\infty, \quad (2.55)$$

$$k_n^\pm \sim \frac{\pm i\pi}{s\beta} n - \frac{MM_x\tilde{\Omega}}{\beta^2} \text{ as } n \rightarrow \infty.$$

Thus by the Weierstrass factorisation theorem (see Thm. 5.14 in Conway (1978, p. 170) or the special case given in Noble (1958, p. 40)) we can express κ in the form

$$\kappa(\alpha) = e^{g(\alpha)} (1 - \alpha M_x / (\tilde{\Omega} M)) \frac{\prod_{m \in \mathbb{Z}} (1 - \alpha / \sigma_m^-) e^{\alpha / \sigma_m^-} \prod_{m \in \mathbb{Z}} (1 - \alpha / \sigma_m^+) e^{\alpha / \sigma_m^+}}{\prod_{n=0}^{\infty} (1 - \alpha / k_n^-) e^{\alpha / k_n^-} \prod_{n=0}^{\infty} (1 - \alpha / k_n^+) e^{\alpha / k_n^+}}$$

where $g(\alpha)$ is some entire function. This suggests a construction of κ^+ as follows

$$\kappa^+(\alpha) = e^{\chi_1(\alpha)} (1 - \alpha M_x / (\tilde{\Omega} M)) \frac{(1 - \alpha / \sigma_0^-)}{(1 - \alpha / k_0^-)} \prod_{n=1}^{\infty} \frac{(1 - \alpha / \sigma_n^-) (1 - \alpha / \sigma_{-n}^-)}{(1 - \alpha / k_n^-)}.$$

Let us denote $a_1 = 2\pi \frac{-d - is\beta}{s^2\beta^2 + d^2}$, $a_2 = 2\pi \frac{d - is\beta}{s^2\beta^2 + d^2}$, $a_3 = -\frac{i\pi}{s\beta}$, then the infinite products in above expression are well-defined because, for any fixed $\alpha \in R^+$, as $n \rightarrow \infty$ the factors have the asymptotic behaviour

$$\begin{aligned} \frac{(1 - \alpha / \sigma_n^-) (1 - \alpha / \sigma_{-n}^-)}{(1 - \alpha / k_n^-)} &\sim \frac{\left(1 - \frac{\alpha}{a_1 n} + \mathcal{O}(n^{-2})\right) \left(1 - \frac{\alpha}{a_2 n} + \mathcal{O}(n^{-2})\right)}{\left(1 - \frac{\alpha}{a_3 n} + \mathcal{O}(n^{-2})\right)} \\ &\sim 1 - \frac{\alpha}{n} \underbrace{\left(\frac{1}{a_1} + \frac{1}{a_2} - \frac{1}{a_3}\right)}_{=0} + \mathcal{O}(n^{-2}) \sim 1 + \mathcal{O}(n^{-2}). \end{aligned}$$

To analyse the behaviour of κ^+ and find a suitable choice of χ_1 , we can use the following result from Noble (1958, p. 128):

Lemma 2.A.1 (Consequence of Stirling's formula). *Let*

$$F(\alpha) = \prod_{n=1}^{\infty} (1 + \alpha/\alpha_n) e^{-\alpha/\beta_n},$$

and suppose that $\alpha_n = an + b + \mathcal{O}(n^{-1})$ and $\beta_n = an + c + \mathcal{O}(n^{-1})$ as $n \rightarrow \infty$. Then we have for any $\epsilon > 0$: As $\alpha \rightarrow \infty$ in $\text{Im } \alpha > -\text{Im}(\alpha_1) + \epsilon$, where α_1 is the root of $F(\alpha)$ with smallest imaginary part, that

$$F(\alpha) \sim C_1 \exp \left(\frac{\alpha}{a} (1 - \gamma) - \left(\frac{\alpha}{a} + \frac{b}{a} + \frac{1}{2} \right) \ln \left(\frac{\alpha}{a} + \frac{b}{a} + 1 \right) + \alpha \sum_{n=1}^{\infty} \left(\frac{1}{an} - \frac{1}{\beta_n} \right) \right),$$

where γ is the Euler-Mascheroni constant.

Applying this lemma to our expression for κ^+ shows, after a few steps of algebra, that:

$$\kappa^+(\alpha) \sim \alpha^{\frac{1}{2}} \exp \left[\chi_1(\alpha) + \alpha \left(\frac{1}{a_1} \ln \left(\frac{a_3}{a_1} \right) + \frac{1}{a_2} \ln \left(\frac{a_3}{a_2} \right) \right) \right] \quad \text{as } \alpha \rightarrow \infty \text{ in } R^+.$$

Thus by choosing

$$\chi_1(\alpha) = -\alpha \left(\frac{1}{a_1} \ln \left(\frac{a_3}{a_1} \right) + \frac{1}{a_2} \ln \left(\frac{a_3}{a_2} \right) \right),$$

we can ensure that $\kappa^+(\alpha) \sim \alpha^{\frac{1}{2}}$ as $\alpha \rightarrow \infty$ in R^+ . Using lemma 2.A.1 one finds $g(\alpha)$ to be a constant and so the equivalent expression for κ^- can easily be found from κ/κ^- , which behaves as $\kappa^-(\alpha) \sim \alpha^{-\frac{1}{2}}$ as $\alpha \rightarrow \infty$ in R^+ .

Finally, we note that in this construction κ^{\pm} are only defined up to a multiplicative constant, and thus in our numerical implementation we fix the factors by requiring that $\kappa^+(0) = \kappa(0)$.

2.B Wiener–Hopf splitting for the trailing edge correction

We have outlined in §2.4.4 how to additively split the term

$$g(\alpha) = \frac{1}{\kappa^-(\alpha)} \left([\tilde{P}_1^+] (\alpha) + [\tilde{P}_3^+] (\alpha) \right),$$

and in this appendix we provide the detailed derivation of this splitting. Although we could remove the poles directly by defining the splitting in terms of Eq. (2.24), in §2.4.4 we exploit the equivalence of this pole removal to the Cauchy splitting Eq. (2.23) to show appropriate decay and analyticity of g_{\pm} for the application of Liouville's theorem. In this

appendix we prove the relevant properties of the Cauchy form of this splitting (including that the Cauchy splitting for g is well-defined).

Firstly, we recall from §2.4.2 that

$$[P_1^+] (\alpha) \propto \frac{\kappa^+(\alpha)}{\alpha - \tilde{\sigma}_\eta^-}, \quad \eta = 0, 1,$$

where we have introduced the notation $A(\alpha) \propto B(\alpha)$ to mean that $A(\alpha) = CB(\alpha)$ for some constant $C \in \mathbb{C}$. Thus taking inverse and half-line transform we find:

$$\begin{aligned} [\tilde{P}_1^+] (\alpha) &\propto e^{-i\alpha} \int_1^\infty \int_{-\infty}^\infty e^{-ix(\alpha'-\alpha)} \frac{1}{\alpha' - \alpha} \frac{\kappa^+(\alpha')}{\alpha - \tilde{\sigma}_\eta^-} d\alpha' dx \\ &\propto e^{-i\alpha} \int_1^\infty \int_{\Gamma_-} e^{-ix(\alpha'-\alpha)} \frac{1}{\alpha' - \alpha} \frac{\kappa^+(\alpha')}{\alpha' - \tilde{\sigma}_\eta^-} d\alpha' dx \\ &\propto \int_{\Gamma_-} e^{-i\alpha'} \frac{1}{\alpha' - \alpha} \frac{\kappa^+(\alpha')}{\alpha' - \tilde{\sigma}_\eta^-} d\alpha', \end{aligned} \quad (2.56)$$

where in the second line we use the properties of κ^+ established in §2.4.3 to change the contour of integration to Γ^- which at its tails behaves like

$$\arg \alpha \sim \begin{cases} -\tilde{\epsilon}, & \operatorname{Re} \alpha \rightarrow +\infty, \\ \pi + \tilde{\epsilon}, & \operatorname{Re} \alpha \rightarrow -\infty, \end{cases}$$

for some $0 < \tilde{\epsilon} \ll 1$. It is possible to change to this contour since we observed in §2.4.3 that κ^+ is analytic and of algebraic behaviour in the domain

$$\mathcal{D} = R^+ \cup \left\{ \alpha \mid \arg \alpha \in \left(-\arctan \left(\frac{s\beta}{d} \right), \arctan \left(\frac{s\beta}{d} \right) \right) \cup \left(\pi - \arctan \left(\frac{s\beta}{d} \right), \pi + \arctan \left(\frac{s\beta}{d} \right) \right) \right\},$$

which in particular contains Γ^- (see Fig. 2.8). Along this contour the integrand is absolutely integrable in the product space and so Fubini's theorem applies, and we can exchange order of integration as we did in the third line, Eq. (2.56).

We note now that the final integral in Eq. (2.56) has integrand with exponential decay along Γ_- and thus there is a constant C such that for all $\alpha \in R^+$:

$$\left| [\tilde{P}_1^+] (\alpha) \right| \leq C |\alpha|^{-1} \int_{\Gamma_-} \left| e^{-i\alpha'} \frac{\kappa^+(\alpha')}{\alpha' - \Omega} \right| d\alpha', \quad \text{i.e. } [\tilde{P}_1^+] (\alpha) = \mathcal{O}(\alpha^{-1}).$$

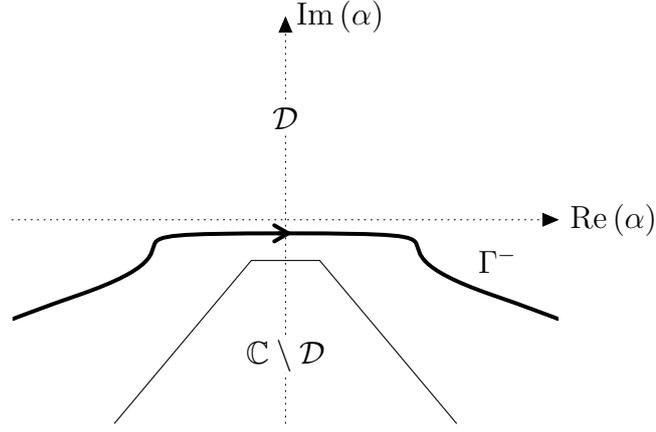


Figure 2.8: The domain \mathcal{D} of algebraic behaviour of $|\kappa^+|$ in the complex α -plane.

As mentioned in §2.4.4, imposing the unsteady Kutta condition at the trailing edge implies that the pressure field must be non-singular at $x = 1$ (see for instance Peake, 1992, p. 274), which means that the above behaviour of $\left[\tilde{P}_1^+\right](\alpha)$ is indeed expected. Imposing the unsteady Kutta condition (condition (vi) in §2.2.1) at the trailing edge to $[p_2]$ and noting that $[p_3](x) = -[p_1](x) - [p_2](x)$, for $x > 1$, we conclude that the pressure jump $[p_3](x)$ must be non-singular at $x = 1$, and it can thus be shown by integration-by-parts that $\left[\tilde{P}_3^+\right](\alpha) = \mathcal{O}(\alpha^{-1})$ as $\alpha \rightarrow \infty, \alpha \in R^+$. As noted in §2.4.4 this implies that there is a strip \mathcal{S} containing the real axis and a constant C such that

$$|g(\alpha)| \leq C (1 + |\operatorname{Re} \alpha|)^{-\frac{1}{2}}, \text{ for all } \alpha \in \mathcal{S},$$

which allows us to apply Thm. B from Noble (1958, p. 13). This yields the additive splitting $g = g_- + g_+$ where:

$$g_+(\alpha) = \frac{1}{2\pi i} \int_{-\infty-ic}^{\infty-ic} \frac{g(\zeta)}{\zeta - \alpha} d\zeta \text{ and } g_-(\alpha) = \frac{1}{2\pi i} \int_{-\infty+ic}^{\infty+ic} \frac{g(\zeta)}{\zeta - \alpha} d\zeta.$$

These functions are analytic in R^+ and R^- respectively (after shrinking the domains R^\pm if necessary), and they are bounded in their respective half-planes, since: For all α , with $\operatorname{Im} \alpha > -\delta$,

$$\left| \int_{-\infty-ic}^{\infty-ic} \frac{g(\zeta)}{\zeta - \alpha} d\zeta \right| \leq C \int_{-\infty}^{\infty} \frac{(1 + |t|)^{-\frac{1}{2}}}{|t - ic - \alpha|} dt \leq \sqrt{2}C \int_{-\infty}^{\infty} \frac{(1 + |t|)^{-\frac{1}{2}}}{|t - \operatorname{Re} \alpha| + |c - \delta|} dt,$$

and so the boundedness of g_+ in R^+ is a consequence of the following lemma:

Lemma 2.B.1. *Let $\epsilon > 0$, then the function $I : \mathbb{R} \rightarrow \mathbb{R}$, defined by*

$$I(y) := \int_{-\infty}^{\infty} \frac{(1 + |t|)^{-\frac{1}{2}}}{|t - y| + \epsilon} dt,$$

is bounded.

Proof. For $y \neq 0$:

$$\begin{aligned}
I'(y) &= \int_{-\infty}^{\infty} \frac{-\operatorname{sgn}(t-y)(1+|t|)^{-\frac{1}{2}}}{(|t-y|+\epsilon)^2} dt \\
&= \int_{-\infty}^y \frac{(1+|t|)^{-\frac{1}{2}}}{(|t-y|+\epsilon)^2} dt - \int_y^{\infty} \frac{(1+|t|)^{-\frac{1}{2}}}{(|t-y|+\epsilon)^2} dt \\
&= \int_{-\infty}^0 \frac{(1+|t-y|)^{-\frac{1}{2}}}{(|t|+\epsilon)^2} dt - \int_0^{\infty} \frac{(1+|t-y|)^{-\frac{1}{2}}}{(|t|+\epsilon)^2} dt \begin{cases} \leq 0, & \text{if } y > 0, \\ \geq 0, & \text{if } y < 0. \end{cases}
\end{aligned}$$

□

The boundedness of g_- in R^- follows analogously. This allows us to solve the trailing-edge correction problem as shown in Eq. (2.27).

2.C Reduction to an infinite algebraic system

In §2.5 we outlined how the scattering problem can be reduced to the solution of an infinite algebraic system, by repeated application of Fourier inversion and change of order of integration. Here we justify these steps rigorously: To derive Eq. (2.30) note that

$$\begin{aligned}
[P_2^-](\alpha') &= \frac{1}{2\pi} \int_{-\infty}^0 \int_{-\infty}^{\infty} e^{ix(\alpha'-\alpha)} e^{i\alpha} [\tilde{P}_2^-](\alpha) d\alpha dx \\
&= \frac{1}{2\pi} \int_{-\infty}^0 \int_{-\infty+i\epsilon}^{\infty+i\epsilon} e^{ix(\alpha'-\alpha)} e^{i\alpha} [\tilde{P}_2^-](\alpha) d\alpha dx \\
&= \frac{-1}{2\pi i} \int_{-\infty+i\epsilon}^{\infty+i\epsilon} \frac{e^{i\alpha\kappa^-}(\alpha)}{\alpha-\alpha'} \sum_{m \in \mathbb{Z}} \frac{1}{\alpha-\tilde{\sigma}_m^+} \operatorname{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_m^+ \right) \left([\tilde{P}_1^+](\tilde{\sigma}_m^+) + [\tilde{P}_3^+](\tilde{\sigma}_m^+) \right) d\alpha,
\end{aligned}$$

where in the second line we changed the contour for some small $\epsilon > 0$ (small enough for the contour to remain within $R^+ \cap R^-$) using Cauchy's theorem and the analyticity of the integrand, and in the third line we used absolute integrability in the product space (along the given contours) to exchange the order of integration by Fubini's theorem, and we substituted $[\tilde{P}_2^-](\alpha)$ using Eq. (2.27). Using the linear asymptotic growth of $\tilde{\sigma}_m^+$ (cf. Eq. (2.54),(2.55)), and the form of residues from Eq. (2.19), we observe that

$$\sum_{m \in \mathbb{Z}} \left| \frac{1}{\alpha-\tilde{\sigma}_m^+} \operatorname{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_m^+ \right) \left([\tilde{P}_1^+](\tilde{\sigma}_m^+) + [\tilde{P}_3^+](\tilde{\sigma}_m^+) \right) \right|$$

is uniformly bounded for $\text{Im}(\alpha) = \epsilon$. Thus we have sufficient decay, of order $O\left(\alpha^{-\frac{3}{2}}\right)$, for

$$\left| \frac{1}{\alpha - \beta} e^{i\alpha \kappa^-}(\alpha) \left| \sum_{m \in \mathbb{Z}} \left| \frac{1}{\alpha - \tilde{\sigma}_m^+} \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_m^+ \right) \left(\left[\tilde{P}_1^+ \right] (\tilde{\sigma}_m^+) + \left[\tilde{P}_3^+ \right] (\tilde{\sigma}_m^+) \right) \right| \right|$$

to be integrable along the given contour. Therefore the dominated convergence theorem applies and we can exchange order of summation and integration to find

$$\begin{aligned} [P_2^-](\alpha') &= \sum_{m \in \mathbb{Z}} \left[\left(\left[\tilde{P}_1^+ \right] (\tilde{\sigma}_m^+) + \left[\tilde{P}_3^+ \right] (\tilde{\sigma}_m^+) \right) \right. \\ &\quad \left. \frac{(-1)}{2\pi i} \int_{-\infty+i\epsilon}^{\infty+i\epsilon} \frac{1}{\alpha - \alpha'} e^{i\alpha \kappa^-}(\alpha) \frac{1}{\alpha - \tilde{\sigma}_m^+} \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_m^+ \right) d\alpha \right]. \end{aligned} \quad (2.57)$$

Thus noting that $B_j = [P_2^-](\tilde{\sigma}_m^-)$ this yields precisely Eq. (2.30). Similarly by considering Eq. (2.28) we find

$$\begin{aligned} \left[\tilde{P}_3^+ \right] (\alpha') &= \frac{1}{2\pi} e^{-i\alpha'} \int_1^\infty \int_{-\infty}^\infty e^{ix(\alpha' - \alpha)} [P_3^+](\alpha) d\alpha dx \\ &= \frac{1}{2\pi} e^{-i\alpha'} \int_1^\infty \int_{-\infty-i\epsilon}^{\infty-i\epsilon} e^{ix(\alpha' - \alpha)} [P_3^+](\alpha) d\alpha dx \\ &= -\frac{1}{2\pi i} \int_{-\infty-i\epsilon}^{\infty-i\epsilon} \frac{1}{\alpha' - \alpha} e^{-i\alpha} [P_3^+](\alpha) d\alpha \\ &= -\frac{1}{2\pi i} \int_{-\infty-i\epsilon}^{\infty-i\epsilon} \frac{e^{-i\alpha \kappa^+}(\alpha)}{\alpha' - \alpha} \sum_{m \in \mathbb{Z}} \frac{1}{\alpha - \tilde{\sigma}_m^-} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) [P_2^-](\tilde{\sigma}_m^-) d\alpha. \end{aligned} \quad (2.58)$$

Here we cannot interchange the integration and summation straight away since $|\kappa^+|$ grows like $|\alpha|^{\frac{1}{2}}$ along this contour. We note, however, that the sum is again uniformly bounded along the given contour and remains so if we change to a contour Γ^- , which at its tails behaves like

$$\arg \alpha \sim \begin{cases} -\tilde{\epsilon}, & \text{Re } \alpha \rightarrow +\infty, \\ \pi + \tilde{\epsilon}, & \text{Re } \alpha \rightarrow -\infty, \end{cases}$$

for some $0 < \tilde{\epsilon} \ll 1$. This change of contour is analogous to the one applied in Appendix 2.B. Along Γ^- we observe exponential decay of the integrand, so absolute convergence on the product space for sum and integrand, so we can exchange the order of summation and integration to find Eq. (2.31).

2.D The finite section method and convergence

Here we provide a more detailed description of the finite section method which we apply to approximately solve Eq. (2.33). We also include a proof of convergence for this method when it is applied to our linear system.

Working on $l^2(\mathbb{Z})$ we shall denote by $P_m \in \mathcal{B}(l^2(\mathbb{Z}))$ the projection onto the coordinates $-m, -m+1, \dots, m-1, m$ and we implicitly identify $P_m(l^2(\mathbb{Z})) \cong \mathbb{C}^{2m+1}$. We employ the following version of the finite section method which is described and analysed for instance by Ben-Artzi et al. (2020): Our approximation of the solution to the system $Lx = b$ is $\Gamma_m(L, b)$ defined by

$$\Gamma_m(L, b) = (P_m L P_m)^{-1} P_m b,$$

which corresponds to truncating the linear system to $2m+1$ entries and solving the resulting system on \mathbb{C}^{2m+1} . The following lemma provides a guarantee for convergence:

Lemma 2.D.1 (Convergence of the finite section method). *If L is invertible as bounded linear operator on $l^2(\mathbb{Z})$ and it has bounded dispersion, i.e.*

$$\lim_{m \rightarrow \infty} \|(I - P_m) L P_m\| = 0, \quad (2.59)$$

then $\lim_{m \rightarrow \infty} \Gamma_m(L, b) = x$.

Proof. The proof is based on the arguments provided in Ben-Artzi et al. (2020, pp. 70–71). Let $x \in l^2(\mathbb{Z})$ be the true solution to $Lx = b$, then we have the following error estimate:

$$\begin{aligned} \|x - \Gamma_m(L, b)\| &\leq \|L^{-1}\| \|b - L\Gamma_m(L, b)\| \\ &\leq \|L^{-1}\| (\|b - P_m L P_m \Gamma_m(L, b)\| + \|(I - P_m) L P_m \Gamma_m(L, b)\|) \\ &\leq \|L^{-1}\| (\|b - P_m b\| + \|(I - P_m) L P_m\| \|x\|) \\ &\quad + \|L^{-1}\| \|(I - P_m) L P_m\| \|x - \Gamma_m(L, b)\|. \end{aligned}$$

Therefore,

$$(1 - \|L^{-1}\| \|(I - P_m) L P_m\|) \|x - \Gamma_m(L, b)\| \leq \|L^{-1}\| (\|b - P_m b\| + \|(I - P_m) L P_m\| \|x\|). \quad (2.60)$$

Now, since L is invertible, we have $\|L^{-1}\| < \infty$. Moreover $b \in l^2(\mathbb{Z})$, thus $\lim_{m \rightarrow \infty} \|b - P_m b\| = 0$. Therefore the result follows by combining Eq. (2.60) and Eq. (2.59). \square

Below we will show that $I - GF$ is Fredholm of index zero with bounded dispersion in the sense of Eq. (2.59). Thus, whenever $I - GF$ is invertible (i.e. away from cascade resonances),

$\Gamma_m(\mathbf{I} - \mathbf{GF}, A)$ converges to the true solution B as $m \rightarrow \infty$ and in the case when $(\mathbf{I} - \mathbf{GF})$ is not invertible the space of eigensolutions (i.e. resonant modes) is finite dimensional. Acoustic resonances in similar cascade geometries have been observed experimentally by Parker (1966) and were analysed by Koch (2009), and we refer the reader to these references for a more detailed treatment. Here we shall mostly be interested to guarantee convergence in the case when $\mathbf{I} - \mathbf{GF}$ is indeed invertible. Given the decay in the entries of \mathbf{F} as shown in Eq. (2.61) we also have $\mathbf{P}_m \mathbf{F} \mathbf{P}_m B \rightarrow C$ as $m \rightarrow \infty$. Thus all of these facts combined mean that the finite section method can be applied to our algebraic system Eq. (2.33)-(2.34) and provides a valid way to approximately solve the scattering problem.

In the following we shall use the convenient notation $A(x) \lesssim B(x)$, for functions $A(x), B(x)$, when there exists a constant $\mathcal{K} > 0$ independent of x such that $A(x) \leq \mathcal{K}B(x)$. To begin with let us look more closely at the coefficients in the linear system: Considering Eq. (2.32) we firstly note that the $e^{-i\alpha}$ term in the integrand decays exponentially along the given contour, thus we can estimate:

$$\begin{aligned} |\mathbf{F}_{jm}| &\lesssim \left| \frac{1}{\alpha - \tilde{\sigma}_j^+} \frac{1}{\alpha - \tilde{\sigma}_m^-} \right| \left| \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) \right| \int_{\Gamma^-} |e^{-i\alpha} \kappa^+(\alpha)| d\alpha \\ &\lesssim j^{-1} m^{-\frac{3}{2}} \int_{\Gamma^-} |e^{-i\alpha} \kappa^+(\alpha)| d\alpha \lesssim (1 + |j|)^{-1} (1 + |m|)^{-\frac{3}{2}}, \end{aligned} \quad (2.61)$$

for some constant that does not depend on j, m . In the above derivation we also used the growth of the residues of $\frac{1}{\kappa^+}$ as we established in §2.4.3. We can change the contour of integration for \mathbf{G}_{jm} to one that is of V shape with tails of the form

$$\arg \alpha \sim \begin{cases} \tilde{\epsilon}, & \text{Re } \alpha \rightarrow +\infty, \\ \pi - \tilde{\epsilon}, & \text{Re } \alpha \rightarrow -\infty, \end{cases}$$

for some $0 < \tilde{\epsilon} \ll 1$. This way we ensure the same exponential decay of the integrand and we can similarly deduce

$$|\mathbf{G}_{jm}| \lesssim (1 + |j|)^{-1} (1 + |m|)^{-\frac{1}{2}}.$$

This allows us to show:

Claim 2.D.2. \mathbf{GF} is a compact operator on $l^2(\mathbb{Z})$.

Proof. It is sufficient to prove that \mathbf{GF} has finite Hilbert–Schmidt norm (Conway, 2010, p.

267):

$$\begin{aligned}
\|\mathbf{GF}\|_{\text{HS}}^2 &= \sum_{j \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} \left| \sum_{m \in \mathbb{Z}} \mathbf{G}_{lm} \mathbf{F}_{mj} \right|^2 \\
&\lesssim \sum_{j \in \mathbb{Z}} \sum_{l \in \mathbb{Z}} \left(\sum_{m \in \mathbb{Z}} (1 + |l|)^{-1} (1 + |m|)^{-\frac{1}{2}} (1 + |m|)^{-1} (1 + |j|)^{-\frac{3}{2}} \right)^2 \\
&\lesssim \left(\sum_{j \in \mathbb{Z}} (1 + |j|)^{-3} \right) \left(\sum_{l \in \mathbb{Z}} (1 + |l|)^{-2} \right) \left(\sum_{m \in \mathbb{Z}} (1 + |m|)^{-\frac{3}{2}} \right)^2 < \infty.
\end{aligned}$$

□

Thus $\mathbf{I} - \mathbf{GF}$ is a compact perturbation of the identity, and as such is Fredholm of index zero.

Finally, the operator has bounded dispersion since we have for any $x \in l^2(\mathbb{Z})$:

$$\begin{aligned}
\|(\mathbf{I} - \mathbf{P}_m)(\mathbf{I} - \mathbf{GF})\mathbf{P}_m x\|_2^2 &= \|(\mathbf{I} - \mathbf{P}_m)\mathbf{GF}\mathbf{P}_m x\|_2^2 \leq \sum_{|k| > m} \left(\sum_{|j| \leq m} \left| \sum_{m \in \mathbb{Z}} \mathbf{G}_{km} \mathbf{F}_{mj} x_j \right| \right)^2 \\
&\leq \|x\|^2 \sum_{|k| > m} \left(\sum_{|j| \leq m} \left| \sum_{m \in \mathbb{Z}} \mathbf{G}_{km} \mathbf{F}_{mj} \right|^2 \right) \\
&\lesssim \|x\|^2 \left(\sum_{|k| > m} (1 + |k|)^{-2} \right) \left(\sum_{|j| \leq m} (1 + |j|)^{-3} \right) \left(\sum_{m \in \mathbb{Z}} (1 + |m|)^{-\frac{3}{2}} \right)^2 \\
&\lesssim \|x\|^2 \underbrace{\sum_{|k| > m+1} (1 + |k|)^{-2}}_{\rightarrow 0 \text{ as } m \rightarrow \infty}.
\end{aligned}$$

Thus $\mathbf{I} - \mathbf{GF}$ indeed satisfies Eq. (2.59), and hence the finite section method converges away from cascade resonances.

2.E Expressions for the solution with acoustic downstream incidence

For completeness we provide a brief account of the Wiener–Hopf analysis as well as expressions for the field away from the cascade structure, when acoustic waves are incident from downstream rather than upstream. The analysis proceeds analogously to the case of upstream incidence as given above, and we omit the repetition of details in the interest of brevity and instead provide just the main formulae of importance to this case.

Boundary conditions and Wiener–Hopf system

We again split $\phi = \phi_1 + \phi_2 + \phi_3$ where each of ϕ_1, ϕ_2, ϕ_3 satisfies Eq. (2.2) on the first cascade cell (i.e. on the space between the first and second blade of the cascade, $\{-\infty < x < \infty, 0 \leq y \leq s\}$), the radiation condition (iii) and the continuity of the blade-normal velocity everywhere (v). For downstream incidence we impose the following pairs of semi-infinite boundary conditions (together with an appropriate distribution of the edge conditions (vi)) on ϕ_j and the corresponding pressures $p_j, j = 1, 2, 3$:

- Trailing-edge interaction with incident field:

$$\begin{aligned} \frac{\partial \phi_1}{\partial y}(x, 0) &= -V e^{-i\tilde{\sigma}_\eta^+ x} \text{ on } x < 1, \\ [p_1](x) &= 0 \text{ on } x > 1, \end{aligned}$$

and ϕ_1 satisfies the Kutta condition at the trailing edge $x = 1$.

- Leading-edge correction:

$$\begin{aligned} \frac{\partial \phi_2}{\partial y}(x, 0) &= 0 \text{ on } x > 0, \\ [p_1 + p_2 + p_3](x) &= 0 \text{ on } x < 0, \end{aligned}$$

and ϕ_2 has the conventional inverse square-root singularity at the leading edge $x = 0$.

- Trailing-edge correction:

$$\begin{aligned} \frac{\partial \phi_3}{\partial y}(x, 0) &= 0 \text{ on } x < 1, \\ [p_2 + p_3](x) &= 0 \text{ on } x > 1, \end{aligned}$$

and ϕ_3 satisfies the Kutta condition at the trailing edge $x = 1$.

Taking the Fourier transform of these boundary conditions and imposing them on Eq. (2.11) we find the following system of coupled Wiener–Hopf equations:

$$\frac{[\tilde{P}_1^-](\alpha)}{\kappa(\alpha)} = \frac{\partial \tilde{\Phi}_1^+}{\partial y}(\alpha, 0) + \frac{iV e^{-i\tilde{\sigma}_\eta^+}}{\alpha - \tilde{\sigma}_\eta^+}, \quad (2.62)$$

$$\frac{\partial \Phi_2^-}{\partial y}(\alpha, 0) = \frac{1}{\kappa(\alpha)} \left([P_2^+](\alpha) - [P_1^-](\alpha) - [P_3^-](\alpha) \right), \quad (2.63)$$

$$\frac{\partial \tilde{\Phi}_3^+}{\partial y}(\alpha, 0) = \frac{1}{\kappa(\alpha)} \left([\tilde{P}_3^-](\alpha) - [\tilde{P}_2^+](\alpha) \right), \quad (2.64)$$

where the half-line Fourier transforms are as defined in §2.3.2.

Solution to the Wiener–Hopf problems

We can solve this system analogously to the analysis in §2.4 and find the following solutions to Eqs. (2.62)-(2.64):

$$\left[\tilde{P}_1^- \right] (\alpha) = iV \kappa^+(\tilde{\sigma}_\eta^+) \frac{\kappa^-(\alpha)}{\alpha - \tilde{\sigma}_\eta^+} e^{-i\tilde{\sigma}_\eta^+}, \quad (2.65)$$

$$\frac{\partial \tilde{\Phi}_1^+}{\partial y} (\alpha, 0) = \frac{-iV e^{-i\tilde{\sigma}_\eta^+}}{\alpha - \tilde{\sigma}_\eta^+} \left(1 - \frac{\kappa^+(\tilde{\sigma}_\eta^+)}{\kappa^+(\alpha)} \right), \quad (2.66)$$

$$\frac{1}{\kappa^+(\alpha)} \left[P_2^+ \right] (\alpha) = \sum_{m \in \mathbb{Z}} \frac{1}{\alpha - \tilde{\sigma}_m^-} \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_m^- \right) \left(\left[P_1^- \right] (\tilde{\sigma}_m^-) + \left[P_3^- \right] (\tilde{\sigma}_m^-) \right), \quad (2.67)$$

$$\frac{1}{\tilde{\kappa}^-(\alpha)} \left[\tilde{P}_3^- \right] (\alpha) = \sum_{m \in \mathbb{Z}} \frac{1}{\alpha - \tilde{\sigma}_m^+} \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_m^+ \right) \left[\tilde{P}_2^+ \right] (\tilde{\sigma}_m^+). \quad (2.68)$$

Reduction to infinite algebraic system

Setting

$$A_m := \left[P_1^- \right] (\tilde{\sigma}_m^-), \quad B_m := \left[\tilde{P}_2^+ \right] (\tilde{\sigma}_m^+), \quad C_m = \left[P_3^- \right] (\tilde{\sigma}_m^-),$$

we can reduce the conditions Eqs. (2.65)-(2.68) to the infinite linear system ($j \in \mathbb{Z}$):

$$\begin{aligned} B_j &= \sum_{m \in \mathbb{Z}} F_{jm} (A_m + C_m), \\ C_j &= \sum_{m \in \mathbb{Z}} G_{jm} B_m, \end{aligned}$$

where G, F are as in Eqs. (2.29) & (2.32) and the forcing is

$$\begin{aligned} A_j &= \frac{-V}{2\pi} \kappa^+(\tilde{\sigma}_\eta^+) \int_{-\infty - i\epsilon_1}^{\infty - i\epsilon_1} \frac{e^{-i\tilde{\sigma}_\eta^+} e^{i\alpha} \kappa^-(\alpha)}{(\alpha - \tilde{\sigma}_j^-)(\alpha - \tilde{\sigma}_\eta^+)} d\alpha \\ &\quad + iV \begin{cases} e^{i\tilde{\sigma}_j^-} e^{-i\tilde{\sigma}_\eta^+} \kappa^-(\tilde{\sigma}_j^-) \kappa^+(\tilde{\sigma}_\eta^+) \frac{1}{\tilde{\sigma}_j^- - \tilde{\sigma}_\eta^+}, & \text{if } \tilde{\sigma}_j^- \text{ is cut-on,} \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Expression for the far field

Note as before the potential at any point is given by the Fourier inversion

$$\phi(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\alpha x} [P] (\alpha) \frac{\cosh(\gamma y) e^{i\sigma + i d \alpha} - \cosh(\gamma(y - s))}{\kappa(\alpha) \sinh(\gamma s) \gamma} d\alpha.$$

Far field downstream

For $s(x-1) > yd$ and $0 \leq y \leq s$ we can close the contour in the lower half plane to find

$$\phi(x, y) = \sum_{m \in \mathbb{Z}} e^{-ix\tilde{\sigma}_m^-} f_m(y) \text{Res} \left(\frac{1}{\kappa}, \tilde{\sigma}_m^- \right) \frac{1}{\gamma(\tilde{\sigma}_m^-)} (-i) [P] (\tilde{\sigma}_m^-), \quad (2.69)$$

where, as in §2.7.2,

$$f_m(y) = \frac{\cosh(\gamma(\tilde{\sigma}_m^-)y) e^{i\sigma + id\tilde{\sigma}_m^-} - \cosh(\gamma(\tilde{\sigma}_m^-)(y-s))}{\sinh(\gamma(\tilde{\sigma}_m^-)s)},$$

which simplifies for acoustic modes to

$$f_m(y) = \begin{cases} e^{\gamma(\tilde{\sigma}_{m-1}^-)y}, & \text{if } m \geq 1, \\ e^{\gamma(\tilde{\sigma}_m^-)y}, & \text{if } m < 0. \end{cases}$$

And for downstream incident acoustic fields we have:

$$\begin{aligned} (-i) [P] (\tilde{\sigma}_m^-) &= V \kappa^+(\tilde{\sigma}_\eta^+) \frac{\kappa^-(\tilde{\sigma}_m^-)}{\tilde{\sigma}_m^- - \tilde{\sigma}_\eta^+} e^{i(\tilde{\sigma}_m^- - \tilde{\sigma}_\eta^+)} - i e^{i\tilde{\sigma}_m^-} \kappa^-(\tilde{\sigma}_m^-) \sum_{j \in \mathbb{Z}} \frac{B_j}{\tilde{\sigma}_m^- - \tilde{\sigma}_j^+} \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_j^+ \right) \\ &+ e^{i\tilde{\sigma}_m^-} \sum_{j \in \mathbb{Z}} (A_j + C_j) \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_j^- \right) \frac{1}{2\pi} \int_{-\infty + i\epsilon_1}^{\infty + i\epsilon_1} \frac{\kappa^+(\alpha)}{\alpha - \tilde{\sigma}_m^-} \frac{e^{-i\alpha}}{\alpha - \tilde{\sigma}_j^-} d\alpha \\ &+ iA_m + iC_m. \end{aligned}$$

Far field upstream

For $sx < yd$ and $0 \leq y \leq s$ we can close the contour in the upper half plane and find:

$$\phi(x, y) = \sum_{m \in \mathbb{Z}} e^{-i\tilde{\sigma}_m^+ x} \text{Res} \left(\frac{1}{\kappa}, \tilde{\sigma}_m^+ \right) \frac{e^{-\gamma(\tilde{\sigma}_m^+)y}}{\gamma(\tilde{\sigma}_m^+)} i [P] (\tilde{\sigma}_m^+), \quad (2.70)$$

where for downstream incident fields we have

$$\begin{aligned} i [P] (\tilde{\sigma}_m^+) &= -i e^{i\tilde{\sigma}_m^+} B_m + i \sum_{j \in \mathbb{Z}} (A_j + C_j) \text{Res} \left(\frac{1}{\kappa^+}, \tilde{\sigma}_j^- \right) \frac{\kappa^+(\tilde{\sigma}_m^+)}{\tilde{\sigma}_m^+ - \tilde{\sigma}_j^-} \\ &+ i \frac{V}{2\pi} \int_{-\infty - i\epsilon_1}^{\infty - i\epsilon_1} \frac{e^{i\alpha}}{\alpha - \tilde{\sigma}_m^+} \frac{\kappa^-(\alpha)}{\alpha - \tilde{\sigma}_\eta^+} e^{-i\tilde{\sigma}_\eta^+} \kappa^+(\tilde{\sigma}_\eta^+) d\alpha \\ &+ \sum_{j \in \mathbb{Z}} \text{Res} \left(\frac{1}{\kappa^-}, \tilde{\sigma}_j^+ \right) \frac{B_j}{2\pi} \int_{-\infty - i\epsilon_1}^{\infty - i\epsilon_1} \frac{e^{i\alpha}}{\alpha - \tilde{\sigma}_m^+} \frac{\kappa^-(\alpha)}{\alpha - \tilde{\sigma}_j^+} d\alpha. \end{aligned}$$

Chapter 3

Acoustic and hydrodynamic power of wave scattering by an infinite cascade of blades in mean flow

3.1 Introduction

In Chapter 2 we derived a solution for the wave scattering problem on a cascade geometry based on the Wiener–Hopf technique. This present chapter is devoted to understanding more closely the energy balance of this system, when the cascade is subjected to an incident acoustic wave from either upstream or downstream. In order to simplify some of the algebraic expressions involved we will focus our attention to the two-dimensional case, i.e. a cascade as pictured in Fig. 2.1a, and we omit spanwise variations in the field, which is equivalent to taking $K_3 = 0$ in the formulation of Chapter 2.

We already indicated in the introductory Chapter 1 that the presence of mean flow leads to interesting aspects in the considerations concerning acoustic energy conservation. This is perhaps somewhat surprising given that the energy of the full flow is conserved: The inviscid continuity and momentum equations for a flow without sources or sinks are

$$\frac{\partial \tilde{\rho}}{\partial t} + \nabla \cdot (\tilde{\rho} \tilde{\mathbf{u}}) = 0, \quad \frac{\partial \tilde{\mathbf{u}}}{\partial t} + \tilde{\mathbf{u}} \cdot \nabla \tilde{\mathbf{u}} + \frac{1}{\tilde{\rho}} \nabla \tilde{p} = 0,$$

where $\tilde{\rho}$, \tilde{p} , $\tilde{\mathbf{u}}$ are the density, pressure and velocity field of the full fluid body. Those two equations can be combined to yield the following energy conservation law for the full flow (here written in the isentropic case) as shown for instance in Myers (1986):

$$\frac{\partial E_t}{\partial t} + \nabla \cdot (\tilde{\mathbf{u}} E_t + \tilde{\mathbf{u}} \tilde{p}) = 0, \quad E_t = \tilde{\rho} \frac{\tilde{u}^2}{2} + \tilde{\rho} \int^{\tilde{p}} \frac{\tilde{p}(\hat{\rho})}{\hat{\rho}^2} d\hat{\rho}.$$

Suppose we then write the flow as $\tilde{p} = p_0 + p + \mathcal{O}(\epsilon^2)$, $\tilde{\rho} = \rho_0 + \rho + \mathcal{O}(\epsilon^2)$, $\tilde{\mathbf{u}} = \mathbf{U} + \mathbf{u} + \mathcal{O}(\epsilon^2)$,

where p_0, ρ_0, \mathbf{U} correspond to some steady mean flow and $p, \rho, \mathbf{u} = \mathcal{O}(\epsilon)$ are first order perturbation quantities – the unsteady fluctuations on the acoustic order. An important question that was subject of a large amount of research in the 1960s and 1970s, amongst others by Cantrell and Hart (1964), Möhring (1970, 1973) and Morfey (1971), was whether an energy conservation law holds that can be expressed in terms of quantities on the acoustic scale alone (i.e. in terms of first order perturbed quantities, without having to calculate the contributions from higher order terms that satisfy more complicated nonlinear equations of motion). Only in the late 1970s, based on exact relations as given by Goldstein (1976) and Myers (1986), the idea took hold that acoustic energy is not conserved – instead the acoustic field can exchange energy with the vortical components of the flow at the same scale. For isentropic flows the exact energy balance for first order quantities takes the following form (non-dimensionalised according to §2.2)

$$\frac{\partial E}{\partial t} + \nabla \cdot \mathbf{I} = \mathbf{U} \cdot ((\nabla \times \mathbf{u}) \times \mathbf{u}) + \rho \mathbf{u} \cdot ((\nabla \times \mathbf{U}) \times \mathbf{U}), \quad (3.1)$$

where

$$E = \frac{p\rho}{2} + \frac{u^2}{2} + \rho \mathbf{U} \cdot \mathbf{u}, \quad \mathbf{I} = (p + \mathbf{U} \cdot \mathbf{u}) (\mathbf{u} + \rho \mathbf{U}).$$

The terms on the right hand side of Eq. (3.1) corresponds to energy exchange between acoustic and vortical components of the field, and its form suggests it may be both positive and negative, meaning sound can be both generated and attenuated through this interaction. The possibility of generation of sound through vorticity has been well known since the pioneering work by Lighthill (1952), and appears for instance in the Ffowcs Williams acoustic analogy of turbulence passing a trailing edge (Ffowcs Williams and Hall, 1970) and in Crighton’s model problem of a vortex passing an edge (Crighton, 1972). However, the possibility of attenuation of sound by vorticity was confirmed experimentally only several years later in the work by Bechert et al. (1977), who found conclusive evidence of sound attenuation in a pure tone exiting a cylindrical nozzle with mean flow. The experiment was studied and explained analytically based on Eq. (3.1) by Howe (1979, 1980). In case of a sharp trailing edge we already discussed in Chapter 2 the appearance of a vortex sheet attached at the edge, and the amount of acoustic energy lost into this wake was first given explicitly by Rienstra (1981), who also provided a full energy balance for an open ended annular duct (Rienstra, 1984). For incompressible flow a similar exchange of energy appears between a region of localised vorticity and the first order perturbation in the flow and this was studied by Arzoumanian (2011) (we outline the connection to this work in §3.3.2).

The sound attenuation through the production of vorticity is highly relevant to fluid mechanical engineering systems, including to turbomachinery, since the attenuation of

sound is a very favourable by-product of the geometry and needs to be taken into account for noise prediction schemes. It is therefore perhaps surprising that in the context of cascades the full power balance for this system has remained largely unstudied in the literature. There is a brief discussion given by Glegg (1999) however this is considering only the acoustic power that is scattered from the cascade, i.e. only the contribution of $\nabla \cdot \mathbf{I}$ in Eq. (3.1) and not accounting for the energy exchange term.

The main novelty described in this chapter is a detailed study of the energy balance in the case of an infinite cascade of blades. In particular, we account for all contributions to the outgoing power – acoustic, hydrodynamic and energy exchange terms. In so doing we provide a physical interpretation of the relevant terms in context of previous literature and its relation to similar expressions found for incompressible flow. Our expressions allow us to conduct numerical experiments from which we can draw a number of interesting conclusions. Specifically, the energy exchange term and the hydrodynamic power account in many parameter settings for a significant portion of the outgoing energy, i.e. lead to significant attenuation in the sound pressure level. We confirm that in the case of a cascade the energy exchange term can be negative (as was found for a single trailing edge by Rienstra (1981)), which means the cascade can harvest energy from the flow and convert it into sound. This leads to additional energy in the scattered acoustic field and in some cases can result in an amplification of the acoustic power, and over-reflection, whereby the scattered acoustic field has a larger amplitude than the incoming acoustic field.

In addition to the numerical experiments, our expressions for the power balance can be used to infer properties of the far-field, such as symmetries with respect to certain angles of incidence and perfect transmission at specific angles of attack. The latter is related to recent work by Porter (2021), who studied perfectly transmitting blade rows and demonstrated that closely spaced blade rows in zero mean flow may be regarded as negative-refraction metamaterials, and we are able to show that some of his conclusions extend to the case when nonzero mean flow is present.

The present chapter is structured as follows: §3.2 provides a very brief recap of the mathematical formulation of the acoustic wave scattering problem which we solved in Chapter 2. We outline the exact modal structure of the field away from the cascade (consisting of radiation acoustic and hydrodynamic modes). In §3.3 we derive the energy balance for this system based on the usual form of conservation of acoustic energy given by Goldstein (1976), and we include a discussion outlining an interpretation of the energy exchange terms and their relation to previous work in §3.3.2. This section includes an explicit form of the time-averaged energy balance in terms of the modal amplitudes of the field. These expressions are used in §3.4 to study symmetries in the field and to give numerical examples to understand the form of each contribution to the energy balance. Concluding remarks are provided in §3.5.

3.2 Problem statement

For completeness we briefly recap the most important properties of the field in this acoustic wave scattering problem, but we refer the reader to Chapter 2 for further details.

Our two-dimensional geometry consists of an infinitely extending cascade of plates located at $\{(x, ns), nd < x < nd + 1\}$, $n \in \mathbb{Z}$. We highlight that the inclusion of a third dimension, with a given spanwise wavenumber as was done in Chapter 2, does not involve any conceptual difficulties but it would complicate some of the algebraic expressions in this chapter. The cascade is located in an inviscid isentropic fluid with nonzero uniform mean flow parallel to the cascade blades. We assume the incidence of harmonic acoustic waves on the cascade with reduced frequency Ω , amplitude I , and normal to the wavefronts that is inclined at an angle φ as shown in Fig. 3.1.

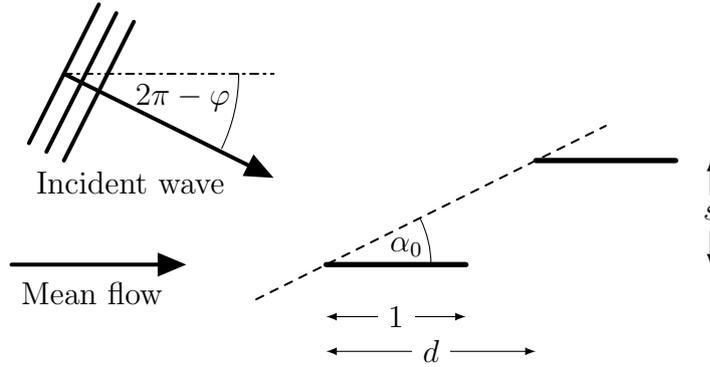


Figure 3.1: Sketch of cascade of blades with incident acoustic wave.

As a result, all quantities describing the perturbed flow are time-harmonic, and indeed can be expressed in terms of a velocity potential $\exp(i\Omega t)\phi(x, y)$ (here $\Omega = \omega l/U_\infty$ is the reduced frequency) which is the sum of incoming wave and scattered potential:

$$\phi(x, y) = Ie^{-ik_x x - ik_y y} + \phi_s(x, y),$$

where

$$k_x = \Omega M \frac{\cos \varphi}{1 + M \cos \varphi}, \quad k_y = \Omega M \frac{\sin \varphi}{1 + M \cos \varphi}. \quad (3.2)$$

The total velocity potential must then satisfy the two-dimensional convected wave equation:

$$\frac{D^2 \phi}{Dt^2} - M^{-2} \Delta \phi = 0,$$

where the non-dimensional material derivative for time-harmonic quantities is given by $D/Dt = i\Omega + \partial_x$. In addition to the equation of motion we impose the boundary and edge conditions (i)-(vi) as described in §2.2. We note that the Bloch condition (iv) becomes for

incident acoustic waves: For all $(x, y) \in \mathbb{R}^2, n \in \mathbb{Z}$:

$$\phi(x, y) = e^{-in\sigma} \phi(x + nd, y + ns),$$

where, according to Eq. (3.2)

$$\sigma = -\frac{\Omega M}{1 + M \cos \varphi} (d \cos \varphi + s \sin \varphi). \quad (3.3)$$

3.2.1 Dispersion relation for free space solutions

In this chapter we are specifically interested in the energy transported to the far-field, thus let us begin by considering the kind of waves that exist in the parts of the geometry away from the cascade structure. We shall see in this section that, as a result of the periodic nature of the geometry, away from the cascade structure the field splits into a discrete set of nodes, the so-called radiation modes (cf. Bragg angles for wave scattering on periodic media). We recall the expression Eq. (2.51)

$$\phi(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\alpha x} [P](\alpha) \frac{\cosh(\gamma y) e^{i\sigma + id\alpha} - \cosh(\gamma(y-s))}{\kappa(\alpha) \sinh(\gamma s) \gamma} d\alpha,$$

which is of the form

$$\phi(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\alpha x} \frac{f(\alpha, y)}{D(\alpha)} \left(\int_0^1 [p](\tilde{x}) d\tilde{x} \right) d\alpha, \quad (3.4)$$

where $D(\alpha) = (i\Omega - i\alpha) (\cos(\sigma + d\alpha) - \cosh \gamma s)$ and $\alpha \mapsto f(\alpha, y)$ is an entire function for any $y \in [0, 1)$. As we saw in §2.7.2 the zeros of $D(\alpha)$ give rise to wave-like modes away from the cascade structure (i.e. modes which are x -harmonic if $sx < dy$ or $s(x-1) > dy$). Therefore the dispersion relation for modal solutions in free space away from the cascade structure is given by

$$D(\alpha) = 0. \quad (3.5)$$

This dispersion relation Eq. (3.5) has a discrete set of solutions, the radiation modes, denoted by $\alpha = \sigma_m^\pm(\Omega)$, $m \in \mathbb{Z}$, and the hydrodynamic mode, with the superscript \pm denoting upstream and downstream travelling modes respectively. The hydrodynamic mode is convected with the dimensional mean-flow speed, leading to the nondimensional

wavenumber $\alpha = \Omega$. We recall that the analytic expressions for the radiation modes are

$$\sigma_m^\pm = \frac{-(s^2 M^2 \Omega + d\sigma + 2d\pi m)}{s^2 \beta^2 + d^2} \mp \frac{\sqrt{(s^2 M^2 \Omega + d\sigma + 2d\pi m)^2 - (s^2 \beta^2 + d^2) ((\sigma + 2\pi m)^2 - s^2 (\Omega^2 M^2))}}{s^2 \beta^2 + d^2}, \quad (3.6)$$

where $\beta^2 = 1 - M^2$. They are called ‘cut-on’ (i.e. propagate to infinity) when $m \in \mathbb{Z}$ is such that

$$(s^2 M^2 (\text{Re } \Omega) + d\sigma + 2d\pi m)^2 - (s^2 \beta^2 + d^2) ((\sigma + 2\pi m)^2 - s^2 ((\text{Re } \Omega)^2 M^2 - K_3^2)) \geq 0,$$

and the modes are evanescent (called ‘cut-off’) otherwise. The dispersion curves for Eq. (3.5) are shown in Fig. 3.2. We recall from Eq. (3.3) that the value of σ depends on the frequency and angle of the incident field, thus it implicitly depends on Ω . This means firstly that in the current formulation the incident acoustic wave always corresponds to σ_0^\pm (\pm corresponding to down- and upstream incidence respectively) and secondly that σ_0^\pm are proportional to Ω , and hence results in the straight-line dispersion curves for σ_0^\pm as observed in Fig. 3.2.

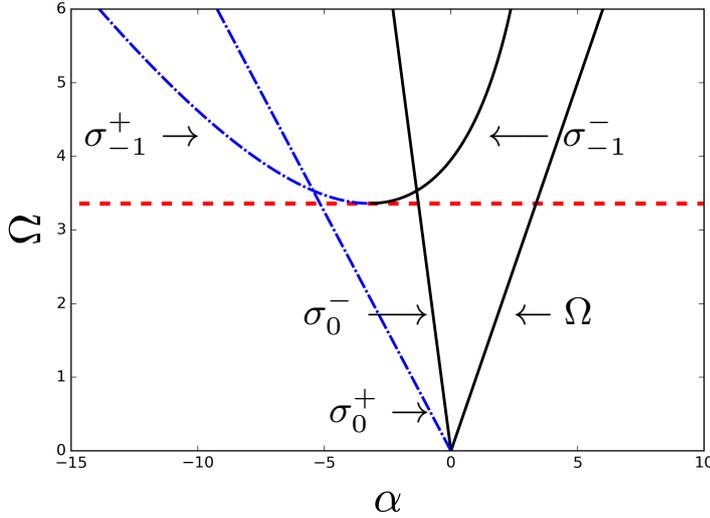


Figure 3.2: Dispersion curves for radiation modes away from the cascade. The dashed red line marks the value of Ω for which σ_{-1}^\pm become cut-on. Upstream travelling modes are marked with dash-dotted blue lines, downstream travelling modes correspond to the solid black lines.

3.2.2 Form of the far-field

We already discussed the solution of the scattering problem in Chapter 2, therefore we take a brief moment here to quickly summarise the shape of the field away from the cascade blades in Tab. 3.1. Note the expressions here specify ϕ in the first cascade cell $0 \leq y < s$ and the field is determined uniquely in the remaining cells by condition (iv). Moreover, the form of the field presented in Tab. 3.1 is the exact form which is valid for each coordinate (x, y) in the range indicated ($sx < dy$, and $s(x - 1) > dy$ respectively). This means the evanescent modes (which, of course, do not contribute to the asymptotic far-field) are included and their contribution is important when evaluating the exchange of acoustic and vortical energy near the trailing edge as discussed in §3.3.

Table 3.1: Expressions for the total field with upstream and downstream incident acoustic waves.

Upstream incidence	Downstream incidence
Incident wavenumber: $k_x = \sigma_0^-$	Incident wavenumber: $k_x = \sigma_0^+$
Upstream ($sx < dy$)	Upstream ($sx < dy$)
$\phi(x, y) = Ie^{-i\sigma_0^- x} e^{\gamma(\sigma_0^-)y}$ $+ \sum_{m \in \mathbb{Z}} U_m e^{-i\sigma_m^+ x} e^{-\gamma(\sigma_m^+)y},$	$\phi(x, y) = \sum_{m \in \mathbb{Z}} U_m e^{-i\sigma_m^+ x} e^{-\gamma(\sigma_m^+)y},$
Downstream ($s(x - 1) > dy$)	Downstream ($s(x - 1) > dy$)
$\phi(x, y) = Be^{-i\Omega x} g(y)$ $+ \sum_{m \in \mathbb{Z}} D_m e^{-i\sigma_m^- x} e^{\gamma(\sigma_m^-)y}.$	$\phi(x, y) = Ie^{-i\sigma_0^+ x} e^{-\gamma(\sigma_0^+)y} + Be^{-i\Omega x} g(y)$ $+ \sum_{m \in \mathbb{Z}} D_m e^{-i\sigma_m^- x} e^{\gamma(\sigma_m^-)y}.$

In Tab. 3.1, U_m, D_m are complex constants, the amplitudes of the radiation acoustic modes upstream and downstream of the cascade respectively, and the complex constant B is the amplitude of the hydrodynamic mode shed from the trailing edge of each blade. Furthermore we have introduced the function

$$g(y) = \frac{e^{i\sigma + id\Omega} \cosh(\Omega y) - \cosh(\Omega(s - y))}{2(\cos(\sigma + d\Omega) - \cosh(s\Omega))}. \quad (3.7)$$

In §3.4 we will evaluate the amplitudes U_m, D_m numerically using the solution developed in Chapter 2.

3.3 An energy balance for the cascade of blades

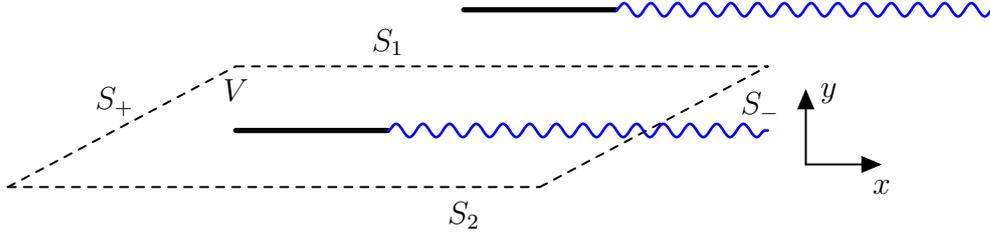


Figure 3.3: Control domain for the energy balance.

We saw in Chapter 2 that the unsteady velocity field \mathbf{u} can be expressed as the gradient of a potential ϕ which is continuous except for a finite jump across the blades and across the wake downstream of the cascade, so that the unsteady vorticity, in terms of this jump, is

$$\boldsymbol{\omega} = \nabla \times \mathbf{u} = -\hat{\mathbf{z}}\delta(y) [\partial_x \phi]. \quad (3.8)$$

For the purpose of describing energy density and flux in this section, we reintroduce a time-dependence in our quantities, such that for instance $\phi = \phi(t, x, y)$. The equation relating the rate of change of acoustic energy, E , to the acoustic energy flux, \mathbf{I} , in uniform mean flow \mathbf{U} is well-known and can be found for instance in Goldstein (1976, p. 41) (stated here in non-dimensionalised form for an isentropic, source-free flow):

$$\frac{\partial E}{\partial t} + \nabla \cdot \mathbf{I} = \mathbf{u} \cdot (\mathbf{U} \times \boldsymbol{\omega}), \quad (3.9)$$

where

$$E = \frac{p\rho}{2} + \left(\frac{|\mathbf{u}|^2}{2} + \rho \mathbf{U} \cdot \mathbf{u} \right), \quad \mathbf{I} = (p + \mathbf{U} \cdot \mathbf{u})(\mathbf{u} + \rho \mathbf{U}),$$

and p, ρ denote the acoustic pressure and density perturbations. Now consider the domain $V = \{(x, y); -s/2 < y < s/2, -X_1 + yd/s < x < X_2 + yd/s\}$, as shown in Fig. 3.3. Integrating Eq. (3.9) over V yields:

$$\frac{d}{dt} \int_V E dx + \int_{\partial V} \mathbf{I} \cdot \mathbf{n} ds = \int_V \mathbf{u} \cdot (\mathbf{U} \times \boldsymbol{\omega}) dx,$$

and taking a time average one can infer, in the absence of external forces and sources, the following balance of energy fluxes for time-harmonic flows:

$$\int_{\partial V} \bar{\mathbf{I}} \cdot \mathbf{n} ds = \int_V \mathbf{u} \cdot (\mathbf{U} \times \boldsymbol{\omega}) dx, \quad (3.10)$$

where on the right hand side the time average $\langle \phi\psi \rangle = \frac{1}{2}\text{Re}[\phi^*\psi]$ is understood (where for $z \in \mathbb{C}$ we denote its complex conjugate by z^*) and $\bar{\mathbf{I}}$ denotes the time averaged acoustic energy flux

$$\bar{\mathbf{I}} = \frac{1}{2}\text{Re}[(p + \mathbf{U} \cdot \mathbf{u})^* (\mathbf{u} + \rho\mathbf{U})].$$

In Eq. (3.10) the term on the left hand side represents the time-averaged acoustic energy flux through the boundary ∂V , and the right hand side arises as a consequence of the possibility of energy conversion between the acoustic and vortical unsteady flow and vice versa, as we mentioned in the introduction §3.1.

3.3.1 Time-averaged energy balance in terms of amplitudes

Let us now find an explicit expression for the contributions in Eq. (3.10) in terms of the modal amplitudes of the field. Noting the continuity of the velocity potential across S_1, S_2 we observe that in Eq. (3.10) the contributions from S_1 and S_2 cancel. Moreover, on the blade we have $\partial\phi/\partial y = 0$ so the contribution from this part of ∂V is zero, and the left hand side of Eq. (3.10) reduces to an integral over S_+ and S_- .

Let us now focus on the time-harmonic case and take the usual time-average. By the expressions for the field from §3.2.2 we have only acoustic contributions to \int_{S_+} and acoustic mixed with hydrodynamic contributions to \int_{S_-} . Of course, the evanescent (cut-off) acoustic modes carry zero energy to the far-field, but the interaction of all the downstream acoustic modes σ_m^- with the hydrodynamic mode along S_- will, in general, involve contributions proportional to $\exp(i((\sigma_m^-)^* - \Omega)X_2)$. It turns out that these contributions cancel with terms from $\int_V \mathbf{u} \cdot (\mathbf{U} \times \boldsymbol{\omega}) dx$. To see this let us demonstrate the idea based on the special case when we have upstream incidence and $d = 0$ (this is such that the algebraic expressions remain within reasonable simplicity). Here the contribution from \int_{S_-} takes the form

$$\begin{aligned} \int_{S_-} \bar{\mathbf{I}} \cdot \mathbf{n} ds &= \frac{1}{2}\text{Re} \left[\int_0^s (-i\Omega\phi(X_2, y))^* (\beta^2\partial_x\phi(X_2, y) - i\Omega M^2\phi(X_2, y)) dy \right] \\ &= \frac{1}{2}\text{Re} \left[i\Omega \int_0^s \left(B^* e^{i\Omega X_2} g^*(y) + \sum_{m \in \mathbb{Z}} D_m^* e^{i(\sigma_m^-)^* X_2} e^{\gamma(\sigma_m^-)^* y} \right) \right. \\ &\quad \left(B e^{-i\Omega X_2} (-i\beta^2\Omega - i\Omega M^2)g(y) \right. \\ &\quad \left. \left. + \sum_{m \in \mathbb{Z}} D_m e^{-i\sigma_m^- X_2} e^{\gamma(\sigma_m^-) y} (-i\sigma_m^- \beta^2 - i\Omega M^2) \right) dy \right]. \end{aligned}$$

The contribution from terms involving both the hydrodynamic and acoustic modes are

thus

$$\begin{aligned} & \frac{1}{2} \operatorname{Re} \left[i\Omega \int_0^s \left(\sum_{m \in \mathbb{Z}} D_m^* e^{i(\sigma_m^-)^* X_2} B e^{-i\Omega X_2} e^{\gamma(\sigma_m^-)^* y} g(y) \right) \right. \\ & \quad \left. (-i\beta^2 \Omega - i\Omega M^2 - i(\sigma_m^-)^* \beta^2 - i\Omega M^2) dy \right] \\ &= \frac{1}{2} \operatorname{Re} \left[-i\Omega \sum_{m \in \mathbb{Z}} D_m^* e^{i(\sigma_m^-)^* X_2} B e^{-i\Omega X_2} (i\beta^2 \Omega + 2i\Omega M^2 + i(\sigma_m^-)^* \beta^2) \int_0^1 e^{\gamma(\sigma_m^-)^* y} g(y) dy \right]. \end{aligned}$$

Using Eq. (3.7) we find

$$\begin{aligned} \int_0^s e^{\gamma(\sigma_m^-)^* y} g(y) dy &= \int_0^s \frac{e^{\gamma(\sigma_m^-)^* y} e^{i\sigma} \cosh(\Omega y) - e^{\gamma(\sigma_m^-)^* y} \cosh(\Omega(s-y))}{2(\cos \sigma - \cosh(s\Omega))} dy \\ &= -\frac{\gamma^*(\sigma_m^-)}{(\gamma^*(\sigma_m^-))^2 - \Omega^2}, \end{aligned}$$

where we used $s\gamma(\sigma_m^-)^* = -i\sigma - 2\pi im$. Thus the contribution on S_- from terms involving both the hydrodynamic and acoustic modes equals

$$\begin{aligned} & \frac{1}{2} \operatorname{Re} \left[-i\Omega \sum_{m \in \mathbb{Z}} D_m^* e^{i(\sigma_m^-)^* X_2} B e^{-i\Omega X_2} \gamma^*(\sigma_m^-) \frac{(-i\beta^2 \Omega - 2i\Omega M^2 - i(\sigma_m^-)^* \beta^2)}{((\sigma_m^-)^*)^2 \beta^2 + 2(\sigma_m^-)^* M^2 \Omega - \Omega^2(1 + M^2)} \right] \\ &= \frac{1}{2} \operatorname{Re} \left[-i\Omega \sum_{m \in \mathbb{Z}} D_m^* B e^{i(\sigma_m^-)^* X_2} e^{-i\Omega X_2} \frac{\gamma^*(\sigma_m^-)}{i(\sigma_m^-)^* - i\Omega} \right]. \end{aligned}$$

In this special case the integral $\int_V \mathbf{u} \cdot (\mathbf{U} \times \boldsymbol{\omega}) dx$, takes the form

$$\begin{aligned} \int_V \mathbf{u} \cdot (\mathbf{U} \times \boldsymbol{\omega}) dx &= \frac{1}{2} \operatorname{Re} \int_1^{X_2} \partial_y \phi(x, 0)^* [\partial_x \phi](x) dx \\ &= \frac{1}{2} \operatorname{Re} \left[-i\Omega \int_1^{X_2} \left(B e^{-i\Omega x} \Omega \frac{\sinh(s\Omega)}{2(\cos(\sigma + d\Omega) - \cosh(s\Omega))} \right. \right. \\ & \quad \left. \left. + \sum_{m \in \mathbb{Z}} D_m \gamma(\sigma_m^-) e^{-i\sigma_m^- x} \right)^* B e^{-i\Omega x} dx \right] \\ &= \frac{1}{2} \operatorname{Re} \left[-i\Omega \int_1^{X_2} \sum_{m \in \mathbb{Z}} D_m^* B \gamma(\sigma_m^-)^* e^{i(\sigma_m^-)^* - i\Omega} x dx \right] \\ &= \frac{1}{2} \operatorname{Re} \left[-i\Omega \sum_{m \in \mathbb{Z}} D_m^* B \gamma(\sigma_m^-)^* \frac{e^{i(\sigma_m^-)^* - i\Omega} X_2 - e^{i(\sigma_m^-)^* - i\Omega}}{i(\sigma_m^-)^* - i\Omega} \right]. \quad (3.11) \end{aligned}$$

Thus one can see that any terms involving $\exp(i((\sigma_m^-)^* - \Omega)X_2)$ cancel exactly in Eq. (3.10). This remains true for downstream incidence and when $d \neq 0$ (though the algebra is more tedious) and it is therefore possible, for X_1, X_2 sufficiently large, to arrive at an exact

energy balance in terms of the modal amplitudes which is independent of X_1, X_2 :

$$P_U + P_D + P_H + \Pi_\omega = P_I, \quad (3.12)$$

where in the above P_U, P_D are the acoustic sound powers that are radiated to $x = \pm\infty$ respectively and P_I is the sound power of the incoming wave. Furthermore, P_H is the power carried by the hydrodynamic mode to $x = +\infty$ and Π_ω is the energy conversion term due to the lift force experienced by particles on the vortex sheet, arising from the right hand side of equation Eq. (3.10). After a few steps of algebra we find the following explicit expressions in terms of the amplitudes, where the aforementioned cancelling contributions proportional to $\exp(i((\sigma_m^-)^* - \Omega)X_2)$ have been dropped (this is equivalent to taking an average over $X_2 \in (1, \infty)$ in the sense $\langle \Pi_\omega \rangle = \lim_{L \rightarrow \infty} L^{-1} \int_0^L \Pi_\omega dX_2$ using the expression Eq. (3.11)):

$$\begin{aligned} P_U &= -\frac{\Omega}{2} \sum_{\substack{m \in \mathbb{Z} \\ \text{cut-on}}} |U_m|^2 (s\Omega M^2 + s\beta^2 \sigma_m^+ - d\gamma(\sigma_m^+)), \\ P_D &= \frac{\Omega}{2} \sum_{\substack{m \in \mathbb{Z} \\ \text{cut-on}}} |D_m|^2 (-s\Omega M^2 - s\beta^2 \sigma_m^- - d\gamma(\sigma_m^-)), \\ P_I &= \begin{cases} \frac{\Omega}{2} |I|^2 (-s\Omega M^2 - s\beta^2 \sigma_0^- - d\gamma(\sigma_0^-)), & \text{if } k_x = \sigma_0^-, \\ \frac{\Omega}{2} |I|^2 (-s\Omega M^2 - s\beta^2 \sigma_0^+ + d\gamma(\sigma_0^+)), & \text{if } k_x = \sigma_0^+. \end{cases} \end{aligned}$$

Furthermore:

$$\begin{aligned} P_H &= \frac{\Omega}{2} |B|^2 \left(s\Omega \int_0^1 |g(st)|^2 dt + \text{Im} \left[d \int_0^1 g(st)^* g'(st) dt \right] \right) \\ &= \frac{\Omega |B|^2}{8(\cos(\sigma + d\Omega) - \cosh(\Omega))^2} \\ &\quad \left(\frac{\Omega s + \cosh(\Omega s) \sinh(\Omega s) - \cos(\sigma + d\Omega)(\Omega s \cosh(\Omega s) + \sinh(\Omega s))}{\Omega s} \right. \\ &\quad \left. - \Omega d \sin(\sigma + d\Omega) \sinh(\Omega s) \right), \end{aligned}$$

and finally,

$$\Pi_\omega = -\frac{\Omega}{2} \text{Im} \left[\sum_{m \in \mathbb{Z}} B^* D_m \gamma(\sigma_m^-) \frac{e^{i\Omega - i\sigma_m^-}}{i\Omega - i\sigma_m^-} + \begin{cases} 0, & \text{if } k_x = \sigma_0^- \\ -B^* I \gamma(\sigma_0^+) \frac{e^{i\Omega - i\sigma_0^+}}{i\Omega - i\sigma_0^+}, & \text{if } k_x = \sigma_0^+ \end{cases} \right].$$

3.3.2 Interpretations of the interaction terms

Before moving on to provide numerical results based on Eq. (3.12), let us provide a few possible ways to interpret the energy exchange terms appearing in Eq. (3.10). These interpretations reflect slightly different points of view taken by various researchers, and we believe there is value in each of these interpretations. For this section we return to considering time-dependent quantities.

Interpretation 1: Rate of work done by a lift force

Howe (1980) and Guo (1991, p. 191) found that the term $\int_V \mathbf{u} \cdot (\mathbf{U} \times \boldsymbol{\omega}) dx$ can be regarded to correspond to the rate of work done by the lift force $\mathbf{U} \times \boldsymbol{\omega}$ experienced by vortex elements in the unsteady velocity field of the sound \mathbf{u} (see also Howe, 1998, p. 407). The interpretation of the term $\mathbf{U} \times \boldsymbol{\omega}$ as a lift force is a well-known result introduced by Prandtl (1918, Eq. (5)); see also Saffman (1993, §3.1).

Interpretation 2: Acoustic energy flux into the wake

This is the point of view taken by Rienstra (1981): Consider two surfaces $\mathcal{W}_\pm := \{(x, y) \mid x > 0, y = 0^\pm\}$ just above/below the wake respectively (i.e. just above/below the blue curves in Fig. 3.3). Then the acoustic energy flux into the region enclosed by those two curves (i.e. into the wake) is given by

$$\begin{aligned} \int_{\mathcal{W}_- + \mathcal{W}_+} \bar{\mathbf{I}} \cdot \mathbf{n} ds &= - \int_1^\infty (p + \mathbf{U} \cdot \mathbf{u}) (\mathbf{u} + \rho \mathbf{U}) \Big|_{y=0^+} \cdot \hat{\mathbf{y}} ds \\ &\quad + \int_1^\infty (p + \mathbf{U} \cdot \mathbf{u}) (\mathbf{u} + \rho \mathbf{U}) \Big|_{y=0^-} \cdot \hat{\mathbf{y}} ds \\ &= - \int_1^\infty [p + \partial_x \phi] \partial_y \phi ds \\ &= - \int_1^\infty [\partial_x \phi] \partial_y \phi ds, \end{aligned}$$

where $\hat{\mathbf{y}} = (0, 1)$. We also have using Eq. (3.8), $-\int_V \mathbf{u} \cdot (\mathbf{U} \times \boldsymbol{\omega}) dx = -\int_1^\infty [\partial_x \phi] \mathbf{n} \cdot \nabla \phi ds$. So we find

$$-\int_V \mathbf{u} \cdot (\mathbf{U} \times \boldsymbol{\omega}) dx = \int_{\mathcal{W}_- + \mathcal{W}_+} \bar{\mathbf{I}} \cdot \mathbf{n} ds,$$

i.e. the energy exchange term can be regarded as the acoustic energy flux into the wake.

Interpretation 3: Change of kinetic energy in the vortical field

This interpretation was given by Howe (1984): Let us focus on the downstream part of the domain, $s(x-1) > dy$, define the vortical component of the field $\mathbf{v} := \nabla\phi_v$, where

$$\phi_v(x, y) = e^{-i\Omega x} g(y), \quad g(y) = \frac{e^{i\sigma + id\Omega} \cosh(\Omega y) - \cosh(\Omega(s-y))}{2(\cos(\sigma + d\Omega) - \cosh(s\Omega))},$$

and denote by ϕ_a the velocity potential of the strictly acoustic contributions (i.e. the modal contributions from σ_m^\pm as shown in Tab. 3.1, whose velocity is continuous everywhere), such that $\mathbf{u} = \nabla\phi_a + \mathbf{v}$. We can write the linearised momentum equation in the form

$$\partial_t \mathbf{v} + \nabla(\partial_t \phi_a + p + \mathbf{U} \cdot (\mathbf{v} + \nabla\phi_a)) = -\boldsymbol{\omega} \times (\mathbf{U} + \mathbf{v} + \nabla\phi_a).$$

Taking the dot product of both sides with $\mathbf{U} + \mathbf{v}$ gives to second order

$$\begin{aligned} \frac{1}{2} \partial_t |\mathbf{v}|^2 + \mathbf{v} \cdot \nabla(\partial_t \phi_a + p + \mathbf{U} \cdot (\mathbf{v} + \nabla\phi_a)) &= -\mathbf{v} \cdot (\boldsymbol{\omega} \times \mathbf{U}) - \mathbf{U} \cdot (\boldsymbol{\omega} \times \mathbf{u}) \\ &= -\nabla\phi_a \cdot (\mathbf{U} \times \boldsymbol{\omega}). \end{aligned}$$

Integration over $\tilde{V} = \{(x, y); -s/2 < y < s/2, 1 + yd/s < x < X_2 + yd/s\}$ and using the divergence theorem with the observation that $\nabla \cdot \mathbf{v} = 0$ we find

$$\begin{aligned} \int_{\tilde{V}} \mathbf{v} \cdot \nabla(\partial_t \phi_a + p + \mathbf{U} \cdot (\mathbf{v} + \nabla\phi_a)) dx &= \int_{\partial\tilde{V}} (\mathbf{n} \cdot \mathbf{v}) (\partial_t \phi_a + p + \mathbf{U} \cdot (\mathbf{v} + \nabla\phi_a)) ds \\ &= \int_{\partial\tilde{V}} (\mathbf{n} \cdot \mathbf{v}) (\partial_t \phi_a + p + \mathbf{U} \cdot \mathbf{u}) ds \\ &= \int_{\partial\tilde{V}} \mathbf{n} \cdot \mathbf{I}_v ds, \end{aligned}$$

where $\mathbf{I}_v = (\mathbf{U} \cdot \mathbf{v}) \mathbf{v}$ is the energy flux of the vortical field. Therefore the rate of change of vortical kinetic energy in \tilde{V} can be expressed as

$$\frac{d}{dt} \int_{\tilde{V}} \frac{1}{2} |\mathbf{v}|^2 dx = \underbrace{- \int_{\partial\tilde{V}} \mathbf{n} \cdot \mathbf{I}_v ds}_{\text{flux of vortical energy out of } \tilde{V}} \quad - \underbrace{\int_{\tilde{V}} \nabla\phi_a \cdot (\mathbf{U} \times \boldsymbol{\omega}) dx}_{\text{exchange of energy with acoustic field}}.$$

Relationship to interface flux

We can also relate the energy exchange term on the right hand side of Eq. (3.10) directly to the ‘interface flux’ that is described by Arzoumanian (2011) for a vortex sheet in an incompressible fluid (and which also occurs when two fluid layers are separated by an elastic sheet as described in Crighton and Oswell (1991)). Indeed, in reality, the vortex sheets are located at a small displacement $y = \eta(t, x)$, which is to ensure they are convected

with the flow (zero flow across the wake):

$$\frac{D\eta}{Dt} = \partial_y \phi.$$

Furthermore, the form of the discontinuity in ϕ is fixed to ensure continuity of pressure (cf. Jones, 1972) such that

$$-\frac{D[\phi]}{Dt} = [p] = 0.$$

We can combine these two observations to show (after a few steps of algebra):

$$\begin{aligned} \int_V \mathbf{u} \cdot (\mathbf{U} \times \boldsymbol{\omega}) \, dx &= \int_V \mathbf{u} \cdot \hat{\mathbf{y}} \delta(y) [\partial_x \phi] \, dx \\ &= \frac{d}{dt} \int_1^{X_2} \eta [\partial_x \phi] \, ds - J_i(X_2) + J_i(1), \end{aligned}$$

where $J_i(x) = \eta [\partial_t \phi]$. We note that the wake is attached to the blades at the trailing edges, $\eta(1) = 0$, thus $J_i(1) = 0$. Taking a time average ensures that, in our time-harmonic flow, the contribution from $\frac{d}{dt} \int_1^{X_2} \eta [\partial_x \phi] \, ds$ vanishes. This allows us to conclude that the time-average of the right hand side in Eq. (3.10) is $-J_i(X_2)$, which is precisely the expression for the interface flux described by Arzoumanian (2011) (given here in non-dimensionalised form), and the analogue for a vortex sheet of the definition by Crighton and Oswell (1991).

3.4 Results and discussion

3.4.1 Symmetries in the field and zero acoustic reflection

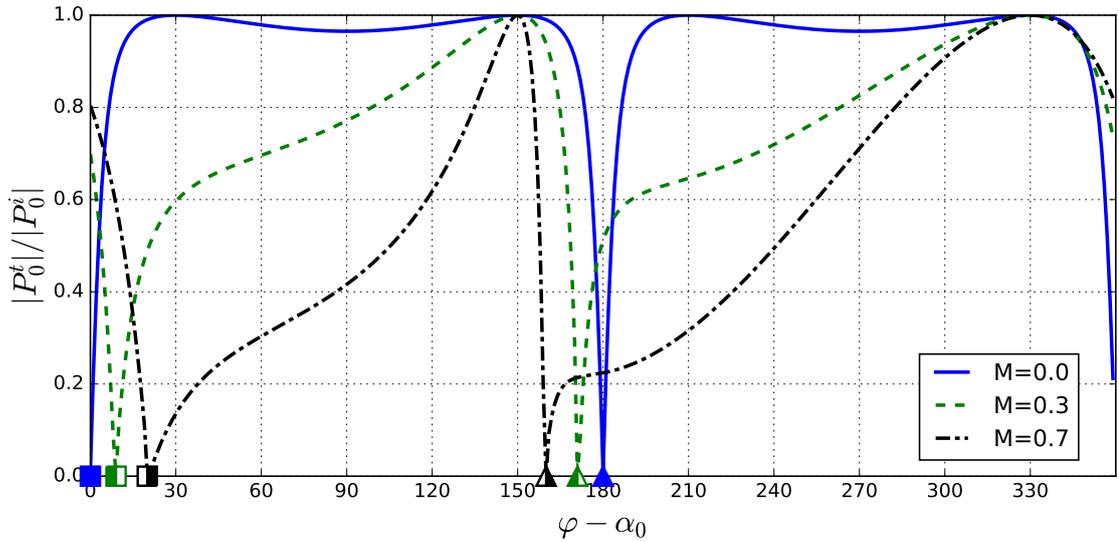
We begin by considering the amplitudes of the first radiating modes U_0, D_0 for various values of φ . In Figs. 3.4 & 3.5 we have plotted these amplitudes as a function of φ . In these graphs we choose $\alpha_0 = \pi/6$, $d^2 + s^2 = 1$, and $\Omega M = \pi/4, 5\pi/4$ respectively, such that the parameters for the two figures differ only in reduced frequency and hence in the number of cut-on radiation modes. In particular in Fig. 3.4 only σ_0^\pm are cut-on, whereas in Fig. 3.5 the cut-on radiation modes are $\sigma_0^\pm, \sigma_{-1}^\pm$.

In order to allow for comparison against earlier work by Koch (1971) (as we showed in Fig. 2.6) we plot the modal pressures defined as follows:

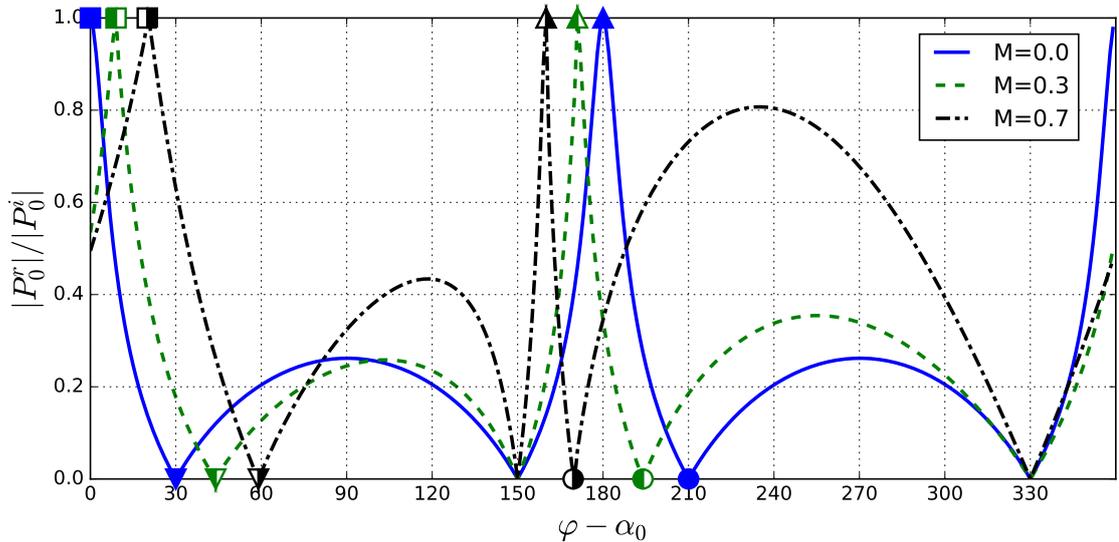
$$P_0^i = \begin{cases} (-i)(\Omega - \sigma_0^-)I, & \text{if } k_x = \sigma_0^-, \\ (-i)(\Omega - \sigma_0^+)I, & \text{if } k_x = \sigma_0^+, \end{cases}$$

$$P_0^t = \begin{cases} (-i)(\Omega - \sigma_0^-)D_0, & \text{if } k_x = \sigma_0^-, \\ (-i)(\Omega - \sigma_0^+)U_0, & \text{if } k_x = \sigma_0^+, \end{cases} \quad P_0^r = \begin{cases} (-i)(\Omega - \sigma_0^+)U_0, & \text{if } k_x = \sigma_0^-, \\ (-i)(\Omega - \sigma_0^-)D_0, & \text{if } k_x = \sigma_0^+. \end{cases}$$

In Figs. 3.4 & 3.5 the transition from upstream to downstream incidence and vice versa is marked with \blacksquare and \blacktriangle respectively (solid, left-half-filled and right-half-filled shapes correspond to $M = 0.0, 0.3$ and 0.7 respectively), such that for each value of M the wave is downstream incident if $\varphi - \alpha_0$ is between \blacksquare and \blacktriangle . Note the region of upstream incidence is fixed by the condition $M \sin \alpha_0 + \sin(\alpha_0 - \varphi) > 0$ (cf. Eq. (2.12)).

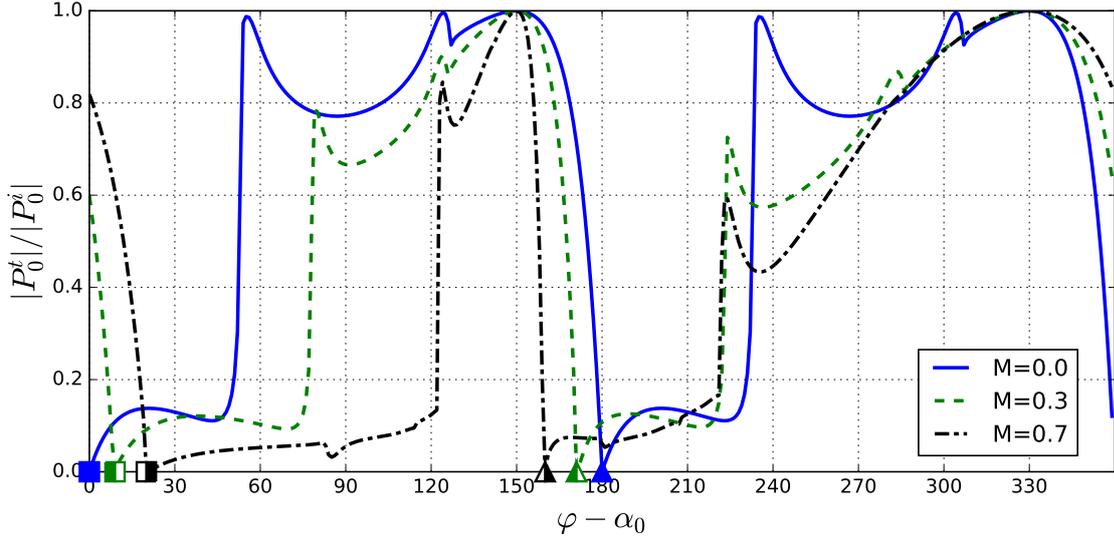


(a) Transmitted modal pressure.

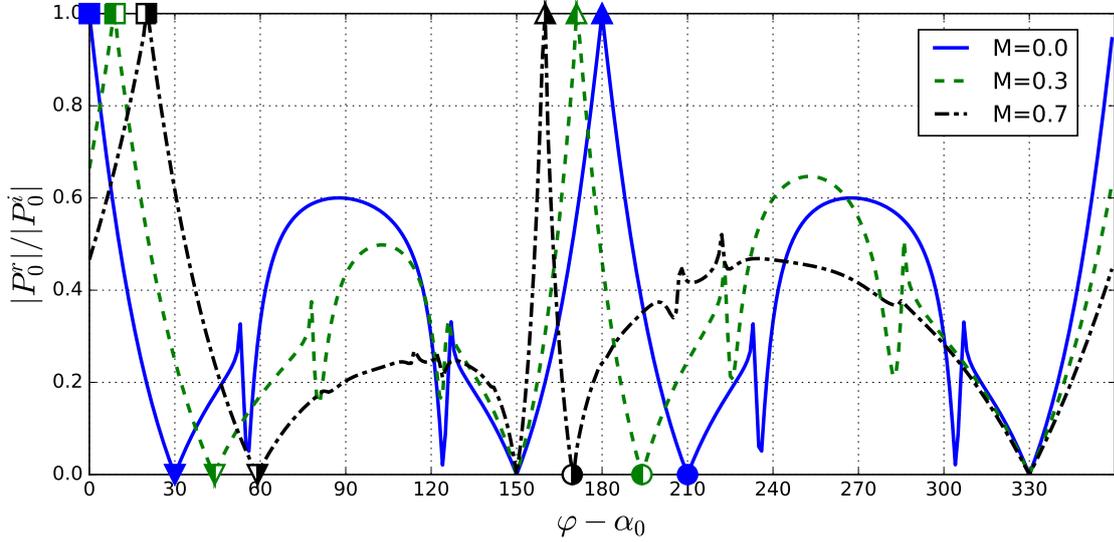


(b) Reflected modal pressure.

Figure 3.4: Relative modal pressures of the first transmitted and reflected mode, with $\alpha_0 = 30^\circ$ and $\Omega M = \pi/4$.



(a) Transmitted modal pressure.



(b) Reflected modal pressure.

Figure 3.5: Relative modal pressures of the first transmitted and reflected mode, with $\alpha_0 = 30^\circ$ and $\Omega M = 5\pi/4$.

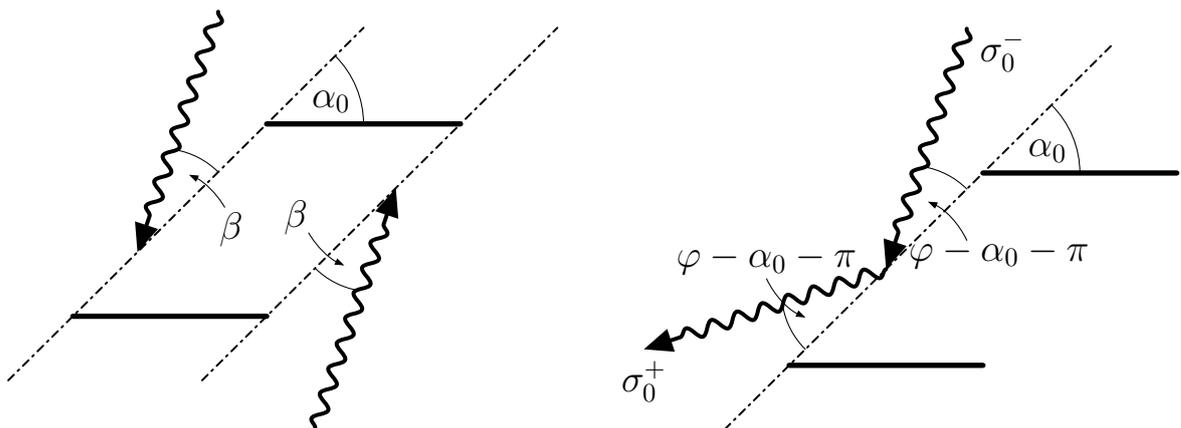
Due to convection, larger values of M result in a wider range of values for $\varphi - \alpha_0$ corresponding to upstream incidence. There are a few interesting features that we can observe in Figs. 3.4 & 3.5:

1. When only σ_0^\pm are cut-on radiation modes (as is the case in Fig. 3.4) and we have zero mean flow ($M = 0$), the reflected and transmitted amplitudes are symmetric about $\varphi = \pi + \alpha_0$, i.e. the amplitudes for $\varphi - \alpha_0 = \pi + \beta$ equal in modulus those for $\varphi - \alpha_0 = \pi - \beta$ for any $\beta \in (0, \pi)$. This is in addition to the obvious symmetry between the amplitudes for $\varphi - \alpha_0 = \beta$ and $\varphi - \alpha_0 = \beta + \pi$. The latter is simply due to the fact that for $M = 0$ there is no distinction between upstream incidence

and downstream incidence as shown in Fig. 3.6a. However a proof of the former requires slightly more careful consideration as is shown below.

2. For any value of $M \in [0, 1), \Omega \in [0, \infty)$ the cascade is perfectly transmitting ($|P_0^t|/|P_0^i| = 1$) when $\varphi = 0, \pi$ (the points $\varphi - \alpha_0 = 330^\circ$ and $\varphi - \alpha_0 = 150^\circ$ respectively), corresponding to incident waves from upstream and downstream respectively propagating perfectly parallel to, and therefore being unimpeded by, the blades.
3. For any choice of M and Ω there are two additional points of zero acoustic reflection ($|P_0^r|/|P_0^i| = 0$), marked with \blacktriangledown and \bullet in Figs. 3.4b & 3.5b. When $M = 0$ these are located at $\varphi = 2\alpha_0, \pi + 2\alpha_0$ and for nonzero M these points shift slightly to increased/decreased values of φ respectively.

These features were observed numerically, for the case of a single cut-on radiation mode σ_0^\pm , by Koch (1971), and studied analytically for the case of zero mean flow, $M = 0$, by Porter (2021). It turns out that point 1 above is indeed a property of the field as long as only a single cut-on propagating mode exits upstream and downstream of the cascade and in Fig. 3.5a we see that the symmetry breaks down for the transmitted amplitude when multiple radiation modes are cut-on. Furthermore, points 2 and 3 are properties of the field for any value of M and Ω . Of course, it is clear that point 2 must hold since for $\varphi = 0, \pi$ the waves are unimpeded by the blades. In this section we provide a proof of features 1 and 3 based entirely on the energy balance – the underlying idea is that there is a direct correspondence between the cascade response for an incident mode σ_0^- from upstream and the response for an incident mode σ_0^+ from downstream.



(a) Standard symmetry in the field when $M = 0$. (b) The angles of propagation for σ_0^\pm ($M = 0$).

Figure 3.6: Special symmetries in the field for zero mean flow.

To understand this correspondence, we need to write the angle of propagation for the zeroth radiating modes σ_0^\pm more explicitly and for this purpose it is useful to describe the

geometry in terms of the stagger angle α_0 and the leading edge separation c such that $s = c \sin \alpha_0$, $d = c \cos \alpha_0$. This allows us to write the inter-blade phase angle, as defined in Eq. (3.3), as

$$\sigma = \frac{-\Omega M}{1 + M \cos \varphi} c \cos(\varphi - \alpha_0),$$

and hence the zeroth radiating mode (cf. Eq. (3.6)) in the form

$$\begin{aligned} \sigma_0^\pm &= \frac{M\Omega (\cos \varphi (1 - M^2 \sin^2 \alpha_0) - M \sin^2 \alpha_0 - \sin \alpha_0 \sin(\alpha_0 - \varphi))}{(1 + M \cos \varphi) (1 - M^2 \sin^2 \alpha_0)} \\ &\mp \frac{M\Omega \sin \alpha_0 |M \sin \alpha_0 + \sin(\alpha_0 - \varphi)|}{(1 + M \cos \varphi) (1 - M^2 \sin^2 \alpha_0)}. \end{aligned}$$

As mentioned above, an incident acoustic wave with $k_x = \Omega M \cos \varphi / (1 + M \cos \varphi)$ is effectively incident from upstream if and only if $M \sin \alpha_0 + \sin(\alpha_0 - \varphi) > 0$ (cf. Eq. (2.12)), and in that case we have

$$\sigma_0^- = \frac{\Omega M \cos \varphi}{1 + M \cos \varphi} = k_x, \quad \sigma_0^+ = \frac{\Omega M \cos \tilde{\varphi}}{1 + M \cos \tilde{\varphi}}, \quad (3.13)$$

where

$$\cos \tilde{\varphi} = \frac{\cos(\varphi - 2\alpha_0) - 2M \sin^2 \alpha_0 - M^2 \cos \varphi \sin^2 \alpha_0}{1 + 2M \sin(\alpha_0 - \varphi) \sin \alpha_0 + M^2 \sin^2 \alpha_0}.$$

In the case of zero mean flow the latter expression reduces to $\cos \tilde{\varphi} = \cos(\varphi - 2\alpha_0)$. If we have downstream incidence, $M \sin \alpha_0 + \sin(\alpha_0 - \varphi) < 0$, the expressions for these two modes are reversed. As mentioned earlier the cascade is perfectly transmitting ($|P_0^t|/|P_0^i| = 1$, $|P_0^r|/|P_0^i| = 0$) when $\varphi = 0, \pi$ due to the incident waves propagating parallel to, and therefore being unimpeded by, the cascade blades. There are two corresponding angles of inclination of the normal to the wavefronts, $\tilde{\varphi}_0, \tilde{\varphi}_\pi$ in the following sense:

- If $\varphi = 0$, then

$$\begin{aligned} \sigma_0^+ &= \frac{\Omega M \cos \tilde{\varphi}_0}{1 + M \cos \tilde{\varphi}_0}, \\ \cos \tilde{\varphi}_0 &= \frac{\cos(2\alpha_0) - 2M \sin^2 \alpha_0 - M^2 \sin^2 \alpha_0}{1 + 2M \sin^2 \alpha_0 + M^2 \sin^2 \alpha_0}, \end{aligned} \quad (3.14)$$

and in particular when $M = 0$ we have $\tilde{\varphi}_0 = 2\alpha_0$.

- If $\varphi = \pi$, then

$$\begin{aligned}\sigma_0^- &= \frac{\Omega M \cos \tilde{\varphi}_\pi}{1 + M \cos \tilde{\varphi}_\pi}, \\ \cos \tilde{\varphi}_\pi &= \frac{\cos(\pi + 2\alpha_0) - 2M \sin^2 \alpha_0 + M^2 \sin^2 \alpha_0}{1 + 2M \sin(\alpha_0 - \pi) \sin \alpha_0 + M^2 \sin^2 \alpha_0},\end{aligned}\tag{3.15}$$

and in particular when $M = 0$ we have $\tilde{\varphi}_\pi = \pi + 2\alpha_0$.

This means that σ_0^+ has a direction of propagation as if it was reflected directly on the cascade front (cf. sketch for $M = 0$ in Fig. 3.6b). So in a sense σ_0^\pm are the usual plane wave modes that one would expect to find for reflection on a flat material that is located parallel to the cascade front.

Symmetry along the cascade face in zero mean flow

We now show the symmetry of the transmission and reflection amplitudes with respect to $\varphi - \alpha_0 = \pi$ in the case of zero mean flow, i.e. property 1 as described above. As can be seen from Fig. 3.5a this symmetry is no longer true for the transmitted amplitude if there is more than a single propagating wave mode upstream and downstream of the cascade. Thus in this section we focus on the case when only σ_0^\pm are cut-on. Let $\varphi = \pi + \alpha_0 + \beta$ and $\tilde{\varphi} = \pi + \alpha_0 - \beta$ for some $\beta > 0$. Then the cascade response for an incident field with normal to the wavefronts that is inclined at an angle φ is

$$\phi_1(x, y) = \begin{cases} I e^{-i\sigma_0^- x} e^{\gamma(\sigma_0^-)y} + R_1 e^{-i\sigma_0^+ x} + \text{cut-off modes}, & sx < dy, \\ T_1 e^{-i\sigma_0^- x} e^{\gamma(\sigma_0^-)y} + \text{cut-off modes}, & sx > dy + 1, \end{cases}$$

and the response for an incident field with normal to the wavefronts that is inclined at an angle $\tilde{\varphi}$ is (according to Eq. (3.13) the radiation modes are the same: $\sigma_0^\pm(\varphi) = \sigma_0^\pm(\tilde{\varphi})$)

$$\phi_2(x, y) = \begin{cases} T_2 e^{-i\sigma_0^+ x} + \text{cut-off modes}, & sx < dy, \\ I e^{-i\sigma_0^+ x} e^{\gamma(\sigma_0^+)y} + R_2 e^{-i\sigma_0^- x} e^{\gamma(\sigma_0^-)y} + \text{cut-off modes}, & sx > dy + 1. \end{cases}$$

By linearity we can add the two fields to produce a valid solution of the scattering problem of the following form (for any arbitrary $z \in \mathbb{C}$):

$$\phi_1 + z\phi_2 = \begin{cases} z I e^{-i\sigma_0^- x} e^{\gamma(\sigma_0^-)y} + (R_1 + zT_2) e^{-i\sigma_0^+ x} + \text{cut-off modes}, & sx < dy, \\ I e^{-i\sigma_0^+ x} e^{\gamma(\sigma_0^+)y} + (T_1 + zR_2) e^{-i\sigma_0^- x} e^{\gamma(\sigma_0^-)y} + \text{cut-off modes}, & sx > dy + 1. \end{cases}$$

This field must satisfy the original equation of motion, and thus the time-averaged energy balance must remain valid (in this case however we have both upstream and downstream

contribution from the incident wave):

$$|I|^2 + |zI|^2 = |R_1 + zT_2|^2 + |T_1 + zR_2|^2, \quad \forall z \in \mathbb{C}. \quad (3.16)$$

Taking $z \rightarrow 0, \infty$ respectively yields

$$|I|^2 = |T_1|^2 + |R_1|^2 = |T_2|^2 + |R_2|^2. \quad (3.17)$$

We can simplify Eq. (3.16) using Eq. (3.17) to find

$$\operatorname{Re}[zR_1^*T_2] = -\operatorname{Re}[zT_1^*R_2], \quad \forall z \in \mathbb{C}.$$

Since this holds for all values of $z \in \mathbb{C}$ we infer that $R_1^*T_2 = T_1^*R_2$ which, together with Eq. (3.17), proves $|R_1| = |R_2|$, $|T_1| = |T_2|$, i.e. transmission and reflection amplitudes are symmetric with respect to $\varphi - \alpha_0 = \pi$.

Zero acoustic reflection in leading radiating mode for subsonic mean flow

Although this aforementioned symmetry is not preserved in the case when $M > 0$, we still find two special angles of inclination of the normal to the wavefronts, $\tilde{\varphi}_0, \tilde{\varphi}_\pi$ as defined in Eqs. (3.14) & (3.15), for which the reflected acoustic field has zero contribution from the zeroth radiating mode (and hence is identically zero in the far-field when there is only one cut-on mode). *Note this property holds irrespective of the number of cut-on modes in the field.* These special angles are marked in Figs. 3.4b, 3.5b & 3.7, with \blacktriangledown and \bullet corresponding to $\tilde{\varphi}_0$ and $\tilde{\varphi}_\pi$ respectively. As can be seen from Fig. 3.7, when $M > 0$, there may be a nonzero contribution from the hydrodynamic mode (which is part of the reflected field in the case $\tilde{\varphi}_0$).

We focus on the case $\tilde{\varphi}_0$ (i.e. downstream incidence) since the case $\tilde{\varphi}_\pi$ can be treated analogously: Let the total velocity potential due to an incident wave with normal to the wavefronts inclined at an angle $\tilde{\varphi}_0$ be denoted by

$$\tilde{\phi}(x, y) = \begin{cases} \sum_{m \in \mathbb{Z}} T_m e^{-i\sigma_m^+ x} e^{-i\gamma(\sigma_m^+) y}, & sx < dy, \\ I e^{-i\sigma_0^+ x} e^{-\gamma(\sigma_0^+) y} + B e^{-i\Omega x} g(y) + \sum_{m \in \mathbb{Z}} R_m e^{-i\sigma_m^- x} e^{\gamma(\sigma_m^-) y}, & sx > dy + 1. \end{cases}$$

We know from our prior discussion that the mode $\sigma_0^- = \sigma_0^-(\tilde{\varphi}_0) = \sigma_0^-(0)$ corresponds to the angle $\varphi = 0$ and that the response for this type of parallel-propagating mode must be of the form

$$\phi(x, y) = I e^{-i\sigma_0^- x} e^{-i\gamma(\sigma_0^-) y} = I e^{-i\sigma_0^- x}, \quad (x, y) \in \mathbb{R}^2.$$

As for the zero mean flow case, we consider the superposition of these two fields, $\tilde{\phi} + z\phi$, which is a valid solution to the scattering problem and hence satisfies the time-averaged energy balance (which must again account for incident fields both upstream and downstream of the cascade):

$$P_U + P_D + P_H + \Pi_\omega = P_I, \quad (3.18)$$

where

$$\begin{aligned} P_U &= - \sum_{\substack{m \in \mathbb{Z} \\ \text{cut-on}}} |T_m|^2 (s\Omega M^2 \beta^{-1} + s\beta\sigma_m^+ - d\beta^{-1}\gamma(\sigma_m^+)), \\ P_D &= |R_0 + zI|^2 (-s\Omega M^2 \beta^{-1} - s\beta\sigma_0^- - d\beta^{-1}\gamma(\sigma_0^-)) \\ &\quad + \sum_{\substack{m \in \mathbb{Z} \setminus \{0\} \\ \text{cut-on}}} |R_m|^2 (-s\Omega M^2 \beta^{-1} - s\beta\sigma_m^- - d\beta^{-1}\gamma(\sigma_m^-)), \\ P_I &= |zI|^2 (-s\Omega M^2 \beta^{-1} - s\beta\sigma_0^- - d\beta^{-1}\gamma(\sigma_0^-)) + |I|^2 (-s\Omega M^2 \beta^{-1} - s\beta\sigma_0^+ + d\beta^{-1}\gamma(\sigma_0^+)), \\ P_H &= |B|^2 \left(s\Omega\beta^{-1} \int_0^1 |g(st)|^2 dt + \text{Im} \left[\beta^{-1}d \int_0^1 g(st)^* g'(st) dt \right] \right), \\ \Pi_\omega &= -\text{Im} \left[B^*(zI + R_0)\gamma(\sigma_0^-) \frac{e^{i\Omega - i\sigma_0^-}}{i\Omega\beta - \beta\sigma_0^-} - B^*I\gamma(\sigma_0^+) \frac{e^{i\Omega - i\sigma_0^+}}{i\Omega\beta - \beta\sigma_0^+} \right. \\ &\quad \left. + \sum_{m \in \mathbb{Z}, m \neq 0} B^*R_m\gamma(\sigma_m^-) \frac{e^{i\Omega - i\sigma_m^-}}{i\Omega\beta - \beta\sigma_m^-} \right]. \end{aligned}$$

One can quickly check that in the current case $\gamma(\sigma_0^-) = 0$, and hence by setting $z \rightarrow 0, +\infty$, and using the resulting equations to cancel appropriate terms in Eq. (3.18), we arrive at

$$\text{Re} [R_0^* z I] (-s\Omega M^2 \beta^{-1} - s\beta\sigma_0^-) = 0 \quad \forall z \in \mathbb{C},$$

which implies $R_0 = 0$, i.e. the leading radiating mode σ_0^- has zero contribution to the reflected far-field. Of course, in general, the remaining reflected amplitudes, R_m with $m \neq 0$, need not vanish when $\varphi = \tilde{\varphi}_0$ or $\varphi = \tilde{\varphi}_\pi$, i.e. sound may be scattered into higher order reflected modes. Thus, while $R_0 = 0$ holds regardless of the choice of frequency and subsonic mean-flow speed, the reflected sound power is truly zero only when a single radiation mode is cut-on.

3.4.2 The significance of the hydrodynamic and energy conversion terms in balancing the energy

We now visualise the terms from the time-averaged energy balance Eq. (3.12), to understand their individual contribution to the overall outgoing power. In order to facilitate comparison against the amplitudes shown in Fig. 3.4 we plot here (and also in §3.4.3) the reflected sound power P_R and the transmitted sound power P_T . These are simply given in terms of upstream and downstream sound powers taking appropriate account for the change of incidence direction as follows (recall that the acoustic wave is incident from upstream if $\sin(\varphi - \alpha_0) < M \sin \alpha_0$):

$$\begin{aligned} P_R &= \begin{cases} P_U, & \text{if } \sin(\varphi - \alpha_0) < M \sin \alpha_0, \\ P_D, & \text{if } \sin(\varphi - \alpha_0) > M \sin \alpha_0, \end{cases} \\ P_T &= \begin{cases} P_D, & \text{if } \sin(\varphi - \alpha_0) < M \sin \alpha_0, \\ P_U, & \text{if } \sin(\varphi - \alpha_0) > M \sin \alpha_0. \end{cases} \end{aligned} \quad (3.19)$$

In Fig. 3.7 we see the contribution of each individual term as a percentage of the overall outgoing power – the reflected and transmitted sound power, hydrodynamic power, and conversion between acoustic energy and kinetic energy in the vortical field. The results in the graph correspond to a stagger angle of $\alpha_0 = \pi/6$, a blade separation of $c = 1$, a Mach number of $M = 0.3$ and a reduced frequency with $\Omega M = \pi/4$.

The curves in Fig. 3.7 are cumulative, so that for instance the value of P_H is given as the difference of the green curve (marked with diamond shapes) and the blue curve (marked with circles) – the value of P_H is highlighted in light-green in the graph. The vertical red dashed lines marked with \blacksquare and \blacktriangle respectively correspond to the values of φ which separate downstream ($\varphi - \alpha_0 \in (9^\circ, 171^\circ)$) and upstream ($\varphi - \alpha_0 \in [0^\circ, 9^\circ) \cup (171^\circ, 360^\circ]$) incident regimes. Note that the roles of reflected and transmitted fields are reversed at these vertical lines, such that for $\varphi - \alpha_0 \in (9^\circ, 171^\circ)$ the reflected sound power P_R represents the downstream sound power P_D and for $\varphi - \alpha_0 \in [0^\circ, 9^\circ) \cup (171^\circ, 360^\circ]$ the reflected sound power represents to the upstream sound power P_U , and vice versa for P_T (cf. Eq. (3.19)). The important point to observe is that the purely acoustic contributions to the outgoing energy are everything below the blue curve (with circles) in the graph. It is clear that for a range of angles of incidence the acoustic contribution only accounts for a fraction of the total outgoing power, indeed sometimes less than half of it, meaning there is significant attenuation of the incident acoustic field. This example is representative for a wide range of numerical experiments that were performed: the acoustic contributions alone account for the significant majority of outgoing power only in very specific cases and

generally the conversion of acoustic energy into vortical kinetic energy plays an important role in balancing the incoming and outgoing energy flux.

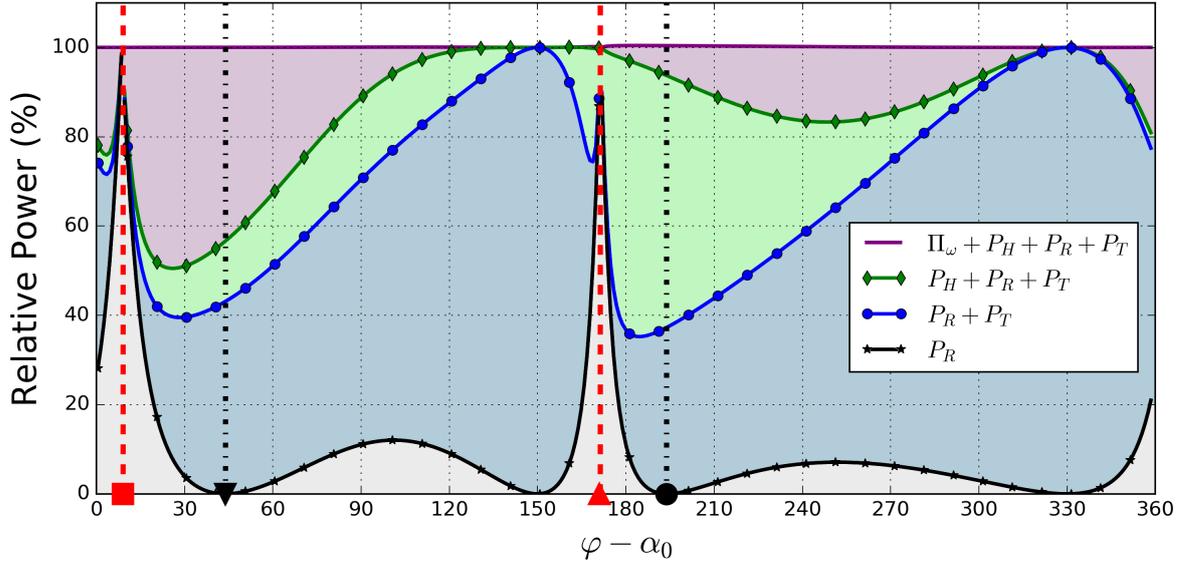


Figure 3.7: The contributions as percentage of the total incoming power; P_U and P_D are respectively the upstream and downstream acoustic power, P_H is the hydrodynamic power and Π_ω is the sound power absorbed by the wake. The vertical red dashed lines (marked with \blacksquare and \blacktriangle) separate downstream and upstream incidence and the symbols \blacktriangledown and \bullet correspond to $\tilde{\varphi}_0$ and $\tilde{\varphi}_\pi$ respectively.

3.4.3 Negative acoustic energy absorption and sound power generation in wave-cascade scattering

It was observed by Rienstra (1981, §4) for a single trailing edge in mean flow that the acoustic energy flux into the wake, Π_ω can take negative values which effectively means the blade harvests energy from the flow to increase the total outgoing acoustic and hydrodynamic power (i.e. $P_R + P_T + P_H$) to be greater than the incident power P_I . In some cases this exchange of energy leads to effective sound power generation when the total outgoing sound power is larger than the incoming acoustic power. We shall see that this effect can also be observed for the cascade of blades, indeed specifically we found the appearance of negative acoustic energy absorption (i.e. acoustic energy emission) by the wake in two scenarios of interest. A particular new contribution in this present section is our detailed study of the size of individual terms $P_R, P_T, P_H, \Pi_\omega$ that constitute the overall outgoing power, which allows us to gain more insight into the consequences of conversion of energy between the acoustic field and the unsteady vortex sheets that takes place at the trailing edge of the cascade blades. One of those insights (which can be seen in Fig. 3.10 and which is described in more detail below) is that energy conversion between vortical and acoustic field can, in some cases, lead to over-reflection on the cascade – i.e.

the reflected wave can be larger in amplitude than the incoming wave $|R_0| > |I|$.

Acoustic energy emission from the wake near modal cut-on

In Fig. 3.8 we plot the relative acoustic power $P_A = P_R + P_T$ and the hydrodynamic power P_H as well as their sum (we have rescaled all quantities by P_I) on a decibel scale (i.e. we plot for example $10 \log_{10}(P_A/P_I)$). The geometry in this case is taken to be $s = 1, d = 0$, and we choose $\varphi = 4\pi/3$ and $M = 0.75$.

Whenever the black curve (marked with stars) rises above 0 in the graph, we have $P_H + P_A > P_I$, i.e. the acoustic energy absorbed by the wake Π_ω must be negative. Interestingly, we observe this phenomenon at those frequencies just before radiation modes σ_m^\pm become cut-on. In the figure these cut-on values of the reduced frequency Ω are highlighted by the red vertical dashed curves – the first one of these corresponds to the red dashed line in the dispersion diagram in Fig. 3.2. The following trend becomes apparent (as was also observed by Glegg (1999, §5.3) for the purely acoustic part of the power): just before a new mode becomes cut-on the hydrodynamic power rises and the acoustic power falls, but sometimes their sum reaches values greater than the incoming power. This is followed by a rapid decay in hydrodynamic power and an increase in acoustic power as the frequency is increased further. Indeed, we may also infer from this figure that for larger frequencies Ω the hydrodynamic effects (both in terms of P_H and Π_ω) play a significant role only close to modal cut-on and that the attenuation of sound is especially pronounced at lower frequencies, which is similar to the observations made by Bechert et al. (1977).

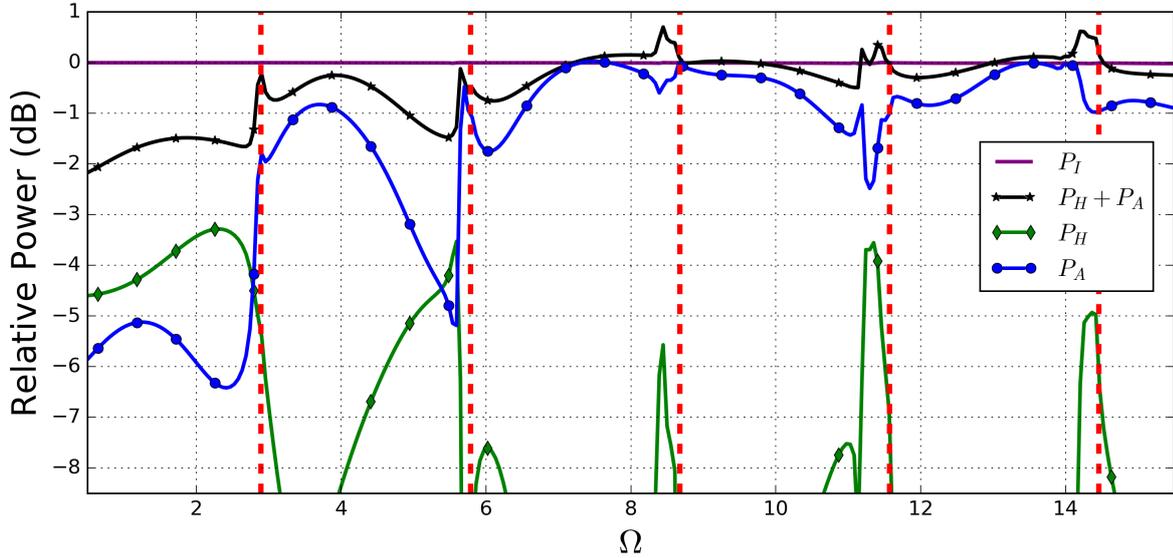


Figure 3.8: The acoustic and hydrodynamic power for varying incident frequency (dB scale, i.e. we plot for example $10 \log_{10}(P_A/P_I)$). The vertical red dashed lines highlight the frequencies where new propagating modes become cut-on.

Sound power generation and over-reflection

Finally, we highlight that even for fixed frequencies the acoustic energy absorbed by the wake Π_ω can become negative for specific values of M and φ .

The effect is particularly pronounced for downstream incidence and large subsonic Mach numbers, and in some cases results in a reflection amplitude that is larger than the incoming amplitude, i.e. $|R_0| > |I|$. In Figs. 3.9 & 3.10 we plot the relevant quantities from the energy balance for various values of M and φ . The cascade geometry in this case is $\alpha_0 = \pi/6, c = 1$ and the incident frequency is $\Omega M = \pi/4$.

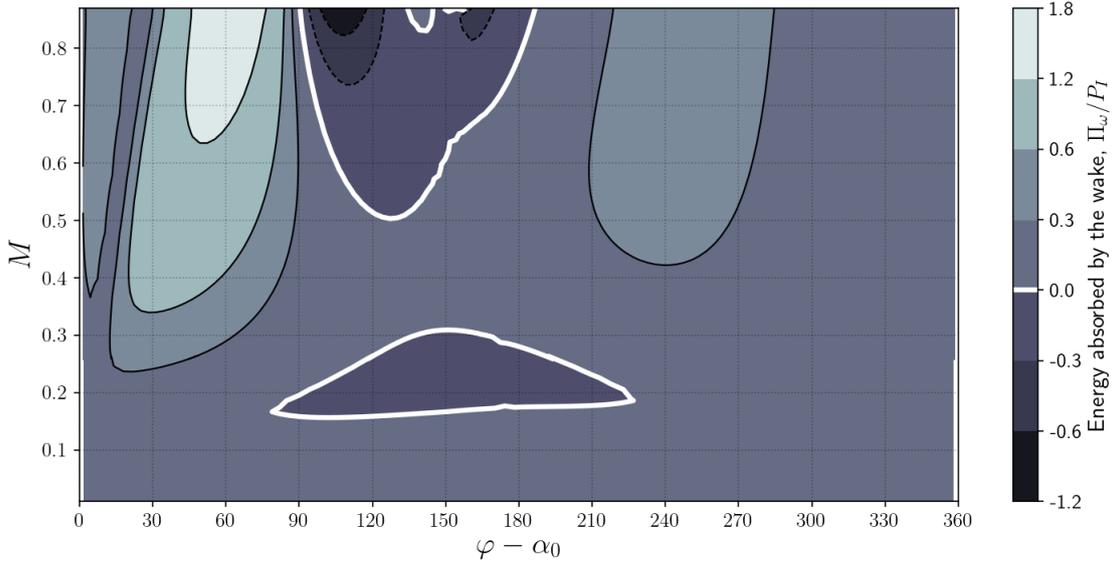


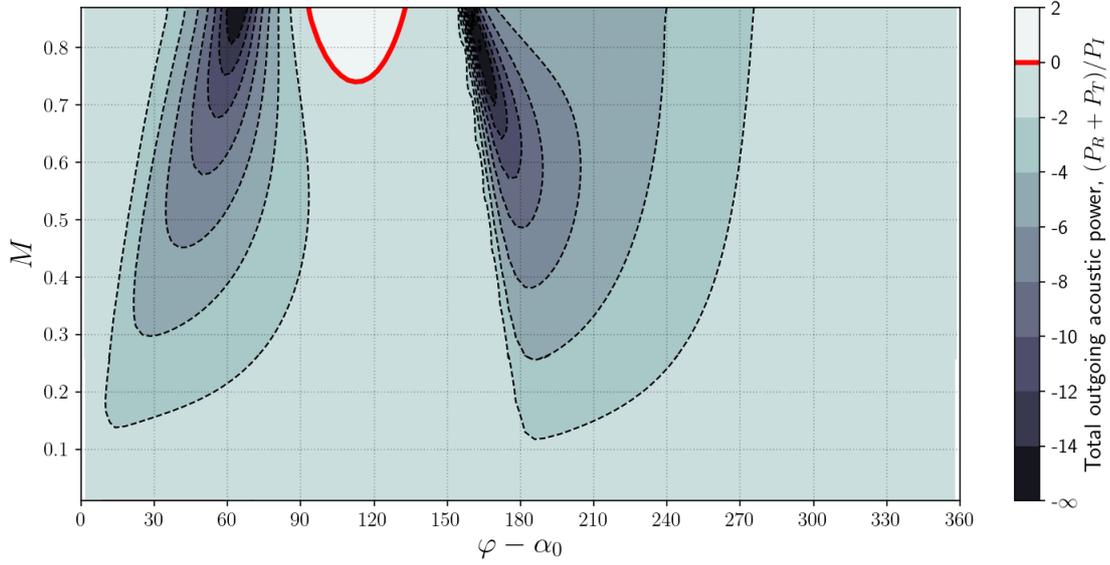
Figure 3.9: Relative acoustic energy absorption by the wake (linear scale), the solid white line marks $\Pi_\omega = 0$.

In Fig. 3.9 we see the energy absorbed by the wake relative to P_I (i.e. Π_ω/P_I) on a linear scale. The solid white lines highlight contours of constant $\Pi_\omega = 0$. We observe the absorbed energy can become negative for a range of angles of incidence, meaning effectively that there is a flux of acoustic energy out of the wake. As is the case for a single trailing edge (cf. Rienstra, 1981, Fig. 3), the majority of this region of acoustic energy emission occurs when the waves are incident from downstream (i.e. when $\sin(\varphi - \alpha_0) > M \sin \alpha_0$). It is precisely in these regions that amplification of the outgoing acoustic power can occur. Indeed looking at Fig. 3.10 (which shows the relative outgoing acoustic power, and the reflected sound power on a decibel scale) we observe that for large Mach numbers, and when $\varphi - \alpha_0 \approx 105^\circ$, this amplification is indeed observed in practice. In fact, we find in Fig. 3.10b that for a part of this parameter region we even have $P_R > P_I$, i.e. $|R_0| > |I|$, which means that the incident wave is over-reflected.

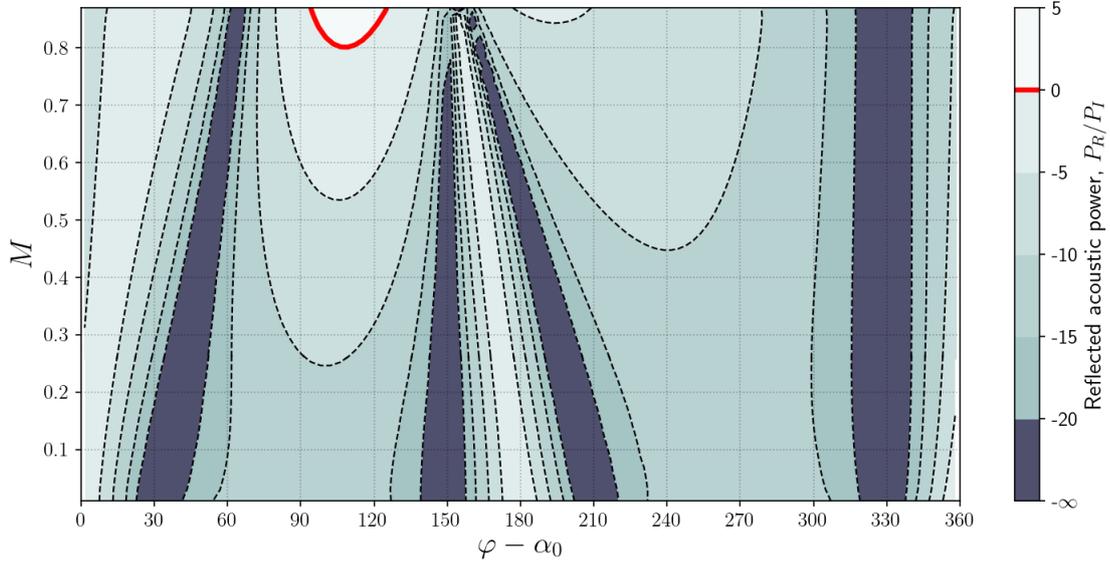
We also observe in Fig. 3.10a that there is significant attenuation of sound over a large range of directions of incidence. This attenuation is due to the conversion of sound into

the wake through the shedding of unsteady vorticity at the trailing edge, and as might be expected the attenuation is found to be greater for larger Mach numbers.

Finally, we highlight that in Fig. 3.10b for every fixed value of M the four dark regions contain isolated zeros of P_R at $\varphi = \tilde{\varphi}_0, \pi, \tilde{\varphi}_\pi, 0$ in this order as described in §3.4.1 (i.e. for any fixed value of M the ratio P_R/P_I takes the value $-\infty$ on the decibel scale at four distinct values of φ).



(a) Total outgoing sound power, where the solid red line marks $P_R + P_T = P_I$.



(b) Reflected sound power, where the solid red line marks $P_R = P_I$.

Figure 3.10: Relative outgoing sound powers (dB scale).

3.5 Conclusions

In this chapter we studied the energy balance for a cascade of flat plates in uniform subsonic mean flow. We showed that the outgoing power is composed of acoustic and hydrodynamic contributions and a term facilitating energy exchange between the wake and acoustic field, all of which, in general, yield a non-negligible contribution to the outgoing energy flux. Specifically, we provided numerical evidence that there is significant attenuation of sound over a large region of Mach numbers and angles of incidence for incoming sound waves, and that this effect is particularly pronounced at low frequencies. We also found numerical evidence of sound power generation when the waves are incident from downstream and showed the generated power can result in a reflected wave with amplitude greater than the incident one.

Finally, we showed that this energy balance can be used to understand symmetries of the field in the angle of inclination of the wave fronts with respect to the cascade face when there is no mean flow as well as the effect of zero acoustic reflection at certain angles of incidence with mean flow.

List of symbols

Symbol	Description
(k_x, k_y)	Non-dimensional wave vector of the incident field.
B	Amplitude of the hydrodynamic mode.
E	Acoustic energy density.
M	Mach number of the uniform mean flow.
P_H	Power carried by the hydrodynamic mode to $x = +\infty$.
P_I	Incident sound power.
P_U, P_D	Acoustic sound powers radiated to $x = \pm\infty$ respectively.
P_j^t, P_j^r	Transmitted and reflected modal pressures.
U_m, D_m	Up-/Downstream amplitudes of acoustic modes in the scattered field.
Ω	Reduced frequency.
Π_ω	Power absorbed by the wake.
α_0	Cascade stagger angle.
β	Prandtl–Glauert number of the mean flow.
η	First order unsteady displacement of the wake.
\mathbf{I}	Acoustic intensity (acoustic energy flux).
\mathbf{U}	Velocity field of the uniform mean flow.
\mathbf{u}	Unsteady velocity field.
$\bar{\mathbf{I}}$	Time averaged acoustic energy flux.
ϕ	Spatial component of unsteady velocity potential.
ρ	Unsteady density field.
σ	Interblade phase angle.
$\tilde{\mathbf{u}}$	Velocity field of the full flow.
$\tilde{\rho}$	Density field of the full flow.
\tilde{p}	Pressure field of the full flow.
φ	Angle of inclination of the normal to the incident wavefronts.
c	Leading edge separation of blades, $c = \sqrt{s^2 + d^2}$.
p	Unsteady pressure field.
z^*	Complex conjugate of $z \in \mathbb{C}$.

Chapter 4

Oversampled collocation methods for boundary element methods in two dimensions

4.1 Introduction

Following the treatment of scattering problems by analytical techniques in the previous two chapters, we shall now focus on the study of numerical methods. Specifically, in this chapter we will study the solution of wave scattering problems in the boundary integral formulation which we described in the introduction of this thesis in §1.1.2.

During the recent decade, the concept of oversampling (i.e. taking more observations than the dimension of the trial space) has found increasing attention in the numerical analysis community as a method to achieve enhanced reconstruction and function approximation. In a range of settings it is now understood that the effects of suboptimal observations can be mitigated and reconstructions stabilised by introducing a sufficient number of additional observations. As one of the first works in this direction we mention Adcock and Hansen (2012), who found that oversampling provides a suitable paradigm for function approximation by sampling from a Riesz basis in a Hilbert space, even when the sampling and trial spaces are distinct. It was then shown by Adcock et al. (2014) that oversampling can be used for equispaced Fourier extensions to achieve superalgebraic convergence in a numerically stable manner (where it is known that no method for the Fourier extension problem can be both numerically stable and exponentially convergent). More recently, oversampling was studied in the context of function approximation using frames by Adcock and Huybrechs (2020) who found it can lead to much-improved accuracy in the approximation and help further mitigate ill-conditioning arising from using a redundant set rather than a basis if an appropriate regularisation is used.

In the present chapter, we are interested in the study of least-squares oversampling in collocation methods for Fredholm integral equations arising in wave scattering problems as described in §1.2.2. The use of collocation methods to solve these types of integral equations is generally speaking a delicate matter. On the one hand, the discretization matrix entries are given by lower-dimensional integrals, which makes the methods easier to implement and permits the use of a wider range of techniques for numerical integration (see for instance Deaño et al. (2017), Gibbs et al. (2020) and the techniques described in Chapter 5 of this thesis). On the other hand, at present no general framework for the convergence analysis of collocation methods exists. This is in stark contrast to Galerkin methods, for which there is a well-known and wide-ranging convergence theory based mainly on the coercivity of associated bilinear forms. Even though a general framework for the convergence analysis of collocation methods is not available, the literature offers a number of deep insights into convergence properties in specific settings. One of the most complete studies, by Arnold and Wendland (1983), provides guarantees for optimal convergence rates for integral equations on 2D smooth Jordan curves for odd degree spline approximations. Their work is based on a coercivity assumption of the integral operator with respect to the inner product on Sobolev spaces $H^j(\Gamma)$, $j \geq 1$, and shows that the corresponding collocation methods are convergent, albeit at a slower rate than the associated Galerkin methods. These estimates were further extended to even degree splines on Jordan curves (subject to a specific pseudo-differential form of the integral operator) by Saranen and Wendland (1985). A unified analysis of spline collocation for strongly elliptic boundary integral and integro-differential operators is given by Arnold and Wendland (1985). Recently, there have also been some advances in the analysis of collocation methods for integral equations on higher dimensional spaces, although these have been restricted to biperiodic spaces where the use of Fourier series is available. These include the work by Arens and Rösch (2016) for integral operators with weakly singular kernels whose singularity can be removed by a transformation to polar coordinates.

In these particular settings it was found that the collocation method converges at a slower rate than the corresponding Galerkin method with a similar trial space. This observation has attracted further research by Sloan (1988), Sloan and Wendland (1989) and Chandler and Sloan (1990), who developed the so-called qualocation method (or ‘quadrature-modified collocation method’), which essentially expresses the inner product in the Galerkin equations through a discrete quadrature rule that is specifically chosen to optimise the convergence rate of the overall method. In those results already a linear amount of appropriately weighted oversampling leads to superconvergence at the same rate as the Galerkin method, which is very promising. However, this approach leads to very specific sampling weights that are tuned to the setting of equispaced sampling and spline spaces on smooth domains in two dimensions.

A further level of discretisation from quolocation methods leads to fully discrete methods, for instance Nyström methods as described by Bremer and Gimbutas (2012) and by Hao et al. (2014) (see also (Meier et al., 2000) for integral equations on unbounded domains), which can be applied to Fredholm integral equations of the second kind. In Nyström methods a discrete representation of the integral operator is chosen with great care and typically tailored quadrature schemes are incorporated to ensure good convergence properties of the overall method. This means that the discretisation needs to be adapted to the type of integral equation, and specifically the singular kernel of the integral operator at hand. Convergence analysis of Nyström methods dates back at least to Anselone (1971) and further details in the case of integral operators with continuous kernel can be found in the books by Anselone (1971), Atkinson (1997) and Kress (1999, Chapter 12). The case of integral operators with weakly singular kernels has also been subject of intense research and many relevant cases are now understood from a convergence theory point of view, such as integral equations arising in two-dimensional wave scattering on smooth obstacles (Kress, 1999, §12.3) and on obstacles with corners (Colton and Kress, 1998, §3.5). Finally, scattering by three-dimensional obstacles has also been studied for instance by Ganesh and Graham (2004), Bruno et al. (2013) and Colton and Kress (1998, §3.6).

In this chapter, we raise the question of how one may improve the convergence properties of collocation methods by introducing oversampling without having to choose the collocation points optimally. Intuitively one might expect that, if used appropriately, any information obtained from additional collocation points can help improve the quality of the approximate solution. One should expect that this is the case even when sampling and basis are not perfectly matched, i.e. when collocation points are chosen suboptimally. This observation was recently verified for a number of practical settings involving wave scattering problems (for instance in the work of Barnett and Betcke (2018), Huybrechs and Olteanu (2019) and Gibbs et al. (2020)).

An example of the favourable convergence properties of oversampled collocation is given in the numerical experiments of Fig. 4.1. Here we plot the error of a boundary element collocation method for the Helmholtz equation measured in a field point of the exterior domain (the point is marked in Fig. 4.1b). In the graph, N denotes the number of basis functions and M is the number of collocation points (more details on the meaning of M, N are provided in §4.2). In this example the basis functions are piecewise linear on an equispaced grid and the collocation points are chosen as an equispaced refinement of this grid. We are specifically interested in the convergence rate of the method if we choose M as a fixed function (linear, or quadratic) of N . We observe in the figure that the error for collocation with linear oversampling ($M = 3N$) initially follows the Galerkin rate, and eventually follows the standard asymptotic collocation rate ($M = N$) with a smaller error constant. We will show that faster-than-linear oversampling may result in

higher asymptotic convergence rates, slightly higher even than Galerkin in this example (when $M = N^2$).

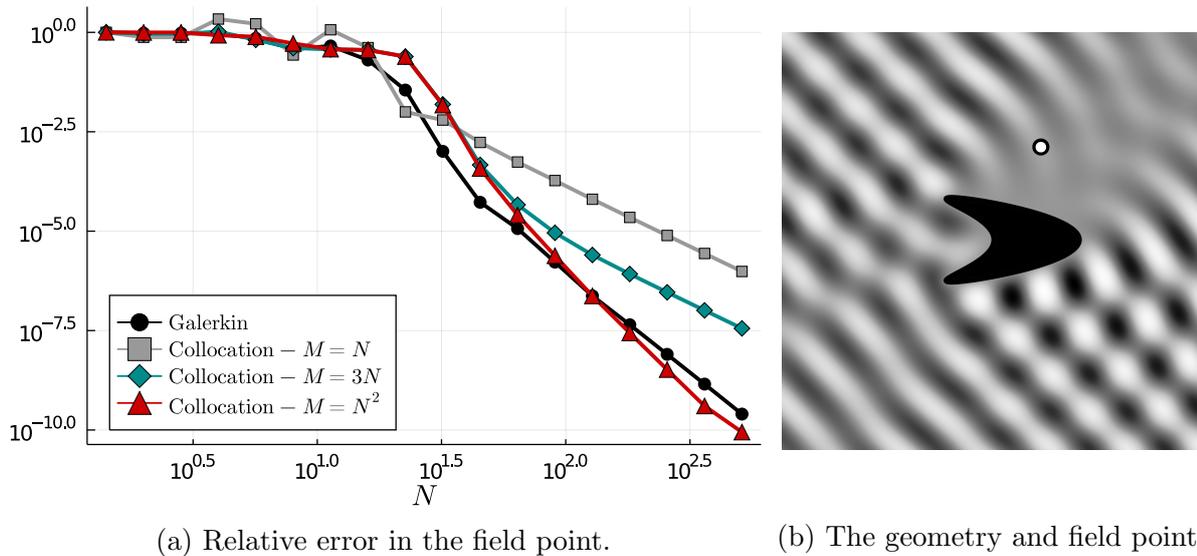


Figure 4.1: The convergence properties of oversampled collocation with equispaced grid points on a smooth scatterer in the single layer potential formulation, with no oversampling ($M = N$), linear oversampling ($M = 3N$), and quadratic oversampling ($M = N^2$). Here M is the number of collocation points, and N is the number of basis functions.

In this chapter, we provide an introductory but rigorous analysis of the least-squares oversampled collocation method, motivated by the favourable performance observed in practice. To the best of our knowledge, this is the first time a rigorous error analysis for oversampled collocation methods is performed in the context of integral equations. Although the results focus on some specific cases for one-dimensional integral equations, we believe the analysis highlights important reasons why least-squares oversampling works (since it provides an approximation to a Bubnov–Galerkin method), and indicates the determining factor for optimal rates of oversampling in the integral equation setting (the quality of approximation in the corresponding trapezoidal rule in the relevant function spaces).

Structure and main results in this chapter

This chapter contains a number of technical results. In order to allow for better clarity, before embarking on the main contents of the chapter we outline both the structure as well as the main ideas of each result in slightly more detail than usual. We introduce the general formulation of the oversampled collocation method together with necessary mathematical assumptions for the later analysis in §4.2. This is followed by a discussion of the principle of oversampling in the context of recent literature in §4.3, and a rigorous convergence analysis of the method in §4.4. In Thm. 4.4.2 we prove convergence of the

oversampled collocation method in the energy space for a broad class of boundary integral operators on smooth domains, subject to using certain regular boundary element spaces that include commonly used basis functions such as piecewise polynomials (Assumption 4.2.2 and Assumption 4.2.3). The convergence is related to a quadrature error estimate in a particular form Eq. (4.31), which can be defined and analysed for arbitrary sets of collocation points. Here, the main point is that the quadrature error may not be small when $M = N$, but it can be made arbitrarily small by increasing M relative to N with minimal assumptions on the points. Though the dimension can be arbitrary, our examples are limited to integral equations on one-dimensional domains. We illustrate the robustness in §4.6 by choosing random collocation points.

In Thm. 4.4.7 we analyse convergence rates and show that the approximate solution to the integral equation computed by an oversampled collocation method may converge to the true solution at optimal rates in a range of Sobolev spaces, so long as we take a sufficient amount of oversampling. This result represents a discrete extension of the so-called Aubin–Nitsche lemma and shows that the lowest possible order of the Sobolev spaces in this range dictates the highest order of convergence of the corresponding error at a field point. This demonstrates superconvergence properties of the oversampled collocation method, though only for domains which are the graph of a smooth parametrisation and with higher than linear oversampling. The latter is, of course, computationally less desirable as it increases the computational complexity of the solver.

Finally, in Thm. 4.4.10, based on earlier results by Sloan (1988), Sloan and Wendland (1989) and Chandler and Sloan (1990) we quantify the impact of the oversampling factor in the computationally more favourable regime of linear oversampling, where $M = JN$ for constant J . The analysis is restricted to integral operators of a specific form, but we note that the integral equations from the single and double layer potential formulations Eqs. (1.16) & (1.17) on smooth domains in two dimensions, using spline basis functions and matching (oversampled) equispaced collocation points, satisfy the assumptions of this theorem. This setting is the most restrictive, but also the most explicit and shows that the error may decay quite rapidly with J . Though the asymptotic convergence rate in the linear oversampling regime remains unchanged compared to standard collocation, the constant involved decays rapidly, for instance like J^{-3} when using linear splines for the single layer integral formulation. This is worthwhile because the cost of the oversampled collocation method increases only linearly with J .

This analysis is followed by a discussion of the results in more general settings in §4.5, including non-equispaced collocation points and Lipschitz domains. These and the aforementioned theoretical results are exhibited on a number of numerical examples in §4.6 and we conclude the chapter with a summary of the main insights in this chapter in §4.7.

4.2 An oversampled collocation method

We first describe the overall set-up and notation before embarking on a more detailed analysis in §4.4. We follow the notation and preliminaries introduced in §1.2.2, but for completeness we repeat a few of the most central assumptions and definitions. We are given an approximation (trial) space S_N , with $\dim S_N = N$, and a domain Γ which we assume to be a plane Jordan curve with a regular parametrisation, i.e. it is the graph of bijective 1-periodic function

$$\begin{aligned} z : [0, 1) &\rightarrow \mathbb{R}^2 \\ t &\mapsto (z_1(t), z_2(t)), \end{aligned} \tag{4.1}$$

which is at least Lipschitz continuous and has $z'(t) \neq 0$ almost everywhere. As mentioned in §1.2.2, in this chapter we shall follow the convention of using plain letters, e.g. x , to denote coordinates in \mathbb{R}^2 (rather than bold-face \mathbf{x} used in the remainder of this thesis). The two main cases of concern in this chapter are when z is a diffeomorphism and when Γ is a polygon (i.e. z is piecewise linear). Here by diffeomorphism we mean that $z : [0, 1) \rightarrow \mathbb{R}^2$ is infinitely differentiable (on the periodic domain $[0, 1)$), z is bijective, and that $z'(t) \neq 0$ for all $t \in [0, 1)$. In this chapter and throughout the thesis we refer to infinitely differentiable functions as ‘smooth’. We consider an integral equation on Γ of the form

$$Vu = f,$$

where $V : H^{s+\alpha}(\Gamma) \rightarrow H^{s-\alpha}(\Gamma)$ is a continuous linear map between two Sobolev spaces on the boundary for some values of $s, \alpha \in \mathbb{R}$. In the integral equation $f \in H^{s-\alpha}(\Gamma)$ is given and $u \in H^{s+\alpha}$ is the unknown function to be determined. In the usual way we call 2α the order of the operator V and its role will become apparent in later sections of this chapter, and the limitations on s are described in §4.2.1. We can rewrite this as an integral equation on the periodic interval $I = [0, 1)$ through the parametrisation z , in the form

$$\tilde{V}\tilde{u} = \tilde{f},$$

where $\tilde{f}(t) = f(z(t))$, $\tilde{u}(t) = u(z(t))z'(t)$ and the action of \tilde{V} on a function $\chi : I \rightarrow \mathbb{C}$ is defined by

$$\tilde{V}\chi(\tau) = V\left(\frac{\chi \circ z^{-1}}{z' \circ z^{-1}}\right)(z(x)), \quad \forall x \in I. \tag{4.2}$$

Note this transformation is well-defined because, by assumption, $z'(t) \neq 0$ almost everywhere. In the case when z is a diffeomorphism Sobolev-regularity properties are preserved under the map z in the sense of Eq. (1.12) and thus \tilde{V} is continuous as a linear map

$\tilde{V} : H^{s+\alpha}(I) \rightarrow H^{s-\alpha}(I)$. This means we can reformulate the integral equation on Γ as an integral equation on $[0, 1)$ in a way which preserves the continuity properties of V regardless of the value of s, α . When z is just Lipschitz (specifically when z is bijective, Lipschitz and $z' \neq 0$ a.e.), then the continuity properties of V are also preserved under this reformulation on $I = [0, 1)$ as long as $|s \pm \alpha| \leq 1$ (cf. Chandler-Wilde, Graham, Langdon and Spence, 2012, Appendix A.3).

Therefore we may, in this chapter, focus on the numerical solution of integral equations on the periodic interval $I = [0, 1)$ in the form (we drop the tilde for simplicity of notation)

$$Vu = f, \quad (4.3)$$

where $V : H^{s+\alpha}(I) \rightarrow H^{s-\alpha}(I)$. Assuming $s - \alpha > 1/2$, such that $H^{s-\alpha}(I) \subset C(I)$, our method of solution is a ‘least-squares oversampled collocation method’, whereby we proceed as follows. Firstly, we choose a sequence of approximation spaces $S_N, N = 1, 2, \dots$, with $\dim S_N = N$ and with the following approximation property: $\forall \xi \in H^{s+\alpha}(I)$ there exists a sequence $\xi_N, N = 1, 2, \dots$, with $a_N \subset S_N$ such that

$$\lim_{N \rightarrow \infty} \xi_N = \xi \quad \text{in } H^{s-\alpha}(I).$$

Then, for each value of N , we choose $M = M(N) \geq N$ distinct collocation points $x_m \in I, m = 1, \dots, M$, and define an approximation $u_N^{(M)}$ to the true solution u in the space S_N as follows. We take a basis $\{\phi_n\}_{n=1}^N$ for S_N (i.e. a spanning set of linearly independent functions) and we expand

$$u_N^{(M)} = \sum_{n=1}^N a_n \phi_n, \quad (4.4)$$

after which the collocation conditions

$$\sum_{n=1}^N a_n \phi_n(x_m) = Vu_N^{(M)}(x_m) = f(x_m), \quad m = 1, \dots, M,$$

provide an overdetermined $M \times N$ linear system for $\mathbf{a} = (a_1, a_2, \dots, a_N)^T$. Motivated by recent results in approximation theory (Adcock and Huybrechs, 2019, 2020), we consider a weighted least-squares solution to this system, such that

$$(\mathbf{G}_{M,N}^* \mathbf{W}_M \mathbf{G}_{M,N}) \mathbf{a} = \mathbf{G}_{M,N}^* \mathbf{f}, \quad (4.5)$$

where $\mathbf{f} = (f(x_1), f(x_2), \dots, f(x_M))^T$, $\mathbf{G}_{M,N}$ is an $M \times N$ matrix with entries given by

$$(\mathbf{G}_{M,N})_{m,n} = V\phi_n(x_m), \quad m = 1, \dots, M, n = 1, \dots, N,$$

and $\mathbf{G}_{M,N}^*$ denotes the conjugate adjoint of $\mathbf{G}_{M,N}$, and \mathbf{W}_M is an $M \times M$ diagonal matrix with entries corresponding to distances between sampling points, i.e.

$$(\mathbf{W}_M)_m = \frac{|x_{m+1} - x_{m-1}|}{2}, \quad m = 1, \dots, M.$$

In the above sum it is understood that $x_{N+1} = x_1$.

This provides us with a sequence $\{u_N^{(M)}\}_{N \in \mathbb{N}}$ (recall that $M = M(N)$ is a function of N) of approximations to the true solution u , and in this chapter we are specifically interested in the convergence properties of this sequence as $N \rightarrow \infty$, when $M = M(N) > N$. Note, this method reduces to the standard collocation method when $M = N$.

4.2.1 Mathematical assumptions

We focus our attention on integral operators V with the following mapping properties:

- V is a continuous linear map

$$V : H^{s+\alpha}(I) \rightarrow H^{s-\alpha}(I) \tag{4.6}$$

for some $\alpha \in \mathbb{R}$ and any $s \in \mathbb{R}$ with $|s| \leq s_0$, some fixed constant s_0 .

- The inverse of V is a well-defined continuous linear map

$$V^{-1} : H^{s-\alpha}(I) \rightarrow H^{s+\alpha}(I) \tag{4.7}$$

for any $s \in \mathbb{R}$ with $|s| \leq s_1$, some fixed constant s_1 .

In the usual way we call 2α the order of the operator. For simplicity we write in the following analysis $H^t = H^t(I)$.

In particular, on $[0, 1)$, we use the following definitions of the Sobolev norm for $f \in L^2$ as we described in §1.2.2:

$$\|f\|_s = \|f\|_{H^s([0,1))} := \left(\sum_{m \in \mathbb{Z}} [m]^{2s} |\hat{f}_m|^2 \right)^{1/2}, \tag{4.8}$$

where

$$\hat{f}_m := \int_0^1 e^{-2\pi imt} f(t) dt, \text{ and } [m] := \begin{cases} 1, & \text{if } m = 0, \\ |m|, & \text{if } m \neq 0. \end{cases} \quad (4.9)$$

Note that the Fourier coefficients of $f \in L^2$ are well-defined since on compact domains $L^2 \subset L^1$. Furthermore, we use the following definition of the L^2 -duality pairing for $f \in H^s, g \in H^{-s} \cap L^2$:

$$\langle f, g \rangle := \sum_{m \in \mathbb{Z}} \overline{\hat{f}_m} \hat{g}_m = \int_0^1 \overline{f(x)} g(x) dx.$$

We also use the following definition for the H^s -inner product which is consistent with the definitions of the Sobolev norms in Eq. (4.8):

$$\langle f, g \rangle_s = \langle f, g \rangle_{H^s([0,1])} = \sum_{m \in \mathbb{Z}} [m]^{2s} \overline{\hat{f}_m} \hat{g}_m. \quad (4.10)$$

By density of L^2 in H^s for any $s \in \mathbb{R}$ these definitions of the Sobolev norms, inner products and pairing extend by continuity to arbitrary elements of H^s , and H^{-s} respectively. Note that in §4.4.2 we will also make use of the expression

$$\|f\|_s = \sup_{\substack{g \in C^\infty \\ \|g\|_{-s}=1}} |\langle f, g \rangle|, \quad (4.11)$$

which follows immediately from Eq. (4.8). Moreover, it will be helpful to introduce the following notation for the operator norm of a map $V : H^s \rightarrow H^t$, with $s, t \in \mathbb{R}$:

$$\|V\|_{s \rightarrow t} = \sup_{\|f\|_s=1} \|Vf\|_t. \quad (4.12)$$

Example 4.2.1 (Integral formulations of the Helmholtz equation, see Eqs. (1.16) & (1.17)). *In §1.2.2 we already introduced the single- and double-layer formulations of the exterior Dirichlet problem of the Helmholtz equation on a Lipschitz domain Ω_+ . Suppose $\Gamma = \partial\Omega_+$ is the image of a map z as described in Eq. (4.1). Let us follow the notation of §1.2.2 and denote by \mathcal{S}, \mathcal{D} the single and double layer operators which are, away from the resonant frequencies mentioned in Thm. 1.2.9, such that the following maps are continuous*

isomorphisms of the relevant spaces

$$\begin{aligned}\mathcal{S} &: H^{s-\frac{1}{2}}(\Gamma) \rightarrow H^{s+\frac{1}{2}}(\Gamma) \\ \frac{1}{2}\mathcal{I} + \mathcal{D} &: H^{s+\frac{1}{2}}(\Gamma) \rightarrow H^{s+\frac{1}{2}}(\Gamma),\end{aligned}$$

for all $s \in \mathbb{R}$ if Ω_+ is smooth, and for all $|s| \leq 1/2$ if Ω_+ is Lipschitz. Thus, away from resonant frequencies, the following holds:

- If z is bijective, infinitely differentiable and with $z'(t) \neq 0, \forall t \in I$, then the integral equations Eqs. (1.16) & (1.17) can be reformulated using the parametrisation z on the domain $I = [0, 1)$ as described in Eq. (4.2) and the resulting integral equations on I are of the form Eq. (4.3) with an integral operator V satisfying assumptions Eqs. (4.6) & (4.7) for all $s \in \mathbb{R}$ with orders $2\alpha = -1$ (for the single layer formulation) and $2\alpha = 0$ (for the double layer formulation) respectively.
- If z is bijective, Lipschitz and with $z' \neq 0$ a.e., then the single layer integral equation Eq. (1.16) can be reformulated on I as above, and the resulting mapped integral operator V satisfies assumptions Eqs. (4.6) & (4.7) for $|s| \leq 1/2$ with $2\alpha = -1$.

We suppose that the sequence of approximation spaces $S_N, N = 1, 2, \dots$, satisfies the following properties: There is a sequence of ‘mesh sizes’ $h = h(N) > 0$ with $\lim_{N \rightarrow \infty} h(N) = 0$ and non-negative real numbers $l, m > 0$ such that $S_N \subset H^{m+1/2}$ and the following assumptions are satisfied:

Assumption 4.2.2 (Approximation property). For fixed $t \leq s \leq l$ and $t < m + \frac{1}{2}$ there exists a constant c such that for any $v \in H^s$, there is a sequence $\chi_N \in S_N$ such that

$$\|v - \chi_N\|_t \leq ch^{s-t}\|v\|_s, \quad \forall N \in \mathbb{N}. \quad (4.13)$$

Moreover for any fixed $\sigma < m + 1/2$ there is such a sequence χ_N which satisfies Eq. (4.13) for all $t > \sigma$, i.e. which can be chosen independently of t .

Assumption 4.2.3 (Inverse property). For given $t \leq s \leq m + \frac{1}{2}$, there exists a constant C such that for all $\chi_N \in S_N$,

$$\|\chi_N\|_s \leq Ch^{t-s}\|\chi_N\|_t, \quad \forall N \in \mathbb{N}.$$

These spaces are typically called (l, m) -regular boundary element spaces with the inverse assumption in the sense of Babuška and Aziz (1972, Section 4.1). Note we followed the convention by Hsiao et al. (2017, p. 38) for one-dimensional domains for the upper bounds $m + 1/2$ on the Sobolev indices in the above assumptions, which is different from the case of approximation spaces for two-dimensional domains to reflect the difference in

Sobolev-regularity of piecewise polynomial functions on subsets of \mathbb{R} and \mathbb{R}^2 . This has no further effect on the remainder of this chapter, but is convenient for the following example.

Example 4.2.4 (Spline spaces on ρ -quasiuniform mesh).

Let $\Delta_N = \{0 = x_1 < x_2 < \dots < x_N < 1\}$, $N = 1, 2, \dots$, be a sequence of ρ -quasiuniform meshes for some $\rho > 0$, i.e. such that for all $N \geq 1$:

$$\max_{1 \leq j \leq N} |x_{j+1} - x_j| \leq \rho \min_{1 \leq j \leq N} |x_{j+1} - x_j|,$$

where it is understood that $x_{N+1} = x_1$. For each N let S_N be the space of all 1-periodic, $d - 1$ times continuously differentiable piecewise polynomials of degree $d \in \mathbb{N}$ subordinate to the partition Δ_N . Then the sequence $S_N, N = 1, \dots$, satisfies all the aforementioned assumptions on our approximation spaces with $h(N) = N^{-1}$, $l = d + 1$ and $m = d$. We call such a sequence of spaces henceforth ‘smoothest splines of degree d ’.

A proof of these properties for splines on uniform meshes can be found in (Saranen and Vainikko, 2013, §13.3) and an extension to quasiuniform meshes is given in (Arnold and Wendland, 1983, pp. 352-353 & pp. 359-360).

4.2.2 From least-squares to a discrete Bubnov–Galerkin method

Oversampled collocation leads to a rectangular linear system and this system is solved in a least-squares sense, recall (4.5). In view of the chosen weights, we will see that these normal equations have a continuous limit as $M \rightarrow \infty$ for fixed N . However, that limit differs from the classical Galerkin method of the same integral equation. The latter leads to the orthogonality conditions

$$\langle \chi_N, V u_N \rangle = \langle \chi_N, f \rangle, \quad \forall \chi_N \in S_N. \quad (4.14)$$

The central observation in the following analysis is that the least-squares system Eq. (4.5) amounts to a discrete Bubnov–Galerkin method instead, in the following sense. Let $\Delta_M = \{0 \leq x_1 < x_2 < \dots < x_M < 1\}$ be the collocation points and let $u_N^{(M)} \in S_N$ be the least-squares collocation approximation to the true solution u in the sense of Eq. (4.4)–(4.5), then we note that Eq. (4.5) is equivalent to

$$\left\langle V \chi_N, V u_N^{(M)} \right\rangle_M = \langle V \chi_N, V u \rangle_M, \quad \forall \chi_N \in S_N. \quad (4.15)$$

Here we defined the discrete inner product to be

$$\langle f, g \rangle_M = \sum_{j=1}^M \frac{|x_{j+1} - x_{j-1}|}{2} \overline{f(x_j)} g(x_j), \quad (4.16)$$

where it is understood that $|x_{N+1} - x_{N-1}| = x_1 + 1 - x_{N-1}$. The ‘discrete orthogonality condition’ Eq. (4.15) plays a central role in the analysis of §4.4. To see the equivalence of Eq. (4.5) and Eq. (4.15) note that the linear system Eq. (4.5) can be written line-for-line as

$$\sum_{m=1}^M \sum_{n=1}^N (\mathbf{G}_{M,N}^*)_{j,m} (\mathbf{W}_M)_m (\mathbf{G}_{M,N})_{m,n} a_n = \sum_{m=1}^M (\mathbf{G}_{M,N}^*)_{j,m} (\mathbf{W}_M)_m f_m, \quad j = 1, \dots, N.$$

Substituting the entries of $\mathbf{G}_{M,N}, \mathbf{W}_M$ yields the equivalent expression

$$\begin{aligned} \sum_{m=1}^M \sum_{n=1}^N \overline{V\phi_j(x_m)} \frac{|x_{m-1} - x_{m+1}|}{2} V\phi_n(x_m) a_n &= \sum_{m=1}^M \overline{V\phi_j(x_m)} \frac{|x_{m-1} - x_{m+1}|}{2} f_m, \quad j = 1, \dots, N, \\ \iff \left\langle V\phi_j, Vu_N^{(M)} \right\rangle_M &= \langle V\phi_j, Vu \rangle_M, \quad j = 1, \dots, N. \end{aligned}$$

Recalling that $\{\phi_n\}_{n=1}^N$ is a basis for S_N we can immediately conclude that the final line is indeed equivalent to Eq. (4.15).

Remark 4.2.5. *Note that, compared to the classical Galerkin orthogonality conditions (4.14), the discrete conditions Eq. (4.15) feature an additional integral operator V in the first argument of the inner product. This is the case both in the left hand side and in the right hand side of the equation: the integral equation is projected using the basis $\{V\chi_N\}_{\chi_N \in S_N}$ rather than $\{\chi_N\}_{\chi_N \in S_N}$. If V has negative order, then the former is smoother than the latter and that underlies some of the differences in convergence rates between Galerkin and the large oversampling limit of the collocation method.*

Remark 4.2.6. *One might also wish to consider a discretisation of the standard Galerkin method Eq. (4.14), which would result in the discrete orthogonality conditions*

$$\left\langle \chi_N, Vu_N^{(M)} \right\rangle_M = \langle \chi_N, Vu \rangle_M, \quad \forall \chi_N \in S_N. \quad (4.17)$$

We call this the ‘modified oversampled collocation method’ because instead of weighted normal equations Eq. (4.5) it can be formulated as

$$(\mathcal{B}_{M,N}^* \mathbf{W}_M \mathbf{G}_{M,N}) \mathbf{a} = \mathcal{B}_{M,N}^* \mathbf{W}_M \mathbf{f},$$

where

$$\begin{aligned} \mathbf{f} &= (f(x_m))_{m=1}^M, & \mathbf{G}_{M,N} &= (V\phi_n(x_m))_{m=1,n=1}^{M,N}, \\ \mathcal{B}_{M,N} &= (\phi_n(x_m))_{m=1,n=1}^{M,N}, & \mathbf{W}_M &= \frac{|x_{m+1} - x_{m-1}|}{2} \end{aligned}$$

For this method, most of the analysis in the following sections can be conducted in a similar

fashion. However, in this case the assumption of invertibility of V is no longer sufficient to guarantee uniform ellipticity of the corresponding discrete forms and it would be necessary to impose a discrete inf-sup assumption on the operator V (similar to what would be required for the guaranteed convergence of the continuous Galerkin method Eq. (4.14)). In this thesis our attention shall be focused on the least squares oversampled collocation method and we will not discuss the modified form in detail.

4.3 Oversampling in collocation methods and in approximation on Hilbert spaces

In this section, we provide a brief discussion concerning the connection between our approach for least-squares oversampling in collocation methods and least-squares oversampling for function reconstruction on Hilbert spaces as described by Adcock and Hansen (2012). It turns out that the underlying principle, which allows oversampling to improve the quality of approximation, is closely related between these two cases. In particular, the mechanism underpinning the success of oversampling in both methods is the following:

Observation 4.3.1. *Least-squares oversampling leads to improved approximation because the discrete least-squares system constitutes an approximation to an exact continuous orthogonality condition on the relevant Hilbert space, see Eqs. (4.21) & (4.22) below.*

4.3.1 Stable function approximation on Hilbert spaces

Let us begin by reviewing the method for stable function approximation on Hilbert spaces described by Adcock and Hansen (2012): Let $T = \{\psi_n\}_{n \in \mathbb{N}}, S = \{\phi_n\}_{n \in \mathbb{N}}$ be two distinct Hilbert bases (i.e. complete linearly independent orthonormal sets) for \mathcal{H} and let $f \in \mathcal{H}$. Our goal is to find the best approximation of f in $S_N = \text{span}\{\phi_n\}_{n=1}^N$ from observations in $T_M = \text{span}\{\psi_m\}_{m=1}^M$, i.e. from $\{\langle \psi_m, f \rangle\}_{m=1}^M$, where $M \geq N$ and $\langle \cdot, \cdot \rangle$ denotes the inner product on \mathcal{H} . Given no prior information about the relation between S and T the samples $\{\langle \psi_m, f \rangle\}_{m=1}^M$ may be suboptimal, and it is challenging to find an optimal strategy for approximating f in this way.

The idea presented by Adcock and Hansen (2012) is to expand $f_N^{(M)} = \sum_{n=1}^N \hat{f}_n^{(M)} \phi_n$ and to use a least-squares solution of an overdetermined linear system of equations in order to determine the coefficients $\hat{f}_n^{(M)}$. The overdetermined linear system is

$$\mathbf{G}_{M,N} \hat{\mathbf{f}}^{(M)} = \mathbf{b}, \quad (4.18)$$

where $\hat{\mathbf{f}}^{(M)} = \left(\hat{f}_1^{(M)}, \dots, \hat{f}_N^{(M)} \right)^T$ and $(\mathbf{G}_{M,N})_{mn} = \langle \psi_m, \phi_n \rangle$, $b_m = \langle \psi_m, f \rangle$, for $m =$

$1, \dots, M, n = 1, \dots, N$. The least-squares solution to Eq. (4.18) is

$$\underbrace{\mathbf{G}_{M,N}^* \mathbf{G}_{M,N}}_{=(\sum_{m=1}^M \langle \phi_i, \psi_m \rangle \langle \psi_m, \phi_j \rangle)} \hat{\mathbf{f}}^{(M)} = \underbrace{\mathbf{G}_{M,N}^* \mathbf{b}}_{=(\sum_{m=1}^M \langle \phi_i, \psi_m \rangle \langle \psi_m, f \rangle)}, \quad (4.19)$$

By Parseval's identity we know that for any $g, h \in \mathcal{H}$

$$\lim_{M \rightarrow \infty} \sum_{m=1}^M \langle g, \psi_m \rangle \langle \psi_m, h \rangle = \langle g, h \rangle. \quad (4.20)$$

Thus the $N \times N$ system in Eq. (4.19) converges, as $M \rightarrow \infty$, to the following limit:

$$\sum_{n=1}^N \langle \phi_j, \phi_n \rangle \hat{f}_n^{(\infty)} = \langle \phi_j, f \rangle, \quad j = 1, \dots, N. \quad (4.21)$$

This is precisely the condition for the orthogonal projection of f onto S_N , i.e.

$$f_N^{(\infty)} = \operatorname{argmin}_{\chi_N \in S_N} \|f - \chi_N\|.$$

This means, by taking $M = M(N)$ sufficiently large, we can always recover the best approximation, even with suboptimal samples in a different basis T . Note, as indicated in Observation 4.3.1, the central reason why this works is because of Parseval's identity Eq. (4.20) which ensures that the least-squares system approximates the appropriate inner product on \mathcal{H} .

4.3.2 Least-squares oversampled collocation method

A similar mechanism ensures the success of least-squares oversampling for collocation methods as introduced in Eq. (4.5). Indeed, recall from Eq. (4.15) that the solution $u_N^{(M)} \in S_N$ of the least-squares collocation method is uniquely characterised by the following (discrete) orthogonality condition

$$\left\langle V \chi_N, V u_N^{(M)} \right\rangle_M = \langle V \chi_N, V u \rangle_M, \quad \forall \chi_N \in S_N,$$

where the discrete inner product $\langle \cdot, \cdot \rangle_M$ is as defined in Eq. (4.16). The main insight is that under a suitably distributed choice of collocation points Δ_M we have an error estimate of the form Eq. (4.26) below, which ensures that these discrete inner products approximate the continuous L^2 -inner product as $M \rightarrow \infty$. In particular, for any fixed N we will show that, under suitable assumptions on Δ_M , we have $\lim_{M \rightarrow \infty} u_N^{(M)} = u_N^{(\infty)}$ for

some $u_N^{(\infty)} \in S_N$ which satisfies

$$\left\langle V\chi_N, Vu_N^{(\infty)} \right\rangle_{L^2} = \langle V\chi_N, Vu \rangle_{L^2}, \quad \forall \chi_N \in S_N. \quad (4.22)$$

These are the Bubnov–Galerkin orthogonality conditions and we know that these yield a quasi-optimal approximation to u in S_N with respect to the norm on the energy space of V :

$$\|u - u_N^{(\infty)}\|_{2\alpha} \leq C \inf_{\chi_N \in S_N} \|u - \chi_N\|_{2\alpha}.$$

Thus the reason oversampling works for collocation methods is that the trapezoidal rule Eq. (4.26) ensures that the least-squares system approximates the L^2 -inner product in the limit $M \rightarrow \infty$. A central question we study in this chapter is how large $M = M(N)$ needs to be, as N increases, in order to ensure $u_N^{(M)}$ converges to u at optimal rates in N .

4.4 Convergence analysis of the oversampled collocation method

As mentioned above the central idea in the analysis is to regard the oversampled collocation method as a discrete version of a Bubnov–Galerkin method, by which we mean the approximation $u_N \in S_N$ which is defined through the continuous orthogonality conditions

$$\langle V\chi_N, Vu_N \rangle_{L^2} = \langle V\chi_N, f \rangle_{L^2}, \quad \forall \chi_N \in S_N. \quad (4.23)$$

A combination of Strang-type estimates (cf. Ciarlet, 2002, Thms. 4.1.1 & 4.2.2) and the error of the trapezoidal rule for the inner product allows us to study the convergence rates of our methods. We begin by focusing on the case when $u, f \in C_{\text{per}}^\infty(I)$ and when assumptions Eqs. (4.6) & (4.7) hold for any $s \in \mathbb{R}$ i.e. when

$$V : H^{s+\alpha} \rightarrow H^{s-\alpha} \quad (4.24)$$

is a continuous isomorphism for all $s \in \mathbb{R}$.

In the case of the boundary integral equations Eqs. (1.16) & (1.17) we have already seen a useful way of assessing the error in Fig. 4.1. Specifically, in the potential formulation of the exterior Dirichlet problem, the solution $\phi : \Omega_+ \rightarrow \mathbb{C}$ is expressed in the form

$$\phi(x) = \int_0^1 k(x, z(t))u(t) dt, \quad x \in \Omega,$$

where $k(x, y)$ equals either $G(x, y)$ or $\partial_{n_y}G(x, y)$ with $G(x, y)$ being the Green's function

of the Helmholtz equation. Note in the above $u : I \rightarrow \mathbb{C}$ denotes the true solution to the integral equation formulated on the boundary as discussed in Example 4.2.1. The kernel $k(x, y)$ is a smooth function away from $x = y$, which means we can estimate, for any $s \in \mathbb{R}$,

$$\begin{aligned} |\phi_N(x) - \phi(x)| &= \left| \int_0^1 k(x, z(t)) (u_N(t) - u(t)) z'(t) dt \right| = \left| \left\langle \overline{k(x, z(\cdot))z'(\cdot)}, u_N - u \right\rangle_{L^2(I)} \right| \\ &\leq \|k(x, z(\cdot))z'(\cdot)\|_s \|u - u_N\|_{-s}, \quad x \in \mathbb{R}^2 \setminus \Gamma. \end{aligned} \quad (4.25)$$

Since $y \mapsto k(x, y) \in C^\infty(\Gamma)$ for all $x \in \mathbb{R}^2 \setminus \Gamma$ the convergence rate of the approximation $\phi_N(x)$ to $\phi(x)$ is governed by the fastest convergence of $\|u - u_N\|_{-s}$ in any Sobolev norm for $s \in \mathbb{R}$. We shall see in the coming two sections that optimal convergence rates can be obtained in low order Sobolev spaces for sufficient amounts of oversampling. It will later become apparent in §4.5.2 that the smoothness of the domain boundary affects the range of Sobolev spaces one can consider, hence limiting the fastest convergence rates that can be achieved.

4.4.1 Strang estimate for convergence on energy space

The first step in the analysis is to show that already a small amount of oversampling is sufficient to guarantee convergence of the least-squares oversampled collocation method on the energy space $H^{2\alpha}$. As before, we let $\Delta_M = \{0 \leq x_1 < \dots < x_M < 1\}$ be the collocation points. As the dimension N of the approximation space $S_N = S_h^{l,m}$ increases, we also increase the number of collocation points $M = M(N)$ in a predefined way and we assume that there is an error estimate for the proximity of the discrete inner product to a continuous L^2 pairing in the following form. For some $r, s > 0$ there is a sequence of constants $\mathcal{E}_{r,s}(\Delta_M) > 0$, such that for any $f, g \in H^{\max\{s,r\}}$ we have

$$|\langle g, f \rangle - \langle g, f \rangle_M| \leq \mathcal{E}_{r,s}(\Delta_M) (\|f\|_r \|g\|_s + \|f\|_s \|g\|_r). \quad (4.26)$$

We will derive such error estimates in a number of settings further on in Lemma 4.4.5 and Lemma 4.5.1.

Proposition 4.4.1. *If Eq. (4.26) holds for s, r with $\max\{r, s\} < m + 1/2 - 2\alpha$, and if $\Delta_M = \Delta_{M(N)}$ is chosen such that*

$$\lim_{N \rightarrow \infty} \mathcal{E}_{r,s}(\Delta_M) h^{-(r+s)} = 0,$$

where $h = h(N)$, then the bilinear forms $(f, g) \mapsto \langle Vf, Vg \rangle_M$ are uniformly S_N -elliptic,

for $N \geq N_0$, some $N_0 > 0$, meaning they satisfy

$$\gamma \|\chi_N\|_{2\alpha}^2 \leq \langle V\chi_N, V\chi_N \rangle_M, \quad \forall \chi_N \in S_N,$$

for some positive constant γ independent of $N \geq N_0$.

Proof. By Eq. (4.24) we know that for any $\chi_N \in S_N$

$$\begin{aligned} \|\chi_N\|_{2\alpha}^2 &\leq \|V^{-1}\|_{0 \rightarrow 2\alpha}^2 (|\langle V\chi_N, V\chi_N \rangle_M| + |\langle V\chi_N, V\chi_N \rangle - \langle V\chi_N, V\chi_N \rangle_M|) \\ &\leq \|V^{-1}\|_{0 \rightarrow 2\alpha}^2 (|\langle V\chi_N, V\chi_N \rangle_M| \\ &\quad + 2\mathcal{E}_{r,s}(\Delta_M) \|V\|_{r+2\alpha \rightarrow r} \|V\|_{s+2\alpha \rightarrow s} \|\chi_N\|_{r+2\alpha} \|\chi_N\|_{s+2\alpha}), \end{aligned}$$

where in the above we used the notation $\|\cdot\|_{0 \rightarrow 2\alpha}$ for the operator norm of a linear map from H^0 to $H^{2\alpha}$ as introduced in (4.12). By the inverse property of S_N (Assumption 4.2.3) we have, whenever $\max\{r, s\} < m + 1/2 - 2\alpha$,

$$\|\chi_N\|_{r+2\alpha} \|\chi_N\|_{s+2\alpha} \leq Ch^{-(r+s)} \|\chi_N\|_{2\alpha}^2,$$

and therefore

$$\left(1 - \tilde{C}\mathcal{E}_{r,s}(\Delta_M)h^{-(r+s)}\right) \|\chi_N\|_{2\alpha}^2 \leq \|V^{-1}\|_{0 \rightarrow 2\alpha}^2 |\langle V\chi_N, V\chi_N \rangle_M|.$$

Hence, since $\lim_{N \rightarrow \infty} \mathcal{E}_{r,s}(\Delta_M)h^{-(r+s)} = 0$, the result follows. \square

Theorem 4.4.2 (Strang-type bound for convergence in $H^{2\alpha}$). *Suppose our trial spaces S_N , are (l, m) -regular boundary element spaces satisfying the inverse property according to Assumptions 4.2.2 & 4.2.3, for some $l, m > 0$. If Eq. (4.26) holds for s, r with $\max\{r, s\} < m + 1/2 - 2\alpha$, and $\Delta_M = \Delta_{M(N)}$ is chosen such that*

$$\lim_{N \rightarrow \infty} \mathcal{E}_{r,s}(\Delta_M)h^{-(r+s)} = 0,$$

then there are constants $N_0, C > 0$ independent of M, N, u such that, for all $N \geq N_0$,

$$\|u_N^{(M)} - u\|_{2\alpha} \leq Ch^{l-2\alpha} \|u\|_l.$$

This means the optimal convergence rate in $H^{2\alpha}$ is achieved.

Proof. We begin with the following estimate on the energy space $H^{2\alpha}$:

$$\|u_N^{(M)} - u\|_{2\alpha} \leq \inf_{\chi_N \in S_N} \left\{ \|u - \chi_N\|_{2\alpha} + \|u_N^{(M)} - \chi_N\|_{2\alpha} \right\}.$$

We can estimate the second term using Prop. 4.4.1 and the discrete orthogonality condition

Eq. (4.15) as follows

$$\begin{aligned}
\gamma \|u_N^{(M)} - \chi_N\|_{2\alpha}^2 &\leq \left| \left\langle V(u_N^{(M)} - \chi_N), V(u_N^{(M)} - \chi_N) \right\rangle_M \right| \\
&= \left| \left\langle V(u - \chi_N), V(u_N^{(M)} - \chi_N) \right\rangle_M \right| \\
&\leq \left| \left\langle V(u - \chi_N), V(u_N^{(M)} - \chi_N) \right\rangle \right| \\
&\quad + \left| \left\langle V(u - \chi_N), V(u_N^{(M)} - \chi_N) \right\rangle - \left\langle V(u - \chi_N), V(u_N^{(M)} - \chi_N) \right\rangle_M \right| \\
&\leq \|V\|_{2\alpha \rightarrow 0}^2 \|u - \chi_N\|_{2\alpha} \|u_N^{(M)} - \chi_N\|_{2\alpha} \\
&\quad + \mathcal{E}_{r,s}(\Delta_M) \|V\|_{r+2\alpha \rightarrow r} \|V\|_{s+2\alpha \rightarrow s} \|u - \chi_N\|_{r+2\alpha} \|u_N^{(M)} - \chi_N\|_{s+2\alpha} \\
&\quad + \mathcal{E}_{r,s}(\Delta_M) \|V\|_{r+2\alpha \rightarrow r} \|V\|_{s+2\alpha \rightarrow s} \|u - \chi_N\|_{s+2\alpha} \|u_N^{(M)} - \chi_N\|_{r+2\alpha}
\end{aligned} \tag{4.27}$$

Using the inverse property of S_N (Assumption 4.2.3) we find if $\max\{r, s\} < M + 1/2 - 2\alpha$,

$$\tilde{\gamma} \|u_N^{(M)} - \chi_N\|_{2\alpha} \leq \|u - \chi_N\|_{2\alpha} + \mathcal{E}_{r,s}(\Delta_M) h^{-s} \|u - \chi_N\|_{r+2\alpha} + \mathcal{E}_{r,s}(\Delta_M) h^{-r} \|u - \chi_N\|_{s+2\alpha}$$

for some $\tilde{\gamma} > 0$ independent of N, Δ_M, u . Therefore, since $\lim_{N \rightarrow \infty} \mathcal{E}_{r,s}(\Delta_M) h^{-(r+s)} = 0$, we have, for N sufficiently large,

$$\tilde{\gamma} \|u_N^{(M)} - \chi_N\|_{2\alpha} \leq \|u - \chi_N\|_{2\alpha} + h^r \|u - \chi_N\|_{r+2\alpha} + h^s \|u - \chi_N\|_{s+2\alpha}.$$

By the approximation property of S_N (Assumption 4.2.2) the result follows. \square

We note that Thm. 4.4.2 implies optimal convergence in Sobolev norms that are of higher order than the energy space by the following argument.

Corollary 4.4.3. *Let $S_N, \Delta_M = \Delta_{M(N)}$ satisfy the assumptions of Thm. 4.4.2, then for all $2\alpha < t < m + 1/2$ we have optimal convergence in the sense that there are constants $C, N_0 > 0$ independent of N, Δ_M, u , such that, for all $N \geq N_0$,*

$$\|u_N^{(M)} - u\|_t \leq Ch^{l-t} \|u\|_l.$$

Proof. Let $2\alpha < t < m + 1/2$, then

$$\begin{aligned}
\|u_N^{(M)} - u\|_t &\leq \|u - \psi_N\|_t + \|\psi_N - u_N^{(M)}\|_t \\
&\leq \|u - \psi_N\|_t + Ch^{2\alpha-t} \|\psi_N - u_N^{(M)}\|_{2\alpha} \\
&\leq \|u - \psi_N\|_t + Ch^{2\alpha-t} \left(\|\psi_N - u\|_{2\alpha} + \|u - u_N^{(M)}\|_{2\alpha} \right),
\end{aligned}$$

and the result follows by Thm. 4.4.2 and by the uniform approximation property in Assumption 4.2.2. \square

Remark 4.4.4. *One can see that the statements in Thm. 4.4.2 and Prop. 4.4.1 are in fact true for much more general settings, including the case of two-dimensional boundary integral equations, i.e. when $\dim \Gamma = 2$, as long as appropriate error estimates for the discrete inner product similar to (4.26) are available. However, for the purpose of this thesis we shall remain in the two-dimensional setting.*

Although the above statements are phrased in a general form we can use them to make concrete predictions. We begin by considering equispaced collocation points $|x_{j+1} - x_j| = 1/M$, for which we have the following error estimate for the L^2 -inner product as in Eq. (4.26):

Lemma 4.4.5 (Error in discrete L^2 inner product for equispaced sampling). *Let $\Delta_M = \{x_m = \tilde{x} + j/M\}_{m=1}^M$ be a set of equispaced collocation points (where it is understood $x + 1 \equiv x$) and fix $r > 1/2$. For $f, g \in H^r([0, 1])$ let*

$$\langle f, g \rangle_M = \frac{1}{M} \sum_{m=1}^M \overline{f(x_m)} g(x_m),$$

then there is a constant $C_{r,s} > 0$ independent of f, g such that, for any $r \geq s > 1/2$,

$$|\langle f, g \rangle - \langle f, g \rangle_M| \leq C_{r,s} M^{-r} \left(\|f\|_r \|g\|_s + \|f\|_s \|g\|_r \right).$$

Proof. For completeness the proof is given in Appendix 4.A. □

If we now choose S_N to be a sequence of smoothest spline spaces of degree d on a quasi-equispaced mesh (cf. Example 4.2.4) we find the following convergence result for the oversampled collocation method. Note in particular that here the collocation points need not match the mesh of the basis functions, i.e. this result reflects the idea that a small amount of oversampling can guarantee convergence even if the collocation points are chosen suboptimally.

Corollary 4.4.6. *If Δ_M are equispaced, S_N are smoothest spline spaces of degree d on a quasi-equispaced mesh, and $M = M(N) \geq N^\beta$ for some $\beta > 1 + \frac{1}{2d+1-4\alpha}$, then there are constants $C, N_0 > 0$ independent of u, β such that, for all $N \geq N_0$,*

$$\|u_N^{(M)} - u\|_{2\alpha} \leq C N^{2\alpha-d-1} \|u\|_{d+1}.$$

4.4.2 Superconvergence and the discrete Aubin–Nitsche lemma

Although the results in Thm. 4.4.2 guarantee convergence of the oversampled collocation method, we have yet to ask at what asymptotic rate we expect this to occur. For continuous Galerkin methods it is possible to prove superconvergence by a duality argument, the

so-called Aubin–Nitsche lemma (see Hsiao et al. (2017, §4.3) and Hsiao and Wendland (1981)). We will demonstrate that a sufficient amount of superlinear oversampling can achieve such superconvergence for the oversampled collocation method as well.

Theorem 4.4.7 (Discrete Aubin–Nitsche lemma). *Let $-l \leq t \leq 0$ and suppose Eq. (4.26) holds for s, r with $\max\{r, s\} < \min\{m + 1/2, -t\} - 2\alpha$, and that $\Delta_M = \Delta_{M(N)}$ is chosen such that*

$$\lim_{N \rightarrow \infty} \mathcal{E}_{r,s}(\Delta_M) h^{2\alpha - t - \max\{r,s\}} = 0.$$

Then there are constants $C, N_0 > 0$ independent of u such that, for all $N \geq N_0$,

$$\|u_N^{(M)} - u\|_{t+4\alpha} \leq Ch^{l-t-4\alpha} \|u\|_l.$$

Proof. Note we have by the expression for the dual norm Eq. (4.11), for all $t \leq 0$,

$$\begin{aligned} \|(V^*V)(u - u_N^{(M)})\|_t &= \sup_{\substack{\psi \in C^\infty \\ \|\psi\|_{-t}=1}} |\langle (V^*V)(u - u_N^{(M)}), \psi \rangle| \\ &= \sup_{\substack{\psi \in C^\infty \\ \|\psi\|_{-t}=1}} |\langle Vu - Vu_N^{(M)}, V\psi \rangle| \\ &\leq \sup_{\substack{\psi \in C^\infty \\ \|\psi\|_{-t}=1}} \inf_{\chi_N \in S_N} \left(|\langle Vu - Vu_N^{(M)}, V\psi - V\chi_N \rangle| + |\langle Vu - Vu_N^{(M)}, V\chi_N \rangle| \right) \\ &= \sup_{\substack{\psi \in C^\infty \\ \|\psi\|_{-t}=1}} \inf_{\chi_N \in S_N} \left(|\langle Vu - Vu_N^{(M)}, V\psi - V\chi_N \rangle| \right. \\ &\quad \left. + |\langle Vu - Vu_N^{(M)}, V\chi_N \rangle - \langle V(u - u_N^{(M)}), V\chi_N \rangle_M| \right) \tag{4.28} \\ &\leq \sup_{\substack{\psi \in C^\infty \\ \|\psi\|_{-t}=1}} \inf_{\chi_N \in S_N} \left(\|V\|_{2\alpha \rightarrow 0}^2 \|u - u_N^{(M)}\|_{2\alpha} \|\psi - \chi_N\|_{2\alpha} \right. \\ &\quad \left. + \mathcal{E}_{r,s}(\Delta_M) \|V\|_{r+2\alpha \rightarrow r} \|V\|_{s+2\alpha \rightarrow s} \|u - u_N^{(M)}\|_{r+2\alpha} \|\chi_N\|_{s+2\alpha} \right. \\ &\quad \left. + \mathcal{E}_{r,s}(\Delta_M) \|V\|_{r+2\alpha \rightarrow r} \|V\|_{s+2\alpha \rightarrow s} \|u - u_N^{(M)}\|_{s+2\alpha} \|\chi_N\|_{r+2\alpha} \right). \tag{4.29} \end{aligned}$$

Here Eq. (4.28) follows from Eq. (4.15), and Eq. (4.29) follows from Eq. (4.26) since $\max\{r, s\} < m + 1/2 - 2\alpha$. We now refer back to the approximation property of the basis spaces S_N (Assumption 4.2.2) which guarantees that for any $\psi \in H^{-t}$ we have $\chi_N \in S_N$, such that

$$\|\psi - \chi_N\|_s < Ch^{-(s+t)} \|\psi\|_{-t}$$

uniformly for all $-l < s < m + 1/2$ and $s \leq -t \leq l$. For this choice of χ_N it also follows

that

$$\begin{aligned}\|\chi_N\|_s &\leq \|\psi - \chi_N\|_s + \|\psi\|_s \\ &\leq \|\psi - \chi_N\|_s + \|\psi\|_{-t} \leq (1 + Ch^{-(s+t)})\|\psi\|_{-t} \leq \tilde{C}\|\psi\|_{-t}\end{aligned}$$

for some $\tilde{C} > 0$ independent of ψ . Thus, choosing χ_N in this way, the right hand side of Eq. (4.29) is bounded by

$$\tilde{C} \left(\|u - u_N^{(M)}\|_{2\alpha} h^{-(2\alpha+t)} + \mathcal{E}_{r,s}(\Delta_M) \left(\|u - u_N^{(M)}\|_{r+2\alpha} + \|u - u_N^{(M)}\|_{s+2\alpha} \right) \right),$$

since $\max\{r, s\} < -t - 2\alpha$. Now we use Corollary 4.4.3 to conclude, for N sufficiently large:

$$\|u - u_N^{(M)}\|_{H^{t+4\alpha}} \leq \tilde{C} h^{l-(4\alpha+t)} (1 + \mathcal{E}_{r,s}(\Delta_M) h^{2\alpha-t-\max\{r,s\}}) \|u\|_{H^l}$$

for some constant \tilde{C} independent of N, Δ_M, u and the result follows. \square

As before, we can use Lemma 4.4.5 to make concrete predictions for equispaced collocation points $\Delta_M = \{\tilde{x} + m/M\}_{m=1}^M$ and approximation spaces S_N consisting of degree d smoothest splines on a quasi-equispaced grid.

Corollary 4.4.8. *If Δ_M are equispaced, S_N are smoothest spline spaces of degree d on a quasi-equispaced mesh, and $M = M(N) \geq N^\beta$ for some $\beta > 2 + \frac{1}{2d+1-4\alpha}$, then there are constants $C, N_0 > 0$ independent of u such that, for all $N \geq N_0$,*

$$\|u_N^{(M)} - u\|_{-d-1+4\alpha} \leq CN^{4\alpha-2d-2} \|u\|_{d+1}.$$

This tells us that just slightly more than quadratic oversampling suffices to achieve the fastest convergence rate in $H^{-d-1+4\alpha}$. We can infer similar results for the convergence rates in $\|\cdot\|_{t+4\alpha}$ for $-d-1 < t \leq 0$.

4.4.3 Exact expression for the error for equispaced spline bases

So far we have tried to keep the analysis fairly general to allow for suboptimal choices of collocation points. In this section, we are interested in understanding the effects of oversampling when the collocation points and approximation spaces are chosen in an optimal way. This is motivated by the desire to understand both the effect of linear oversampling and the exact rates of superlinear oversampling leading to improvements in the convergence rate of the method. Ultimately we expect that, with sufficient oversampling, the least-squares oversampled collocation method recovers the convergence rate of an associated Bubnov–Galerkin method, and we shall see that quadratic oversampling is

sufficient in this present setting to achieve this. Specifically, in this section we choose S_N to consist of smoothest splines of degree d on an equispaced grid $0 < x_1 < x_2 < \dots < x_N = 1$, with $x_n = n/N, n = 1, \dots, N$, and we let the corresponding collocation points be given by the following mesh refinement (cf. sketch in Fig. 4.2)

$$\Delta_M = \left\{ \frac{l + \xi_j}{N} \mid l = 1, \dots, N, j = 1, \dots, J \right\}, \quad \xi_j = j/J, \quad (4.30)$$

such that $M = JN, J \in \mathbb{N}$, and the discrete inner product is given by

$$\langle f, g \rangle_M = \frac{1}{N} \sum_{n=1}^N \frac{1}{J} \sum_{j=1}^J \overline{f\left(\frac{l + \xi_j}{N}\right)} g\left(\frac{l + \xi_j}{N}\right). \quad (4.31)$$

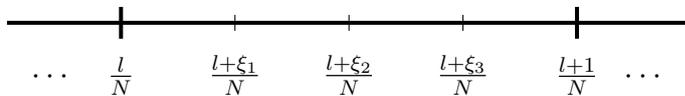


Figure 4.2: Sketch of the mesh refinement for $J = 4$.

The approach we take here was first described by Sloan (1988) (see also the work by Chandler and Sloan (1990) and Sloan (1992)) as a way to study generalised quadrature rules for the inner product in the Galerkin orthogonality conditions. In the following, we adapt the results from Chandler and Sloan (1990) to our setting, with the main difference between our case and theirs being that our test functions are of the form $V\chi_N$ as opposed to χ_N . While Chandler and Sloan (1990) were mainly focused on constructing specific quadrature rules similar to Eq. (4.31) which keep J fixed as N increases, the main novelty for this section is to use their arguments to understand the behaviour when $J = J(N)$ varies with N .

In order to facilitate this, we need to make a slightly stronger assumption on V , namely that it can be expressed in the form

$$V = V_0 + \mathcal{K}, \quad (4.32)$$

where V_0 is a constant multiple of an operator that has a Fourier series representation in the form

$$V_0 g(x) = \sum_{m \in \mathbb{Z}} [m]^{2\alpha} \hat{g}_m e^{2\pi i m x}, \quad (4.33)$$

and \mathcal{K} is a compact perturbation, whose form we will describe in more detail below. In the above we used the notation $[m]$ which we introduced in Eq. (4.9). This means that we assume V_0 is a pseudo-differential operator whose action maps every Fourier mode to a

constant multiple of itself. We will consider two types of compact perturbations:

- $\mathcal{K} : H^s \rightarrow H^t$ is continuous for any choice of $s, t \in \mathbb{R}$. Any integral operator of the form $u \mapsto \int_0^1 g(x, y)u(y) dy$ with $g \in C_{\text{per}}^\infty([0, 1) \times [0, 1))$ has the continuity properties of \mathcal{K} . In particular, the operators $\mathcal{S}, \frac{1}{2}\mathcal{I} + \mathcal{D}$ for smooth boundaries Γ with $k = 0$ (i.e. for Laplace's equation) are of the form $V_0 + \mathcal{K}$ (Sloan, 1992, §3), where $\alpha = -1/2, 0$ respectively.
- \mathcal{K} arising from $\tilde{\mathcal{S}} = V_0 + \mathcal{K}$ on $[0, 1)$ which is the integral operator that comes from reformulating \mathcal{S} for smooth domains and $k > 0$ through the parametrisation $z : [0, 1) \rightarrow \mathbb{R}^2$. The properties of this class are described in Lemma 4.4.9.

Lemma 4.4.9. *The integral operator $\tilde{\mathcal{S}}$ on $[0, 1)$ which arises when we reformulate \mathcal{S} for smooth domains through the parametrisation $z : [0, 1) \rightarrow \mathbb{R}^2$ has the form Eq. (4.32), with $\alpha = -1/2$ and \mathcal{K} that satisfies: $\mathcal{K}V_0^{-1} : H^s \rightarrow H^{s+1.25}$ is continuous for all $s \in \mathbb{R}$ and the action of \mathcal{K} can be expressed in terms of Fourier modes as follows for every $u \in L^2$:*

$$\mathcal{K}V_0^{-1}u(x) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} k_{m,n} \hat{u}_m e^{2\pi i n x}, \quad x \in [0, 1)$$

where the series converges absolutely uniformly and for every $s, t \in \mathbb{R}$ there is a constant C_s such that $\forall m, n \in \mathbb{Z}$:

$$|k_{m,n}| \leq C_{s,t} \left((1 + |m|)^{s+1} (1 + |n|)^t + (1 + |m|) \sum_{l \in \mathbb{Z}} (1 + |m - l|)^{-3} (1 + |l|)^t (1 + |m - n - l|)^s \right). \quad (4.34)$$

Proof. These properties were shown by Kress and Sloan (1993), whose original proof of the continuity property for $s \rightarrow s + 1$ can be easily adapted to show continuity for $s \rightarrow s + 1.25$. For completeness these arguments are given in Appendix 4.D. \square

To begin with, we shall focus on the case when $V = V_0$, since in this case we are able to find an exact expression for the Fourier coefficients of the error $u_N^{(M)} - u$ and use this to derive tight estimates on the convergence rate as shown in Thm. 4.4.10 below. These results can then be extended to the case $V = V_0 + \mathcal{K}$ using a perturbation argument as follows:

- When $\mathcal{K} : H^s \rightarrow H^t$ is continuous for any choice of $s, t \in \mathbb{R}$ the results are valid verbatim and are proved in Appendix 4.F.
- When \mathcal{K} satisfies the properties described in Lemma 4.4.9, the conclusion of the theorem holds when $M \geq N^2$ and this is stated and proved in Appendix 4.G.

Theorem 4.4.10. *If the method satisfies the consistency condition $d > 2\alpha$, then it converges satisfying the following error estimate,*

$$\|u_N^{(M)} - u\|_{4\alpha-(d+1)} \leq C (M^{-(d+1)+2\alpha} + N^{-2(d+1)+4\alpha}) \|u\|_{d+1},$$

where $C > 0$ is a constant depending on d, α , but independent of N, M .

This means that the method converges like

$$u - u_N^{(M)} = \mathcal{O}(M^{-(d+1)+2\alpha} + N^{-2(d+1)+4\alpha}),$$

and hence the fastest possible rate is achieved for $M = N^2$. Moreover, the result also implies that linear oversampling, $M = JN$ for a constant $J \in \mathbb{N}$, leads to an improvement of the error by a factor of $J^{-(d+1)+2\alpha}$, so long as $J \lesssim N$. For the single layer potential with linear splines this means that the error decreases by a factor of J^{-3} which is indeed observed in practice, as we show in Fig. 4.4.

Remark 4.4.11. *We mentioned in §4.2.2 that one might also consider a modified oversampled collocation method which is defined through a modified discrete orthogonality condition Eq. (4.17) of the form*

$$\left\langle \chi_N, V u_N^{(M)} \right\rangle_M = \langle \chi_N, V u \rangle_M, \quad \forall \chi_N \in S_N, \quad (4.35)$$

and which corresponds to a discrete version of a standard Galerkin method. The analysis in this present section can be applied in an analogous way for this modified oversampled collocation method, and one can prove (under the same consistency assumption $d > 2\alpha$) that there is $C > 0$ such that:

$$\|u_N^{(M)} - u\|_{2\alpha-(d+1)} \leq C (M^{-(d+1)+2\alpha} + N^{-2(d+1)+4\alpha}) \|u\|_{d+1}. \quad (4.36)$$

We highlight also that one can show a similar convergence estimate as Eq. (4.36) in the case when $V = V_0 + \mathcal{K}$, where V_0 takes the form shown in Eq. (4.33) and $\mathcal{K} : H^s \rightarrow H^{s-2\alpha+1}$ is continuous as a bounded linear map for all $s, t \in \mathbb{R}$. This follows from a standard argument that is described for instance in Arnold and Wendland (1985, §3) and reproduced for completeness in Appendix 4.E.

Proof of Thm. 4.4.10. To begin with we focus on the case $V = V_0$. An argument extending this result to the case $V = V_0 + \mathcal{K}$ is provided in Appendix 4.F. We follow Chandler and Sloan (1990, §2) and Sloan (1992, §7) and introduce a convenient basis for S_N (where we

write $\Lambda_N = \{\mu \in \mathbb{Z} : -N/2 < \mu \leq N/2\}$ and $\Lambda_N^* = \Lambda_N \setminus \{0\}$:

$$\psi_\mu(x) = \begin{cases} 1, & \mu = 0, \\ \sum_{k \equiv \mu(N)} (\mu/k)^{d+1} e^{2\pi i k x}, & \mu \in \Lambda_N^*. \end{cases} \quad (4.37)$$

Here and in the following we write $k \equiv \mu(N)$ to mean k is congruent to μ modulo N . If $d = 0$ the Fourier series in the above expression and in the following analysis is not absolutely convergent and has to be understood as the limit of the symmetric partial sums (cf. Sloan, 1992, p. 326). We highlight that under this convention our results and proofs apply to the case $d = 0$ in the same manner. Note also that ψ_μ , as defined in Eq. (4.37), is indeed a smoothest spline of the given degree d since its Fourier coefficients satisfy the appropriate recurrence relation,

$$k^{d+1} \hat{v}_k = \mu^{d+1} \hat{v}_\mu, \quad \text{if } k \equiv \mu(N).$$

In many ways the basis $\{\psi_\mu\}_{\mu \in \Lambda_N}$ behaves like a Fourier basis, in particular

$$\psi_\mu(x + n/N) = e^{2\pi i \mu n/N} \psi_\mu(x),$$

which allows us to treat the leading order terms in the oversampled collocation system exactly: let us write our oversampled collocation approximation as

$$u_N^{(M)} = \sum_{\nu \in \Lambda_N} a_\nu \psi_\nu,$$

and let the true solution to Eq. (4.3) be $u(x) = \sum_{m \in \mathbb{Z}} \hat{u}_m \exp(2\pi i m x)$. The discrete orthogonality conditions Eq. (4.15) are

$$\sum_{\nu \in \Lambda_N} \langle V \psi_\mu, V \psi_\nu \rangle_M a_\nu = \langle V \psi_\mu, V u \rangle_M, \quad \mu = 1, \dots, N. \quad (4.38)$$

One can then show, after a few steps of algebra which we provide in Appendix 4.C, that for $\mu \in \Lambda_N$:

$$a_\mu = \begin{cases} \frac{1}{J} \sum_{j=1}^J \sum_{n \equiv 0(N)} [n]^{2\alpha} \hat{u}_n \exp\left(\frac{n}{N} \xi_j\right), & \text{if } \mu = 0, \\ D\left(\frac{\mu}{N}\right)^{-1} \frac{1}{J} \sum_{j=1}^J \sum_{n \equiv \mu(N)} \left[\frac{n}{\mu}\right]^{2\alpha} \exp\left(2\pi i \frac{n-\mu}{N} \xi_j\right) \hat{u}_n \left(1 + \overline{\Omega\left(\xi_j, \frac{\mu}{N}\right)}\right), & \text{if } \mu \neq 0, \end{cases} \quad (4.39)$$

where

$$D(y) = \frac{1}{J} \sum_{j=1}^J |1 + \Omega(\xi_j, y)|^2, \quad \Omega(\xi, y) = |y|^{d+1-\beta} \sum_{l \neq 0} \frac{1}{|l+y|^{d+1-\beta}} e^{2\pi i l \xi}.$$

In consequence it follows that

$$a_\mu - \hat{u}_\mu = \begin{cases} Z_N, & \text{if } \mu = 0, \\ -\frac{E(\mu/N)}{D(\mu/N)}\hat{u}_\mu + R_N(\mu), & \text{if } \mu \neq 0, \end{cases} \quad (4.40)$$

where

$$\begin{aligned} Z_N &= \sum_{\substack{n \in \mathbb{Z} \\ n \neq 0}} [nM]^{2\alpha} \hat{u}_{nM}, \\ E(y) &= |y|^{d+1-2\alpha} \sum_{l \neq 0} \frac{1}{|lJ + y|^{d+1-2\alpha}} + \frac{1}{J} \sum_{j=1}^J |\Omega(\xi_j, y)|^2, \\ R_N(\mu) &= D \left(\frac{\mu}{N} \right)^{-1} \left(\sum_{k \neq 0} \left[\frac{\mu + kM}{\mu} \right]^{2\alpha} \hat{u}_{\mu+kM} \right. \\ &\quad \left. + \sum_{k \neq 0} \left[\frac{\mu + kN}{\mu} \right]^{2\alpha} \hat{u}_{\mu+kN} \left| \frac{\mu}{N} \right|^{d+1-2\alpha} \sum_{\substack{l \equiv k(J) \\ l \neq 0}} \left| \frac{1}{l + \mu/N} \right|^{d+1-2\alpha} \right). \end{aligned}$$

The details of this derivation require only very minor modification compared to the discussion in Chandler and Sloan (1990, §2), but for completeness the arguments are provided in Appendix 4.C. We show in Appendix 4.C in Eq. (4.64) that

$$D(y) \geq 1, \quad \forall y \in [-1/2, 1/2],$$

for any choice $J = J(N) : \mathbb{N}_{>0} \rightarrow \mathbb{N}_{>0}$. This means the oversampled collocation system is well-posed and stable for any integer rate of oversampling $M = J(N)N$. The expression for the error in the small frequency Fourier modes in Eq. (4.40) determines the fastest possible rate of convergence in any Sobolev norm, because the following ‘projection’ P_N onto the low-frequencies,

$$P_N : f \mapsto \sum_{\mu \in \Lambda_N} \hat{f}_\mu \psi_\mu(x),$$

satisfies (Saranen and Vainikko, 2013, Eq. (13.11))

$$\|f - P_N f\|_s \leq C_s N^{s-t} \|f\|_t, \quad \forall s + 1/2 < t \leq d + 1. \quad (4.41)$$

We can use the expressions in Eq. (4.40) analogously to Sloan (1992, Eqs. (7.38)-(7.39))

and show that, subject to the consistency assumption $d > 2\alpha$,

$$\begin{aligned} |Z_N|^2 + \sum_{\mu \in \Lambda_N^*} |\mu|^{2(4\alpha-d-1)} |R_N(\mu)|^2 & \quad (4.42) \\ & \leq C_{\alpha,r} \left(M^{-2r+4\alpha} \|u\|_r^2 + N^{-2r-2(d+1)+8\alpha} \|u\|_r^2 \right), \quad \forall r > 2\alpha + 1/2. \end{aligned}$$

Thus, by taking r arbitrarily large, we can ensure an arbitrarily fast rate of decay in the terms arising from $Z_N, R_N(\mu)$. Therefore, the leading order error term is due to $E(\mu/N)$, and we can in fact estimate this one as well:

$$\left| \frac{E(\mu/N)}{D(\mu/N)} \hat{u}_\mu \right|^2 \leq C_{\alpha,d} \left(|M|^{4\alpha-2(d+1)} |\mu|^{2(d+1)-4\alpha} + |N|^{8\alpha-4(d+1)} |\mu|^{4(d+1)-8\alpha} \right) |\hat{u}_\mu|^2, \quad (4.43)$$

if $d > 2\alpha$. Thus combining Eqs. (4.42) & (4.43) we have:

$$\begin{aligned} \|u_N^{(M)} - P_N u\|_{4\alpha-(d+1)}^2 & = \sum_{\mu \in \Lambda_N} [\mu]^{2(4\alpha-(d+1))} |a_\mu - \hat{u}_\mu|^2 \left(1 + \sum_{l \neq 0} \left[\frac{\mu}{\mu + lN} \right]^{4(d+1)-8\alpha} \right) \\ & \leq C \sum_{\mu \in \Lambda_N} [\mu]^{2(4\alpha-(d+1))} |a_\mu - \hat{u}_\mu|^2 \quad (4.44) \\ & \leq \tilde{C} \left(M^{4\alpha-2(d+1)} \|u\|_{2\alpha}^2 + N^{8\alpha-4(d+1)} \|u\|_{d+1}^2 \right. \\ & \quad \left. + M^{-2r+4\alpha} \|u\|_r^2 + N^{-2r-2(d+1)+8\alpha} \|u\|_r^2 \right), \quad \forall r > 2\alpha + 1/2. \end{aligned} \quad (4.45)$$

Taking $r = d + 1$ and combining this with the projection estimate Eq. (4.41) yields the desired bound

$$\|u_N^{(M)} - u\|_{4\alpha-(d+1)} \leq C \left(M^{2\alpha-(d+1)} + N^{-2(d+1)+4\alpha} \right) \|u\|_{d+1}.$$

□

4.5 Oversampled collocation in specific settings

In the previous section, we focused our attention mainly on smooth boundaries with equispaced collocation points, but in this section we aim to demonstrate that a slight amount of oversampling can stabilise the oversampled collocation method even in the case of Lipschitz domains and the case of highly sub-optimal choices of collocation points. This is with a view to the possible advantages of oversampled collocation methods in general settings/geometries, particularly in three dimensions, where an optimal choice of collocation points may not be immediately obvious.

4.5.1 Non-equispaced sampling points

We begin by examining non-equispaced collocation points. Let us consider a general sequence of collocation points

$$\Delta_M = \Delta_{M(N)} = \{0 \leq x_1 < \cdots < x_M < 1\}.$$

We now assume nothing more than the requirement that the maximum spacing of consecutive collocation points,

$$\text{dist}(\Delta_M) = \max_{1 \leq j \leq M} |x_{j+1} - x_j|,$$

reduces to 0 in a certain way as $N \rightarrow \infty$. Here it is again understood that $x_{N+1} = x_1$ and the distance is measured on the periodic domain $[0, 1)$. We have the following error estimate for the discrete inner product.

Lemma 4.5.1. *Let $\Delta_M(N)$ be as above and fix $r > 5/2, s > 1/2$. Then there is a constant $C_{r,s} > 0$, independent of Δ_M , such that for any $f, g \in H^{\max\{r,s\}}$:*

$$|\langle f, g \rangle - \langle f, g \rangle_M| \leq C_{r,s} M \text{dist}(\Delta_M)^3 (\|f\|_r \|g\|_s + \|f\|_s \|g\|_r). \quad (4.46)$$

Proof. The error estimate is based on Morrey's inequality and the well-known error-expression for the trapezoidal rule for C^2 -functions. For completeness, a proof is included in Appendix 4.B. \square

The estimate in Eq. (4.46) is precisely of the form Eq. (4.26), which allows us to apply Thms. 4.4.2 & 4.4.7 to prove the following result.

Corollary 4.5.2. *If for some $\epsilon > 0$ we have $5/2 + \epsilon \leq m + 1/2 - 2\alpha$, and $\Delta_M = \Delta_{M(N)}$ is chosen such that*

$$\lim_{N \rightarrow \infty} \text{dist}(\Delta_M)^3 h^{-3-\epsilon} M = 0, \quad (4.47)$$

then there are constants $C, N_0 > 0$ independent of u such that, for all $N \geq N_0$

$$\|u_N^{(M)} - u\|_{H^{2\alpha}} \leq Ch^{l-2\alpha} \|u\|_{H^l}.$$

If in addition we have for some $-l \leq t \leq 5/2 + \epsilon + 2\alpha$

$$\lim_{N \rightarrow \infty} \text{dist}(\Delta_M)^3 h^{2\alpha-t-5/2-\epsilon} = 0, \quad (4.48)$$

then there are constants $C, N_0 > 0$ independent of u such that, for all $N \geq N_0$

$$\|u_N^{(M)} - u\|_{H^{t+4\alpha}} \leq Ch^{l-t-4\alpha} \|u\|_{H^l}.$$

These results guarantee convergence of oversampled collocation methods even when the collocation points are very badly chosen. A sufficient condition for success is that, as N increases, the collocation points are distributed sufficiently uniformly to provide a good approximation to the L^2 -inner product in the sense of Eqs. (4.47) & (4.48). Indeed one may see Corollary 4.5.2 as confirmation that, in settings when standard collocation fails, a certain amount of oversampling can help resolve the convergence issues. We will see this to be the case in practice in §4.6.2.

Remark 4.5.3. *Although this is not a main focus of this thesis we highlight that through a similar argument to Prop. 4.4.1 it can be seen that conditioning of the weighted normal equations Eq. (4.5) depends (if sufficiently many collocation points are taken) mainly on the properties of the basis functions S_N and the integral operator V (specifically on the order 2α of V). Hence, for a sufficient amount of oversampling, the conditioning of the linear system Eq. (4.5) is closely related to the conditioning of the corresponding linear system arising in the Bubnov–Galerkin method Eq. (4.23).*

4.5.2 Lipschitz domains

So far we focused on the case when V satisfies assumptions Eqs. (4.6) & (4.7) for any $s \in \mathbb{R}$ and when $u, f \in C^\infty(I)$. As mentioned in Example 4.2.1 this includes the case of single and double layer formulations for the exterior Dirichlet problem on smooth domains. In this section we aim to show on a more specific case that our results in Thm. 4.4.2 and Corollary 4.4.6 extend to wave scattering problems on Lipschitz domains. In the process of extending our results to less regular boundaries, we must pay attention to the limited continuity properties of the integral operators when the problem is reformulated on $I = [0, 1)$ using the Lipschitz parametrisation z .

The continuity properties of the integral operators need to be considered on a case-by-case basis. For that reason and for simplicity, we shall focus on the single layer integral equation on the boundary of a Lipschitz domain (e.g. a polygon) for the exterior Dirichlet problem of the Helmholtz equation. This means, in the above notation, that the parametrisation $z : [0, 1) \rightarrow \Gamma$ as in Eq. (4.1) is a Lipschitz function with Lipschitz inverse (piecewise linear in the case of polygonal boundaries) and our original integral equation on Γ is

$$\mathcal{S}u = f, \tag{4.49}$$

where the single layer operator \mathcal{S} is as defined §1.2.2.

We recall from Thm. 1.2.9 (see also Torres and Welland, 1993, p. 1466) that

$$\mathcal{S} : H^{s+1/2}(\Gamma) \rightarrow H^{s-1/2}(\Gamma)$$

for $|s| \leq 1/2$ is bounded linear, and furthermore that \mathcal{S} is invertible as a mapping between the spaces indicated above for $|s| \leq 1/2$ as long as the wavenumber k is not in a discrete set of resonant values. As we note in §1.2.2 (see also Chandler-Wilde, Graham, Langdon and Spence, 2012, Appendix A.3), the reformulation of the integral equation on $I = [0, 1)$ using the parametrisation z as described in Eq. (4.2) preserves the continuity properties of the integral operator in the limited regime $|s \pm 1/2| \leq 1$. In particular, as long as k is not a resonant wavenumber, the integral operator $\tilde{\mathcal{S}}$ defined by Eq. (4.2) is a continuous isomorphism, with continuous inverse as a map

$$\tilde{\mathcal{S}} : H^{s+1/2}(I) \rightarrow H^{s-1/2}(I).$$

To distinguish between the integral equation on Γ and its reformulation on I for this section only we shall keep the tilde notation, i.e. the integral equation on I is

$$\tilde{\mathcal{S}}\tilde{u} = \tilde{f}.$$

We now look back at the proofs of Prop. 4.4.1 and Thm. 4.4.2 and notice that in all estimates the only norm bounds on V and its inverse appeared as

$$\|V^{-1}\|_{2\alpha \rightarrow 0}, \|V\|_{r+2\alpha \rightarrow r}, \|V\|_{s+2\alpha \rightarrow s},$$

which means that as long as $0 \leq r, s \leq 1$ these terms are still bounded in the present case. Furthermore, in the upper bound Eq. (4.27) any term of the form $\|u - \chi_N\|_{H^t}$ has $t \leq 0$ (by virtue of $r, s \leq 1 = -2\alpha$). Thus, the conclusions of these two statements remain true for the present case.

Corollary 4.5.4. *Let $V = \tilde{\mathcal{S}}$, suppose Eq. (4.26) holds for s, r with $0 \leq r, s < \min\{m + 3/2, 1\}$, and that $\Delta_M = \Delta_{M(N)}$ is chosen such that*

$$\lim_{N \rightarrow \infty} \mathcal{E}_{r,s}(\Delta_M) h^{-(r+s)} = 0,$$

then for any $t \geq 0$ there are constants $C, N_0 > 0$ independent of \tilde{u} such that, for all $N \geq N_0$:

$$\|u_N^{(M)} - \tilde{u}\|_{H^{-1}} \leq Ch^{\min\{t, 0\}+1} \|\tilde{u}\|_{H^t}. \quad (4.50)$$

This means the optimal convergence rate in H^{-1} is achieved.

To make a more specific conclusion we can combine this with the estimate in Lemma 4.4.5 to show that for equispaced sampling on smoothest spline spaces convergence in H^{-1} is achieved if

$$M = M(N) = N^\beta, \quad \text{for some } \beta > 3/2.$$

It is well-known that for smooth boundary conditions on polygonal scatterers, the solution u to Eq. (4.49) has specific corner singularities (see Chandler-Wilde and Langdon, 2007, Thm. 2.3). This means its regularity for arbitrary boundary conditions in $C_0^\infty(\mathbb{R}^2)|_\Gamma$ is restricted to

$$u \in H^s(\Gamma) \quad \text{if and only if } s < -1/2 + \pi/\max_j \Omega_j, \quad (4.51)$$

where $\Omega_j, j = 1, \dots, 2K$, are the interior and exterior angles of the polygon. Since $-1/2 + \pi/\max_j \Omega_j < 1/2$ the continuity property is preserved under z , i.e.

$$\tilde{u} = u \circ z \in H^s(I) \quad \text{if and only if } s < -1/2 + \pi/\max_j \Omega_j,$$

thus placing a limit on the maximum possible value of t in Eq. (4.50). We shall see in §4.6.3 that these predicted rates are indeed close to the observed convergence rates in numerical experiments.

4.6 Numerical results

In this section, we confirm the aforementioned theoretical results using several numerical experiments. In the following examples we consider the two types of integral operators introduced in Example 4.2.1 arising in the single and double layer potential formulations for the exterior Dirichlet problem of the Helmholtz equation as examples of integral operators with orders $2\alpha = -1$ and $2\alpha = 0$ respectively. As discussed in §4.2 we solve these integral equations numerically by reformulating them on the interval $I = [0, 1)$ using the parametrisation z as in Eq. (4.2). To avoid confusing the two formulations we keep the tilde notation to refer to quantities defined on I and recall the formulations:

$$\begin{aligned} Vu &= f, & u, f &: \Gamma \rightarrow \mathbb{C}, \\ \tilde{V}\tilde{u} &= \tilde{f}, & \tilde{u}, \tilde{f} &: I \rightarrow \mathbb{C}, \end{aligned} \quad (4.52)$$

with the bijective correspondence $\tilde{u} = u \circ z, \tilde{f} = f \circ z$. We denote by $\tilde{u}_N^{(M)}$ the approximations to \tilde{u} arising from the oversampled collocation method applied to Eq. (4.52).

4.6.1 Smooth domains with equispaced sampling

We begin by verifying that the results in Thm. 4.4.10 are indeed tight and accurately predict the effect of oversampling for equispaced grids and matching collocation points as defined in Eq. (4.30). In the first example, we present results for the single layer potential formulation

$$\mathcal{S}u = f,$$

where u represents the normal derivative of the field $\partial_n \phi$ on the boundary Γ . We choose $\Gamma = \{|x| = 1\}$, the unit circle, and wavenumber $k = 4.2$. For this domain an exact reference solution u in terms of Bessel functions is available (cf. Morse and Feshbach (1953, pp. 501–504) and Weisstein (2002, p. 1332)). Therefore, we can evaluate the error directly in the Sobolev norms $\|\cdot\|_{H^s}$, which we compute using the expression in terms of Fourier coefficients as in Eq. (4.8). We note that for practical purposes these Fourier coefficients can be approximated efficiently using the fast Fourier transform (FFT). In this example the approximation spaces S_N are taken to be smoothest spline spaces of degree $d = 1$ subordinate to an equispaced mesh on $[0, 1)$ (cf. Example 4.2.4). In Fig. 4.3 we display the Sobolev error $\|\tilde{u} - \tilde{u}_N\|_{H^s(I)}$ and the error in a field point $|S(\tilde{u} \circ z^{-1} - \tilde{u}_N^{(M)} \circ z^{-1})(x)|$, the latter of which converges at the fastest rate of any Sobolev error as per Eq. (4.25). We consider the following approximations to \tilde{u} :

- H^s -projection: The orthogonal projection of \tilde{u} onto S_N with respect to the inner product $\langle \cdot, \cdot \rangle_{H^s(I)}$ which we defined in Eq. (4.8). The projection is computed by expanding $\tilde{u}_N = \sum_{n=1}^N a_n \phi_n$ in a basis $\{\phi_n\}_{n=1}^N$ for S_N and solving the following linear system for the coefficients $a_n, n = 1, \dots, N$:

$$\sum_{n=1}^N a_n \langle \phi_m, \phi_n \rangle_{H^s(I)} = \langle \phi_m, \tilde{u} \rangle_{H^s(I)}, \quad m = 1, \dots, N, \quad (4.53)$$

where the inner product $\langle \cdot, \cdot \rangle_{H^s(I)}$ is computed from the Fourier coefficients of the functions using Eq. (4.10).

- Galerkin method: The solution of the continuous Galerkin equations on Γ

$$\langle \chi_N \circ z^{-1}, V(\tilde{u}_N \circ z^{-1}) \rangle_{L^2(\Gamma)} = \langle \chi_N \circ z^{-1}, f \rangle_{L^2(\Gamma)}, \quad \forall \chi_N \in S_N.$$

The Galerkin approximation is computed by expanding $\tilde{u}_N = \sum_{n=1}^N a_n \phi_n$ in a basis $\{\phi_n\}_{n=1}^N$ for S_N and solving the following linear system for the coefficients

$a_n, n = 1, \dots, N,$

$$\sum_{n=1}^N a_n \langle \phi_m \circ z^{-1}, V(\phi_n \circ z^{-1}) \rangle_{L^2(\Gamma)} = \langle \phi_m \circ z^{-1}, f \rangle_{L^2(\Gamma)}, \quad m = 1, \dots, N. \quad (4.54)$$

The L^2 -inner products are computed using adaptive quadrature in Julia (in particular the `HCubature.jl` package).

- Collocation – $M = N$: The standard collocation method at equispaced points, $x_m = m/M$.
- Collocation – $M = N \lceil N^{1/2} \rceil$, $M = N^2$: The oversampled collocation method at equispaced points with the appropriate rates of oversampling.

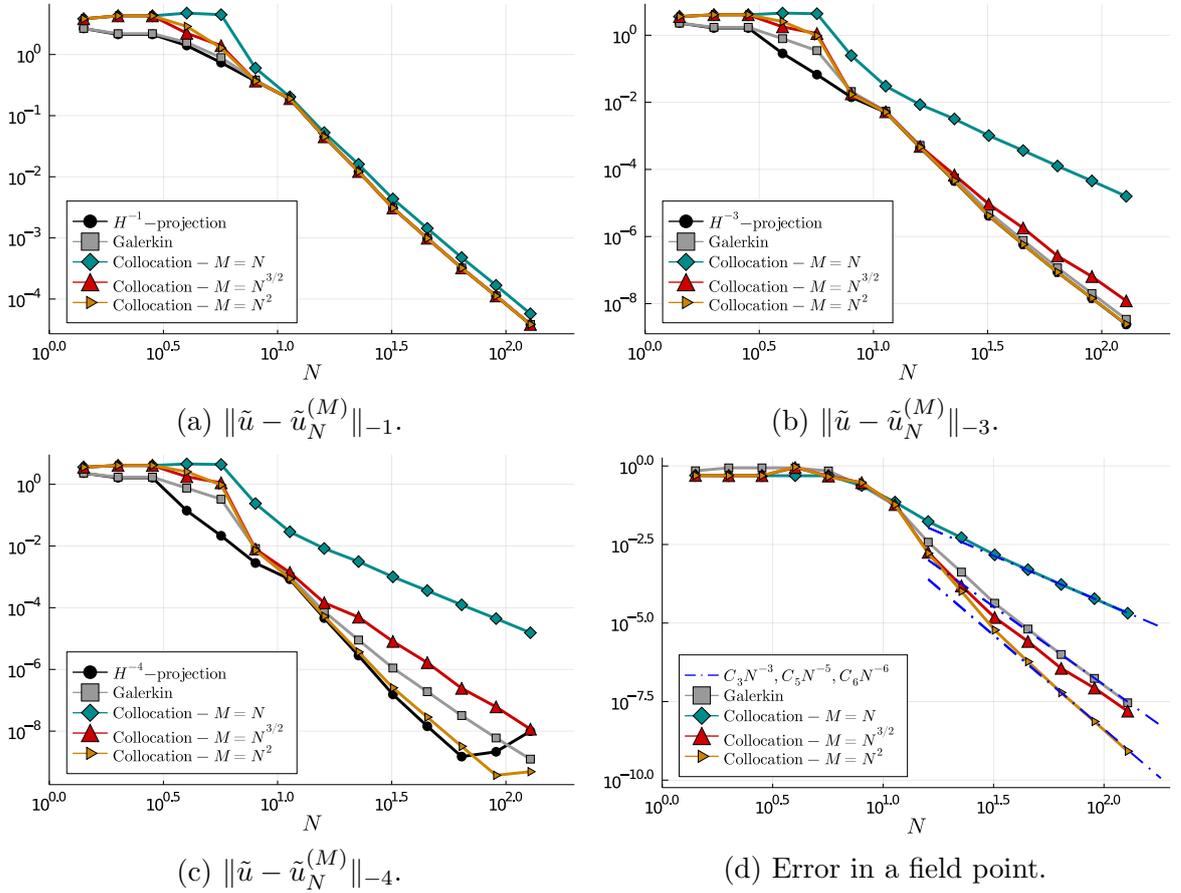


Figure 4.3: Error in the numerical method for a smooth circular scatterer, using the single layer potential formulation and linear splines. In (d) the error in a field point shows the fastest convergence rate in any Sobolev norm, the blue dash-dotted lines indicate, from top to bottom, $C_3 N^{-3}$, $C_5 N^{-5}$ and $C_6 N^{-6}$ respectively, for some appropriately chosen constants C_3, C_5 and C_6 .

We use smoothest splines of degree 1 which means, in terms of earlier notation, $d = 1$, $2\alpha = -1$. Recall that the single layer operator \mathcal{S} is of the form $\mathcal{S} = \mathcal{S}_0 + \mathcal{K}$, where

\mathcal{S}_0 has the form Eq. (4.33) and \mathcal{K} is an integral operator satisfying the properties in Lemma 4.4.9. Thus we can refer to the results in §4.4.3, where we showed that for this case the convergence rates of the oversampled collocation method are

$$\mathcal{O}(N^{-2(d+1)+4\alpha}) = \mathcal{O}(N^{-6}),$$

as long as $M \geq N^2$, and that the fastest possible rate is attained in $H^{4\alpha-(d+1)} = H^{-4}$. Indeed, the same reasoning as in the proof of Thm. 4.4.10 suggests that one might be able to prove the finer result that the convergence order for the oversampled collocation method in H^t , $4\alpha - (d+1) \leq t \leq 2\alpha$, is

$$\mathcal{O}(M^{-(d+1)+2\alpha} + N^{\min\{t-(d+1), -2(d+1)+4\alpha\}}) = \mathcal{O}(M^{-3} + N^{\min\{t-2, -6\alpha\}}).$$

In Fig. 4.3 we observe that the predictions of Thm. 4.4.10 match the numerical experiments very well:

- In the energy space, H^{-1} , the optimal convergence rates are achieved for any choice of $M = JN$, $J = J(N) \geq 1$, in particular for the standard collocation method ($J = 1$), which reflects the results from Arnold and Wendland (1983).
- In the space H^{-3} the Galerkin method achieves its fastest rate of convergence $\mathcal{O}(N^{-5})$ (see e.g. Hsiao et al., 2017, Lemma 4.7), as does the oversampled collocation method with $M = N \lceil N^{1/2} \rceil$. Indeed as we expect, we also see that the standard collocation method converges at a slower rate than oversampled collocation and Galerkin in this norm.
- Finally, in the space H^{-4} the oversampled collocation method with $M = N^2$ converges at the optimal rate $\mathcal{O}(N^{-6})$ as predicted by the results in §4.4.3, whereas all of the other methods, including Galerkin, converge at slower rates. We note that the Bubnov–Galerkin method would also converge at this same fastest rate $\mathcal{O}(N^{-6})$. This can be seen, for instance, from the error analysis in (Hsiao et al., 2017, Lemma 4.7) applied to the operator V^*V .

Remark 4.6.1. *We note in Fig. 4.3c that the convergence rates of our methods appear to level off around $N \approx 10^2$. This is a result of numerical errors due to the ill-conditioning in the H^{-4} -projection matrix and the need to compute a large number of Fourier coefficients of the error function to very high accuracy to accurately compute the Sobolev norm.*

Of course, faster than linear oversampling is unlikely to yield methods which are overall computationally competitive, though it cannot be ruled out a priori that structured low-rank approximations to the system matrix in Eq. (4.15) may (partially) offset the increased dimension of the matrix. Still, we are interested in the quantification of the

benefit of linear oversampling. We highlight that Thm. 4.4.10 actually has a consequence for linear oversampling as well.

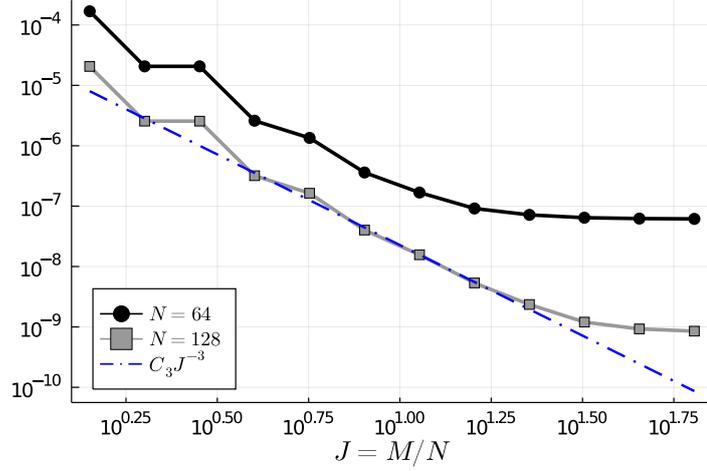


Figure 4.4: The effect of linear oversampling for the same experiment as in Fig. 4.3. Here we plot the error in a field point as a function of $J = M/N$ for $N = 64$ and $N = 128$. For linear splines ($d = 1$), Eq. (4.43) predicts $O(J^{-(d+1)+2\alpha}) = O(J^{-3})$ behaviour, for $1 \leq J \lesssim N$.

Indeed, if we fix N sufficiently large, and then choose a range of constants $J \in \mathbb{N}$ with $1 \leq J \lesssim N$, the theorem predicts that, by taking $M = JN$, we can achieve a decay in the error of order $\mathcal{O}(J^{-(d+1)+2\alpha}) = \mathcal{O}(J^{-3})$ as we increase J in the initial regime $J \lesssim N$. This means, whilst linear oversampling improves the overall error just by a constant, the improvement is cubic in J and so still worthwhile. The result is observed in practice as shown in Fig. 4.4.

4.6.2 Suboptimal choice of collocation points

Having understood the effect of oversampling for an optimal choice of collocation points on smooth domains (i.e. equispaced collocation points which refine the grid of the spline basis spaces), we turn our attention to the case of suboptimal choices of these points. For the present examples we consider the double layer formulation of the Helmholtz equation on a smooth domain, i.e.

$$V = \frac{1}{2}\mathcal{I} + \mathcal{D},$$

which has order $2\alpha = 0$. We now consider two examples for the interior Dirichlet problem on the kite shape shown in Fig. 4.1, which is parametrised by

$$z : t \mapsto (-\sin(2\pi t) - \cos(4\pi t), \cos(2\pi t)), \quad t \in [0, 1].$$

In both cases we plot the error in a field point for the interior field, which is given by

$$\begin{aligned} \left| \phi(x) - \phi_N^{(M)}(x) \right| &= \left| D \left(\tilde{u} \circ z^{-1} - \tilde{u}_N^{(M)} \circ z^{-1} \right) (x) \right| \\ &= \left| \int_0^1 \frac{\partial G}{\partial n_y}(x, y) \Big|_{y=z(t)} \left(\tilde{u}(t) - \tilde{u}_N^{(M)}(t) \right) z'(t) dt \right|, \end{aligned}$$

and captures the optimal convergence properties of the method in any Sobolev norm as we demonstrated in Eq. (4.25). We solve the interior Dirichlet problem with the field point $x = (0.1, 0.2)$, wavenumber $k = 5$ and plane wave boundary conditions

$$\phi|_{\Gamma}(x_1, x_2) = e^{i \cos \theta x_1 + i \sin \theta x_2},$$

with $\theta = 0$, for which an exact reference solution of the interior problem is given precisely in terms of a plane wave $\phi(x) = \phi(x_1, x_2) = \exp(i \cos \theta x_1 + i \sin \theta x_2)$.

The first example concerns linear splines ($d = 1$) on an equispaced mesh, but we take collocation points that are slightly offset. In particular we take

$$\Delta_M = \{0.5/N + m/M \mid m = 1, \dots, M\},$$

i.e. for $M = N$ the collocation points are the midpoints of the spline mesh and for higher rates of oversampling the collocation points are shifted by $0.5/N$. The results are shown in Fig. 4.5.

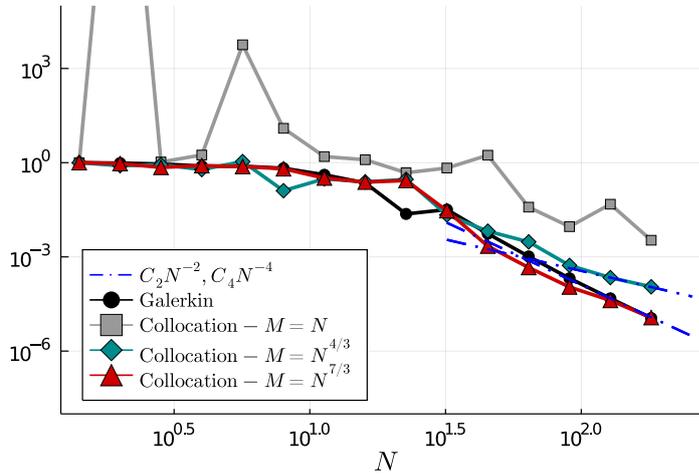


Figure 4.5: Error in an interior field point $\left| D \left(\tilde{u} \circ z^{-1} - \tilde{u}_N^{(M)} \circ z^{-1} \right) (x) \right|$ for wave scattering on a smooth domain, with equispaced points that are offset from the equispaced spline mesh.

We find that, as guaranteed by Corollary 4.4.6, there is convergence at rate $\mathcal{O}(N^{-2})$ for slightly more than linear oversampling $M = N^\beta$, with $\beta = 1 + \frac{1}{2d+1-4\alpha} = 4/3$ and convergence at rate $\mathcal{O}(N^{-4})$ for slightly more than quadratic oversampling $M = N^\beta$, with

$\beta = 2 + \frac{1}{2d+1-4\alpha} = 7/3$. These two rates are indicated with the blue dash-dotted lines in Fig. 4.5.

In contrast, the standard (midpoint) collocation method using odd degree splines for an integral operator of even order is known to be unstable (Sloan, 1992, p. 327). This is observed to some extent by the erratic behaviour of the method in Fig. 4.5, which is particularly bad for small values of N . We highlight that this example exhibits that already a small amount of oversampling can help enhance the robustness of the method towards the choice of collocation points.

The positive effect of oversampling is even more noticeable in our **second example** in Fig. 4.6, where we choose the collocation points in a highly suboptimal way: we draw the points independently from a uniform random distribution:

$$y_m \sim U[0, 1), \quad m = 1, \dots, M, \quad (4.55)$$

and $\{x_m\}_{m=1}^M = \{y_m\}_{m=1}^M$ with $0 \leq x_1 < x_2 < \dots < x_M < 1$. In this example we take smoothest splines of degree $d = 2$ (again on an equispaced mesh) to satisfy the assumptions of Corollary 4.5.2, but other than that consider the same integral equation as in the previous example of Fig. 4.5. As before, in §4.5.1, we let $\text{dist}(\Delta_M) = \max_{1 \leq m \leq M} |x_{m+1} - x_m|$. One can then show that (see e.g. Holst, 1980, Thm. 2.2) the expected maximum distance in the collocation points is

$$\mathbb{E}[\text{dist}(\Delta_M)] = \frac{1}{M-1} \sum_{m=1}^{M-1} \frac{1}{m} \leq \frac{\log(M-1)}{M-1}. \quad (4.56)$$

We recall from Corollary 4.5.2 that convergence in the energy space $H^{2\alpha} = H^0$ (since for the double layer potential $2\alpha = 0$) is guaranteed if $\Delta_M = \Delta_{M(N)}$ is chosen such that, for some $\epsilon > 0$,

$$\lim_{N \rightarrow 0} \text{dist}(\Delta_M)^3 M N^{3+\epsilon} = 0. \quad (4.57)$$

Based on Eq. (4.56) we expect Eq. (4.57) to be the case whenever $M \geq N^\beta, \beta > 4/3$. However, due to the log-term in Eq. (4.56) this convergence may only occur for rather large values of N . As such, in the present example we increase the amount of oversampling (still in the form Eq. (4.55)) for our experiment to $M = N \lceil N^{1.5} \rceil \geq N^{2.5}$.

The results are shown in Fig 4.6. In Fig. 4.6b one can see the quantity $\text{dist}(\Delta_M)^3 M N^3$ which tends to zero for the oversampled collocation method, but which diverges for the case $M = N$, i.e. for the standard collocation method. In Fig. 4.6a we see the error in the field point $x = (0.1, 0.2)$ for the Galerkin method, the collocation method with $M = N$ and for the oversampled collocation method with $M = N^{5/2}$. We notice that the standard

collocation method, with random collocation points as per Eq. (4.55), does not converge. In contrast the oversampled collocation method exhibits convergence and the predicted asymptotic rate $\mathcal{O}(N^{2\alpha-(d+1)}) = \mathcal{O}(N^{-3})$ starts to emerge for large N , although, due to the random nature of the samples, the convergence behaviour is slightly more unsteady than in previous examples. In the plot this asymptotic rate is highlighted by the upper blue dash-dotted line. Of course, as expected from standard convergence theory for the method (see e.g. Hsiao et al., 2017, Lemma 4.7), the Galerkin method converges at the rate $\mathcal{O}(N^{2\alpha-2(d+1)}) = \mathcal{O}(N^{-6})$ since the random sampling plays no role in the Galerkin method, just in the collocation method.

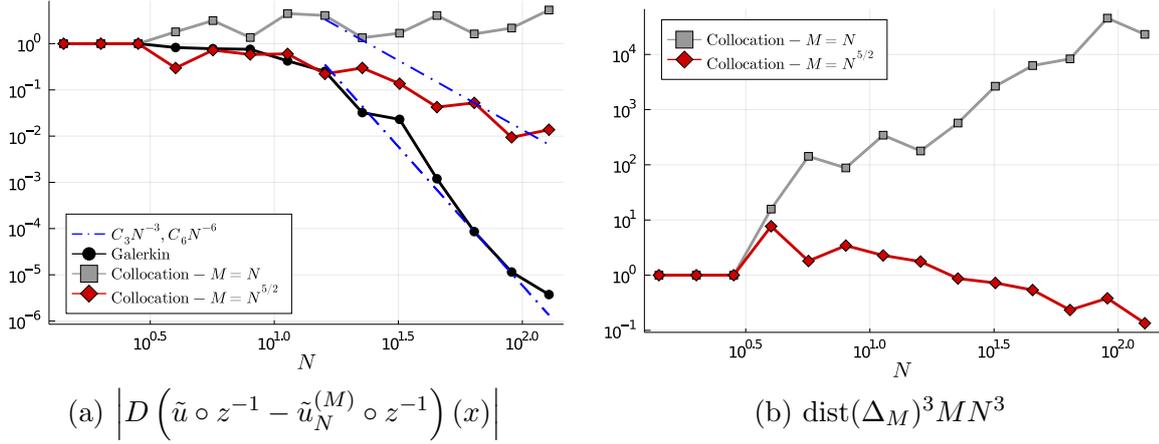


Figure 4.6: Double layer formulation of the interior Dirichlet problem for the Helmholtz equation. The sampling points are drawn uniformly at random $x_m \sim U[0, 1]$.

4.6.3 Polygonal scatterers

Our final numerical example concerns a polygonal domain. Specifically, we consider the single layer formulation for the exterior scattering problem on a pentagonal domain as shown in Fig. 4.8b. We solve the integral equation using a smoothest spline basis of degree 1 based on the piecewise parametrisation $z : [0, 1) \rightarrow \Gamma$ of the pentagon. We showed in Corollary 4.5.4 that taking

$$M = N^\beta, \quad \beta > 3/2,$$

is sufficient to guarantee convergence at the optimal rate in the energy space H^{-1} . For the present example we offset the collocation points by a quarter of the spline basis mesh, i.e.

$$\Delta_M = \left\{ 0.25/N + m/M \mid m = 1, \dots, M \right\},$$

to emphasize that the results are not dependent on an optimal choice of collocation points, even for polygonal domains. The geometry is a regular polygon of side length $2 \sin \frac{2\pi}{5}$ and

we take $k = 10$.

In Fig. 4.7 the error for standard collocation and oversampled collocation methods are compared to the Galerkin method and the optimal rate of convergence provided by the H^{-1} -projection, which is

$$\mathcal{O}(N^{-1/2-\pi/\max_j \Omega_j}) = \mathcal{O}(N^{-9/8})$$

according to Eq. (4.51). We observe that all three methods follow the optimal rate very closely, albeit the oversampled collocation method does so with a smaller error constant than the standard collocation. Of course, it seems even the standard collocation method converges, which suggests that already smaller amounts of oversampling may be beneficial and guarantee convergence.

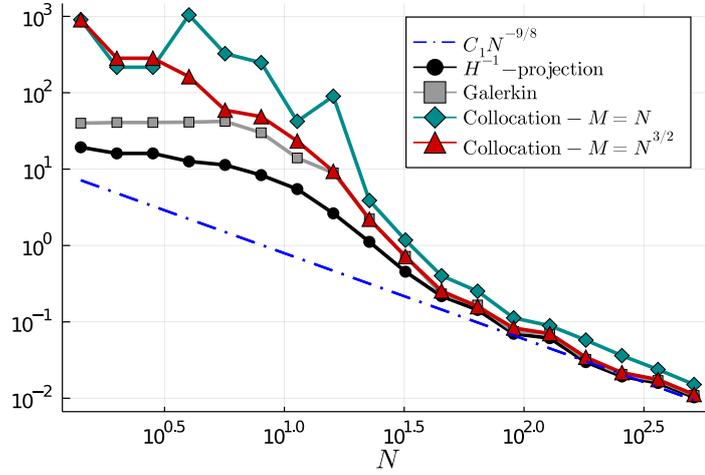


Figure 4.7: Error in the numerical method $\|\tilde{u} - \tilde{u}_N^{(M)}\|_{-1}$ for a regular pentagonal scatterer, using the single layer potential (order $2\alpha = -1$) and linear splines ($d = 1$), with offset equispaced collocation points.

Although there is no guarantee in the polygonal case that this energy space provides the lowest Sobolev order for which optimal convergence holds (since polygonal boundaries are slightly more regular than general Lipschitz ones), we can, in Fig. 4.8a, look at the convergence rates of the method in a field point. This means we plot

$$\left| S(\tilde{u} \circ z^{-1} - \tilde{u}_N^{(M)} \circ z^{-1})(x) \right|.$$

The blue dash-dotted line again indicates the optimal convergence rate in H^{-1} , which we recall is $\mathcal{O}(N^{-9/8})$. It appears that the convergence rate in the field point is very close to this rate, and we also observe again that, even though standard and oversampled collocation methods appear to converge at similar rates, the latter does so with a smaller constant and improved stability for small values of N .

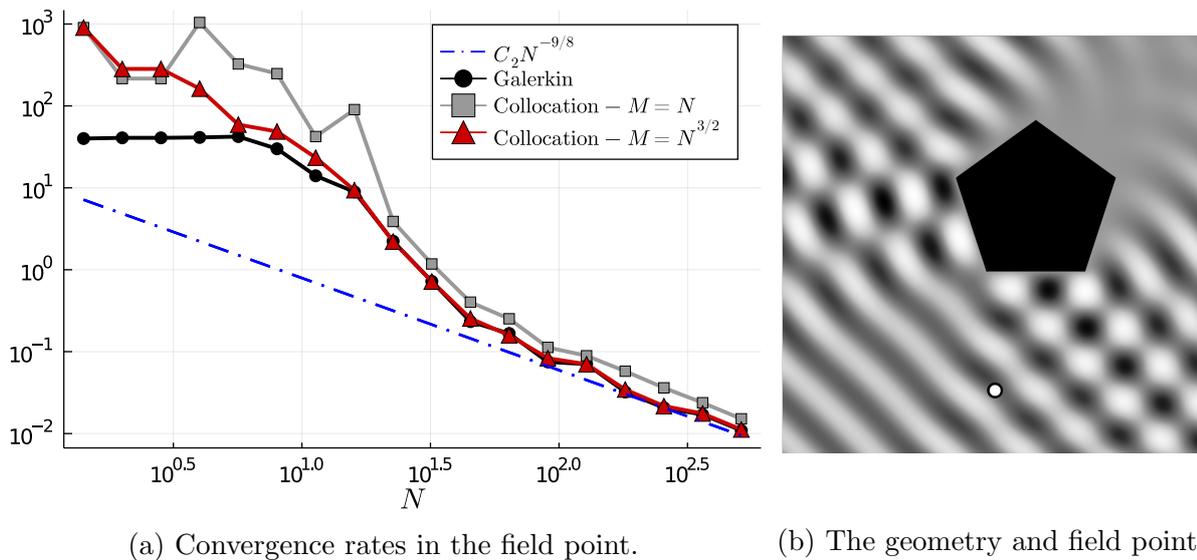


Figure 4.8: Single layer formulation of the exterior Dirichlet problem for the Helmholtz equation on a pentagonal scatterer.

4.7 Conclusions

In the present chapter, we considered an ‘oversampled collocation method’ for Fredholm integral equations, i.e. a collocation method for which the number of collocation points is greater than the dimension of the trial space. Specifically, in the method of consideration, the approximation is given as a weighted least-squares solution to the overdetermined linear system arising from this oversampling process. Our goal was to understand whether this process can be used to enhance the approximation in the collocation setting, with a particular emphasis on avoiding having to choose the collocation points in a very specific, optimal way.

We sought to understand the properties of the oversampled collocation method both through rigorous analysis and numerical examples and our conclusions are twofold. Firstly, we showed that for sufficient amounts of (superlinear) oversampling the convergence rate of the collocation method can be improved using oversampling. Indeed, it is well-known in the literature (Arnold and Wendland, 1983, 1985) that spline collocation on quasi-equispaced grids converges, when they are stable at the rate $\mathcal{O}(N^{-(d+1)+2\alpha})$ where N and d are the dimension and degree of the smoothest spline approximation space respectively and 2α is the order of the integral operator. We showed that due to the relationship of the oversampled collocation method with a Bubnov–Galerkin method, superlinear sampling can double the convergence rate to $\mathcal{O}(N^{-2(d+1)+4\alpha})$. Our main results in this direction are Thms. 4.4.7 & 4.4.10. The former was formulated in a very general framework, based on a general selection of collocation points and regular boundary element spaces in the sense of Babuška and Aziz (1972), and provides a sufficient rate of oversampling to guarantee these

improved convergence rates. Thm. 4.4.10 is specific to settings where the integral operator is a perturbation of an operator of certain pseudo-differential form which allowed us to give an exact expression for the error and to show that in the corresponding setting the fastest convergence rate is achieved in the oversampled collocation method when $M \propto N^2$. Of course, oversampling at a quadratic rate may not be favourable in practice but the same results show that in some cases even linear oversampling at rate $M = JN$ can decrease the error of the method by a factor of J^{-3} , which seems certainly worthwhile, given that the cost of the oversampled collocation method grows only linearly with J .

Our second conclusion is that oversampling can help to enhance the robustness of the method to a suboptimal choice of collocation points. The main result in this direction is Thm. 4.4.2 which provides a convergence guarantee for the oversampled collocation method for a very general choice of collocation points and trial spaces. Indeed this result quantifies a sufficient amount of oversampling that guarantees convergence of the method even for highly suboptimal choices of collocation points. This result was exhibited specifically in Fig. 4.6 where a suboptimal choice of collocation points lead the standard collocation method to diverge, whilst the oversampled version converges. Additionally, in Fig. 4.5 and Fig. 4.8 we provided numerical evidence that shows, especially for small values of N (i.e. the initial range in the convergence plots), oversampling has a significant stabilising effect on the error of the collocation method.

List of symbols

Symbol	Description
2α	Order of the integral operator V .
$H^s(\Gamma)$	Boundary Sobolev space of order s .
I	The periodic interval $[0, 1)$.
M	Number of collocation points; $M > N$ means oversampling.
N	Dimension of the trial space.
S, D	Single and double layer potentials.
$S_N = S_h^{d,m}$	Regular boundary element spaces of dimension N .
V_0	Pseudodifferential component of the operator V .
V	Integral operator of order 2α .
Δ_M	Set of M collocation points.
Γ	The domain of the integral equations, $\Gamma \subset \mathbb{R}^2$.
Λ_N^*	Set of indices Λ_N excluding $\nu = 0$, i.e. $\Lambda_N^* = \Lambda_N \setminus \{0\}$.
Λ_N	Set of indices $\nu \in \mathbb{Z}$ with $-N/2 < \nu \leq N/2$.
Ω	Domain in \mathbb{R}^2 representing the scatterer.
β	Algebraic rate of oversampling, $M = N \lceil N^{\beta-1} \rceil$.
u	Exact solution to the integral equation.
$\langle \cdot, \cdot \rangle_M$	Discrete inner product resulting from M collocation points.
$\langle \cdot, \cdot \rangle_{L^2}$	Continuous L^2 inner product.
$\mathcal{E}_{r,s}(\Delta_M)$	Generic error constant in the error of the discrete inner product.
\mathcal{H}	A generic Hilbert space.
\mathcal{K}^*	Continuous adjoint of the map \mathcal{K} .
\mathcal{K}	Perturbation component of the integral operator, $V = V_0 + \mathcal{K}$.
\mathcal{S}, \mathcal{D}	Single and double layer operators.
$\text{dist}(\Delta_M)$	Maximum spacing of consecutive collocation points.
\bar{z}	Complex conjugate of $z \in \mathbb{C}$.
d	Degree of the piecewise polynomials in the spline basis.
h	Maximum mesh spacing of trial space mesh.
$u_N^{(M)}$	Oversampled collocation approximation to the solution.
u_N	Galerkin approximation to the solution.

4.A Error estimate in the discrete inner product

Here we describe the proof of Lemma 4.4.5. We recall the statement of the Lemma:

Lemma 4.A.1 (Error in discrete L^2 inner product for equispaced sampling). *Let $\Delta_M = \{x_m = \tilde{x} + j/M\}_{m=1}^M$ be a set of equispaced collocation points (where it is understood $x + 1 \equiv x$) and fix $r > 1/2$. For $f, g \in H^r([0, 1])$ let*

$$\langle f, g \rangle = \frac{1}{M} \sum_{m=1}^M \overline{f(x_m)} g(x_m),$$

then there is a constant $C_{r,s} > 0$ independent of f, g such that, for any $r \geq s > 1/2$,

$$|\langle f, g \rangle - \langle f, g \rangle_M| \leq C_{r,s} M^{-r} \left(\|f\|_r \|g\|_s + \|f\|_s \|g\|_r \right).$$

Since \tilde{x} introduces a simple phase shift in all Fourier modes we may, without loss of generality, for the purpose of proving this result assume $\tilde{x} = 0$. Let $\mathcal{Q}_M[f] = 1/M \sum_{m=0}^{M-1} f(m/M)$ be the trapezoidal rule, then we have the following well-known result.

Lemma 4.A.2. *If $f \in L^2([0, 1])$ then*

$$\int_0^1 f(x) dx - \mathcal{Q}_M[f] = \sum_{j \neq 0} \hat{f}_{jM}.$$

Proof.

$$\mathcal{Q}_M[f] = \frac{1}{M} \sum_{m=0}^{M-1} \sum_{k \in \mathbb{Z}} e^{2\pi i m k / M} \hat{f}_k = \frac{1}{M} \sum_{k \in \mathbb{Z}} \hat{f}_k \sum_{m=0}^{M-1} e^{2\pi i m k / M} = \sum_{j \neq 0} \hat{f}_{jM}.$$

□

This implies the following estimate.

Corollary 4.A.3. *For all $t > 1/2$, there is a constant $C_t > 0$ (independent of M), such that*

$$\left| \int_0^1 f(x) dx - \mathcal{Q}_M[f] \right| \leq C_t M^{-t} \|f\|_{H^t}.$$

In particular from Corollary 4.A.3 we have that for any $t > 1/2$:

$$|\langle f, g \rangle - \langle f, g \rangle_M| \leq C_t M^{-t} \|f \bar{g}\|_{H^t}.$$

Proof of Lemma 4.4.5. We observe, using the notation \lesssim when there is a constant independent of u, w implicit in the inequality, and by \lesssim_s we mean that this implicit constant

may depend on s ,

$$\begin{aligned}
\|uw\|_r^2 &= \left| \sum_{m \in \mathbb{Z}} \hat{u}_{-m} \hat{w}_m \right|^2 + \sum_{n \in \mathbb{Z}} |n|^{2r} \left| \sum_{m \in \mathbb{Z}} \hat{u}_{n-m} \hat{w}_m \right|^2 \\
&\lesssim \left(\sum_{m \in \mathbb{Z}} |\hat{u}_{-m}|^2 \right) \left(\sum_{m \in \mathbb{Z}} |\hat{w}_m|^2 \right) + \sum_{n \in \mathbb{Z}} \left| \sum_{m \in \mathbb{Z}} |n-m|^r \hat{u}_{n-m} \hat{w}_m + \sum_{m \in \mathbb{Z}} |m|^r \hat{u}_{n-m} \hat{w}_m \right|^2 \\
&\lesssim \|u\|_0^2 \|w\|_0^2 + \sum_{n \in \mathbb{Z}} \left| \sum_{m \in \mathbb{Z}} |n-m|^r \hat{u}_{n-m} \hat{w}_m \right|^2 + \left| \sum_{m \in \mathbb{Z}} |m|^r \hat{u}_{n-m} \hat{w}_m \right|^2, \tag{4.58}
\end{aligned}$$

where in Eq. (4.58) we used that for $t \geq 1/2$ there is a constant $C = C_t > 0$ such that

$$(|m| + |n|)^t \leq C(|m|^t + |n|^t), \quad \forall m, n \in \mathbb{Z}.$$

By the discrete Minkowski integral inequality

$$\begin{aligned}
\sum_{n \in \mathbb{Z}} \left| \sum_{m \in \mathbb{Z}} |n-m|^r \hat{u}_{n-m} \hat{w}_m \right|^2 &\leq \left(\sum_{m \in \mathbb{Z}} \left(\sum_{n \in \mathbb{Z}} |n-m|^{2r} |\hat{u}_{n-m}|^2 |\hat{w}_m|^2 \right)^{\frac{1}{2}} \right)^2 \\
&\leq \|u\|_r^2 \left(\sum_{m \in \mathbb{Z}} |\hat{w}_m|^2 \right) \lesssim_s \|u\|_r^2 \|w\|_s^2, \quad \text{any } s > 1/2.
\end{aligned} \tag{4.59}$$

The final term in Eq. (4.58) can be bounded similarly and this concludes the proof of Lemma 4.4.5. \square

4.B Error estimate for the discrete inner product non-uniform collocation points

Here we provide a proof of Lemma 4.5.1. We recall the statement of the lemma:

Lemma 4.B.1. *Let $\Gamma, \Delta_M(N)$ be as above and fix $r > 5/2, s > 1/2$. Then there is a constant $C_{r,s} > 0$, independent of Δ_M , such that for any $f, g \in H^{\max\{r,s\}}$:*

$$|\langle f, g \rangle - \langle f, g \rangle_M| \leq C_{r,s} M \text{dist}(\Delta_M)^3 (\|f\|_{H^r} \|g\|_{H^s} + \|f\|_{H^s} \|g\|_{H^r}).$$

Proof. For $f \in C^2$ we have the well-known estimate

$$\left| \int_a^b f(x) dx - \frac{b-a}{2} (f(b) + f(a)) \right| \leq \frac{(b-a)^3}{12} f''(\xi), \quad \text{some } \xi \in (a, b). \tag{4.60}$$

This means in particular for any choice of quadrature points $0 = x_1 < x_2 < \dots < x_M < 1$

that the trapezoidal rule with weights,

$$\int_0^1 f(x) dx \approx \sum_{m=1}^M \frac{|x_{j+1} - x_{j-1}|}{2} f(x_j),$$

has an error of the form

$$\left| \int_0^1 f(x) dx - \sum_{m=1}^M \frac{|x_{j+1} - x_{j-1}|}{2} f(x_j) \right| \leq \frac{M}{12} \sup_{x \in [0,1]} |f''(x)| \max_{1 \leq j \leq N} |x_{j+1} - x_j|^3.$$

We can translate this to Sobolev spaces using Morrey's inequality: For any $l \in \mathbb{N}$

$$\begin{aligned} |f^{(l)}(x)| &= \left| \sum_{m \in \mathbb{Z}} (2\pi i m)^l e^{2\pi i m x} \hat{f}_m \right| \leq \sum_{m \in \mathbb{Z}} |2\pi m|^l |\hat{f}_m| \\ &\leq (2\pi)^l \left(\sum_{m \in \mathbb{Z}} |m|^{-2s} \right)^{1/2} \left(\sum_{m \in \mathbb{Z}} |m|^{2s+2l} |\hat{f}_m|^2 \right)^{1/2} \lesssim_s \|f\|_{H^{l+s}} \quad \text{any } s > 1/2, \end{aligned}$$

where \lesssim_s indicates an implicit constant independent of f but dependent on s . Thus we have for any $f \in H^r$, $r > 5/2$,

$$\left| \int_0^1 f(x) dx - \sum_{m=1}^M \frac{|x_{j+1} - x_{j-1}|}{2} f(x_j) \right| \lesssim_r M \max_{1 \leq j \leq N} |x_{j+1} - x_j|^3 \|f\|_{H^r}.$$

Following through the same steps Eqs. (4.58) & (4.59) as in Appendix 4.A we find

$$|\langle f, g \rangle - \langle f, g \rangle_M| \lesssim_{r,s} M \max_{1 \leq j \leq N} |x_{j+1} - x_j|^3 (\|f\|_{H^r} \|g\|_{H^s} + \|f\|_{H^s} \|g\|_{H^r}),$$

for any $r > 5/2$, $s > 1/2$. □

4.C Derivation of exact error expression for equi-spaced grids

Here we provide the derivation of Eqs. (4.39) & (4.40). The arguments are analogous to the discussion in Chandler and Sloan (1990, §2) with very minor modifications to adapt to our notation and the discrete Bubnov–Galerkin setting. Given the pseudo-differential

form Eq. (4.33) of V , its action on the basis ψ_μ is quickly determined to be

$$\begin{aligned} V\psi_\mu &= \sum_{m \equiv \mu(N)} [m]^{2\alpha} \left(\frac{\mu}{m}\right)^{d+1} e^{2\pi i m x} = [\mu]^{2\alpha} \sum_{m \equiv \mu(N)} \left[\frac{m}{\mu}\right]^{2\alpha} \left(\frac{\mu}{m}\right)^{d+1} e^{2\pi i m x} \\ &= [\mu]^{2\alpha} e^{2\pi i \mu x} \left(1 + \Omega\left(Nx, \frac{\mu}{N}\right)\right), \end{aligned}$$

where

$$\Omega(\xi, y) = |y|^{d+1-2\alpha} \sum_{l \neq 0} \frac{1}{|l+y|^{d+1-2\alpha}} e^{2\pi i l \xi}.$$

Thus we can write the discrete inner product Eq. (4.31) coming from our collocation points as follows

$$\langle V\psi_\mu, V\psi_\nu \rangle_M = \begin{cases} 0, & \text{if } \mu \neq \nu, \\ 1, & \text{if } \mu = \nu = 0, \\ [\mu]^{4\alpha} \frac{1}{J} \sum_{j=1}^J \left|1 + \Omega\left(\xi_j, \frac{\mu}{N}\right)\right|^2, & \text{if } \mu = \nu \neq 0. \end{cases}$$

Similarly, we can compute

$$\langle V\psi_\mu, \exp(2\pi i n \cdot) \rangle_M = \begin{cases} 0, & \text{if } \mu \neq \nu(N), \\ \frac{1}{J} \sum_{j=1}^J \exp(2\pi i l \xi_j), & \text{if } n = lN, \mu = 0, \\ \frac{1}{J} \sum_{j=1}^J \exp(2\pi i l \xi_j) [\mu]^{2\alpha} \left(1 + \overline{\Omega\left(\xi_j, \frac{\mu}{N}\right)}\right), & \text{if } n = \mu + lN, \mu \neq 0. \end{cases}$$

Thus we have for a general u :

$$\begin{aligned} \langle V\psi_\mu, Vu \rangle_M &= \sum_{m \in \mathbb{Z}} [m]^{2\alpha} \hat{u}_m \langle V\psi_\mu, \exp(2\pi i m \cdot) \rangle_M \\ &= \begin{cases} \frac{1}{J} \sum_{j=1}^J \sum_{n \equiv 0(N)} [n]^{2\alpha} \hat{u}_n \exp\left(2\pi i \frac{n}{N} \xi_j\right), & \text{if } \mu = 0, \\ [\mu]^{2\alpha} \frac{1}{J} \sum_{j=1}^J \sum_{n \equiv \mu(N)} \exp\left(2\pi i \frac{n-\mu}{N} \xi_j\right) [n]^{2\alpha} \hat{u}_n \left(1 + \overline{\Omega\left(\xi_j, \frac{\mu}{N}\right)}\right), & \text{if } \mu \neq 0. \end{cases} \end{aligned}$$

Hence the linear system Eq. (4.38) for the coefficients a_μ of $u_N^{(M)}$ in the basis ψ_μ ($u_N^{(M)} = \sum_{\nu \in \Lambda_N} a_\nu \psi_\nu$) is diagonal and we find

$$a_\mu = \begin{cases} \frac{1}{J} \sum_{j=1}^J \sum_{n \equiv 0(N)} [n]^{2\alpha} \hat{u}_n \exp\left(\frac{n}{N} \xi_j\right), & \text{if } \mu = 0, \\ D\left(\frac{\mu}{N}\right)^{-1} \frac{1}{J} \sum_{j=1}^J \sum_{n \equiv \mu(N)} \left[\frac{n}{\mu}\right]^{2\alpha} \exp\left(2\pi i \frac{n-\mu}{N} \xi_j\right) \hat{u}_n \left(1 + \overline{\Omega\left(\xi_j, \frac{\mu}{N}\right)}\right), & \text{if } \mu \neq 0, \end{cases} \quad (4.61)$$

where

$$D(y) = \frac{1}{J} \sum_{j=1}^J |1 + \Omega(\xi_j, y)|^2. \quad (4.62)$$

As in §4.4.3 we let the true solution to Eq. (4.3) be $u(x) = \sum_{m \in \mathbb{Z}} \hat{u}_m \exp(2\pi i m x)$. Thus, simplifying Eq. (4.61) we find the required expressions Eq. (4.40),

$$a_\mu - \hat{u}_\mu = \begin{cases} P_N, & \text{if } \mu = 0, \\ -\frac{E(\mu/N)}{D(\mu/N)} \hat{u}_\mu + R_N(\mu), & \text{if } \mu \neq 0, \end{cases}$$

where:

$$\begin{aligned} P_N &= \frac{1}{J} \sum_{j=1}^J \sum_{\substack{n \equiv 0(N) \\ n \neq 0}} [n]^{2\alpha} \hat{u}_n \exp(2\pi i n \xi_j / N) \\ E(y) &= \frac{1}{J} \sum_{j=1}^J \Omega(\xi_j, y) \left(1 + \overline{\Omega(\xi_j, y)}\right) \\ R_N(\mu) &= D\left(\frac{\mu}{N}\right)^{-1} \frac{1}{J} \sum_{j=1}^J \sum_{\substack{n \equiv \mu(N) \\ n \neq \mu}} \left[\frac{n}{\mu}\right]^{2\alpha} \exp\left(2\pi i \frac{n - \mu}{N} \xi_j\right) \hat{u}_n \left(1 + \overline{\Omega\left(\xi_j, \frac{\mu}{N}\right)}\right). \end{aligned}$$

We can now use the fact that $\xi_j = j/J$ and the identity

$$\frac{1}{J} \sum_{j=1}^J \exp(2\pi i m j / J) = \begin{cases} 1, & m \equiv 0(J), \\ 0, & m \not\equiv 0(J), \end{cases} \quad (4.63)$$

to further simplify the above expressions:

$$\begin{aligned} D(y) &= 1 + \frac{1}{J} \sum_{j=1}^J |\Omega(\xi_j, y)|^2 + 2\operatorname{Re} \left(\frac{1}{J} \sum_{j=1}^J \Omega(\xi_j, y) \right) \\ &= 1 + \frac{1}{J} \sum_{j=1}^J |\Omega(\xi_j, y)|^2 + 2|y|^{d+1-2\alpha} \sum_{l \neq 0} \frac{1}{|lJ + y|^{d+1-2\alpha}} \\ &\geq 1, \quad \forall y \in [-1/2, 1/2]. \end{aligned} \quad (4.64)$$

Similarly we find

$$Z_N = \sum_{\substack{n \in \mathbb{Z} \\ n \neq 0}} [nJN]^{2\alpha} \hat{u}_{nJN} = \sum_{\substack{n \in \mathbb{Z} \\ n \neq 0}} [nM]^{2\alpha} \hat{u}_{nM},$$

$$E(y) = |y|^{d+1-2\alpha} \sum_{l \neq 0} \frac{1}{|lJ + y|^{d+1-2\alpha}} + \frac{1}{J} \sum_{j=1}^J |\Omega(\xi_j, y)|^2,$$

where we made extensive use of the trigonometric identity Eq. (4.63). Finally, we simplify the expression for $R_N(\mu)$

$$\begin{aligned} R_N(\mu) &= D \left(\frac{\mu}{N} \right)^{-1} \left(\sum_{k \neq 0} \left[\frac{\mu + kM}{\mu} \right]^{2\alpha} \hat{u}_{\mu+kM} \right. \\ &\quad \left. + \sum_{\substack{n \equiv \mu(N) \\ n \neq \mu}} \left[\frac{n}{\mu} \right]^{2\alpha} \hat{u}_n \left| \frac{\mu}{N} \right|^{d+1-2\alpha} \frac{1}{J} \sum_{j=1}^J \sum_{l \neq 0} \frac{1}{|l + \mu/N|^{d+1-2\alpha}} \exp \left(2\pi i \left(\frac{n - \mu}{N} - l \right) \frac{j}{J} \right) \right) \\ &= D \left(\frac{\mu}{N} \right)^{-1} \left(\sum_{k \neq 0} \left[\frac{\mu + kM}{\mu} \right]^{2\alpha} \hat{u}_{\mu+kM} \right. \\ &\quad \left. + \sum_{k \neq 0} \left[\frac{\mu + kN}{\mu} \right]^{2\alpha} \hat{u}_{\mu+kN} \left| \frac{\mu}{N} \right|^{d+1-2\alpha} \sum_{\substack{l \equiv k(J) \\ l \neq 0}} \left| \frac{1}{l + \mu/N} \right|^{d+1-2\alpha} \right). \end{aligned}$$

4.D Pseudodifferential form of the single layer operator

Here we provide a proof of Lemma 4.4.9. These properties were proved by Kress and Sloan (1993) and we reproduce this proof with minor adaptations to our present setting. We recall the statement of the lemma:

Lemma 4.D.1. *The integral operator $\tilde{\mathcal{S}}$ on $[0, 1)$ which arises when we reformulate \mathcal{S} for smooth domains through the parametrisation $z : [0, 1) \rightarrow \mathbb{R}^2$ have the form Eq. (4.32), with $\alpha = -1/2$ and \mathcal{K} that satisfies: $\mathcal{K}V_0^{-1} : H^s \rightarrow H^{s+1.25}$ is continuous for all $s \in \mathbb{R}$ and the action of \mathcal{K} can be expressed in terms of Fourier modes as follows for every $u \in L^2$:*

$$\mathcal{K}V_0^{-1}u(x) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} k_{m,n} \hat{u}_m e^{2\pi i n x}, \quad x \in [0, 1)$$

where the series converges absolutely uniformly and for every $s, t \in \mathbb{R}$ there is a constant

C_s such that $\forall m, n \in \mathbb{Z}$:

$$|k_{m,n}| \leq C_{s,t} \left((1 + |m|)^{s+1} (1 + |n|)^t + (1 + |m|) \sum_{l \in \mathbb{Z}} (1 + |m - l|)^{-3} (1 + |l|)^t (1 + |m - n - l|)^s \right). \quad (4.65)$$

Proof. Let us consider the form of the operator arising from re-parametrising \mathcal{S} using Eq. (4.2). Write $z(t) = (z_1(t), z_2(t))$ then $\tilde{\mathcal{S}}$ is given by

$$\tilde{\mathcal{S}}u(t) = \int_0^1 \frac{i}{4} H_0^{(1)} \left(k \sqrt{(z_1(t) - z_1(\tau))^2 + (z_2(t) - z_2(\tau))^2} \right) u(\tau) d\tau$$

For notational simplicity we drop the constant factor $i/4$ henceforth (this does not alter the analysis in any way). Thus the integral operator takes the form

$$\tilde{\mathcal{S}}u(t) = \int_0^1 K(t, \tau) u(\tau) d\tau.$$

Kress and Sloan (1993, pp. 207-208) showed that the integral kernel can be written in the form

$$K(t, \tau) = \ln \left(\frac{4}{e} \sin^2(\pi(t - \tau)) \right) + K_1(t, \tau) \sin^2(\pi(t - \tau)) \ln \left(\frac{4}{e} \sin^2(\pi(t - \tau)) \right) + K_2(t, \tau),$$

where K_1 and K_2 are infinitely differentiable, i.e. $K_1, K_2 \in C_{\text{per}}^\infty([0, 1] \times [0, 1])$. We split the integral operator $\tilde{\mathcal{S}}$ therefore in the following way $\tilde{\mathcal{S}} = V_0 + \mathcal{K}$ (again dropping the constant factor $i/4$ for notational simplicity), where

$$\begin{aligned} V_0 u(t) &= \int_0^1 \ln \left(\frac{4}{e} \sin^2(\pi(t - \tau)) \right) u(t) dt, \\ \mathcal{K} u(t) &= \int_0^1 \left(K_1(t, \tau) \sin^2(\pi(t - \tau)) \ln \left(\frac{4}{e} \sin^2(\pi(t - \tau)) \right) + K_2(t, \tau) \right) u(\tau) d\tau. \end{aligned}$$

The operator V_0 is precisely of the form Eq. (4.33) (cf. Kress and Sloan, 1993, Eq. (3.11)) with $\alpha = -1/2$, specifically

$$V_0 u(t) = - \sum_{m \in \mathbb{Z}} [m]^{-1} \hat{u}_m e^{2\pi i m x}. \quad (4.66)$$

Moreover we have (Kress and Sloan, 1993, Eq. (3.12)) for any $m \in \mathbb{Z}$:

$$\int_0^1 \sin^2(\pi\tau) \ln \left(\frac{4}{e} \sin^2(\pi\tau) \right) e^{2\pi i m \tau} d\tau = \frac{1}{4} (2[m]^{-1} - [m+1]^{-1} - [m-1]^{-1}) =: \gamma_m.$$

In particular, for $|m| \geq 2$, we have

$$\gamma_m = \frac{1}{2(1-m^2)|m|}. \quad (4.67)$$

We begin by showing that $\mathcal{K} : H^{-2.25} \rightarrow L^2$ is bounded: For this we follow Kress and Sloan (1993) and let

$$a_m(t) = \int_0^1 K_1(t, \tau) u(\tau) e^{-2\pi i m \tau} d\tau.$$

This implies

$$|\mathcal{K}_1 u(t)|^2 \leq \sum_{m \in \mathbb{Z}} [m]^{9/2} |\gamma_m|^2 \sum_{m \in \mathbb{Z}} \frac{|a_m(t)|^2}{[m]^{9/2}} = c \|K_1(t, \cdot) u\|_{-2.25}^2$$

where

$$\mathcal{K}_1 u(t) = \int_0^1 K_1(t, \tau) \sin^2(\pi(t - \tau)) \ln \left(\frac{4}{e} \sin^2(\pi(t - \tau)) \right) u(\tau) d\tau.$$

This is true because for $|m|$ sufficiently large we have $[m]^{9/2} |\gamma_m|^2 \lesssim [m]^{-3/2}$. Thus we have

$$\|\mathcal{K}_1 u\|_0 \leq c \sup_{0 \leq t \leq 1} \|K_1(t, \cdot) u\|_{-2.25},$$

hence $\mathcal{K}_1 : H^{-2.25} \rightarrow L^2$ is bounded. Since $K_2 \in C_{\text{per}}^\infty([0, 1) \times [0, 1))$ the same holds true for \mathcal{K} . We can then follow the arguments in (Kress and Sloan, 1993, p. 209) analogously to show by differentiation of $\mathcal{K}u(t)$ that the boundedness remains true for $\mathcal{K} : H^{s-2.25} \rightarrow H^s$ for all $s \in \mathbb{N}$ the result for arbitrary $s \in \mathbb{R}$ then follows by interpolation and duality as in (Kress and Sloan, 1993, p. 209).

Let us now prove Eq. (4.65). Because $K_1, K_2 \in C_{\text{per}}^\infty([0, 1) \times [0, 1))$ we can express them in terms of their bivariate Fourier series in the form

$$K_j(t, \tau) = \sum_{m, n \in \mathbb{Z}} \hat{\eta}_{m, n}^{(j)} e^{2\pi i m t} e^{2\pi i n \tau}, \quad j = 1, 2, \quad (4.68)$$

with coefficients that decay spectrally fast, i.e. which are such that for every $s, t \in \mathbb{R}$ there are constants $C_{s, t}^{(1)}, C_{s, t}^{(2)} > 0$, such that

$$|\hat{\eta}_{m, n}^{(j)}| \leq C_{s, t}^{(j)} (1 + |m|)^s (1 + |n|)^t, \quad \forall m, n \in \mathbb{Z}.$$

Using this observation we can write the action of \mathcal{K} on a Fourier mode $\phi_l(t) := \exp(2\pi i l t)$

as

$$\begin{aligned}\mathcal{K}\phi_l(t) &= \sum_{m,n \in \mathbb{Z}} e^{2\pi i m t} \int_0^1 \left(\eta_{m,n}^{(1)} \sin^2(\pi(t-\tau)) \ln \left(\frac{4}{e} \sin^2(\pi(t-\tau)) \right) + \eta_{m,n}^{(2)} \right) e^{2\pi i n \tau} e^{2\pi i l \tau} d\tau \\ &= \sum_{m,n \in \mathbb{Z}} e^{2\pi i m t} \left(e^{2\pi i(n+l)t} \eta_{m,n}^{(1)} \gamma_{n+l} + \eta_{m,-l}^{(2)} \right).\end{aligned}$$

Define

$$\tilde{k}_{l,\xi} := \eta_{\xi,-l}^{(2)} + \sum_{n+l+m=\xi} \eta_{m,n}^{(1)} \gamma_{n+l}$$

Then we have, by Eq. (4.67) & (4.68), for any $s, t \in \mathbb{R}$

$$\begin{aligned}|\tilde{k}_{l,\xi}| &\leq C_{s,t} \left((1+|\xi|)^t (1+|l|)^s + \sum_{n+l+m=\xi} (1+|n+l|)^{-3} (1+|m|)^s (1+|n|)^t \right) \\ &\leq C_{s,t} \left((1+|\xi|)^t (1+|l|)^s + \sum_{n \in \mathbb{Z}} (1+|l+n|)^{-3} (1+|\xi-l-n|)^s (1+|n|)^t \right).\end{aligned}\quad (4.69)$$

Combining Eqs. (4.66) & (4.69) yields the desired bound Eq. (4.65). \square

4.E A perturbation argument for modified over-sampled collocation

Here we outline the compact perturbation argument referenced in Remark 4.4.11 which is standard in the relevant literature for the analysis of collocation methods, see for instance Arnold and Wendland (1985, §3). Suppose we proved the following a priori estimate for some $s < t$:

Lemma 4.E.1. *If $a_N^{(M)} \in S_N$ is such that*

$$\left\langle \chi_N, V_0 a_N^{(M)} \right\rangle_M = \langle \chi_N, V_0 \tilde{a} \rangle_M \quad \forall \chi_N \in S_N, \quad (4.70)$$

for some $\tilde{a} \in H^t$, then we have $\|\tilde{a} - a_N^{(M)}\|_s \lesssim N^{s-t} \|\tilde{a}\|_t$.

Here, and in the remainder of this appendix, we use the notation $A(\tilde{a}, a_N) \lesssim B(\tilde{a})$ to mean there is a constant $C > 0$ independent of N, \tilde{a}, a_N such that $A(\tilde{a}, a_N) \leq CB(\tilde{a})$. Under the perturbation $V = V_0 + \mathcal{K}$, where $\mathcal{K} : H^p \rightarrow H^{p-2\alpha+1}$ is continuous for all $p \in \mathbb{R}$, the orthogonality conditions Eq. (4.35) become

$$\left\langle \chi_N, (V_0 + \mathcal{K}) u_N^{(M)} \right\rangle_M = \langle \chi_N, (V_0 + \mathcal{K}) u \rangle_M \quad \forall \chi_N \in S_N. \quad (4.71)$$

The trick, given by Arnold and Wendland (1985), is then to write Eq. (4.71) in the equivalent form

$$\left\langle \chi_N, V_0 u_N^{(M)} \right\rangle_M = \left\langle \chi_N, V_0 \left(V_0^{-1}(V_0 + \mathcal{K})u - V_0^{-1}\mathcal{K}u_N^{(M)} \right) \right\rangle_M \quad \forall \chi_N \in S_N,$$

which means that by Eq. (4.70) we have

$$\left\| \left(V_0^{-1}(V_0 + \mathcal{K})u - V_0^{-1}\mathcal{K}u_N^{(M)} \right) - u_N^{(M)} \right\|_s \lesssim N^{s-t} \left\| V_0^{-1}(V_0 + \mathcal{K})u - V_0^{-1}\mathcal{K}u_N^{(M)} \right\|_t.$$

Simplifying both sides we have

$$\left\| V_0^{-1}(V_0 + \mathcal{K}) \left(u - u_N^{(M)} \right) \right\|_s \lesssim N^{s-t} \left\| u + V_0^{-1}\mathcal{K} \left(u - u_N^{(M)} \right) \right\|_t.$$

Now, by the pseudo-differential form Eq. (4.33), $V_0 : H^{s+2\alpha} \rightarrow H^s$ is continuous and by the assumptions on V , we have $V = V_0 + \mathcal{K} : H^{s+2\alpha} \rightarrow H^s$ is invertible, thus

$$\left\| u - u_N^{(M)} \right\|_s \lesssim_s \left\| V_0^{-1}(V_0 + \mathcal{K}) \left(u - u_N^{(M)} \right) \right\|_s,$$

and, by continuity of $\mathcal{K} : H^{t-1} \rightarrow H^{t-2\alpha}$ and of $V_0^{-1} : H^{t-2\alpha} \rightarrow H^t$,

$$\left\| u + V_0^{-1}\mathcal{K} \left(u - u_N^{(M)} \right) \right\|_t \leq \|u\|_t + C \left\| u - u_N^{(M)} \right\|_{t-1}.$$

By the uniform approximation property (Assumption 4.2.2) and the inverse property (Assumption 4.2.3) we can find constants $\tilde{C}, \tilde{C} > 0$ such that

$$\begin{aligned} \left\| u - u_N^{(M)} \right\|_{t-1} &\leq \|u - \chi_N\|_{t-1} + \left\| \chi_N - u_N^{(M)} \right\|_{t-1} \\ &\leq \tilde{C}N^{-1}\|u\|_t + \tilde{C}N^{t-1-s} \left\| \chi_N - u_N^{(M)} \right\|_s \\ &\leq \tilde{C}N^{-1}\|u\|_t + \tilde{C}N^{t-1-s} \left(\left\| u - u_N^{(M)} \right\|_s + \|\chi_N - u\|_s \right) \\ &\leq \tilde{C}N^{-1}\|u\|_t + N^{t-s-1} \left\| u - u_N^{(M)} \right\|_s \end{aligned}$$

Thus, in summary, there is a constant $C > 0$ independent of N, M, u such that

$$\left\| u - u_N^{(M)} \right\|_s (1 - CN^{-1}) \lesssim N^{s-t} \|u\|_t.$$

Finally, since $(1 - CN^{-1}) \rightarrow 1$ as $N \rightarrow \infty$, we find for some $N_0 > 0$ and all $N \geq N_0$

$$\left\| u - u_N^{(M)} \right\|_s \lesssim N^{s-t} \|u\|_t,$$

i.e. the conclusion of Lemma 4.E.1 still holds when V_0 is replaced by $V_0 + \mathcal{K}$.

4.F A perturbation argument for least-squares over-sampled collocation (Laplace)

Our goal is to develop a perturbation approach for the discrete Bubnov-Galerkin method similar to the approach given in Appendix 4.E. This Bubnov-Galerkin case requires a few additional steps which, to the best of our knowledge, are not yet available in the literature. Specifically, recalling Thm. 4.4.10, we will prove the following statement.

Theorem 4.F.1. *If $u_N^{(M)} \in S_N$ is such that*

$$\left\langle (\mathcal{I} + \mathcal{K})V_0\chi_N, (\mathcal{I} + \mathcal{K})V_0u_N^{(M)} \right\rangle_M = \langle (\mathcal{I} + \mathcal{K})V_0\chi_N, (\mathcal{I} + \mathcal{K})V_0u \rangle_M, \quad \forall \chi_N \in S_N, \quad (4.72)$$

where $\mathcal{I} + \mathcal{K} : H^s \rightarrow H^s$ is a continuous isomorphism for all $s \in \mathbb{R}$ and $\mathcal{K} : H^s \rightarrow H^t$ is continuous for all $s, t \in \mathbb{R}$, then we still have the estimate Eq. (4.36), i.e. there are constants $M_0, N_0, C > 0$ independent of N, u, M such that, for all $M \geq M_0, N \geq N_0$,

$$\|u_N^{(M)} - u\|_{2\alpha-(d+1)} \leq C(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)})\|u\|_{d+1}. \quad (4.73)$$

Note our formulation of this theorem is equivalent to the perturbation $V = V_0 + \mathcal{K}$ simply by writing

$$V = (\mathcal{I} + \mathcal{K}V_0^{-1})V_0,$$

since, by the pseudodifferential form of V_0 , the map $\mathcal{K}V_0^{-1} : H^s \rightarrow H^t$ is still continuous for all $s, t \in \mathbb{R}$, and $V : H^{s+2\alpha} \rightarrow H^s$ being invertible for all $s \in \mathbb{R}$ is equivalent to $\mathcal{I} + \mathcal{K}V_0^{-1} : H^s \rightarrow H^s$ being invertible for all $s \in \mathbb{R}$. In order to prove this statement in a similar fashion to the perturbation argument given in Appendix 4.E we need to find a way to take the ‘discrete adjoint’ of \mathcal{K} with respect to $\langle \cdot, \cdot \rangle_M$. Specifically, we would like to formulate the orthogonality conditions Eq. (4.72) in a form similar to

$$\left\langle V_0\chi_N, (\mathcal{I} + \mathcal{K}^*)(\mathcal{I} + \mathcal{K})V_0u_N^{(M)} \right\rangle_M = \langle V_0\chi_N, (\mathcal{I} + \mathcal{K}^*)(\mathcal{I} + \mathcal{K})V_0u \rangle_M \quad \forall \chi_N \in S_N, \quad (4.74)$$

where by \mathcal{K}^* we have denoted the continuous adjoint map corresponding to \mathcal{K} , which is a continuous map $\mathcal{K}^* : H^{-t} \rightarrow H^{-s}$ for all $s, t \in \mathbb{R}$. We note that Eq. (4.74) would be exactly equivalent to Eq. (4.72) if we were to replace $\langle \cdot, \cdot \rangle_M$ by the exact L^2 -inner product $\langle \cdot, \cdot \rangle_{L^2}$. However, the discrete nature of $\langle \cdot, \cdot \rangle_M$ prevents this exact equivalence, and so we need to find a way to account for the error incurred in $u_N^{(M)}$ when we choose to solve Eq. (4.74) instead of Eq. (4.72).

In order to do so let us introduce the following bilinear form ϵ :

$$\epsilon(\chi_N, \tilde{b}) := \left\langle V_0 \chi_N, \mathcal{K}^* V_0 \tilde{b} \right\rangle_M - \left\langle \mathcal{K} V_0 \chi_N, V_0 \tilde{b} \right\rangle_M.$$

Note that the continuity properties of \mathcal{K} allow us to represent the map \mathcal{K} by its action on the Fourier basis, i.e. letting $k_{mn} = \langle \exp(2\pi i n \cdot), \mathcal{K} \exp(2\pi i m \cdot) \rangle_{L^2}$ we have for any $u \in L^2$

$$(\mathcal{K}u)(x) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} k_{mn} \hat{u}_m e^{2\pi i n x},$$

and the series converges absolutely uniformly, since (by continuity of $\mathcal{K} : H^s \rightarrow H^t, \forall s, t \in \mathbb{R}$) for every $s, t \in \mathbb{R}$ there is $C_{s,t} > 0$ such that

$$|k_{mn}| \leq C_{s,t} (1 + |m|)^{-s} (1 + |n|)^{-t}, \quad \forall m, n \in \mathbb{Z}. \quad (4.75)$$

Similarly, \mathcal{K}^* is represented by the conjugate transpose of these values, i.e.

$$\langle \exp(2\pi i n \cdot), \mathcal{K}^* \exp(2\pi i m \cdot) \rangle_{L^2} = \overline{k_{nm}}$$

where \bar{y} denotes the complex conjugate of $y \in \mathbb{C}$. The following is the central new a priori estimate facilitating the proof of Thm. 4.F.1.

Lemma 4.F.2. *Suppose $a_N^{(M)} \in S_N$ satisfies*

$$\left\langle V_0 \chi_N, V_0 a_N^{(M)} \right\rangle_M = \langle V_0 \chi_N, V_0 \tilde{v} \rangle_M + \epsilon(\chi_N, b - c_N), \quad \forall \chi_N \in S_N, \quad (4.76)$$

for some $\tilde{v}, b \in H^{d+1}$ and a sequence of spline functions $c_N \in S_N, N \in \mathbb{N}$. Then there is a constant $C > 0$ independent of N, M, \tilde{v}, b, c_N such that for all $M \geq N > 0$:

$$\begin{aligned} \|\tilde{v} - a_N^{(M)}\|_{4\alpha-(d+1)} &\leq C (M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)}) (\|\tilde{v}\|_{d+1} + \|b\|_{d+1}) \\ &\quad + CN^{-1} \|b - c_N\|_{4\alpha-(d+1)}. \end{aligned}$$

Proof of Lemma 4.F.2. We recall that for the basis $\{\psi_\mu\}_{\mu \in \Lambda_N}$ of S_N , which we introduced in Eq. (4.37), the orthogonality conditions Eq. (4.76) are represented by a diagonal matrix system for the coefficients a_μ of the expansion $a_N^{(M)} = \sum_{\mu \in \Lambda_N} a_\mu \psi_\mu$. The linear system arising from using this basis in Eq. (4.76) therefore differs from the linear system Eq. (4.38) only in the right hand side, and more specifically only by terms of the form

$$\epsilon(\psi_\mu, b - c_N).$$

Thus let us express those terms in a more explicit way: Let us write \hat{b}_n, \hat{c}_n for the Fourier coefficients of b, c_N respectively, noting that for all $n \in \Lambda_N^*, l \in \mathbb{Z}$, we have

$\hat{c}_{n+lN} = n^{d+1}/(n+lN)^{d+1}\hat{c}_n$. Let us focus on the case $\mu \in \Lambda_N^*$ first, and consider $\mu = 0$ after this initial calculation. We can compute

$$\mathcal{K}V_0\psi_\mu = \sum_{n \in \mathbb{Z}} \sum_{m \equiv \mu(N)} [m]^{2\alpha} \left(\frac{\mu}{m}\right)^{d+1} k_{mn} e^{2\pi i n x}.$$

Thus

$$\begin{aligned} & \langle \mathcal{K}V_0\psi_\mu, V_0(b - c_N) \rangle_M \\ &= \sum_{n \in \mathbb{Z}} \sum_{m \equiv \mu(N)} \sum_{p \in \mathbb{Z}} [m]^{2\alpha} \left(\frac{\mu}{m}\right)^{d+1} \overline{k_{mn}} [p]^{2\alpha} (\hat{b}_p - \hat{c}_p) \langle \exp(2\pi i n \cdot), \exp(2\pi i p \cdot) \rangle_M \\ &= \sum_{n \in \mathbb{Z}} \sum_{m \equiv \mu(N)} \sum_{p \equiv n(M)} [m]^{2\alpha} \left(\frac{\mu}{m}\right)^{d+1} \overline{k_{mn}} [p]^{2\alpha} (\hat{b}_p - \hat{c}_p). \end{aligned}$$

Similarly we find

$$\begin{aligned} & \langle V_0\psi_\mu, \mathcal{K}^*V_0b \rangle_M \\ &= \sum_{n \in \mathbb{Z}} \sum_{m \in \mathbb{Z}} [m]^{2\alpha} (\hat{b}_m - \hat{c}_m) \overline{k_{nm}} \left\langle \sum_{p \equiv \mu(N)} [p]^{2\alpha} \left(\frac{\mu}{p}\right)^{d+1} \exp(2\pi i p \cdot), \exp(2\pi i n \cdot) \right\rangle_M \\ &= \sum_{p \equiv \mu(N)} \sum_{n \equiv p(M)} \sum_{m \in \mathbb{Z}} [p]^{2\alpha} \left(\frac{\mu}{p}\right)^{d+1} [m]^{2\alpha} (\hat{b}_m - \hat{c}_m) \overline{k_{nm}}. \end{aligned}$$

Therefore we have

$$\begin{aligned} \epsilon(\psi_\mu, b - c_N) &= \langle V_0\psi_\mu, \mathcal{K}^*V_0(b - c_N) \rangle_M - \langle \mathcal{K}V_0\psi_\mu, V_0(b - c_N) \rangle_M \\ &= \sum_{p \equiv \mu(N)} \sum_{n \equiv p(M)} \sum_{m \in \mathbb{Z}} [p]^{2\alpha} \left(\frac{\mu}{p}\right)^{d+1} [m]^{2\alpha} (\hat{b}_m - \hat{c}_m) \overline{k_{nm}} \\ &\quad - \sum_{n \in \mathbb{Z}} \sum_{m \equiv \mu(N)} \sum_{p \equiv n(M)} [m]^{2\alpha} \left(\frac{\mu}{m}\right)^{d+1} \overline{k_{mn}} [p]^{2\alpha} (\hat{b}_p - \hat{c}_p) \\ &= \sum_{p \equiv \mu(N)} \sum_{n \equiv p(M)} \sum_{m \in \mathbb{Z}} [p]^{2\alpha} \left(\frac{\mu}{p}\right)^{d+1} [m]^{2\alpha} (\hat{b}_m - \hat{c}_m) \overline{k_{nm}} \\ &\quad - \sum_{m \in \mathbb{Z}} \sum_{n \equiv \mu(N)} \sum_{p \equiv m(M)} [n]^{2\alpha} \left(\frac{\mu}{n}\right)^{d+1} \overline{k_{nm}} [p]^{2\alpha} (\hat{b}_p - \hat{c}_p). \end{aligned}$$

We can then extract the low-frequency terms in both sums which are found to cancel:

$$\begin{aligned}
\epsilon(\psi_\mu, b - c_N) &= \sum_{p \equiv \mu(N)} \sum_{m \in \mathbb{Z}} [p]^{2\alpha} \left(\frac{\mu}{p}\right)^{d+1} [m]^{2\alpha} (\hat{b}_m - \hat{c}_m) \overline{k_{pm}} \\
&\quad - \sum_{m \in \mathbb{Z}} \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left(\frac{\mu}{n}\right)^{d+1} \overline{k_{nm}} [m]^{2\alpha} (\hat{b}_m - \hat{c}_m) \\
&\quad + \sum_{p \equiv \mu(N)} \sum_{n \equiv p(M), n \neq p} \sum_{m \in \mathbb{Z}} [p]^{2\alpha} \left(\frac{\mu}{p}\right)^{d+1} [m]^{2\alpha} (\hat{b}_m - \hat{c}_m) \overline{k_{nm}} \\
&\quad - \sum_{m \in \mathbb{Z}} \sum_{n \equiv \mu(N)} \sum_{\substack{p \equiv m(M) \\ p \neq m}} [n]^{2\alpha} \left(\frac{\mu}{n}\right)^{d+1} \overline{k_{nm}} [p]^{2\alpha} (\hat{b}_p - \hat{c}_p) \\
&= \underbrace{\sum_{p \equiv \mu(N)} \sum_{n \equiv p(M), n \neq p} \sum_{m \in \mathbb{Z}} [p]^{2\alpha} \left(\frac{\mu}{p}\right)^{d+1} [m]^{2\alpha} (\hat{b}_m - \hat{c}_m) \overline{k_{nm}}}_{=: A_1} \\
&\quad - \underbrace{\sum_{m \in \mathbb{Z}} \sum_{n \equiv \mu(N)} \sum_{\substack{p \equiv m(M) \\ p \neq m}} [n]^{2\alpha} \left(\frac{\mu}{n}\right)^{d+1} \overline{k_{nm}} [p]^{2\alpha} (\hat{b}_p - \hat{c}_p)}_{=: A_2}.
\end{aligned}$$

In what follows we will bound the remaining two terms A_1, A_2 individually. For this (and the remainder of this appendix) we shall again make use of the notation \lesssim to indicate an implicit constant in the inequality, which is in all cases independent of $a_N^{(M)}, \tilde{v}, b, c_N, N, M$, though it may sometimes depend on other parameters in the inequalities. Where this is of relevance we will indicate this dependence by a subscript, for instance \lesssim_r .

$$\begin{aligned}
|A_1| &\leq \left| \sum_{n \equiv \mu(M), n \neq \mu} [\mu]^{2\alpha} (\mathcal{K}V_0(\widehat{b - c_N}))_n \right| \\
&\quad + \left| \sum_{p \equiv \mu(N), p \neq \mu} \sum_{\substack{n \equiv p(M), n \neq p}} [p]^{2\alpha} \left(\frac{\mu}{p}\right)^{d+1} (\mathcal{K}V_0(\widehat{b - c_N}))_n \right| \\
&\lesssim_r [\mu]^{2\alpha} M^{-r} \|\mathcal{K}V_0(b - c_N)\|_r \\
&\quad + \sum_{l \neq 0} \sum_{\substack{n \equiv \mu + lN(M) \\ n \neq \mu + lN}} [\mu + lN]^{2\alpha} \left| \frac{\mu}{\mu + lN} \right|^{d+1} |(\mathcal{K}V_0(\widehat{b - c_N}))_n| \\
&\lesssim_r [\mu]^{2\alpha} M^{-r} \|\mathcal{K}V_0(b - c_N)\|_r \\
&\quad + \sum_{l \neq 0} [\mu + lN]^{2\alpha} \left| \frac{\mu}{\mu + lN} \right|^{d+1} \sum_{n \equiv \mu + lN(M)} |(\mathcal{K}V_0(\widehat{b - c_N}))_n|
\end{aligned}$$

for any $r > 1/2$. Therefore,

$$\begin{aligned}
|A_1| &\lesssim_r [\mu]^{2\alpha} M^{-r} \|\mathcal{K}V_0(b - c_N)\|_r \\
&\quad + \sum_{l \neq 0} [\mu + lN]^{2\alpha} \left| \frac{\mu}{\mu + lN} \right|^{d+1} C_r \|\mathcal{K}V_0(b - c_N)\|_r,
\end{aligned} \tag{4.77}$$

for any $r > 1/2$. For the second term we have

$$\begin{aligned}
|A_2| &= \left| \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left(\frac{\mu}{n}\right)^{d+1} \sum_{m \in \mathbb{Z}} \sum_{\substack{p \equiv m(M) \\ p \neq m}} \overline{k_{nm}}(V_0(\widehat{b - c_N}))_p \right| \\
&\leq \underbrace{\left| \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left(\frac{\mu}{n}\right)^{d+1} \sum_{m \in \Lambda_N} \sum_{\substack{p \equiv m(M) \\ p \neq m}} \overline{k_{nm}}(V_0(\widehat{b - c_N}))_p \right|}_{=: A_{21}} \\
&\quad + \underbrace{\left| \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left(\frac{\mu}{n}\right)^{d+1} \sum_{m \notin \Lambda_N} \sum_{\substack{p \equiv m(M) \\ p \neq m, p \notin \Lambda_N}} \overline{k_{nm}}(V_0(\widehat{b - c_N}))_p \right|}_{=: A_{22}}.
\end{aligned}$$

Let us estimate A_{21}, A_{22} separately: Firstly, using the continuity properties of \mathcal{K} , i.e. Eq. (4.75) we find for any $s_1, t_1 \in \mathbb{Z}$:

$$\begin{aligned}
A_{21} &\leq \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} \sum_{m \in \Lambda_N} \sum_{\substack{p \equiv m(M) \\ p \neq m}} |\overline{k_{nm}}| \left| (V_0(\widehat{b - c_N}))_p \right| \\
&\lesssim_{s_1, t_1} \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} \sum_{m \in \Lambda_N} \sum_{\substack{p \equiv m(M) \\ p \neq m}} (1 + |n|)^{-s_1} (1 + |m|)^{-t_1} \left| (V_0(\widehat{b - c_N}))_p \right|.
\end{aligned}$$

Moreover, we have for $t_1 > -4\alpha + d + 3/2$:

$$\begin{aligned}
&\sum_{m \in \Lambda_N} (1 + |m|)^{-t_1} \sum_{\substack{p \equiv m(M) \\ p \neq m}} [p]^{2\alpha} \left| (\widehat{b - c_N})_p \right| \\
&= \sum_{m \in \Lambda_N} (1 + |m|)^{-t_1} \sum_{l \neq 0} [m + lM]^{2\alpha} \left| \hat{b}_{m+lM} - \left(\frac{m}{m + lM}\right)^{d+1} \hat{c}_m \right| \\
&\leq \sum_{m \in \Lambda_N} (1 + |m|)^{-t_1} \sum_{l \neq 0} [m + lM]^{2\alpha} \left(|\hat{b}_{m+lM}| + |\hat{c}_m| \left| \frac{m}{m + lM} \right|^{d+1} \right) \\
&\lesssim \sum_{m \in \Lambda_N} (1 + |m|)^{-t_1} \sum_{l \neq 0} [m + lM]^{2\alpha - (d+1)} [m + lM]^{d+1} |\hat{b}_{m+lM}| \\
&\quad + M^{2\alpha - (d+1)} \sum_{m \in \Lambda_N} (1 + |m|)^{-t_1} |m|^{d+1} |\hat{c}_m|.
\end{aligned} \tag{4.78}$$

Therefore, for $t_1 > -4\alpha + d + 3/2$,

$$\begin{aligned}
& \sum_{m \in \Lambda_N} (1 + |m|)^{-t_1} \sum_{\substack{p \equiv m(M) \\ p \neq m}} [p]^{2\alpha} \left| (\widehat{b - c_N})_p \right| \\
& \lesssim_{t_1} \left(\sum_{m \in \Lambda_N} (1 + |m|)^{-t_1} \sum_{l \neq 0} [m + lM]^{4\alpha - 2(d+1)} \right)^{\frac{1}{2}} \|b\|_{d+1} \\
& \quad + M^{2\alpha - (d+1)} \|c_N\|_{4\alpha - (d+1)} \\
& \lesssim_{t_1} M^{2\alpha - (d+1)} (\|b\|_{d+1} + \|c_N\|_{4\alpha - (d+1)}) \\
& \lesssim_{t_1} M^{2\alpha - (d+1)} (\|b\|_{d+1} + \|b - c_N\|_{4\alpha - (d+1)} + \|b\|_{4\alpha - (d+1)}) \\
& \lesssim_{t_1} M^{2\alpha - (d+1)} (2\|b\|_{d+1} + \|b - c_N\|_{4\alpha - (d+1)}),
\end{aligned} \tag{4.79}$$

where in Eqs. (4.78) & (4.79) we used that $|m| \leq N/2 \leq M/2$ for all $m \in \Lambda_N$ and the consistency condition $d > 2\alpha$, which implies that $d + 1 - 2\alpha > 1$. In the final line we also relied on the consistency condition $d > 2\alpha$, which implies $d + 1 > 4\alpha - (d + 1)$. Therefore, we found, for $t_1 > -4\alpha + d + 3/2$:

$$A_{21} \lesssim_{s_1, t_1} \left(\sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} (1 + |n|)^{-s_1} \right) M^{2\alpha - (d+1)} (2\|b\|_{d+1} + \|b - c_N\|_{4\alpha - (d+1)}) \tag{4.80}$$

For A_{22} we have by Eq. (4.75)

$$\begin{aligned}
A_{22} & \leq \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} \sum_{m \notin \Lambda_N} \sum_{\substack{p \equiv m(M) \\ p \neq m}} |\overline{k_{nm}}| |(V_0(\widehat{b - c_N}))_p| \\
& \lesssim_{s_2, t_2} \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} \sum_{m \notin \Lambda_N} \sum_{\substack{p \equiv m(M) \\ p \neq m}} (1 + |n|)^{-s_2} (1 + |m|)^{-t_2} |(V_0(\widehat{b - c_N}))_p| \\
& \lesssim_{s_2, t_2} \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} (1 + |n|)^{-s_2} \sum_{m \notin \Lambda_N} (1 + |m|)^{-t_2} \sum_{p \in \mathbb{Z}} |(V_0(\widehat{b - c_N}))_p| \\
& \lesssim_{s_2, t_2, \delta} \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} (1 + |n|)^{-s_2} \sum_{m \notin \Lambda_N} (1 + |m|)^{-t_2} \|V_0(b - c_N)\|_{-2\alpha + d + 1/2 - \delta},
\end{aligned}$$

where by the consistency assumption $d > 2\alpha$ we were able to choose δ with $0 < \delta < d - 2\alpha$ which implies $\sum_{p \in \mathbb{Z}} [p]^{2(-2\alpha + d + 1/2 - \delta)} < \infty$. By the approximation property of splines spaces (Assumption 4.2.2) we can choose $\chi_N \in S_N$ such that for some $c > 0$ and for all $t < d + 1/2$:

$$\|b - \chi_N\|_t \leq c\|b\|_{d+1}.$$

This allows us to estimate, using the inverse property of smoothest splines (Assumption 4.2.3),

$$\begin{aligned}
\|V_0(b - c_N)\|_{-2\alpha+d+1/2-\delta} &\lesssim \|b - c_N\|_{d+1/2-\delta} \lesssim (\|b - \chi_N\|_{d+1/2-\delta} + \|\chi_N - c_N\|_{d+1/2-\delta}) \\
&\lesssim (N^{-1/2-\delta}\|b\|_{d+1} + N^{2d+3/2-\delta-4\alpha}\|\chi_N - c_N\|_{4\alpha-(d+1)}) \\
&\lesssim (N^{-1/2-\delta}\|b\|_{d+1} + N^{2d+3/2-\delta-4\alpha}\|b - \chi_N\|_{4\alpha-(d+1)} \\
&\quad + N^{2d+3/2-\delta-4\alpha}\|b - c_N\|_{4\alpha-(d+1)}) \\
&\lesssim (N^{-1/2-\delta}\|b\|_{d+1} + N^{2d+3/2-\delta-4\alpha}\|b - c_N\|_{4\alpha-(d+1)}).
\end{aligned}$$

Furthermore we have, whenever $t_2 > 1$,

$$\sum_{m \notin \Lambda_N} (1 + |m|)^{-t_2} \leq \tilde{C}_{t_2} N^{1-t_2}$$

for some constant $\tilde{C}_{t_2} > 0$ independent of N . We can combine these estimates to show that, when $t_2 = 3/2 - \delta + 2(d+1) - 4\alpha$,

$$A_{22} \lesssim_{s_2, t_2} \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} (1 + |n|)^{-s_2} (N^{4\alpha-2(d+1)}\|b\|_{d+1} + N^{-1}\|b - c_N\|_{4\alpha-(d+1)}). \quad (4.81)$$

This means combining Eqs. (4.77), (4.80) & (4.81) gives a bound on $\epsilon(\psi_\mu, b - c_N)$, whenever $\mu \in \Lambda_N^*$. A similar bound for $\mu = 0$ can be found by simply replacing the sums over $\sum_{n \equiv \mu(N)}$ by the unique choice $n = 0$ with no summation. We now understand how the linear system for a_μ is affected by the perturbation $\epsilon(\psi_\mu, b - c_N)$, and we can use a similar procedure as in the proof of Thm. 4.4.10 to understand how this affects $\|\tilde{v} - a_N^{(M)}\|_{4\alpha-(d+1)}$. Specifically, since the linear system is diagonal we have following analogue of Eq. (4.61)

$$a_\mu = a'_\mu + [\mu]^{-4\alpha} D \left(\frac{\mu}{N} \right)^{-1} \epsilon(\psi_\mu, b - c_N), \quad \mu \in \Lambda_N,$$

where

$$a'_\mu = \begin{cases} \frac{1}{J} \sum_{j=1}^J \sum_{n \equiv 0(N)} [n]^{2\alpha} \hat{v}_n \exp\left(\frac{n}{N} \xi_j\right), & \text{if } \mu = 0, \\ D \left(\frac{\mu}{N} \right)^{-1} \frac{1}{J} \sum_{j=1}^J \sum_{n \equiv \mu(N)} \left[\frac{n}{\mu} \right]^{2\alpha} \exp\left(2\pi i \frac{n-\mu}{N} \xi_j\right) \hat{v}_n \left(1 + \overline{\Omega\left(\xi_j, \frac{\mu}{N}\right)}\right), & \text{if } \mu \neq 0, \end{cases}$$

and $D(y)$ is as defined in Eq. (4.62). We recall from Eq. (4.64) that $D(\mu/N) \geq 1, \forall \mu \in \Lambda_N$, and note that therefore the coefficients a_μ are affected by the perturbation $\epsilon(\psi_\mu, b - c_N)$

as follows

$$|a_\mu - a'_\mu| \leq [\mu]^{-4\alpha} |\epsilon(\psi_\mu, b - c_N)|.$$

To estimate the norm $\|a_N^{(M)} - \tilde{v}\|_{4\alpha-(d+1)}$ we recall the proof of Thm. 4.4.10 and specifically the estimate Eq. (4.44), which implies

$$\|a_N^{(M)} - P_N \tilde{v}\|_{4\alpha-(d+1)}^2 \leq C \sum_{\mu \in \Lambda_N} [\mu]^{2(4\alpha-(d+1))} |a_\mu - \hat{v}_\mu|^2.$$

Therefore, compared to Thm. 4.4.10, the additional term in the Sobolev error arising from the perturbation $\epsilon(\psi_\mu, b - c_N)$ is bounded above by

$$C \sum_{\mu \in \Lambda_N} [\mu]^{2(4\alpha-(d+1))} |a_\mu - a'_\mu|^2 \leq C \sum_{\mu \in \Lambda_N} [\mu]^{-2(d+1)} |\epsilon(\psi_\mu, b - c_N)|^2,$$

where $C > 0$ is independent of $\tilde{v}, a_N^{(M)}, M, N, b, c_N$. We can now apply the estimates derived in Eqs. (4.77), (4.80) & (4.81), to find there are constants $C_r, C_s > 0$ such that for any $r > 1/2$ and any $s > 0$

$$\begin{aligned} C \sum_{\mu \in \Lambda_N} [\mu]^{-2(d+1)} |\epsilon(\psi_\mu, b - c_N)|^2 &\leq C_r \sum_{\mu \in \Lambda_N} \left([\mu]^{2\alpha-(d+1)} M^{-r} + \sum_{l \neq 0} [\mu + lN]^{2\alpha-(d+1)} \right)^2 \\ &\quad \|\mathcal{KV}_0(b - c_N)\|_r^2 \\ &\quad + C_s \sum_{\mu \in \Lambda_N} \left(\sum_{n \equiv \mu(N)} [n]^{2\alpha-(d+1)} (1 + |n|)^{-s} \right)^2 \\ &\quad \left(M^{2\alpha-(d+1)} \|b\|_{d+1} + M^{2\alpha-(d+1)} \|b - c_N\|_{4\alpha-(d+1)} \right. \\ &\quad \left. + N^{4\alpha-2(d+1)} \|b\|_{d+1} + N^{-1} \|b - c_N\|_{4\alpha-(d+1)} \right)^2. \end{aligned}$$

Given that $2\alpha - d < 0$, we have $\sum_{l \neq 0} [\mu + lN]^{2\alpha-(d+1)} \lesssim N^{-1}$. Moreover $\mathcal{K} : H^{2\alpha-(d+1)} \rightarrow H^r$ is continuous, and so we find for $r > 1$ and $s > 1$:

$$\begin{aligned} \sum_{\mu \in \Lambda_N} [\mu]^{-2(d+1)} |\epsilon(\psi_\mu, b - c_N)|^2 &\lesssim N^{-2} \|b - c_N\|_{4\alpha-(d+1)}^2 \\ &\quad + (M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)})^2 \|b\|_{d+1}^2. \end{aligned}$$

Combining this with the estimates Eqs. (4.45) & (4.41) yields

$$\begin{aligned} \|a_N^{(M)} - \tilde{v}\|_{4\alpha-(d+1)} &\leq C(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)}) (\|\tilde{v}\|_{d+1} + \|b\|_{d+1}) \\ &\quad + CN^{-1} \|b - c_N\|_{4\alpha-(d+1)}, \end{aligned}$$

for some constant $C > 0$ independent of $a_N^{(M)}, \tilde{v}, N, M, b, c_N$, which completes the proof of Lemma 4.F.2. \square

We can now use Lemma 4.F.2 to prove Thm. 4.F.1 in a manner closely related to the argument given in Appendix 4.E.

Proof of Thm. 4.F.1. We proceed in two steps: Firstly we show that a perturbation of the test functions $V_0\chi_N \mapsto (\mathcal{I} + \mathcal{K})V_0\chi_N$ yields a similar error estimate as in Thm. 4.4.10 and then we proceed to perturb the operator V_0 on the right hand side of the orthogonality conditions.

Claim 4.F.3. *Suppose $a_N^{(M)} \in S_N$ satisfies*

$$\left\langle (1 + \mathcal{K})V_0\chi_N, V_0a_N^{(M)} \right\rangle_M = \left\langle (1 + \mathcal{K})V_0\chi_N, V_0\tilde{a} \right\rangle_M, \quad \forall \chi_N \in S_N, \quad (4.82)$$

where \mathcal{K} satisfies the assumptions of Thm. 4.F.1. Then there are constants $C, M_0, N_0 > 0$ independent of $\tilde{a}, a_N^{(M)}, M, N$ such that for $M \geq M_0, N \geq N_0$:

$$\|a_N^{(M)} - \tilde{a}\|_{4\alpha-(d+1)} \leq C(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)})\|\tilde{a}\|_{d+1}. \quad (4.83)$$

Proof. To begin with we note the conditions Eq. (4.82) are equivalent to

$$\left\langle V_0\chi_N, (\mathcal{I} + \mathcal{K}^*)V_0a_N^{(M)} \right\rangle_M = \left\langle V_0\chi_N, (\mathcal{I} + \mathcal{K}^*)V_0\tilde{a} \right\rangle_M + \epsilon(\chi_N, a_N^{(M)} - \tilde{a}), \quad \forall \chi_N \in S_N.$$

This can be equivalently written as

$$\left\langle V_0\chi_N, V_0a_N^{(M)} \right\rangle_M = \left\langle V_0\chi_N, V_0 \left(\tilde{a} + V_0^{-1}\mathcal{K}^*V_0(\tilde{a} - a_N^{(M)}) \right) \right\rangle_M + \epsilon(\chi_N, a_N^{(M)} - \tilde{a}), \quad \forall \chi_N \in S_N.$$

Therefore, Lemma 4.F.2 applies and shows that

$$\begin{aligned} & \|V_0^{-1}(\mathcal{I} + \mathcal{K}^*)V_0(\tilde{a} - a_N^{(M)})\|_{4\alpha-(d+1)} \\ & \leq C(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)})\|\tilde{a} + V_0^{-1}\mathcal{K}^*V_0(\tilde{a} - a_N^{(M)})\|_{d+1} \\ & \quad + C(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)})\|\tilde{a}\|_{d+1} \\ & \quad + CN^{-1}\|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)}. \end{aligned}$$

Noting that $V_0 : H^{4\alpha-(d+1)} \rightarrow H^{4\alpha-(d+1)}$ is a continuous isomorphism, and that $\mathcal{I} + \mathcal{K}^* : H^{2\alpha-(d+1)} \rightarrow H^{2\alpha-(d+1)}$ is invertible, there is a constant $\tilde{C} > 0$ such that

$$\begin{aligned} \|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)} & \leq \tilde{C}(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)})\|\tilde{a}\|_{d+1} \\ & \quad + \tilde{C}(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)} + N^{-1})\|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)}. \end{aligned}$$

Equivalently,

$$\begin{aligned} \left(1 - \tilde{C} \left(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)} + N^{-1}\right)\right) \|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)} \\ \leq \tilde{C} \left(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)}\right) \|\tilde{a}\|_{d+1}. \end{aligned}$$

The we conclude that, for M, N sufficiently large, the estimate Eq. (4.83) holds. \blacksquare

Having proved Claim 4.F.3 we can proceed to prove Thm. 4.F.1 in an analogous way to the arguments in Appendix 4.E: Suppose $u_N^{(M)} \in S_N$ satisfies

$$\left\langle (\mathcal{I} + \mathcal{K})V_0\chi_N, (\mathcal{I} + \mathcal{K})V_0u_N^{(M)} \right\rangle_M = \left\langle (\mathcal{I} + \mathcal{K})V_0\chi_N, (\mathcal{I} + \mathcal{K})V_0u \right\rangle_M, \quad \forall \chi_N \in S_N.$$

These conditions are equivalent to

$$\left\langle (\mathcal{I} + \mathcal{K})V_0\chi_N, V_0u_N^{(M)} \right\rangle_M = \left\langle (\mathcal{I} + \mathcal{K})V_0\chi_N, V_0 \left(u + V_0^{-1}\mathcal{K}V_0(u - u_N^{(M)})\right) \right\rangle_M, \quad \forall \chi_N \in S_N.$$

Thus by Claim 4.F.3 we have for some $C > 0$

$$\begin{aligned} \|V_0^{-1}(\mathcal{I} + \mathcal{K})V_0 \left(u_N^{(M)} - u\right)\|_{4\alpha-(d+1)} \\ \leq C \left(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)}\right) \|u + V_0^{-1}\mathcal{K}V_0(u - u_N^{(M)})\|_{d+1}. \end{aligned}$$

We note that by continuity of $\mathcal{K} : H^{2\alpha-(d+1)} \rightarrow H^{d+1-2\alpha}$ we again have

$$\|u + V_0^{-1}\mathcal{K}V_0(u - u_N^{(M)})\|_{d+1} \lesssim \|u\|_{d+1} + \|u - u_N^{(M)}\|_{4\alpha-(d+1)}.$$

Moreover, by the assumptions on V_0, \mathcal{K} the map $V_0^{-1}(\mathcal{I} + \mathcal{K})^{-1}V_0 : H^{4\alpha-(d+1)} \rightarrow H^{4\alpha-(d+1)}$ is bounded and, therefore, we have, for some $\tilde{C} > 0$ independent of $u, u_N^{(M)}, M, N$,

$$\left(1 - \tilde{C} \left(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)}\right)\right) \|u - u_N^{(M)}\|_{4\alpha-(d+1)} \leq \tilde{C} \left(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)}\right) \|u\|_{d+1}.$$

Thus, similarly to the proof given in Appendix 4.E, we conclude for N, M sufficiently large the estimate Eq. (4.73) holds, hence completing the proof of Thm. 4.F.1. \square

4.G A perturbation argument for least-squares over-sampled collocation (Single Layer Helmholtz)

In this appendix we will weaken some of the continuity assumptions used in Appendix 4.F in order to show that the following result holds also for the integral operators arising from the single layer formulation of Helmholtz equation on a smooth curve. Specifically,

recalling Thm. 4.4.10, we will prove the following statement. We note that in this appendix we have $\alpha = -1/2$ throughout as we are working with single layer operator.

Theorem 4.G.1. *If $u_N^{(M)} \in S_N$ is such that*

$$\left\langle (\mathcal{I} + \mathcal{K})V_0\chi_N, (\mathcal{I} + \mathcal{K})V_0u_N^{(M)} \right\rangle_M = \langle (\mathcal{I} + \mathcal{K})V_0\chi_N, (\mathcal{I} + \mathcal{K})V_0u \rangle_M, \quad \forall \chi_N \in S_N, \quad (4.84)$$

where V_0 has the form Eq. (4.33) with $\alpha = -1/2$ and $\mathcal{I} + \mathcal{K} : H^s \rightarrow H^s$ is a continuous isomorphism for all $s \in \mathbb{R}$ and $\mathcal{K} : H^s \rightarrow H^{s+1.25}$ with a representation in the form

$$\mathcal{K}u(x) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} k_{m,n} \hat{u}_m e^{2\pi i n x}, \quad x \in [0, 1)$$

for every $u \in L^2$ and $k_{m,n}$ satisfy the estimate Eq. (4.34) for all $s, t \in \mathbb{R}$, then we still have the estimate Eq. (4.36), i.e. there are constants $N_0, C > 0$ independent of N, u, M such that, for all $M \geq N^2$ and $N \geq N_0$,

$$\|u_N^{(M)} - u\|_{2\alpha-(d+1)} \leq CN^{4\alpha-2(d+1)} \|u\|_{d+1}. \quad (4.85)$$

As in Appendix 4.F this formulation of the theorem is equivalent to the perturbation $V = V_0 + \mathcal{K}$ simply by writing

$$V = (\mathcal{I} + \mathcal{K}V_0^{-1})V_0,$$

and by Lemma 4.4.9 the assumptions of this theorem are satisfied by the integral operator arising from a single layer formulation on a smooth curve Γ . We will follow the proof of Thm. 4.F.1 in Appendix 4.F very closely. In particular we will again introduce the bilinear form ϵ :

$$\epsilon(\chi_N, \tilde{b}) := \left\langle V_0\chi_N, \mathcal{K}^*V_0\tilde{b} \right\rangle_M - \left\langle \mathcal{K}V_0\chi_N, V_0\tilde{b} \right\rangle_M.$$

Similarly to Lemma 4.F.2 we will rely on a crucial a priori estimate in order to facilitate the proof of Thm. 4.G.1.

Lemma 4.G.2. *Suppose $a_N^{(M)} \in S_N$ satisfies*

$$\left\langle V_0\chi_N, V_0a_N^{(M)} \right\rangle_M = \langle V_0\chi_N, V_0\tilde{v} \rangle_M + \epsilon(\chi_N, b - c_N), \quad \forall \chi_N \in S_N, \quad (4.86)$$

for some $\tilde{v}, b \in H^{d+1}$ and a sequence of spline functions $c_N \in S_N$, $N \in \mathbb{N}$. Given $r, r_1, r_2 > 1/2$ and $t > 0$, there is a constant $C > 0$ independent of N, M, \tilde{v}, b, c_N such that

for all $M \geq N > 0$:

$$\begin{aligned}
\|\tilde{v} - a_N^{(M)}\|_{4\alpha-(d+1)} &\leq C(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)})\|\tilde{v}\|_{d+1} \\
&\quad + C(M^{-r_1}\|b - c_N\|_{r_1-2.25} + N^{2\alpha-(d+1)+1/2}\|b - c_N\|_{r_2-2.25}) \\
&\quad + CM^{2\alpha-(d+1)}(\|b\|_{d+1} + \|b - c_N\|_{d-1}) \\
&\quad + C(N^{-t} + N^{2\alpha-(d+1)-1/2})N^{-r}\|b - c_N\|_{r+2\alpha}.
\end{aligned}$$

Proof of Lemma 4.G.2. We closely follow the proof of Lemma 4.F.2. Let us write \hat{b}_n, \hat{c}_n for the Fourier coefficients of b, c_N respectively, noting that for all $n \in \Lambda_N^*, l \in \mathbb{Z}$, we have $\hat{c}_{n+lN} = n^{d+1}/(n+lN)^{d+1}\hat{c}_n$. We first recall that for $\mu \in \Lambda_N^*$ we can write the bilinear form ϵ evaluated at $\psi_\mu, b - c_N$ in the following way

$$\begin{aligned}
\epsilon(\psi_\mu, b - c_N) &= \underbrace{\sum_{p \equiv \mu(N)} \sum_{n \equiv p(M), n \neq p} \sum_{m \in \mathbb{Z}} [p]^{2\alpha} \left(\frac{\mu}{p}\right)^{d+1} [m]^{2\alpha} (\hat{b}_m - \hat{c}_m) \overline{k_{nm}}}_{=: A_1} \\
&\quad - \underbrace{\sum_{m \in \mathbb{Z}} \sum_{n \equiv \mu(N)} \sum_{\substack{p \equiv m(M) \\ p \neq m}} [n]^{2\alpha} \left(\frac{\mu}{n}\right)^{d+1} \overline{k_{nm}} [p]^{2\alpha} (\hat{b}_p - \hat{c}_p)}_{=: A_2}.
\end{aligned}$$

In what follows we will bound the remaining two terms A_1, A_2 individually. For this (and the remainder of this appendix) we shall again make use of the notation \lesssim to indicate an implicit constant in the inequality, which is in all cases independent of $a_N^{(M)}, \tilde{v}, b, c_N, N, M$, though it may sometimes depend on other parameters in the inequalities. Where this is of relevance we will indicate this dependence by a subscript, for instance \lesssim_r . Note that it is specifically these estimates which differ from similar steps in the proof of Lemma 4.F.2, because in the present case \mathcal{K} satisfies slightly weaker continuity assumptions. Firstly we have

$$\begin{aligned}
|A_1| &\leq \left| \sum_{n \equiv \mu(M), n \neq \mu} [\mu]^{2\alpha} (\mathcal{K}V_0(\widehat{b - c_N}))_n \right| \\
&\quad + \left| \sum_{p \equiv \mu(N), p \neq \mu} \sum_{n \equiv p(M), n \neq p} [p]^{2\alpha} \left(\frac{\mu}{p}\right)^{d+1} (\mathcal{K}V_0(\widehat{b - c_N}))_n \right|.
\end{aligned}$$

Therefore,

$$\begin{aligned}
|A_1| &\lesssim_{r_1} [\mu]^{2\alpha} M^{-r_1} \|\mathcal{K}V_0(b - c_N)\|_{r_1} \\
&\quad + \sum_{l \neq 0} \sum_{\substack{n \equiv \mu + lN(M) \\ n \neq \mu + lN}} [\mu + lN]^{2\alpha} \left| \frac{\mu}{\mu + lN} \right|^{d+1} \left| (\mathcal{K}V_0(\widehat{b - c_N}))_n \right| \\
&\lesssim_{r_1} [\mu]^{2\alpha} M^{-r_1} \|\mathcal{K}V_0(b - c_N)\|_{r_1} \\
&\quad + \sum_{l \neq 0} [\mu + lN]^{2\alpha} \left| \frac{\mu}{\mu + lN} \right|^{d+1} \sum_{n \equiv \mu + lN(M)} \left| (\mathcal{K}V_0(\widehat{b - c_N}))_n \right|
\end{aligned}$$

for any $r_1 > 1/2$. Therefore,

$$\begin{aligned}
|A_1| &\lesssim_{r_1, r_2} [\mu]^{2\alpha} M^{-r_1} \|\mathcal{K}V_0(b - c_N)\|_{r_1} \\
&\quad + \sum_{l \neq 0} [\mu + lN]^{2\alpha} \left| \frac{\mu}{\mu + lN} \right|^{d+1} \|\mathcal{K}V_0(b - c_N)\|_{r_2}, \tag{4.87}
\end{aligned}$$

for any $r_1, r_2 > 1/2$. For the second term we have

$$\begin{aligned}
|A_2| &= \left| \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left(\frac{\mu}{n} \right)^{d+1} \sum_{m \in \mathbb{Z}} \sum_{\substack{p \equiv m(M) \\ p \neq m}} \overline{k_{nm}} (V_0(\widehat{b - c_N}))_p \right| \\
&\leq \underbrace{\left| \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left(\frac{\mu}{n} \right)^{d+1} \sum_{m = -3N/4}^{3N/4} \sum_{\substack{p \equiv m(M) \\ p \neq m}} \overline{k_{nm}} (V_0(\widehat{b - c_N}))_p \right|}_{=: A_{21}} \\
&\quad + \underbrace{\left| \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left(\frac{\mu}{n} \right)^{d+1} \sum_{|m| > 3N/4} \sum_{\substack{p \equiv m(M) \\ p \neq m, p \notin \Lambda_N}} \overline{k_{nm}} (V_0(\widehat{b - c_N}))_p \right|}_{=: A_{22}}.
\end{aligned}$$

Let us estimate A_{21}, A_{22} separately: Firstly, using the bounds Eq. (4.34) of the Fourier series representation of \mathcal{K} we have for any $s_1, t_1 \in \mathbb{R}$:

$$A_{21} \leq \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} \sum_{m = -3N/4}^{3N/4} \sum_{\substack{p \equiv m(M) \\ p \neq m}} |\overline{k_{nm}}| \left| (V_0(\widehat{b - c_N}))_p \right|.$$

Therefore

$$A_{21} \lesssim_{s_1, t_1} \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} \sum_{m=-3N/4}^{3N/4} \sum_{\substack{p \equiv m(M) \\ p \neq m}} ((1+|n|)^{-s_1+1} (1+|m|)^{-t_1} + (1+|n|) \sum_{l \in \mathbb{Z}} (1+|n-l|)^{-3} (1+|l|)^{-t_1} (1+|m-n-l|)^{-s_1}) \left| (V_0(\widehat{b-c_N}))_p \right|.$$

Moreover, we have for $t_1 > -4\alpha + d + 3/2$:

$$\begin{aligned} & \sum_{m=-3N/4}^{3N/4} (1+|m|)^{-t_1} \sum_{\substack{p \equiv m(M) \\ p \neq m}} [p]^{2\alpha} \left| (\widehat{b-c_N})_p \right| \\ &= \sum_{m \in \Lambda_N} (1+|m|)^{-t_1} \sum_{l \neq 0} [m+lM]^{2\alpha} \left| \hat{b}_{m+lM} - \left(\frac{m}{m+lM} \right)^{d+1} \hat{c}_m \right| \\ &\leq \sum_{m \in \Lambda_N} (1+|m|)^{-t_1} \sum_{l \neq 0} [m+lM]^{2\alpha} \left(|\hat{b}_{m+lM}| + |\hat{c}_m| \left| \frac{m}{m+lM} \right|^{d+1} \right) \\ &\lesssim \sum_{m \in \Lambda_N} (1+|m|)^{-t_1} \sum_{l \neq 0} [m+lM]^{2\alpha-(d+1)} [m+lM]^{d+1} |\hat{b}_{m+lM}| \\ &\quad + M^{2\alpha-(d+1)} \sum_{m \in \Lambda_N} (1+|m|)^{-t_1} |m|^{d+1} |\hat{c}_m|. \end{aligned} \tag{4.88}$$

Therefore, for $t_1 > -4\alpha + d + 3/2$,

$$\begin{aligned} & \sum_{m \in \Lambda_N} (1+|m|)^{-t_1} \sum_{\substack{p \equiv m(M) \\ p \neq m}} [p]^{2\alpha} \left| (\widehat{b-c_N})_p \right| \\ &\lesssim_{t_1} \left(\sum_{m \in \Lambda_N} (1+|m|)^{-t_1} \sum_{l \neq 0} [m+lM]^{4\alpha-2(d+1)} \right)^{\frac{1}{2}} \|b\|_{d+1} \\ &\quad + M^{2\alpha-(d+1)} \|c_N\|_{4\alpha-(d+1)} \\ &\lesssim_{t_1} M^{2\alpha-(d+1)} (\|b\|_{d+1} + \|c_N\|_{4\alpha-(d+1)}) \\ &\lesssim_{t_1} M^{2\alpha-(d+1)} (\|b\|_{d+1} + \|b-c_N\|_{4\alpha-(d+1)} + \|b\|_{4\alpha-(d+1)}) \\ &\lesssim_{t_1} M^{2\alpha-(d+1)} (2\|b\|_{d+1} + \|b-c_N\|_{4\alpha-(d+1)}), \end{aligned} \tag{4.89}$$

where in Eqs. (4.88) & (4.89) we used that $|m| \leq N/2 \leq M/2$ for all $m \in \Lambda_N$ and the consistency condition $d > 2\alpha$, which implies that $d+1-2\alpha > 1$. In the final line we also relied on the consistency condition $d > 2\alpha$, which implies $d+1 > 4\alpha - (d+1)$. Next we

observe that for $t_1 > 3, s_1 > -1$:

$$\begin{aligned}
& \sum_{l' \in \mathbb{Z}} (1 + |m - l'|)^{-3} (1 + |l'|)^{-t_1} (1 + |m - n - l'|)^{-s_1} \leq \sum_{l' \in \mathbb{Z}} (1 + |m - l'|)^{-3} (1 + |l'|)^{-t_1} \\
& \leq (1 + |m|/2)^{-3} \sum_{|l'| \leq m/2} (1 + |l'|)^{-t_1} + (1 + |m|/2)^{-t_1} \sum_{|l'| > n/2} (1 + |m - l'|)^{-3} \\
& \lesssim (1 + |m|)^{-3}.
\end{aligned}$$

Moreover, when $|n| \geq N$ and $|m| < 3N/4$ we have in addition the following estimate

$$\begin{aligned}
& \sum_{l' \in \mathbb{Z}} (1 + |m - l'|)^{-3} (1 + |l'|)^{-t_1} (1 + |m - n - l'|)^{-s_1} \leq \sum_{l' \in \mathbb{Z}} (1 + |l'|)^{-t_1} (1 + |m - n - l'|)^{-s_1} \\
& \leq (1 + |n|/8)^{-s_1} \sum_{|l'| < n/8} (1 + |l'|)^{-t_1} + (1 + |n|/8)^{-s_1} \sum_{|l'| > n/8} (1 + |m - n - l'|)^{-s_1} \\
& \lesssim (1 + |n|)^{-\min\{s_1, t_1\}}.
\end{aligned}$$

Therefore we have, setting the summation index to $l' = m - n - l$, if $|n| < N$:

$$\begin{aligned}
& \sum_{m=-3N/4}^{3N/4} \sum_{\substack{p \equiv m(M) \\ p \neq m}} (1 + |n|) \sum_{l' \in \mathbb{Z}} (1 + |m - l'|)^{-3} (1 + |m - n - l'|)^{-t_1} (1 + |l'|)^{-s_1} \left| (V_0 \widehat{(b - c_N)})_p \right| \\
& \lesssim (1 + |n|) \sum_{m=-3N/4}^{3N/4} (1 + |m|)^{-3} \sum_{\substack{p \equiv m(M) \\ p \neq m}} [p]^{2\alpha} \left| (\widehat{b - c_N})_p \right| \\
& \lesssim_{t_1} N (1 + |n|)^{-2} \left(\sum_{l \neq 0} [m + lM]^{4\alpha - 2(d+1)} \right)^{\frac{1}{2}} \|b\|_{d+1} \\
& \quad + (1 + |n|) M^{2\alpha - (d+1)} \|c_N\|_{d-1} \\
& \lesssim_{t_1} (1 + |n|) M^{2\alpha - (d+1)} (\|b\|_{d+1} + \|c_N\|_{d-1}) \tag{4.90} \\
& \lesssim_{t_1} (1 + |n|) M^{2\alpha - (d+1)} (\|b\|_{d+1} + \|b - c_N\|_{d-1} + \|b\|_{d-1}) \\
& \lesssim_{t_1} (1 + |n|) M^{2\alpha - (d+1)} (2\|b\|_{d+1} + \|b - c_N\|_{d-1}),
\end{aligned}$$

and, if $|n| \geq N$,

$$\begin{aligned}
& \sum_{m=-3N/4}^{3N/4} \sum_{\substack{p \equiv m(M) \\ p \neq m}} (1 + |n|) \sum_{l' \in \mathbb{Z}} (1 + |m - l'|)^{-3} (1 + |m - n - l'|)^{-t_1} (1 + |l'|)^{-s_1} \left| (V_0 \widehat{(b - c_N)})_p \right| \\
& \lesssim_{t_1, s_1} |n|^{-\min\{s_1, t_1\} + 1} M^{2\alpha - (d+1)} (2\|b\|_{d+1} + \|b - c_N\|_{d-1}),
\end{aligned}$$

and so on the whole we estimate

$$A_{21} \lesssim_{s_1, t_1} \left(\sum_{n \equiv \mu(N), |n| < N} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} (1 + |n|) + \sum_{n \equiv \mu(N), |n| \geq N} [n]^{-\min\{s_1, t_1\}} \left| \frac{\mu}{n} \right|^{d+1} \right) M^{2\alpha - (d+1)} (2\|b\|_{d+1} + \|b - c_N\|_{d-1}). \quad (4.91)$$

For A_{22} we have the following upper bound:

$$\begin{aligned} A_{22} &\leq \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} \sum_{|m| \geq 3N/4} \sum_{\substack{p \equiv m(M) \\ p \neq m, p \notin \Lambda_N}} |\overline{k_{nm}}| |(V_0(\widehat{b - c_N}))_p| \\ &\lesssim_{s_2, t_2} \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} \sum_{|m| \geq 3N/4} \sum_{\substack{p \equiv m(M) \\ p \neq m, p \notin \Lambda_N}} \left((1 + |n|)^{-s+1} (1 + |m|)^{-t} \right. \\ &\quad \left. (1 + |n|) \sum_{l \in \mathbb{Z}} (1 + |n - l|)^{-3} (1 + |l|)^{-t} (1 + |n - m - l|)^{-s} \right) |(V_0(\widehat{b - c_N}))_p|. \end{aligned}$$

Firstly, analogously to Eq. (4.81) we have for $t > 3/2 - \delta + 2(d+1) - 4\alpha$,

$$\sum_{|m| \geq 3N/4} \sum_{\substack{p \equiv m(M) \\ p \neq m, p \notin \Lambda_N}} (1 + |m|)^{-t} |(V_0(\widehat{b - c_N}))_p| \lesssim_{s, t} (N^{4\alpha - 2(d+1)} \|b\|_{d+1} + N^{-1} \|b - c_N\|_{4\alpha - (d+1)}).$$

Secondly, let m' be the unique integer with $m' \in \Lambda_N$ and $m' \equiv m(N)$, then for an arbitrary set of Fourier coefficients \hat{f}_p and any $r > 1/2$

$$\sum_{p \equiv m(M), p \neq m, p \notin \Lambda_N} |\hat{f}_p| \leq \sum_{p \equiv m'(N), p \neq m'} |\hat{f}_p| \lesssim N^{-r} \|f\|_r.$$

Therefore

$$\begin{aligned} &\sum_{|m| \geq 3N/4} \sum_{\substack{p \equiv m(M) \\ p \neq m, p \notin \Lambda_N}} \sum_{l \in \mathbb{Z}} (1 + |n - l|)^{-3} (1 + |l|)^{-t} (1 + |n - m - l|)^{-s} |(V_0(\widehat{b - c_N}))_p| \\ &\lesssim_r \sum_{|m| \geq 3N/4} \sum_{l \in \mathbb{Z}} (1 + |n - l|)^{-3} (1 + |l|)^{-t} (1 + |m - n - l|)^{-s} N^{-r} \|V_0(b - c_N)\|_r. \end{aligned}$$

Now in addition we can estimate

$$\begin{aligned}
& \sum_{n \equiv \mu(N)} [n]^{2\alpha} \left| \frac{\mu}{n} \right|^{d+1} (1 + |n|) \sum_{|m| \geq 3N/4} \sum_{l \in \mathbb{Z}} (1 + |n - l|)^{-3} (1 + |l|)^{-t} (1 + |m - n - l|)^{-s} \\
& \lesssim [\mu]^{2\alpha} (1 + |\mu|) \sum_{|m| \geq 3N/4} \sum_{l \in \mathbb{Z}} (1 + |\mu - l|)^{-3} (1 + |l|)^{-t} (1 + |\mu - m - l|)^{-s} \\
& + \sum_{j \neq 0} \left| \frac{\mu}{\mu + jN} \right|^{d+1} \sum_{|m| \geq 3N/4} \sum_{l \in \mathbb{Z}} (1 + |\mu + jN - l|)^{-3} (1 + |l|)^{-t} (1 + |\mu + jN - m - l|)^{-s}
\end{aligned}$$

Note that $\mu - m > N/4$ thus the first term on the right hand side is, for s, t sufficiently large

$$\lesssim [\mu]^{2\alpha} N \sum_{|m| \geq 3N/4} \sum_{l \in \mathbb{Z}} (1 + |l|)^{-t} (1 + |\mu - m - l|)^{-s} \lesssim_{s,t} [\mu]^{2\alpha} N^{-t+1}$$

We can also estimate the second term as follows (note $2\alpha = -1$):

$$\begin{aligned}
& \sum_{j \neq 0} \left| \frac{\mu}{\mu + jN} \right|^{d+1} \sum_{|m| \geq 3N/4} \sum_{l \in \mathbb{Z}} (1 + |\mu + jN - l|)^{-3} (1 + |l|)^{-t} (1 + |\mu + jN - m - l|)^{-s} \\
& \lesssim \sum_{j \neq 0} \left| \frac{\mu}{\mu + jN} \right|^{d+1} (1 + |\mu + jN|)^{-3} \sum_{|m| \geq 3N/4} \\
& \lesssim \sum_{j \neq 0} \left| \frac{\mu}{\mu + jN} \right|^{d+1} (1 + |\mu + jN|)^{-3} N \lesssim [\mu]^{d+1} N^{-(d+1)-2}
\end{aligned}$$

So overall we estimate

$$A_{22} \lesssim ([\mu]^{2\alpha} N^{-t+1} + [\mu]^{d+1} N^{-(d+1)-2}) N^{-r} \|b - c_N\|_{r+2\alpha}. \quad (4.92)$$

A similar bound for $|\epsilon(\psi_\mu, b - c_N)|^2$ in the case $\mu = 0$ can be found by simply replacing the sums over $\sum_{n \equiv \mu(N)}$ by the unique choice $n = 0$ with no summation. We can now combine Eqs. (4.87), (4.91) & (4.92) to find the following estimate for $r, r_1, r_2 > 1/2, t > 0$ and some constant $C = C_{r_1, r_2, r, t}$:

$$\begin{aligned}
\sum_{\mu \in \Lambda_N} [\mu]^{-2(d+1)} |\epsilon(\psi_\mu, b - c_N)|^2 &\leq C \sum_{\mu \in \Lambda_N} [\mu]^{4\alpha-2(d+1)} M^{-2r_1} \|\mathcal{K}V_0(b - c_N)\|_{r_1}^2 \\
&+ C \sum_{\mu \in \Lambda_N} \left(\sum_{l \neq 0} [\mu + lN]^{2\alpha-(d+1)} \right)^2 \|\mathcal{K}V_0(b - c_N)\|_{r_2}^2 \\
&+ C \sum_{\mu \in \Lambda_N} \left(\sum_{n \equiv \mu(N), |n| < N} [n]^{-(d+1)} + \sum_{n \equiv \mu(N), |n| \geq N} [n]^{-t} \right)^2 \\
&\quad (M^{2\alpha-(d+1)} (2\|b\|_{d+1} + \|b - c_N\|_{d-1}))^2 \\
&+ C \sum_{\mu \in \Lambda_N} [\mu]^{-2(d+1)} ([\mu]^{4\alpha} N^{-2t} + [\mu]^{2(d+1)} N^{4\alpha-2(d+1)-2}) N^{-2r} \|b - c_N\|_{r+2\alpha}^2.
\end{aligned}$$

Therefore we have by the continuity properties of \mathcal{K} , for N sufficiently large,

$$\begin{aligned}
\sum_{\mu \in \Lambda_N} [\mu]^{-2(d+1)} |\epsilon(\psi_\mu, b - c_N)|^2 &\lesssim (M^{-2r_1} \|b - c_N\|_{r_1-2.25}^2 + N^{4\alpha-2(d+1)+1} \|b - c_N\|_{r_2-2.25}^2) \\
&+ M^{4\alpha-2(d+1)} (\|b\|_{d+1}^2 + \|b - c_N\|_{d-1}^2) \\
&+ (N^{-2t} + N^{4\alpha-2(d+1)-1}) N^{-2r} \|b - c_N\|_r^2.
\end{aligned}$$

Similar to the proof of Lemma 4.F.2 we can combine this estimate with the estimates Eqs. (4.45) & (4.41) to bound the solution of the linear system Eq. (4.86) in the following way

$$\begin{aligned}
\|a_N^{(M)} - \tilde{v}\|_{4\alpha-(d+1)} &\leq C(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)}) \|\tilde{v}\|_{d+1} \\
&+ C(M^{-r_1} \|b - c_N\|_{r_1-2.25} + N^{2\alpha-(d+1)+1/2} \|b - c_N\|_{r_2-2.25}) \\
&+ CM^{2\alpha-(d+1)} (\|b\|_{d+1} + \|b - c_N\|_{d-1}) \\
&+ C(N^{-t} + N^{2\alpha-(d+1)-1/2}) N^{-r} \|b - c_N\|_{r+2\alpha}
\end{aligned}$$

for some constant $C > 0$ independent of $a_N^{(M)}$, \tilde{v} , N , M , b , c_N , which completes the proof of Lemma 4.G.2. \square

We can now use Lemma 4.G.2 to prove Thm. 4.G.1 in a manner closely related to the proof of Thm. 4.F.1.

Proof of Thm. 4.G.1. We proceed in two steps: Firstly we show that a perturbation of the test functions $V_0\chi_N \mapsto (\mathcal{I} + \mathcal{K})V_0\chi_N$ yields a similar error estimate as in Thm. 4.4.10 and then we proceed to perturb the operator V_0 on the right hand side of the orthogonality conditions.

Claim 4.G.3. Suppose $a_N^{(M)} \in S_N$ satisfies

$$\left\langle (1 + \mathcal{K})V_0\chi_N, V_0a_N^{(M)} \right\rangle_M = \langle (1 + \mathcal{K})V_0\chi_N, V_0\tilde{a} \rangle_M, \quad \forall \chi_N \in S_N, \quad (4.93)$$

where \mathcal{K} satisfies the assumptions of Thm. 4.G.1. Then there are constants $C, N_0 > 0$ independent of $\tilde{a}, a_N^{(M)}$ such that for $M \geq N^2, N \geq N_0$:

$$\|a_N^{(M)} - \tilde{a}\|_{4\alpha-(d+1)} \leq CN^{4\alpha-2(d+1)}\|\tilde{a}\|_{d+1}. \quad (4.94)$$

Proof. To begin with we note the conditions Eq. (4.82) are equivalent to

$$\left\langle V_0\chi_N, (\mathcal{I} + \mathcal{K}^*)V_0a_N^{(M)} \right\rangle_M = \langle V_0\chi_N, (\mathcal{I} + \mathcal{K}^*)V_0\tilde{a} \rangle_M + \epsilon(\chi_N, a_N^{(M)} - \tilde{a}), \quad \forall \chi_N \in S_N.$$

This can be equivalently written as

$$\left\langle V_0\chi_N, V_0a_N^{(M)} \right\rangle_M = \left\langle V_0\chi_N, V_0 \left(\tilde{a} + V_0^{-1}\mathcal{K}^*V_0(\tilde{a} - a_N^{(M)}) \right) \right\rangle_M + \epsilon(\chi_N, a_N^{(M)} - \tilde{a}), \quad \forall \chi_N \in S_N.$$

Therefore, Lemma 4.G.2 applies and shows that

$$\begin{aligned} & \|V_0^{-1}(\mathcal{I} + \mathcal{K}^*)V_0(\tilde{a} - a_N^{(M)})\|_{4\alpha-(d+1)} \\ & \leq C \left(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)} \right) \|\tilde{a} + V_0^{-1}\mathcal{K}^*V_0(\tilde{a} - a_N^{(M)})\|_{d+1} \\ & + C(M^{-r_1}\|\tilde{a} - a_N^{(M)}\|_{r_1-2.25} + N^{2\alpha-(d+1)+1/2}\|\tilde{a} - a_N^{(M)}\|_{r_2-2.25}) \\ & + CM^{2\alpha-(d+1)} \left(\|\tilde{a}\|_{d+1} + \|\tilde{a} - a_N^{(M)}\|_{d-1} \right) \\ & + C(N^{-t} + N^{2\alpha-(d+1)-1/2}) N^{-r} \|\tilde{a} - a_N^{(M)}\|_{r+2\alpha} \end{aligned} \quad (4.95)$$

Let $r_1 = d + 1.25$ then we have (using the usual estimates exploiting Assumptions 4.2.2 & 4.2.3):

$$M^{-r_1}\|\tilde{a} - a_N^{(M)}\|_{r_1-2.25} \lesssim M^{-(d+5/4)} \left(N^{-2}\|\tilde{a}\|_{d+1} + N^{2(d+1)-4}\|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)} \right)$$

For the second term we take $r_2 = 3/4$ then we have similarly

$$\begin{aligned} & N^{2\alpha-(d+1)+1/2}\|\tilde{a} - a_N^{(M)}\|_{r_2-2.25} \\ & \lesssim N^{2\alpha-(d+1)+1/2} \left(N^{-3/2-(d+1)}\|\tilde{a}\|_{d+1} + N^{-3/2-4\alpha+(d+1)}\|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)} \right) \\ & \lesssim N^{4\alpha-2(d+1)}\|\tilde{a}\|_{d+1} + N^{-1/2}\|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)} \end{aligned}$$

In a similar manner (again using Assumptions 4.2.2 & 4.2.3) we can estimate:

$$\begin{aligned}
M^{2\alpha-(d+1)}\|\tilde{a} - a_N^{(M)}\|_{d-1} &\lesssim M^{2\alpha-(d+1)} \left(\|\tilde{a} - \chi_N\|_{d-1} + \|\chi_N - a_N^{(M)}\|_{d-1} \right) \\
&\lesssim M^{2\alpha-(d+1)} \left(N^{-2}\|\tilde{a}\|_{d+1} + N^{-4\alpha-2d}\|\chi_N - a_N^{(M)}\|_{4\alpha-(d+1)} \right) \\
&\lesssim M^{2\alpha-(d+1)}N^{-2}\|\tilde{a}\|_{d+1} + M^{2\alpha-(d+1)}N^{-4\alpha-2d}\|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)}.
\end{aligned}$$

Finally, we also have for $t > -2\alpha + d + 3/2$ and any $r > 1/2$:

$$\begin{aligned}
&(N^{-t} + N^{2\alpha-(d+1)-1/2}) N^{-r}\|\tilde{a} - a_N^{(M)}\|_{r+2\alpha} \\
&\lesssim N^{2\alpha-(d+1)-1/2-r} \left(N^{r+2\alpha-(d+1)}\|\tilde{a}\|_{d+1} + N^{r+2\alpha-4\alpha+(d+1)}\|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)} \right) \\
&\lesssim N^{4\alpha-2(d+1)-1/2}\|\tilde{a}\|_{d+1} + N^{-1/2}\|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)}
\end{aligned}$$

Therefore the estimate Eq. (4.95) simplifies for N sufficiently large to

$$\begin{aligned}
&\|V_0^{-1}(\mathcal{I} + \mathcal{K}^*)V_0(\tilde{a} - a_N^{(M)})\|_{4\alpha-(d+1)} \\
&\leq C \left(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)} \right) \|\tilde{a} + V_0^{-1}\mathcal{K}^*V_0(\tilde{a} - a_N^{(M)})\|_{d+1} \\
&\quad + C \left(M^{-(d+5/4)}N^{-2} + N^{4\alpha-2(d+1)} + M^{2\alpha-(d+1)} \right) \|\tilde{a}\|_{d+1} \\
&\quad + C \left(M^{-(d+5/4)}N^{2(d+1)-4} + N^{-1/2} + M^{2\alpha-(d+1)}N^{-4\alpha+2(d+1)}N^{-2} \right) \|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)}
\end{aligned}$$

Noting that $V_0 : H^{4\alpha-(d+1)} \rightarrow H^{4\alpha-(d+1)}$ is a continuous isomorphism, and that $\mathcal{I} + \mathcal{K}^* : H^{2\alpha-(d+1)} \rightarrow H^{2\alpha-(d+1)}$ is invertible, there is a constant $\tilde{C} > 0$ such that

$$\begin{aligned}
\|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)} &\leq \tilde{C} \left(M^{2\alpha-(d+1)} + N^{4\alpha}M^{-(d+1)-1/4} + N^{4\alpha-2(d+1)} \right) \|\tilde{a}\|_{d+1} \\
&\quad + \tilde{C} \left(M^{2\alpha-(d+1)} + N^{4\alpha-2(d+1)} + M^{-(d+5/4)}N^{2(d+1)-4} \right. \\
&\quad \left. + N^{-1/2} + M^{2\alpha-(d+1)}N^{-4\alpha+2(d+1)}N^{-2} \right) \|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)}.
\end{aligned}$$

Letting $M \geq N^2$ we find

$$\begin{aligned}
\|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)} &\leq \tilde{C}N^{4\alpha-2(d+1)}\|\tilde{a}\|_{d+1} \\
&\quad + \tilde{C} \left(N^{4\alpha-2(d+1)} + N^{-9/2} + N^{-1/2} + N^{-2} \right) \|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)}.
\end{aligned}$$

Equivalently, for N sufficiently large, there is a constant $\tilde{C} > 0$ such that,

$$\left(1 - \tilde{C} \left(N^{4\alpha-2(d+1)} + N^{-1/2} \right) \right) \|\tilde{a} - a_N^{(M)}\|_{4\alpha-(d+1)} \leq \tilde{C}N^{4\alpha-2(d+1)}\|\tilde{a}\|_{d+1}.$$

The we conclude that, for $M \geq N^2$ and N sufficiently large, the estimate Eq. (4.94)

holds. ■

Having proved Claim 4.G.3 we can proceed to prove Thm. 4.G.1 in an analogous way to the arguments in Appendix 4.E: Suppose $u_N^{(M)} \in S_N$ satisfies

$$\left\langle (\mathcal{I} + \mathcal{K})V_0\chi_N, (\mathcal{I} + \mathcal{K})V_0u_N^{(M)} \right\rangle_M = \left\langle (\mathcal{I} + \mathcal{K})V_0\chi_N, (\mathcal{I} + \mathcal{K})V_0u \right\rangle_M, \quad \forall \chi_N \in S_N.$$

These conditions are equivalent to

$$\left\langle (\mathcal{I} + \mathcal{K})V_0\chi_N, V_0u_N^{(M)} \right\rangle_M = \left\langle (\mathcal{I} + \mathcal{K})V_0\chi_N, V_0 \left(u + V_0^{-1}\mathcal{K}V_0(u - u_N^{(M)}) \right) \right\rangle_M, \quad \forall \chi_N \in S_N.$$

Thus by Claim 4.G.3 we have for some $C > 0$

$$\begin{aligned} \|V_0^{-1}(\mathcal{I} + \mathcal{K})V_0 \left(u_N^{(M)} - u \right)\|_{4\alpha-(d+1)} \\ \leq CN^{4\alpha-2(d+1)} \|u + V_0^{-1}\mathcal{K}V_0(u - u_N^{(M)})\|_{d+1}. \end{aligned}$$

We note that by continuity of $\mathcal{K} : H^{d+1-2\alpha-1} \rightarrow H^{d+1-2\alpha}$ we again have (using the Assumptions 4.2.2 & 4.2.3)

$$\begin{aligned} \|u + V_0^{-1}\mathcal{K}V_0(u - u_N^{(M)})\|_{d+1} &\lesssim \|u\|_{d+1} + \|u - u_N^{(M)}\|_d \\ &\lesssim N^{-1}\|u\|_{d+1} + N^{2(d+1)-4\alpha-1}\|u - u_N^{(M)}\|_{4\alpha-(d+1)}. \end{aligned}$$

Moreover, by the assumptions on V_0, \mathcal{K} the map $V_0^{-1}(\mathcal{I} + \mathcal{K})^{-1}V_0 : H^{4\alpha-(d+1)} \rightarrow H^{4\alpha-(d+1)}$ is bounded and, therefore, we have, for some $\tilde{C} > 0$ independent of $u, u_N^{(M)}, M, N$,

$$\left(1 - \tilde{C}N^{-1}\right) \|u - u_N^{(M)}\|_{4\alpha-(d+1)} \leq \tilde{C}N^{4\alpha-2(d+1)}\|u\|_{d+1}.$$

Thus, we conclude for $M \geq N^2$ and N sufficiently large the estimate Eq. (4.85) holds, hence completing the proof of Thm. 4.G.1. □

Chapter 5

Recursive moment computation in Filon methods

5.1 Introduction

In Chapter 4 we saw how efficient (oversampled) collocation methods may be constructed for wave scattering problems on compact obstacles. An essential step in the implementation of these methods is the assembly of the collocation matrix entries, which are given as integrals of the kernel of the layer operators (i.e. the Green's function of the Helmholtz equation or its normal derivative) against the basis functions of the trial space (denoted by S_N in Chapter 4). As the frequency increases the spline bases discussed in the previous chapter are no longer able to capture the solution efficiently (essentially because the dimension of the trial space would have to increase linearly with the frequency of the incident wave in order to preserve the quality of approximation). As we discussed in §1.1.2, a way to overcome this is by introducing oscillatory elements in the trial space, hence leading to so-called hybrid numerical-asymptotic boundary integral methods. In this setting the trial space consists of functions with predefined oscillations, and the degrees of freedom are limited to the approximation of slowly varying amplitudes. However this brings the challenge that basis functions are supported on a range much larger than a single wavelength and therefore the assembly of the collocation matrix requires the integration of these oscillatory basis functions against the singular, oscillatory Green's function (see for instance Chandler-Wilde, Graham, Langdon and Spence, 2012, §4), creating a demand for the efficient approximation of highly-oscillatory integrals.

While our motivation in this thesis comes from wave scattering, we note that highly oscillatory integrals arise in the simulation of many physical systems involving high-frequency phenomena, including quantum physics and laser dynamics (Iserles et al., 2019*a,b*), and in electronic engineering (Condon et al., 2009). Although efficient methods for the computation of certain highly oscillatory integrals had been discovered as early

as the first half of the twentieth century by Louis Napoleon George Filon (1930), as we mentioned in §1.2.3, and thorough research over the past decades has led to an immense increase in efficiency and applicability of such methods (Deaño et al., 2017), many open problems remain. Amongst them is the so-called ‘moment-problem’, i.e. the computation of Filon quadrature moments, which can be given in simple explicit form only for isolated instances of interpolation bases and oscillators. This hinders the direct application of Filon methods to several important classes of integrals which involve complicated oscillatory kernel functions.

In the present chapter, we address this problem by providing a method for the construction of recursive relations satisfied by the Filon quadrature moments, which leads to a very efficient strategy for finding the moments in a range of settings. This is a continuation of extensive work on Filon methods over the past two decades which was started by Iserles (2004, 2005) and Iserles and Nørsett (2004) who were the first to provide a detailed asymptotic error analysis and an extension of the ideas presented by Filon (1930) specifically describing the favourable asymptotic properties of Filon methods in the high-frequency regime. Improvements to the Filon method by reducing the asymptotic error were introduced by Iserles and Nørsett (2005) through including information about the derivative values of the amplitude function f , and resulted in the development of the extended Filon method which was studied in greater detail by Gao and Iserles (2017*a,b*). With the goal to understand the quadrature error uniformly in the frequency, it was shown by Melenk (2010) that the error analysis of Filon methods for non-stationary oscillators can, in essence, be reduced to the study of the interpolation error of the amplitude function at the relevant quadrature points. Both Melenk (2010) (based on analyticity properties in a neighbourhood of the domain of integration) and Domínguez et al. (2011) (based on the regularity of the amplitude in certain periodic Sobolev spaces) use this observation to provide error estimates that are explicit in the frequency of oscillations as well as the number of interior quadrature points in the non-stationary case.

The interest in extending these Filon methods from simple linear oscillators to more general kernels has led to work by Olver (2006, 2007), who described a moment-free version of the Filon method that is applicable to algebraic singularities and stationary points. A different type of approach involves recursive moment computation, which has been successfully applied to a number of individual cases in the context of Clenshaw–Curtis interior points: for integrals involving Bessel functions of linear arguments by Piessens and Branders (1983), for linear oscillators by Domínguez et al. (2011), and for linear oscillators with logarithmic singularities in the form $\log|x - \alpha|$ by Domínguez (2014). In a related context of computing indefinite integrals over oscillatory and singular functions using a Levin-type method, Keller (1999, 2007) described a method for the recursive computation of Chebyshev coefficients for functions that satisfy a linear differential

equation with polynomial coefficients. This is based on earlier work by Lewanowicz (1991) on the recursive computation of Jacobi coefficients of special functions satisfying similar differential equations. The ideas underpinning these final three studies, namely that the null space of certain differential operators can be related to expansions in a Hilbert basis whose coefficients satisfy recurrences, are closely related to Thm. 5.3.2 in this chapter.

In recent years, high-frequency wave scattering has provided strong motivation for further advances in the development of highly oscillatory quadrature. Domínguez et al. (2013) constructed a composite (graded) version of the Filon method that can be applied to arbitrary algebraic and logarithmic singularities, and which has been successfully applied to hybrid numerical-asymptotic methods in wave scattering by Chandler-Wilde, Graham, Langdon and Spence (2012), Kim (2012) and Parolin (2015). While this method is already significantly better than traditional quadrature, the flexibility of this graded method comes at the price of losing some of the favourable asymptotic properties of Filon methods, and we will see in the present chapter how this may be overcome in certain cases by the construction of a direct Filon method for the corresponding integrals. An alternative approach to computing highly oscillatory integrals is the method of numerical steepest descent which was introduced by Huybrechs and Vandewalle (2006). Numerical steepest descent has recently been applied by Gibbs et al. (2020) to wave scattering problems on multiple screens (see Gibbs (2020*a,b*)). These results serve as a motivation and reference for the application of our methods to a collocation method in high-frequency wave scattering in §5.5.5.

The structure and main contributions of this chapter are as follows. We begin with a general description of the extended Filon method as introduced by Gao and Iserles (2017*a,b*) in §5.2 building on what we have discussed in §1.2.3. This sets the scene for the ‘moment-problem’ in Filon methods, and specifically for our first main result, Thm. 5.3.2, which we prove in §5.3. The theorem provides a set of sufficient conditions for Filon moments to satisfy recurrences, and is based on the observation that many relevant interpolation bases are in fact a (scaled) Hilbert basis of a weighted L^2 -space and that several relevant oscillators satisfy differential equations of special form. Following two instructive examples, we focus on applying this methodology to Filon–Clenshaw–Curtis methods in §5.4 & §5.5.

Specifically, in §5.4 we construct a direct Filon–Clenshaw–Curtis method for integrals with either a stationary point or an algebraic singularity. In this section our main results are the rigorous stability analysis for the corresponding moment recurrences in Thms. 5.4.1 & 5.4.2. Although we focus our attention to the initial stability regime (which is most relevant for practical computations) we also indicate how one may use Oliver’s algorithm (Oliver, 1968) for the stable computation of the tail. The computation of the tail of the recurrence is not studied in detail since it is mostly of theoretical interest, as in practice

the computational advantage of Filon methods over classical quadrature exists only when the number of required moments is smaller than the frequency of the oscillator, $N \lesssim \omega$. Numerical examples are included in §5.4.4 demonstrating the advantage of this direct application of the Filon method over composite versions.

The second major application of our methodology is described in §5.5, where we consider the direct construction of Filon methods for hybrid numerical-asymptotic collocation methods for high-frequency wave scattering on a screen. The first step in this construction is the proof of a Filon paradigm in Prop. 5.5.2, which is a simple, but non-trivial result describing the asymptotic behaviour of the integral over the combination of a linear exponential oscillator and a Hankel function, the latter of which has a frequency dependent singularity in the domain of integration. This facilitates the study of error estimates that are explicit in both frequency and number of interior points in Corollary 5.5.3. Although the stability analysis for the relevant moment recurrences is non-tractable for theoretical study, the recursive moment evaluation provides an enormous improvement in practical computations when combined with the expressions for initial moments found in Lemma 5.5.4. We evaluate the practical performance of the method based on an example of a hybrid numerical-asymptotic method describing the scattering of a Gaussian beam by a finite screen in §5.5.5.

Our results are summarised in the concluding remarks in §5.6.

5.2 The extended Filon method

We begin with a review of the extended Filon method as introduced by Iserles and Nørsett (2005) and Gao and Iserles (2017*a,b*) based on the following generic form of a one-dimensional oscillatory integral:

$$I_\omega[f] = \int_a^b f(x)h_\omega(x) dx, \quad -\infty < a < b < \infty.$$

Here the kernel function $h_\omega(x)$ is an ω -oscillatory function and Filon quadrature methods are designed to approximate $I_\omega[f]$ with good accuracy and at uniform cost when $\omega \gg 1$. One may think of the example $h_\omega(x) = \exp(i\omega g(x))$ with a suitable choice of g , although in §5.5 we will also study the possibility when $h_\omega(x)$ is expressed in terms of certain special functions. Iserles and Nørsett (2005) observed that for $h_\omega(x) = \exp(i\omega g(x))$ with $g'(x) \neq 0, x \in (a, b)$, the asymptotic expansion of $I_\omega[f]$ for large ω depends only on the values $\mathcal{S} = \{f^{(j)}(a), f^{(j)}(b) | j = 0, 1, \dots\}$, if $f \in C^\infty([a, b])$, and thus proposed to construct a Filon quadrature method by computing $\mathcal{Q}_\omega^{[\nu, \mathcal{S}]}[f] := I_\omega[p]$, where p is an interpolating

polynomial of degree $2s + \nu + 1$ satisfying the Hermite-type interpolation conditions

$$p^{(j)}(\pm 1) = f^{(j)}(\pm 1), \quad j = 0, \dots, s, \quad \text{and} \quad p(c_l) = f(c_l), \quad l = 1, \dots, \nu, \quad (5.1)$$

for some specified interior interpolation points $a = c_0 < c_1 < \dots < c_\nu < c_{\nu+1} = b$. Since the asymptotic behaviour of $I_\omega[f]$ is determined by the values in \mathcal{S} , one can show that the asymptotic error of this quadrature method is

$$|\mathcal{Q}_\omega^{[\nu, s]}[f] - I_\omega[f]| = \mathcal{O}(\omega^{-s-2}), \quad \omega \rightarrow \infty,$$

i.e. it can be made to decay at an arbitrary algebraic rate in ω so long as f possesses a sufficient number of derivatives on $[a, b]$ (cf. Eq. (5.16) and Prop. 5.5.2). This idea extends more generally also to oscillators with stationary points and to higher dimensions: as long as the derivative values of p match those of f on a specified finite set of points (consisting of stationary points, hidden stationary points, and singularities) up to a certain order, the asymptotic error of the quadrature method $\mathcal{Q}_\omega[f] := I_\omega[p]$ decays in ω . For more details we refer the reader to (Deaño et al., 2017, §4).

5.2.1 The Achilles' heel of Filon methods: Moment computation

In practice the interpolation problem Eq. (5.1) is solved by finding the coefficients of p with respect to a given set of interpolation basis functions $\{\phi_n\}_{n=0}^{\nu+2s+1}$. These are typically polynomials on finite intervals, but we shall see an example below where ϕ_n can be taken to be Fourier modes (see Example 5.3.3). Although formally there is no difference in the particular choice of finite degree polynomial basis, the conditioning and (as explained in §5.2.2) the speed of the interpolation algorithm are affected. Depending on the choice of interior nodes $c_j, j = 1, \dots, \nu$, it is therefore often advantageous to express the interpolating polynomial in terms of its coefficients with respect to a specific basis of orthogonal polynomials. Following the solution of the interpolation problem Eq. (5.1) the polynomial p is thus expressed in the form

$$p(x) = \sum_{n=0}^{\nu+2s+1} p_n \phi_n(x),$$

and we compute the Filon quadrature as $\mathcal{Q}_\omega^{[\nu, s]}[f] = I_\omega[p] = \sum_{n=0}^{\nu+2s+1} p_n I_\omega[\phi_n]$. Hence an important step in the Filon method is to evaluate the Filon quadrature moments $\{I_\omega[\phi_n]\}_{n=0}^{\nu+2s+1}$. For a general oscillator $h_\omega(x)$ this task is extremely tricky, mainly because explicit expressions for the moments are only rarely available, or given in terms of special functions that are computationally expensive to evaluate. Thus it is fair to say that the computation of moments is the ‘Achilles’ heel’ of Filon methods.

5.2.2 Fast interpolation at Filon–Clenshaw–Curtis points

Before embarking on a more thorough study of recursive moment computation, let us consider the choice of interior interpolation points in greater detail. A good choice of interior points c_n is determined by a number of competing goals: accuracy for $\omega \gg 1$, uniform accuracy, simplicity of coefficients, and, for large values of ν , minimization of computation cost for the interpolation problem. This aspect was studied in Deaño et al. (2017, §4.2), and it was first shown by Domínguez et al. (2011) that for finite intervals $[a, b]$ a particularly interesting choice of intermediate interpolation points are shifted Clenshaw–Curtis points: Suppose without loss of generality $a = -1, b = 1$, then the Clenshaw–Curtis points (in this case no shift is required since $[-1, 1]$ is the standard domain of Clenshaw–Curtis points) are given by

$$c_n = \cos\left(\frac{n\pi}{\nu+1}\right), \quad n = 0, \dots, \nu+1. \quad (5.2)$$

It is well-known that for $s = 0$ the solution of the interpolation problem Eq. (5.1) can be expressed as a finite linear combination of Chebyshev polynomials T_n using a single application of a Discrete Cosine Transform, DCT-I, bringing the cost of the interpolation part of the Filon method to just $\mathcal{O}(\nu \log \nu)$ operations. The work by Domínguez et al. (2011) is particularly close to the present considerations in this chapter as it proposed a recursive approach for computing the Chebyshev moments $I_\omega[T_n]$ for an exponential oscillator with linear phase, $\exp(i\omega x)$. Our present results can be seen as a generalisation of this recursive approach. The advantageous interpolation properties of interior Chebyshev points carry over to nonzero values of s , as was shown by Gao and Iserles (2017a): the interpolation problem Eq. (5.1) for p in a basis of Chebyshev polynomials can be solved very efficiently by a single application of DCT-I, and the solution of a small auxiliary linear system of size $2s \times 2s$, bringing the overall cost of interpolation to $\mathcal{O}(\nu \log \nu + \nu s + s^3)$. For full details on this procedure we refer the reader to the original work (Gao and Iserles, 2017a).

As we indicated in §1.2.3, it is possible for the asymptotic behaviour of the integral $I_\omega[f]$ to also depend on interior derivative values of f , for instance in the presence of a stationary point of h_ω . Subject to a few minor modifications, which we describe in Appendix 5.A, the aforementioned procedure can also be used to solve efficiently the following interpolation problem: Compute q , the unique polynomial of degree $3s + \nu + 1$ such that

$$\begin{aligned} q^{(j)}(0) &= f^{(j)}(0), \quad q^{(j)}(\pm 1) = f^{(j)}(\pm 1), \quad j = 0, \dots, s, \\ \text{and} \quad q(c_n) &= f(c_n), \quad n = 1, \dots, \nu, \end{aligned} \quad (5.3)$$

where c_n are as in Eq. (5.2) and we take ν to be odd, to ensure that $0 = c_{(\nu-1)/2}$ is amongst the interior points. Indeed, as was the case for Eq. (5.1), the Chebyshev coefficients of q can again be found in $\mathcal{O}(\nu \log \nu + \nu s + s^3)$ operations. We will use this result when considering integrals with stationary points and algebraic singularities at the origin in §5.4.

5.3 Recursive moment computation in Filon methods

As remarked above, a crucial step in the Filon method is the accurate and efficient computation of the quadrature moments $I_\omega[\phi_n]$. In this section, we present a constructive result that can be used to find recurrences for these quadrature moments in a range of settings, by regarding them as the coefficients in a Hilbert basis formed by the (appropriately scaled) interpolation basis. This generally yields a very efficient way for computing Filon moments, provided initial conditions for the recurrence can be found.

We note that a related result for the recursive computation of Chebyshev coefficients of functions satisfying linear ODEs with polynomial coefficients is given in Lemma 2.4 by Keller (2007) and we also highlight similar work by Lewanowicz (1991) for the recursive computation of Jacobi coefficients of functions satisfying linear ODEs with polynomial coefficients. The general constructive result is given in Thm. 5.3.2, and we shall see some applications to relevant problems in the sequel. We provide a rigorous stability analysis of some of these types of recurrences for Filon–Clenshaw–Curtis methods in §5.4.

In order to obtain a recurrence for the Filon moments, it is useful to consider them as coefficients of the oscillatory kernel function with respect to a Hilbert basis of some weighted L^2 -space. Let us quickly recap the following definition:

Definition 5.3.1. *Let $-\infty < a < b < \infty$ and let $W : [a, b] \rightarrow [0, \infty)$ be a non-negative measurable weight function with $W \neq 0$ almost everywhere. We define the weighted space $L^2([a, b], W)$ as*

$$L^2([a, b], W) := \left\{ f : [a, b] \rightarrow \mathbb{C} \mid f \text{ measurable, } \|f\|_{L^2}^2 := \int_a^b |f(x)|^2 W(x) dx < \infty \right\}.$$

For two functions $f, g \in L^2([a, b], W)$ we can then define the weighted L^2 -inner product, denoted by $(\cdot, \cdot)_{L^2}$ where the weight function is to be understood from context, with the convention

$$(g, f)_{L^2} := \int_a^b \overline{g(x)} f(x) W(x) dx,$$

note the complex conjugate of g is taken in this definition to ensure this is indeed an inner product on the complex vector space $L^2([a, b], W)$.

Theorem 5.3.2. Let $\{\phi_n\}_{n \in \mathcal{I}}$ ($\mathcal{I} = \mathbb{N}$ or \mathbb{Z}) be a Hilbert basis, i.e. a complete orthonormal set, of $L^2([a, b], W)$ for some non-negative weight function $W : [a, b] \rightarrow [0, \infty)$, with $W \neq 0$ almost everywhere and $-\infty < a < b < \infty$. Suppose further $h_\omega \in L^2([a, b], W)$, and consider the moments expressed in the form

$$\sigma_n = \int_a^b \phi_n(x) h_\omega(x) W(x) dx, \quad n \in \mathcal{I}.$$

If there is a linear differential operator \mathcal{L}_ω^* such that for all $m, n \in \mathcal{I}$:

$$\mathcal{L}_\omega^* \overline{\phi_n} \in L^2([a, b], W), \quad (5.4)$$

$$(\mathcal{L}_\omega^* \overline{\phi_n}, h_\omega)_{L^2} = 0, \quad (5.5)$$

and such that the action of \mathcal{L}_ω^* on the conjugate of the basis functions is given by a banded (infinite) matrix $[B_{mn}]_{m, n \in \mathcal{I}}$ with bandwidth k , i.e.

$$\mathcal{L}_\omega^* \overline{\phi_n} = \sum_{m \in \mathcal{I}} B_{nm} \overline{\phi_m}, \quad \forall n \in \mathcal{I}. \quad (5.6)$$

Then the moments satisfy a $2k - 1$ -term recurrence relation, $\sum_{n \in \mathcal{I}} B_{nm}^T \sigma_n = 0$, for $m \in \mathcal{I}$.

Proof. Since $\{\phi_n\}_{n \in \mathcal{I}}$ is a Hilbert basis, the set of complex conjugates $\{\overline{\phi_n}\}_{n \in \mathcal{I}}$ is also a Hilbert basis for $L^2([a, b], W)$. Moreover, since $h_\omega \in L^2([a, b], W)$, we have (in both cases $\mathcal{I} = \mathbb{N}, \mathbb{Z}$):

$$h_\omega = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} (\overline{\phi_n}, h_\omega)_{L^2} \overline{\phi_n}.$$

Thus, for any $m \in \mathcal{I}$, we can write by Eqs. (5.4) & (5.5)

$$0 = (\mathcal{L}_\omega^* \overline{\phi_m}, h_\omega)_{L^2} = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} (\mathcal{L}_\omega^* \overline{\phi_m}, \overline{\phi_n})_{L^2} (\overline{\phi_n}, h_\omega)_{L^2}$$

By Eq. (5.6) we can express the right hand side in terms of the entries of the transpose of B :

$$0 = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} B_{nm}^T (\overline{\phi_n}, h_\omega)_{L^2} = \lim_{N \rightarrow \infty} \sum_{|n| \leq N} B_{nm}^T \sigma_n.$$

Since B is banded, the above sum over n involves only finitely many non-zero terms and therefore convergences as $N \rightarrow \infty$ simply to the desired recurrence

$$\sum_{n \in \mathcal{I}} B_{nm}^T \sigma_n.$$

□

Note in several important cases we can initiate these recurrences with initial conditions that are given in terms of special functions, or, alternatively, in terms of simple integrals that can be approximated efficiently (for instance exponentially decaying integrals as in Lemma 5.5.4).

We will demonstrate based on a few examples that a simple way of constructing \mathcal{L}_ω^* is to take the adjoint (via integration by parts) of a differential operator \mathcal{L}_ω which maps h_ω to zero and whose action on the basis functions can be represented by a banded matrix. We note that the choice of \mathcal{L}_ω^* is not unique, but in practice it is often possible to spot a simple choice by inspection, leading to a low-order recurrence. Let us begin by illustrating the result with a simple example where recurrences for moments are already well-known.

Example 5.3.3. Consider $f \in L_2[0, 2\pi] \cap C_{\text{per}}[0, 2\pi]$ and the oscillatory integral

$$I_\omega[f] := \int_0^{2\pi} e^{i\omega \cos x} f(x) dx.$$

In this case a natural interpolation basis is the Fourier basis $\left\{ \frac{1}{\sqrt{2\pi}} e^{inx} \right\}_{n \in \mathbb{Z}}$, which has good interpolation properties on equispaced points, and which is also a Hilbert basis for $L^2[0, 2\pi]$. The oscillator in the weighted space is $h_\omega(x) = e^{i\omega \cos x}$ and the weight is $W(x) = 1$. We note h_ω satisfies the following differential equation

$$\mathcal{L}_\omega h_\omega = 0, \quad \mathcal{L}_\omega = \frac{d}{dx} + i\omega \sin x.$$

Moreover,

$$\mathcal{L}_\omega \overline{\phi_n} = \left(\frac{d}{dx} + i\omega \sin x \right) \frac{1}{\sqrt{2\pi}} e^{-inx} = \frac{\omega}{2} \overline{\phi_{n-1}} - in \overline{\phi_n} - \frac{\omega}{2} \overline{\phi_{n+1}} =: \sum_{m \in \mathbb{Z}} B_{nm} \overline{\phi_m}. \quad (5.7)$$

Because we work on a periodic domain we can integrate by parts to note for any $f, g \in C_{\text{per}}^1$

$$\int_0^{2\pi} (\mathcal{L}_\omega^* f)(x) g(x) dx = \int_0^{2\pi} f(x) (\mathcal{L}_\omega g)(x) dx, \quad \mathcal{L}_\omega^* := -\frac{d}{dx} + i\omega \sin x.$$

Thus, in particular we have by Eq. (5.7)

$$(\mathcal{L}_\omega^* \overline{\phi_n}, \overline{\phi_m})_{L^2} = (\overline{\phi_n}, \mathcal{L}_\omega \overline{\phi_m})_{L^2} = \sum_{j \in \mathbb{Z}} B_{mj} (\overline{\phi_n}, \overline{\phi_j})_{L^2} = B_{mn}, \quad (5.8)$$

i.e. the action of \mathcal{L}_ω^* on the conjugate of the basis functions is indeed given by a banded infinite matrix. Thus we note that $h_\omega, W, \{\phi_n\}_{n \in \mathbb{Z}}$ and \mathcal{L}_ω^* satisfy the assumptions of Thm. 5.3.2 and we deduce that the Filon moments, $\sigma_n = I_\omega[\phi_n]$ must satisfy the following

recurrence

$$-\frac{\omega}{2}\sigma_{n-1} - i n \sigma_n + \frac{\omega}{2}\sigma_{n+1} = 0. \quad (5.9)$$

This recurrence provides a highly efficient way of computing the moments, and we have actually recovered a well-known relation: In the present case the moments can be expressed in terms of Bessel functions of the first kind, J_n (cf. the integral expression Abramowitz and Stegun, 1965, Eq. 9.1.21)

$$\sigma_n = \sqrt{2\pi} e^{i\pi n/2} J_n(\omega),$$

and the recurrence Eq. (5.9) is equivalent to the Bessel recurrence satisfied by J_n (Abramowitz and Stegun, 1965, Eq. 9.1.27).

The next example concerns a case where, to the best of our knowledge, recurrences are not yet readily available in the literature:

Example 5.3.4. For our second example we consider an integral over $[-1, 1]$ with a quadratic oscillator,

$$I_\omega[f] := \int_{-1}^1 e^{i\omega x^2} f(x) dx.$$

One possible choice of interpolation basis is the use of Legendre polynomials (when $s = 0$ in Eq. (5.3)). This choice is guided by the idea that interpolating f at Legendre points optimises the order of the method when $\omega = 0$, as described by Deaño et al. (2017, §4.2.1). Thus, we choose $\phi_n = \tilde{P}_n := \sqrt{n + \frac{1}{2}} P_n$, $n = 0, \dots$, where P_n are Legendre polynomials with the standard normalisation $P_n(1) = 1$ and \tilde{P}_n are scaled such that they form a Hilbert basis for $L^2([-1, 1])$, i.e. with $W(x) = 1$. The oscillator $h_\omega(x) = \exp(i\omega x^2)$ satisfies

$$\mathcal{L}_\omega h_\omega = 0, \quad \mathcal{L}_\omega = (x^2 - 1) \frac{d}{dx} - 2xi\omega(x^2 - 1).$$

We have chosen \mathcal{L}_ω specifically with the following two identities in mind (Abramowitz and Stegun, 1965, Eqs. 22.8.5 & 22.7.10):

$$\frac{x^2 - 1}{n} \frac{d}{dx} P_n(x) = \frac{n+1}{2n+1} P_{n+1}(x) - \frac{n+1}{2n+1} P_{n-1}(x), \quad n \geq 1, \quad (x^2 - 1) \frac{d}{dx} P_0(x) = 0, \quad (5.10)$$

$$x P_n(x) = \frac{n+1}{2n+1} P_{n+1}(x) + \frac{n}{2n+1} P_{n-1}(x), \quad n \geq 1, \quad x P_0(x) = P_1(x). \quad (5.11)$$

The identities Eqs. (5.10) & (5.11) ensure that the action of \mathcal{L}_ω on the basis $\left\{ \tilde{P}_n \right\}_{n=0}^\infty$ can be represented by a banded infinite matrix. We also note that by simple integration by

parts, for any $f, g \in C^1([-1, 1])$:

$$\int_{-1}^1 (\mathcal{L}_\omega^* f)(x)g(x) dx = \int_{-1}^1 f(x) (\mathcal{L}_\omega g)(x) dx,$$

$$\text{where } \mathcal{L}_\omega^* f(x) := -\frac{d}{dx} ((x^2 - 1)f(x)) - 2xi\omega(x^2 - 1)f(x).$$

The boundary terms from integration by parts at $x = \pm 1$ are zero by construction of \mathcal{L}_ω . Analogously to Eq. (5.8) we thus conclude that the action of \mathcal{L}_ω^* on the basis functions is represented by a banded infinite matrix and hence that Thm. 5.3.2 can be applied. In fact, one may use Eqs. (5.10) & (5.11) in an analogous way to find a suitable differential operator for any oscillator of the form $\exp(i\omega q(x))$ when $q(x)$ is a polynomial. Thm. 5.3.2 allows us to construct, after a few steps of algebra, the following recurrence,

$$0 = -\frac{2i(n-2)(n-1)n\omega}{\sqrt{2n-5}(2n-3)(2n-1)\sqrt{2n+1}} \tilde{\sigma}_{n-3}$$

$$+ \frac{n(2i\omega(n^2-3) + (2n-3)(2n+3)(n-1))}{(2n-3)\sqrt{2n-1}\sqrt{2n+1}(2n+3)} \tilde{\sigma}_{n-1}$$

$$+ \frac{(n+1)(2i\omega(n^2+2n-1) - (2n-1)(n+2)(2n+5))}{(2n-1)\sqrt{2n+1}\sqrt{2n+3}(2n+5)} \tilde{\sigma}_{n+1}$$

$$- \frac{2i(n+1)(n+2)(n+3)\omega}{\sqrt{2n+1}(2n+3)(2n+5)\sqrt{2n+7}} \tilde{\sigma}_{n+3},$$

valid for $n \geq 3$, where the moments are $\sigma_n = I_\omega[\tilde{P}_n]$, $n \geq 0$. Additionally, the first column of the matrix representation \mathbf{B}_{nm} of \mathcal{L}_ω gives rise to the extra condition

$$\sigma_4 = \sqrt{\frac{5}{9}} \frac{5(21i + 2\omega)}{24\omega} \sigma_2 + \sqrt{\frac{1}{9}} \frac{7}{12} \sigma_0,$$

which means that the moments σ_{2n} , $n \geq 0$, can be computed from just two initial conditions for which we have the following expressions:

$$\sigma_0 = \frac{e^{i\frac{\pi}{4}}}{\sqrt{2\omega}} \left(\gamma \left(\frac{1}{2}, -i\omega \right) \right), \quad \sigma_2 = \sqrt{\frac{5}{2}} \left[\frac{3}{2} \frac{e^{i\omega}}{i\omega} - \left(\frac{3}{4i\omega} + \frac{1}{2} \right) \frac{e^{i\frac{\pi}{4}}}{\sqrt{\omega}} \left(\gamma \left(\frac{1}{2}, -i\omega \right) \right) \right],$$

where $\gamma(\cdot, \cdot)$ is the lower incomplete Gamma function (Abramowitz and Stegun, 1965, Eq. 6.5.2). Of course, for all moments of odd order, $\sigma_{2n+1} = 0$, $n \geq 0$, since $\exp(i\omega x^2)$ is an even function. We note that many efficient methods exist for computing the incomplete Gamma function (Cody, 1976, Gautschi, 1979), so the above expressions constitute a suitable way of initiating the recurrence.

In a similar spirit to Example 5.3.4, one may choose other orthogonal polynomials as interpolation basis (and their zeros as corresponding interior nodes) in an attempt to maximize the order of the Filon quadrature when $\omega = 0$ and to ensure the interpolation

problem is well-conditioned. In many cases one can use a similar approach to the above and exploit the three-term recurrence of orthogonal polynomials to extract recurrences for the moments in this manner.

5.3.1 Recursive moment computation for Filon–Clenshaw–Curtis methods

In the remainder of this chapter we shall focus our attention to Filon–Clenshaw–Curtis methods, motivated by fast interpolation properties as described in §5.2.2. We have seen in the previous section how one may construct a suitable differential operator \mathcal{L}_ω^* to apply Thm. 5.3.2 through integration by parts.

In the remainder of this chapter we focus in particular on the application of this theorem to Chebyshev interpolation bases. To outline the overall recipe for constructing recurrences for the Chebyshev–Filon moments, let us consider an integral of the form

$$\tilde{I}_\omega[f] = \int_{-1}^1 f(x)\tilde{h}_\omega(x) dx,$$

where, to begin with, we take $\tilde{h}_\omega(x) \in C^1([-1, 1])$. The enormous speed up achieved by the use of Clenshaw–Curtis points in the Filon method makes those an excellent choice for interior interpolation points even when the weight function W does not match the original integral. This is in a similar spirit to classical quadrature where Clenshaw–Curtis points can be preferable to optimal Legendre points as noted by Trefethen (2008). Thus we need to compute the moments

$$\tilde{I}_\omega[\phi_n] = \sqrt{s_n} \int_{-1}^1 T_n(x)\tilde{h}_\omega(x) dx.$$

To ensure the interpolation basis is a Hilbert basis such that we can apply a methodology similar to Thm. 5.3.2 it is thus appropriate to write $\tilde{I}_\omega[f]$ in the form

$$\tilde{I}_\omega[\phi_n] = \int_{-1}^1 f(x)h_\omega(x)\frac{dx}{\sqrt{1-x^2}}, \quad (5.12)$$

where $h_\omega(x) = \sqrt{1-x^2}\tilde{h}_\omega(x)$. We then construct \mathcal{L}_ω^* based on the following observations: Let us follow the convention $T_{-n}(x) = T_n(x)$, then one can show the following using standard trigonometric identities.

Lemma 5.3.5 (Eqs. 22.7.4 & 22.8.3 in Abramowitz and Stegun (1965)). *For all $n \in \mathbb{Z}$:*

$$xT_n(x) = \frac{1}{2}T_{n-1}(x) + \frac{1}{2}T_{n+1}(x), \quad \text{and} \quad (1-x^2)T_n'(x) = \frac{n}{2}T_{n-1}(x) - \frac{n}{2}T_{n+1}(x).$$

In particular, the actions of $x, (1-x^2)d/dx$ on $\{\sqrt{s_n} T_n\}_{n=0}^\infty$ are both of bandwidth 3.

These operators are such that boundary terms in the integration by parts argument are zero. Indeed, consider the operator $\mathcal{L} = (1-x^2)d/dx$ and suppose h_ω is $C^1([a, b])$. Then by simple integration by parts we have:

$$\begin{aligned}
& \int_{-1}^1 T_n(x) \mathcal{L} \left(\sqrt{1-x^2} h_\omega(x) \right) \frac{dx}{\sqrt{1-x^2}} \\
&= \lim_{\epsilon \rightarrow 0^+} \int_{-1+\epsilon}^{1-\epsilon} T_n(x) (1-x^2) \frac{d}{dx} \left(\sqrt{1-x^2} h_\omega(x) \right) \frac{dx}{\sqrt{1-x^2}} \\
&= \lim_{\epsilon \rightarrow 0^+} \left[T_n(x) (1-x^2) h_\omega(x) \right]_{-1+\epsilon}^{1-\epsilon} \\
&\quad - \lim_{\epsilon \rightarrow 0^+} \int_{-1+\epsilon}^{1-\epsilon} \frac{d}{dx} \left(\sqrt{1-x^2} T_n(x) \right) \left(\sqrt{1-x^2} h_\omega(x) \right) \frac{dx}{\sqrt{1-x^2}} \tag{5.13} \\
&= \int_{-1}^1 \frac{d}{dx} \left(\sqrt{1-x^2} T_n(x) \right) \left(\sqrt{1-x^2} h_\omega(x) \right) \frac{dx}{\sqrt{1-x^2}} \\
&= \int_{-1}^1 \mathcal{L}^* T_n(x) \left(\sqrt{1-x^2} h_\omega(x) \right) \frac{dx}{\sqrt{1-x^2}},
\end{aligned}$$

meaning if we can find an operator \mathcal{L}_ω^* which is a linear combination of $(1-x^2)d/dx$ and x^n , such that

$$\mathcal{L}_\omega h_\omega = 0, \quad \mathcal{L}_\omega \overline{\phi_n} = \sum_{m \in \mathbb{N}} \mathbf{B}_{nm} \overline{\phi_m},$$

for a banded matrix \mathbf{B} , then we can construct a suitable adjoint operator \mathcal{L}_ω^* similar to Examples 5.3.3 & 5.3.4, which satisfies the assumptions of Thm. 5.3.2. In fact the same integration by parts argument can be applied if \tilde{h}_ω satisfies an ordinary differential equation with polynomial coefficients that has a simple singularity in the interior of the domain. We shall demonstrate the principle on two types of integrals in greater detail: Integrals with stationary points/algebraic singularities in §5.4, and integrals involving Hankel functions and hybrid numerical-asymptotic basis functions in §5.5.

5.4 Application to integrals with algebraic singularities and stationary points

Consider the case of algebraic singularities or stationary points at $x = 0$, i.e. integrals of the form

$$I_{\omega}^{(1)}[f] = \int_{-1}^1 f(x)e^{i\omega x^r} dx, \quad r \in \mathbb{N}, r \geq 2,$$

and

$$I_{\omega}^{(2)}[f] = \int_{-1}^1 \operatorname{sgn}(x)|x|^{\alpha}e^{i\omega x} f(x) dx, \quad \alpha \in (-1, 1).$$

By using the simple change of variable $y = x^r$ the integral $I_{\omega}^{(1)}[f]$ can be brought into the form $I_{\omega}^{(2)}[f]$. In fact, by the inverse function theorem, an integral with a general oscillator $h_{\omega}(x) = \exp(i\omega g(x))$ with $g^{(j)}(0) = 0, j = 0, \dots, r, g^{(r+1)}(0) \neq 0, g'(x) \neq 0, \forall x \neq 0$, can also be brought into the above forms, by substituting $g(x) = y^r$, or, equivalently, as noted by Olver (2007), by choosing an interpolation basis that is in the span of $\{\operatorname{sgn}(x)g'(x)|g(x)|^{(n+1-r)/r}\}_{n=0}^{\nu+2s+1}$. These types of integrals were considered in the Filon context by Olver (2007) and Domínguez et al. (2013). To illustrate the main ideas we focus on the integral $I_{\omega}^{(2)}[f]$. Here the natural basis described by Olver (2006, 2007) essentially reduces to a monomial interpolation basis, x^n , and the central observation is that its moments can be expressed explicitly in terms of the lower incomplete gamma function γ : For any $n \geq 0$,

$$I_{\omega}^{(2)}[x^n] = (-i\omega)^{-1-n-\alpha} \gamma(1+n+\alpha, -i\omega) + (i\omega)^{-1-n-\alpha} \gamma(1+n+\alpha, i\omega). \quad (5.14)$$

This approach is particularly suitable when only a small number ν of interior interpolation points and hence moments are required. However, if we choose ν at moderate or large size relative to s the cost of interpolating with standard polynomials increases rapidly. Yet resolving to fast interpolation at Clenshaw–Curtis points is seemingly prevented by the well-known exponential instability of computing $I_{\omega}^{(2)}[T_n]$ through directly expanding T_n in terms of x^n (cf. Fokas et al., 2013, Eq. (3.5)). Domínguez et al. (2013) approached the problem from a slightly different perspective, which also applies to integrals of the form $I_{\omega}^{(2)}[f]$. Choosing a mesh $-1 = x_{-M} < \dots < x_{-1} < x_0 = 0 < x_1 < \dots < x_M = 1$ that is graded towards the algebraic singularity/stationary point at $x_0 = 0$, the method presented by Domínguez et al. (2013) evaluates the integral by splitting it into integrals on each subinterval $[x_l, x_{l+1}]$, $-M \leq l \leq M - 1$, and approximates those as follows:

- The integrals on each subinterval $[x_l, x_{l+1}]$, $-M \leq l \leq -2, 1 \leq l \leq M - 1$ (and $l = -1, 0$ when $\alpha \geq 0$), are approximated using the classical non-singular Filon method;

- The integrals over $[x_{-1}, 0]$ and $[0, x_1]$ are approximated by 0 when $-1 < \alpha < 0$.

While this approach is quite flexible, as it avoids having to know the exact type of singularity at $x = 0$, this flexibility comes at the price of asymptotic sub-optimality. Indeed, we know from the method of stationary phase, for any $-1 < \alpha < 0$, and $\epsilon > 0$ fixed:

$$\int_{-\epsilon}^{\epsilon} \operatorname{sgn}(x)|x|^{\alpha} e^{i\omega x} f(x) dx \sim I_{\omega}^{(2)}[f], \quad \text{as } \omega \rightarrow \infty, \quad (5.15)$$

i.e. we have asymptotic concentration near the singularity. Thus the method proposed by Domínguez et al. (2013) leads, as ω increases, to an absolute error that is of the same size as the original integral (this is demonstrated in practical examples in §5.4.4).

A resolution of these two approaches can be found based on the methodology from Thm. 5.3.2: We can apply a direct version of the Filon method to integrals of the form $I_{\omega}^{(2)}[f]$ and still interpolate at Clenshaw–Curtis points as in Eq. (5.3), by computing the Chebyshev moments $I_{\omega}^{(2)}[\mathbb{T}_n]$ accurately and efficiently using a recurrence initialised with exact expressions in Eq. (5.14). Indeed, one can easily check using Watson’s lemma (Bender and Orszag, 2013, pp. 263–265) that, if q satisfies the interpolation conditions Eq. (5.3) and $f \in C^{s+2}([-1, 1])$, the direct Filon quadrature $\mathcal{Q}_{\omega}^{[\nu, s]}[f] := I_{\omega}^{(2)}[q]$ satisfies

$$|I_{\omega}^{(2)}[f] - \mathcal{Q}_{\omega}^{[\nu, s]}[f]| = \mathcal{O}(\omega^{-(s+2)-\min\{0, \alpha\}}), \quad \omega \rightarrow \infty. \quad (5.16)$$

We note that an explicit derivation of Eq. (5.16) shows that there is $C_{\alpha, s} > 0$, dependent on α, s , such that

$$|I_{\omega}^{(2)}[f] - \mathcal{Q}_{\omega}^{[\nu, s]}[f]| \leq C_{\alpha, s} \min_{0 \leq j \leq s} \omega^{-j-2} (\omega^{-\alpha} \|f^{(j)} - q^{(j)}\|_{L^{\infty}([-1, 1])} + \|f^{(j+1)} - q^{(j+1)}\|_{L^{\infty}([-1, 1])}), \quad (5.17)$$

thus allowing us to account for the dependency on ν by studying the quality of interpolation of f by q . This principle is the same as for a non-stationary oscillator as described by Melenk (2010, §1.1). In the interest of brevity we omit the details of this derivation here, and instead refer to an analogous argument for a Hankel oscillator which we provide in Prop. 5.5.2 and Corollary 5.5.3. Having understood that the direct application of a Filon method truly matches the asymptotic behaviour of the integral it remains to compute the moments $\sigma_n = I_{\omega}^{(2)}[\phi_n] = I_{\omega}^{(2)}[\sqrt{s_n} \mathbb{T}_n]$. In line with Eq. (5.12) we write $h_{\omega}(x) = \sqrt{1-x^2} \operatorname{sgn}(x)|x|^{\alpha} \exp(i\omega x)$. We can now follow the recipe of Thm. 5.3.2: A suitable differential operator is given by

$$\mathcal{L}_{\omega} = x(1-x^2) \frac{d}{dx} + x^2 - \alpha(1-x^2) - i\omega x(1-x^2),$$

which is such that $\mathcal{L}_\omega h_\omega(x) = 0$ pointwise for all $x \neq 0$ (and including $x = 0$ when $\alpha \geq 0$). Moreover, in the same way as in Eq. (5.13) we can take \mathcal{L}_ω^* the adjoint of \mathcal{L}_ω using integration by parts (this time also excluding a small neighbourhood of $x = 0$) thus ensuring the assumptions of Thm. 5.3.2 are satisfied. After a few steps of algebra this results in the following recurrence, where for ease of notation we introduced $\tilde{\sigma}_n := \sigma_n/\sqrt{s_n}$ and follow the convention $\sigma_{-n} = \sigma_n$:

$$\begin{aligned} \tilde{\sigma}_{n-3} + \frac{2(-(n-3) + \alpha)}{i\omega} \tilde{\sigma}_{n-2} - \tilde{\sigma}_{n-1} + \frac{4-4\alpha}{i\omega} \tilde{\sigma}_n \\ - \tilde{\sigma}_{n+1} + \frac{2(n+3 + \alpha)}{i\omega} \tilde{\sigma}_{n+2} + \tilde{\sigma}_{n+3} = 0, \quad \forall n \in \mathbb{Z}. \end{aligned} \quad (5.18)$$

This means, in particular, that the initial values $\tilde{\sigma}_0, \tilde{\sigma}_1 = \tilde{\sigma}_{-1}, \tilde{\sigma}_2 = \tilde{\sigma}_{-2}$ are sufficient in order to compute all the moments using Eq. (5.18). For those we have the explicit expressions in terms of the lower incomplete gamma function based on Eq. (5.14):

$$\tilde{\sigma}_0 = I_\omega^{(2)}[x^0], \quad \tilde{\sigma}_1 = I_\omega^{(2)}[x^1], \quad \tilde{\sigma}_1 = 2I_\omega^{(2)}[x^2] - I_\omega^{(2)}[x^0].$$

5.4.1 Stability analysis of the recurrences

In this section we seek to understand the stability of moment recurrences for the integrals $I_\omega^{(1)}, I_\omega^{(2)}$. Similar to work by Piessens and Branders (1983) and Domínguez et al. (2011) we find there is a balance between N , the number of required moments, and ω , the frequency of oscillation, which results in two regions of different behaviour:

- **The initial regime, when $N \ll \omega$:** Here the recurrences lead at worst to algebraic instabilities, which are moderate relative to the decay of the interpolation coefficients for sufficiently smooth f . We provide rigorous results for two cases of interest in §5.4.2 below.
- **The tail of the recurrences, when $N \gg \omega$:** This is mostly of theoretical interest, since in practice, if we require $N \sim \omega$ moments, Gaussian quadrature applied to the full integral will be of comparable cost and error to the Filon method thus there is no necessity to resort to Filon methods in this regime. Nevertheless, in §5.4.3 we provide an indication of the behaviour of the recurrences in this regime which can be used in practice to compute their tail stably using Oliver's algorithm (Oliver, 1968).

5.4.2 Stability results for the initial regime

We note, firstly, that for a linear phase function, i.e. integrals of the form $\int_{-1}^1 f(x)e^{i\omega x} dx$, it was shown by Domínguez et al. (2011) that (an equivalent) recurrence to the one found through the application of Thm. 5.3.2 is algebraically stable for $n < \omega$. Our first stability

result shows that a very similar analysis can be applied to the case of a simple stationary point at $x = 0$, i.e. integrals of the form $I_\omega[f] = \int_{-1}^1 f(x)e^{i\omega x^2} dx$. Although this case is covered by the recurrence Eq. (5.18), we can find a simpler version by noting that $\mathcal{L}_\omega h_\omega = 0$, where

$$h_\omega(x) = \sqrt{1-x^2}e^{i\omega x^2}, \quad \mathcal{L}_\omega = (1-x^2)\frac{d}{dx} - i\omega 2x(1-x^2) + x.$$

The resulting recurrence satisfied by the moments is (again writing $\tilde{\sigma}_n = \sigma_n/\sqrt{s_n}$ and $\tilde{\sigma}_{-n} = \tilde{\sigma}_n$):

$$\tilde{\sigma}_{n-3} + \left(-1 - \frac{2(n-2)}{i\omega}\right)\tilde{\sigma}_{n-1} + \left(-1 + \frac{2(n+2)}{i\omega}\right)\tilde{\sigma}_{n+1} + \tilde{\sigma}_{n+3} = 0, \quad n \in \mathbb{Z}, \quad (5.19)$$

and we see, as we observed in Example 5.3.4, by the symmetry of the kernel, $\tilde{\sigma}_{2n+1} = 0$, $n \geq 0$. We also have the following initial conditions in terms of the lower incomplete Gamma function:

$$\tilde{\sigma}_0 = \frac{e^{i\frac{\pi}{4}}}{\sqrt{\omega}} \left(\gamma \left(\frac{1}{2}, -i\omega \right) \right), \quad \tilde{\sigma}_2 = 2\frac{e^{i\omega}}{i\omega} - \left(1 + \frac{1}{i\omega} \right) \frac{e^{i\frac{\pi}{4}}}{\sqrt{\omega}} \left(\gamma \left(\frac{1}{2}, -i\omega \right) \right).$$

Theorem 5.4.1. *Suppose the moments $\check{\tilde{\sigma}}_n$ are computed using Eq. (5.19) with slightly perturbed initial conditions: $\check{\tilde{\sigma}}_0 = \tilde{\sigma}_0 + \epsilon_0$, $\check{\tilde{\sigma}}_2 = \tilde{\sigma}_2 + \epsilon_2$, for some $|\epsilon_0|, |\epsilon_2| < \epsilon$. Then, for any n with $2n+1 < \omega$,*

$$|\check{\tilde{\sigma}}_{2n} - \tilde{\sigma}_{2n}| < \frac{8n\omega^{\frac{1}{2}}}{3(\omega^2 - (2n+1)^2)^{\frac{1}{4}}} \left(2 + \frac{1}{\omega} \right) \epsilon.$$

Proof. The proof of this result is new, though similar to Domínguez et al. (2011, Thm. 5.1), and can be found in Appendix 5.B. \square

This result tells us that we can reliably compute the first N moments from two initial conditions using Eq. (5.19), and, as long as $N < C\omega$ for some constant $C < 1$, any error in initial conditions grows no faster than linearly in N .

We now consider the recurrence Eq. (5.18) for a general value of $\alpha > -1$: We can still guarantee at worst linear growth of initial perturbations, but this time our rigorous analysis applies to the slightly narrower regime $N+1 < \min\{C\sqrt{\omega}, \omega\}$ for some $C > 0$.

Theorem 5.4.2. *Suppose the moments $\check{\tilde{\sigma}}_n$ are computed using Eq. (5.18) with the perturbed initial conditions $\check{\tilde{\sigma}}_0 = \tilde{\sigma}_0 + \epsilon_0$, $\check{\tilde{\sigma}}_1 = \check{\tilde{\sigma}}_{-1} = \tilde{\sigma}_1 + \epsilon_1$, $\check{\tilde{\sigma}}_2 = \check{\tilde{\sigma}}_{-2} = \tilde{\sigma}_2 + \epsilon_2$, $|\epsilon_j| < \epsilon$ for some $\epsilon > 0$, and assume $\check{\tilde{\sigma}}_3 = \check{\tilde{\sigma}}_{-3}$. Then, whenever $n+1 < \min\{C\sqrt{\omega}, \omega\}$ for a given $C > 0$,*

we have

$$|\check{\sigma}_n - \tilde{\sigma}_n| \leq \epsilon \frac{(K_0 + nK_1)}{2} \left(\frac{K_2 \omega^{\frac{1}{2}}}{K_2 \omega^{\frac{1}{2}} - 1} \exp\left(\frac{C}{K_2 - \omega^{-\frac{1}{2}}}\right) + 1 \right),$$

where the constants K_0, K_1, K_2 are independent of n and are given by

$$K_0 = \frac{2\sqrt{\omega}}{\sqrt{\omega - C^2}}, \quad K_1 = \frac{\omega + 2 + |\alpha|}{\sqrt{\omega^2 - C^2\omega}}, \quad K_2 = \frac{(\omega - C^2)^{\frac{1}{4}}}{\omega^{\frac{1}{4}} \sqrt{2|\alpha| + 2}}.$$

Proof. The proof of this result is given in Appendix 5.C. □

Note that the above constants have simple limits as $\omega \rightarrow \infty$ which means that for ω sufficiently large we can simplify the upper bound:

Corollary 5.4.3. *For any $\delta > 0, C > 0$, there is $\omega_0 > 0$ such that whenever the assumptions of Thm. 5.4.2 are satisfied, and $\omega \geq \omega_0$, the error is bounded above by*

$$|\check{\sigma}_n - \tilde{\sigma}_n| \leq \epsilon \frac{((2 + \delta) + (1 + \delta)n)}{2} \left(\exp\left(C\sqrt{2|\alpha| + 2}(1 + \delta)\right) + 1 \right).$$

5.4.3 Change of behaviour of homogeneous solutions and Oliver's algorithm

The above results suggest that, as n increases, there will be change in the behaviour of homogeneous solutions to Eq. (5.18) and Eq. (5.19) and that, for sufficiently large n , some of the solutions will exhibit super-algebraic growth. Understanding when exactly this transition occurs for general recurrences with non-constant coefficients is an open problem, however, based on numerical experiments, we find that the following heuristic argument provides a reasonably accurate practical indication of the location of this change of behaviour: Our Ansatz is that the change from algebraic to super-algebraic regime occurs when $n \propto \omega$. Thus we let $n = C_{n,\omega}\omega$ in Eq. (5.18), set $\lambda = \sigma_{n+1}/\sigma_n$ and we assume that for $-3 \leq j \leq 3$:

$$\frac{\sigma_{n+j}}{\sigma_n} \sim \lambda^j \quad \text{as } n \rightarrow \infty.$$

Plugging into Eq. (5.19) and matching the leading order terms in n yields the condition

$$\lambda^{-3} + (-1 + 2iC_{n,\omega})\lambda^{-1} + (-1 - 2iC_{n,\omega})\lambda + \lambda^3 = 0,$$

which has solutions $\lambda = \pm 1, \pm \sqrt{iC_{n,\omega} \pm \sqrt{1 - C_{n,\omega}^2}}$. Thus these solutions for λ have modulus no larger than one if and only if $n/\omega = C_{n,\omega} < 1$. This prediction matches our rigorous result in Thm. 5.4.1 which showed no larger than linear growth in that regime. For

$n/\omega = C_{n,\omega} > 1$, two of those values of λ have modulus greater than one, thus indicating that there may be two out of six linearly independent solutions to Eq. (5.19) that exhibit super-algebraic growth in this regime.

A similar heuristic argument can be applied to Eq. (5.18), which reduces to the condition

$$\lambda^{-3} + (2iC_{\omega,n} + \mathcal{O}(\omega^{-1}))\lambda^{-2} - \lambda^{-1} - \lambda + (-2iC_{\omega,n} + \mathcal{O}(\omega^{-1}))\lambda^2 + \lambda^3 = 0. \quad (5.20)$$

The solutions are now $\lambda = i^j, j = 1, \dots, 4$, and $\lambda = iC_{n,\omega} \pm \sqrt{1 - C_{n,\omega}^2}$. This means we expect algebraic behaviour in the regime $n/\omega = C_{n,\omega} < 1$, which suggests that the results in Thm. 5.4.2 might extend to larger values of n than we are currently able to prove. This near-linear growth until $n \approx \omega$ is indeed observed in practice as we show in Fig. 5.1. Finally, when $n/\omega = C_{n,\omega} > 1$, one of the solutions for λ has modulus greater than one, which indicates that we might expect to have one out of six linearly independent solutions exhibiting super-algebraic growth in this regime. Of course, the moments σ_n decay algebraically as $n \rightarrow \infty$ for any fixed ω . This suggests that the tail (i.e. moments with $n > \omega$) can be computed stably using Oliver's algorithm (Oliver, 1968) with five initial and one endpoint value, the latter of which can be approximated by an asymptotic expression for σ_n as $n \rightarrow \infty$, as was done for a linear oscillator by Domínguez et al. (2011). Numerical experiments support this observation, but in the interest of brevity those are omitted from this thesis. Instead we shall provide numerical evidence that supports the above argument of algebraic stability for Eq. (5.18) when $n < C\omega$ for some $C < 1$.

5.4.4 Numerical examples and comparison to previous work

One way to verify this numerically is by expressing the recurrence Eq. (5.18) in the equivalent form

$$\mathbf{x}_{N+1} = \mathbf{A}_N \mathbf{x}_N = \mathbf{A}_N \mathbf{A}_{N-1} \cdots \mathbf{A}_0 \mathbf{x}_0 =: \prod_{n=1}^N \mathbf{A}_n \mathbf{x}_1, \quad \forall N \in \mathbb{Z}, \quad (5.21)$$

where $\mathbf{x}_n = (x_{n+2}, \dots, x_{n-3})^T$ and $\mathbf{A}_n, n \geq 0$, are 6×6 matrices given by

$$\mathbf{A}_n = \begin{pmatrix} -\frac{2(n+3+\alpha)}{i\omega} & 1 & \frac{4\alpha-4}{i\omega} & 1 & \frac{2(n-3)-\alpha}{i\omega} & -1 \\ & & \mathbf{l}_5 & & & 0 \end{pmatrix}, \quad (5.22)$$

where \mathbf{l}_5 is the 5×5 identity matrix. The growth of an arbitrary homogeneous solution to Eq. (5.18) with given initial conditions x_{-2}, \dots, x_3 is then bounded above by the norm $\left\| \prod_{n=1}^N \mathbf{A}_n \right\| \|\mathbf{x}_1\|$. In Figure 5.1 we plot this quantity for various values of ω and α . As guaranteed by Thm. 5.4.2, the initial regime exhibits only linear growth, which changes

to super-algebraic growth near $n \approx \omega$ as has been predicted by the heuristic argument Eq. (5.20).

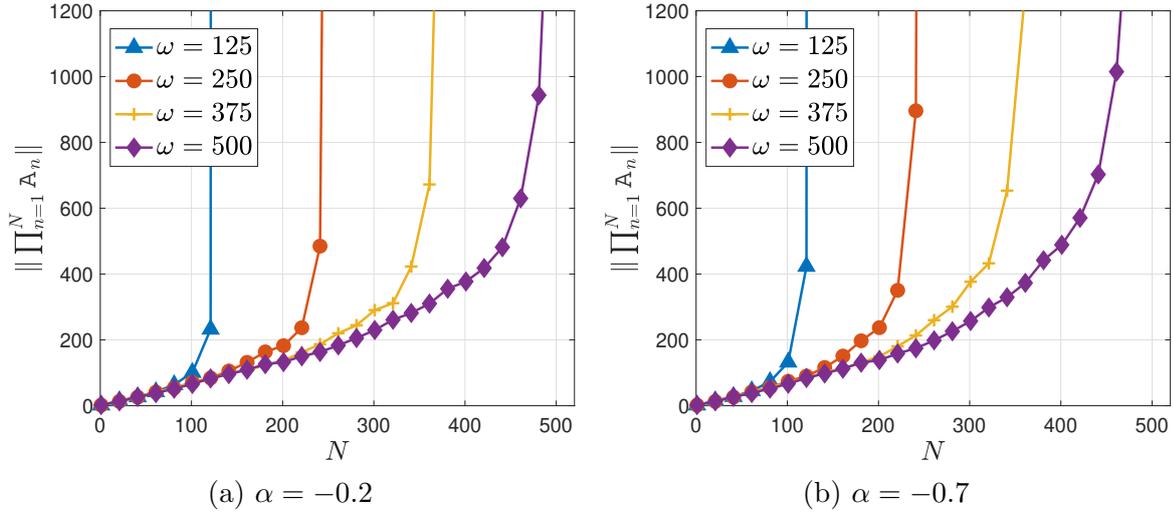


Figure 5.1: The growth of solutions to Eq. (5.18) as measured by $\left\| \prod_{n=1}^N \mathbf{A}_j \right\|$. The initial linear growth transitions to super-algebraic growth nearby $n \approx \omega$ as predicted heuristically in Eq. (5.20).

In our second numerical example we evaluate the practical performance of the direct Filon method with recursive moment computation for $I_\omega^{(2)}$. In this example we let $\alpha = -0.25$ and $f(x) = x/(1+x^2) + 1/(1+x^4)$, i.e. we approximate the integral

$$I_\omega^{(2)}[f] = \int_{-1}^1 \operatorname{sgn}(x)|x|^{-0.25} e^{i\omega x} \left(\frac{x}{1+x^2} + \frac{1}{1+x^4} \right) dx.$$

We begin by considering the absolute error of the direct Filon method

$$\mathcal{E}_{[s,\nu]}[f] = |I_\omega^{(2)}[f] - \mathcal{Q}_\omega^{[s,\nu]}[f]|,$$

where $\mathcal{Q}_\omega^{[s,\nu]}[f] = I_\omega^{(2)}[q]$, and q satisfies Eq. (5.3). Figure 5.2a shows this absolute error for the range $\omega \in [16, 400]$ with fixed values of $\nu = 5, s = 0, 1, 2$. The black dash-dotted curves correspond to the asymptotic orders $\mathcal{O}(\omega^{-k-2+0.25})$ and confirm Eq. (5.16).

In Fig. 5.2b we compare the efficiency of our direct method (for $s = 0$) with the composite Filon method described by Domínguez et al. (2013) and with a simple graded Clenshaw–Curtis approach. As we described in §5.4, both of the latter methods define a mesh that is graded towards the singularity at $x = 0$ with $x_j = \operatorname{sgn}(j)|j/M|^r$ for $j = -M, \dots, M$, and split the integral into its contribution on each of the intervals $[x_j, x_{j+1}]$ proceeding as follows:

- The integrals over $[x_j, x_{j+1}]$, $j = -M, \dots, -2, 1, \dots, M$, are approximated by the classical Filon method as in Domínguez et al. (2011) for the composite Filon method,

and by Clenshaw–Curtis quadrature in the graded Clenshaw–Curtis case. These ‘sub-methods’ on each $[x_j, x_{j+1}]$ come with an additional parameter ν , where $\nu + 2$ is the number of quadrature points on $[x_j, x_{j+1}]$ analogously to Eq. (5.1) with $s = 0$.

- For both methods the integral on $[x_{-1}, x_0]$ and $[x_0, x_1]$ is approximated by zero.

These contributions are then summed to provide an overall approximation to $I_\omega^{(2)}[f]$.

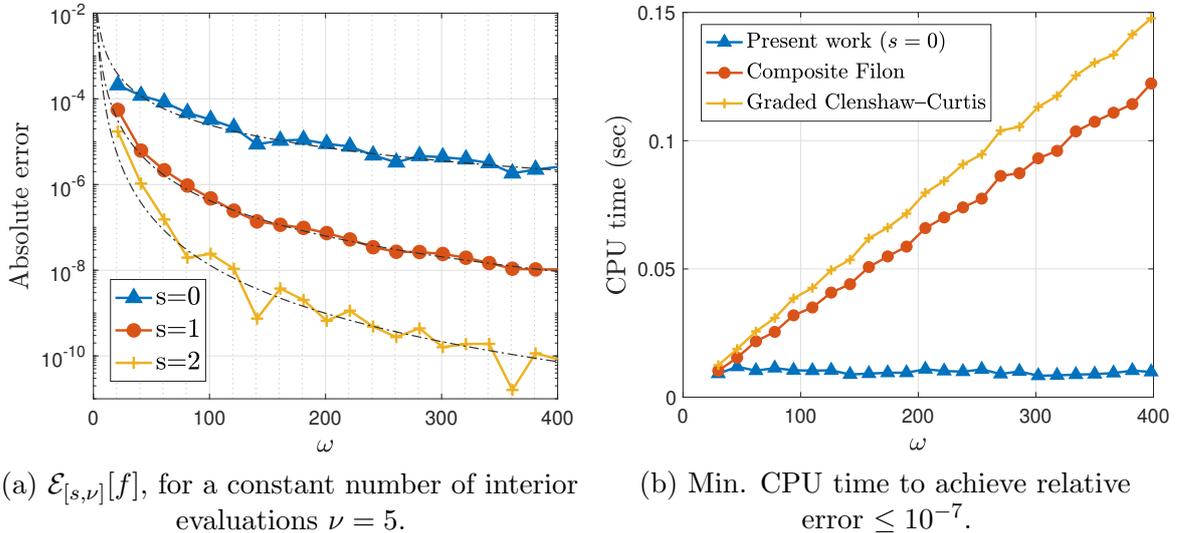


Figure 5.2: Comparison of our direct Filon method using recursive moment computation, with the literature. On the left we show the absolute error for our present method, and on the right a comparison of the minimum CPU time necessary to achieve a fixed relative accuracy using our present work in comparison to the composite Filon method¹ and a graded Clenshaw–Curtis method.

In the figure we compare the minimum CPU time each of the methods required in order to compute the integral to a fixed relative error of 10^{-7} . According to Domínguez et al. (2013), if we choose $r > (\nu + 2)/(1 + \alpha)$ the composite Filon method converges as $M \rightarrow \infty$. Thus we fix $\nu = 4, r = 8.1$ and proceed by increasing M from $M = 10$ until we reach the desired relative error with a certain choice $M_0(\omega)$. The CPU time plotted in Fig. 5.2b is the time the method took to evaluate the integral with the fixed setting $M = M_0(\omega)$. We repeat the process for each frequency ω and proceed similarly for the graded Clenshaw–Curtis method. According to Eq. (5.17) the direct Filon–Clenshaw–Curtis method converges as $\nu \rightarrow \infty$ and so for this case we start with $\nu = 3$ and proceed by increasing ν until we achieve the desired relative error with some $\nu_0(\omega)$ before plotting the CPU time it took to evaluate the method with $\nu = \nu_0(\omega)$ and repeating the process for each frequency. All experiments were performed on a single core of an Intel Core i5-8500 CPU.

¹We used the original code kindly provided to us by Victor Domínguez for the integration on each subinterval, but the grading is due to our own implementation based on the description in Domínguez et al. (2013).

While this example is certainly no complete parametric study of the convergence properties of all three methods, and especially the absolute value of the timings depends significantly on the specific implementation and CPU used, the point to take away is the clear overall trend in the cost as ω increases: As expected the classical graded Clenshaw–Curtis method requires the fastest, linear, increase in cost. While the composite Filon method performs better, it is, as explained in Eq. (5.15), still required to increase the cost with frequency, since it does not fully match the asymptotic behaviour of the integral near the singularity. Finally, the direct application of the Filon method achieves the approximation, as expected, at completely frequency-independent cost.

In all of the above numerical examples the reference solution for the true integral was computed with a graded Clenshaw–Curtis method with $M = 3000$, $\nu = 10$, $r = 40$.

5.5 Application to high-frequency wave scattering

As a final application of our method for recursive moment computation we consider integrals arising in hybrid numerical-asymptotic methods for high-frequency wave-scattering on a screen. As will be explained in more detail in §5.5.5 the integrals of relevance in this context are of the form

$$\begin{aligned} I_{\omega,\beta}^{(3)}[f] &= 2 \int_0^1 f(2x-1) H_0^{(1)}(\omega x) e^{i\omega\beta x} dx \\ &= \int_{-1}^1 f(x) H_0^{(1)}\left(\omega \frac{x+1}{2}\right) e^{i\omega\beta(x+1)/2} dx, \end{aligned} \tag{5.23}$$

where $\beta \in \mathbb{R}$, $\beta \neq -1$, and $H_0^{(1)}$ is the Hankel function of first kind and order zero (Abramowitz and Stegun, 1965, Eq. 9.1.3). As a first step in constructing a suitable direct Filon method, we aim to understand the asymptotic properties of the integral $I_{\omega,\beta}^{(3)}$. In order to do so we recall the following property of $H_0^{(1)}$.

Lemma 5.5.1 (Phase extraction of $H_0^{(1)}$, see Chandler-Wilde, Graham, Langdon and Spence (2012, Lemma 4.6)). *Let $h_0(z) := \exp(-iz)H_0^{(1)}(z)$, then for each $n \geq 0$ there is a constant C_n such that*

$$\left| \frac{d^n}{dz^n} h_0(z) \right| \leq C_n \begin{cases} \max\{1 + \log(1/z), z^{-n}\}, & z \in (0, 1], \\ z^{-(n+1/2)}, & z \in [1, \infty). \end{cases}$$

With this control on the oscillations in $h_0(\omega x)$ we can proceed to show:

Proposition 5.5.2 (Filon paradigm for $I_{\omega,\beta}^{(3)}$). *For any $k \in \mathbb{N}$ there is a constant $C_k > 0$ such that for all $\beta \in \mathbb{R}$, $\beta \neq -1$, $\omega \geq 1$ and any function $\tilde{f} \in C^{k+2}[0, 1]$ with $\tilde{f}^{(j)}(\pm 1) = 0$*

for $j = 0, \dots, k$:

$$\left| I_{\omega, \beta}^{(3)}[\tilde{f}] \right| \leq C_k \left(\omega^{-(k+2)} \|\tilde{f}^{(k+1)}\|_{\infty} \frac{|\beta + 1|^{k+3} - 1}{|\beta + 1| - 1} + \omega^{-(k+2)} \log \omega \|\tilde{f}^{(k+2)}\|_{\infty} |\beta + 1|^{-(k+2)} \right).$$

Proof. The proof is given in Appendix 5.E. \square

This means that the Filon–Clenshaw–Curtis rule, $\mathcal{Q}_{\omega}^{[s, \nu]}[f] := I_{\omega, \beta}^{(3)}[p]$ with p satisfying Eq. (5.1), has the asymptotic error

$$\left| I_{\omega, \beta}^{(3)}[f] - \mathcal{Q}_{\omega}^{[s, \nu]}[f] \right| = \mathcal{O}(\omega^{-(k+2)} \log \omega), \quad \omega \rightarrow \infty.$$

Moreover Prop. 5.5.2 allows us to understand the ν -dependency of the quadrature error through the quality of approximation of f by the interpolating polynomial p . We already mentioned in §1.1.2 that there are a number of ways to estimate $\|f^{(j)} - p^{(j)}\|_{\infty}$ including via the Hermite interpolation formula as was done in the Filon context for non-stationary oscillators by Melenk (2010), by relating the error to the regularity of f in periodic Sobolev norms on $[0, 2\pi]$ via the change of variable $x = \cos \theta$, this approach was taken for linear oscillators by Domínguez et al. (2011), or via the optimal estimates using the Peano kernel theorem given by Shadrin (1995). We shall use the latter approach: Define the nodal polynomial for the interpolation problem Eq. (5.1) as $\tilde{r}(x) = (x^2 - 1)^{s+1} \prod_{j=1}^{\nu} (x - c_j)$ then we have the following bounds (where the constants are optimal over $f \in C^{(\nu+2s+1)}([-1, 1])$):

$$\|f^{(j)} - p^{(j)}\|_{\infty} \leq \frac{\|\tilde{r}^{(j)}\|_{\infty}}{(\nu + 2s + 1)!} \|f^{(\nu+2s+1)}\|_{\infty}. \quad (5.24)$$

We can combine Eq. (5.24) with Prop. 5.5.2 and the trivial estimate

$$\left| I_{\omega, \beta}^{(3)}[f] - \mathcal{Q}_{\omega}^{[s, \nu]}[f] \right| \leq \omega^{-1} \|\mathbf{H}_0^{(1)}\|_{L^1([0, \omega])} \|f - p\|_{\infty},$$

to find:

Corollary 5.5.3. *For any $s \in \mathbb{N}$ there is $C_s > 0$ such that for all $f \in C^{\infty}([-1, 1])$, $\nu \in \mathbb{N}$, $\beta \neq 0$, $\omega \geq 1$:*

$$\left| I_{\omega, \beta}^{(3)}[f] - \mathcal{Q}_{\omega}^{[s, \nu]}[f] \right| \leq \min \left\{ \omega^{-1} \|\mathbf{H}_0^{(1)}\|_{L^1([0, \omega])} \|\tilde{r}\|_{\infty}, C_k \omega^{-(s+2)} \left(\|\tilde{r}^{(s+1)}\|_{\infty} \frac{|\beta + 1|^{s+3} - 1}{|\beta + 1| - 1} + \log \omega \|\tilde{r}^{(s+2)}\|_{\infty} |\beta + 1|^{-(s+2)} \right) \right\} \frac{\|f^{(\nu+2s+1)}\|_{\infty}}{(\nu + 2s + 1)!}.$$

5.5.1 Recursive moment computation

The results in Prop. 5.5.2 and Corollary 5.5.3 guarantee convergence of the direct Filon method $\mathcal{Q}_\omega^{[p,s]}[f] = I_{\omega,\beta}^{(3)}[p]$, where p satisfies Eq. (5.1). Thus it remains to compute the corresponding quadrature moments $\sigma_n := \sqrt{s_n} I_{\omega,\beta}^{(3)}[T_n]$. The oscillatory kernel of $I_{\omega,\beta}^{(3)}$ with respect to the Chebyshev weight is

$$h_\omega(x) = \sqrt{1-x^2} H_0^{(1)} \left(\omega \frac{x+1}{2} \right) e^{i\omega\beta(x+1)/2}.$$

The Hankel function satisfies Bessel's equation $(x^2(d/dx)^2 + xd/dx + x^2)H_1^{(0)}(x) = 0$ (Abramowitz and Stegun, 1965, Eq. 9.1.1). Thus a change of variable and multiplication by $(1-x)$ (to ensure the equation involves a combination of operators from Lemma 5.3.5) shows $\mathcal{L}_\omega h_\omega = 0$ for

$$\begin{aligned} \mathcal{L}_\omega = & \left((1-x^2) \frac{d}{dx} \right)^2 + i(\beta\omega(x^2-1) - i(3x+1)) \left((1-x^2) \frac{d}{dx} \right) \\ & - \frac{1}{4} \left((\beta^2-1)\omega^2(x^2-1)^2 - 2i\beta\omega(x-1)(x+1)^2 - 4(x^2+x+1) \right). \end{aligned}$$

By Lemma 5.3.5, the action of \mathcal{L}_ω on the basis $\phi_n = \sqrt{s_n} T_n$ has a banded matrix representation. We can use the same integration by parts argument as in Eq. (5.13) to construct \mathcal{L}_ω^* satisfying all assumptions of Thm. 5.3.2 holds and we find, after a few steps of algebra, the following recursive relation satisfied by the moments valid for all $n \in \mathbb{Z}$, where we again use the notation $\tilde{\sigma}_n = \sigma_n/\sqrt{s_n}$ and $\tilde{\sigma}_{-n} = \tilde{\sigma}_n$,

$$\begin{aligned} (1-\beta^2)\tilde{\sigma}_{n-4} - \frac{4i\beta(2n-7)}{\omega}\tilde{\sigma}_{n-3} + \frac{16(n-3)^2 + 8i\beta\omega - 4(1-\beta^2)\omega^2}{\omega^2}\tilde{\sigma}_{n-2} \\ + \frac{-32(n-2) + 4i\beta(6n-7)\omega}{\omega^2}\tilde{\sigma}_{n-1} + \frac{32(3-n^2) - 16i\beta\omega + 6(1-\beta^2)\omega^2}{\omega^2}\tilde{\sigma}_n \\ + \frac{32(n+2) - 4i\beta(6n+7)\omega}{\omega^2}\tilde{\sigma}_{n+1} + \frac{16(n+3)^2 + 8i\beta\omega - 4(1-\beta^2)\omega^2}{\omega^2}\tilde{\sigma}_{n+2} \\ + \frac{4i\beta(2n+7)}{\omega}\tilde{\sigma}_{n+3} + (1-\beta^2)\tilde{\sigma}_{n+4} = 0. \end{aligned} \tag{5.25}$$

5.5.2 Initial conditions

Since the recurrence Eq. (5.25) holds when $n = 0, 1, 2, 3$, four initial conditions $\tilde{\sigma}_0, \dots, \tilde{\sigma}_4$ are sufficient to start the moment computation (and just three initial conditions suffice when $\beta = 1$). We begin by proving an expression that allows for efficient and accurate computation of $\tilde{\sigma}_0$:

Lemma 5.5.4. For $\omega > 0, \beta \neq -1$:

$$\begin{aligned} \tilde{\sigma}_0 = & -\frac{2i}{\pi} \omega^{-1} e^{i(\beta+1)\omega} \int_0^\infty \frac{1}{\sqrt{t}} \frac{e^{-t\omega}}{\sqrt{2i-t}(1+\beta+it)} dt \\ & + \frac{2}{\pi} \omega^{-1} \begin{cases} \frac{2}{\sqrt{1-\beta^2}} \arctan \frac{\sqrt{1-\beta^2}}{1+\beta}, & |\beta| < 1, \\ \frac{1}{\sqrt{\beta^2-1}} \ln \frac{1+\beta+\sqrt{\beta^2-1}}{1+\beta-\sqrt{\beta^2-1}}, & |\beta| > 1, \\ 1, & \beta = 1. \end{cases} \end{aligned} \quad (5.26)$$

Proof. We have the following integral expression (DLMF, 2021, Eq. 10.9.10)

$$H_0^{(1)}(z) = -\frac{2i}{\pi} \int_0^\infty e^{iz \cosh t} dt, \quad \forall 0 < \arg z < \pi.$$

Thus we have

$$\begin{aligned} \omega \sigma_0 &= \lim_{\epsilon \rightarrow 0^+} \int_0^\omega H_0^{(1)}(x+i\epsilon) e^{i\beta x} dx = -\frac{2i}{\pi} \lim_{\epsilon \rightarrow 0^+} \int_0^\omega \left(\int_0^\infty e^{(ix-\epsilon) \cosh t} dt \right) e^{i\beta x} dx \\ &= -\frac{2}{\pi} \lim_{\epsilon \rightarrow 0^+} \int_0^\infty e^{-\epsilon \cosh t} \frac{e^{i\omega(\cosh t+\beta)} - 1}{\cosh t + \beta} dt = -\frac{2}{\pi} \int_0^\infty \frac{e^{i\omega(\cosh t+\beta)} - 1}{\cosh t + \beta} dt \\ &= -\frac{2}{\pi} \int_1^\infty \frac{e^{i\omega(y+\beta)}}{y+\beta} \frac{dy}{\sqrt{y^2-1}} + \frac{2}{\pi} \int_0^\infty \frac{1}{\cosh t + \beta} dt, \end{aligned}$$

where in the final line we used the change of variable $y = \cosh t$. The second integral can be evaluated explicitly (Gradshteyn and Ryzhik, 2000, Eq. 3.513.2) when $\beta \neq \pm 1$ and by taking an appropriate limit as $\beta \rightarrow 1^+$, which is justified by the dominated convergence theorem, we can also deduce its value for $\beta = 1$. For the first integral the decay of the integrand is sufficient so that we can deform the contour of integration to $y = 1 + it$ (noting that the square root singularity at $y = 1$ can be dealt with by excluding a small neighbourhood of the origin during the change of contour). Combining those observations yields precisely the expression Eq. (5.26). \square

If we set $g_\beta(t) = (\sqrt{2i-t}(1+\beta+it))^{-1}$ we can write the remaining integral in Eq. (5.26) in the form

$$-\frac{2i}{\pi} \omega^{-1} \int_0^\infty \frac{1}{\sqrt{t}} \frac{e^{i(\beta+1)\omega} e^{-t\omega}}{\sqrt{2i-t}(1+\beta+it)} dt = -\frac{2i}{\pi} \omega^{-\frac{3}{2}} e^{i(\beta+1)\omega} \int_0^\infty g_\beta\left(\frac{t}{\omega}\right) \frac{1}{\sqrt{t}} e^{-t} dt.$$

Differentiation with respect to β (justified by the dominated convergence theorem) allows us to find similar expressions for the remaining initial conditions $\tilde{\sigma}_j, j = 1, 2, 3$. For completeness these are provided in Appendix 5.D. Thus we found expressions for the initial

moments in terms of simple functions and integrals of the form

$$\int_0^\infty f\left(\frac{t}{\omega}\right) \frac{1}{\sqrt{t}} e^{-t} dt = \sqrt{\omega} \int_{-\infty}^\infty f(z^2) e^{-\omega z^2} dz,$$

which have an exponentially decaying integrand (with faster exponential decay in z as ω increases). An efficient method for evaluating these integrals using modified trapezium rules can be found in the thesis of Al Azah (2017) and the work by Al Azah and Chandler-Wilde (2021). We shall not provide a detailed study of the evaluation, but note that we can write

$$g_\beta\left(\frac{t^2}{\omega}\right) = \omega^{\frac{3}{2}} \frac{1}{\sqrt{2i\omega - t^2}} \frac{1}{(1 + \beta)\omega + it^2},$$

which means the integrand becomes nearly singular when $\omega, (1 + \beta)\omega \ll 1$. Thus for practical purposes we restrict the use of Eq. (5.26) to the case when $\omega, (1 + \beta)\omega \geq 1$. Of course, the kernel of $I_{\omega, \beta}^{(3)}[f]$ takes the form $h_0(\omega x) \exp(i\omega(1 + \beta)x)$ (cf. Lemma 5.5.1) which means it is only highly oscillatory when $\omega(1 + \beta) \gg 1$, and so in the case when Eq. (5.26) is near singular we do not need to use Filon methods for the approximation of $I_{\omega, \beta}^{(3)}[f]$ to begin with. The same holds true for the analogous expressions for $\tilde{\sigma}_j, j = 1, 2, 3$, as given in Appendix 5.D.

5.5.3 Behaviour of homogeneous solutions and initial stability

Note that we can write the recurrence Eq. (5.25) in the form

$$\begin{aligned} & (1 - \beta^2)(\tilde{\sigma}_{n-4} - 4\tilde{\sigma}_{n-2} + 6\tilde{\sigma}_n - 4\tilde{\sigma}_{n+2} + \tilde{\sigma}_{n+4}) \\ &= \frac{n^2}{\omega^2}(-16)(\tilde{\sigma}_{n-2} - 2\tilde{\sigma}_n + \tilde{\sigma}_{n+2}) \\ & \quad + \frac{n}{\omega} 8i\beta(\tilde{\sigma}_{n-3} - 3\tilde{\sigma}_{n-1} + 3\tilde{\sigma}_{n+1} - \tilde{\sigma}_{n+3}) \\ & \quad + \frac{n}{\omega^2} 32(-3\tilde{\sigma}_{n-2} + \tilde{\sigma}_{n-1} - \tilde{\sigma}_{n+1} - 3\tilde{\sigma}_{n+2}) \\ & \quad + \frac{1}{\omega} 4i\beta(-7\tilde{\sigma}_{n-3} - 2\tilde{\sigma}_{n-2} + 7\tilde{\sigma}_{n-1} + 4\tilde{\sigma}_n + 7\tilde{\sigma}_{n+1} - 2\tilde{\sigma}_{n+2} - 7\tilde{\sigma}_{n+3}) \\ & \quad + \frac{1}{\omega^2}(-32)(3\tilde{\sigma}_{n-2} + 2\tilde{\sigma}_{n-1} + 2\tilde{\sigma}_n + 2\tilde{\sigma}_{n+1} + 3\tilde{\sigma}_{n+2}). \end{aligned} \tag{5.27}$$

A similar argument to the proof of Thm. 5.4.2 shows that, whenever $\beta \neq 1$, if ω is sufficiently large compared to n , the solutions to Eq. (5.27) grow no faster than algebraically with n .

For the case $\beta = 1$ the recurrence reduces to seven terms and takes the form

$$\begin{aligned}
& (-2n + 7)\tilde{\sigma}_{n-3} + 2\tilde{\sigma}_{n-2} + (6n - 7)\tilde{\sigma}_{n-1} + (-4)\tilde{\sigma}_n \\
& \quad + (-6n - 7)\tilde{\sigma}_{n+1} + 2\tilde{\sigma}_{n+2} + (2n + 7)\tilde{\sigma}_{n+3} \\
& = -\frac{-12n + 28 + 4(n - 2)(n - 1)}{i\omega}\tilde{\sigma}_{n-2} - \frac{-8n + 16}{i\omega}\tilde{\sigma}_{n-1} - \frac{-8n^2 + 24}{i\omega}\tilde{\sigma}_n \\
& \quad - \frac{8n + 16}{i\omega}\tilde{\sigma}_{n+1} - \frac{12n + 28 + 4(n + 1)(n + 2)}{i\omega}\tilde{\sigma}_{n+2}. \tag{5.28}
\end{aligned}$$

Here we can understand the behaviour of the recurrence operator on the left hand side by substituting $\xi_n = (2n + 3)\tilde{\sigma}_{n+2} + \tilde{\sigma}_n - (2n + 1)\tilde{\sigma}_{n-2}$, which is such that

$$\begin{aligned}
& (-2n + 7)\tilde{\sigma}_{n-3} + 2\tilde{\sigma}_{n-2} + (6n - 7)\tilde{\sigma}_{n-1} \\
& \quad + (-4)\tilde{\sigma}_n + (-6n - 7)\tilde{\sigma}_{n+1} + 2\tilde{\sigma}_{n+2} + (2n + 7)\tilde{\sigma}_{n+3} = \xi_{n-2} - 2\xi_n + \xi_{n+2},
\end{aligned}$$

and shows, by a simple discrete variation of constants argument, that if ω is sufficiently large compared to n then the solutions to Eq. (5.28) grow at most linearly in n . In both cases $\beta \neq 1$ and $\beta = 1$ the solutions to Eq. (5.25) thus have algebraic behaviour in the initial regime. We find from numerical experiments that this behaviour changes as n increases for fixed ω , and that some of the solutions exhibit super-algebraic growth for n sufficiently large, thus leading to instability in Eq. (5.25). To understand where this transition occurs we follow the procedure described in §5.4.3. We suspect the change of behaviour occurs when $n \propto \omega$. Thus we let $n = C_{n,\omega}\omega$ and make the Ansatz $\tilde{\sigma}_{n+j}/\tilde{\sigma}_n = \lambda^j$, $-4 \leq j \leq 4$, which, when substituted into Eq. (5.25), results in the following condition at leading order in ω :

$$\begin{aligned}
& (1 - \beta^2)\lambda^{-4} - 8i\beta C_{n,\omega}\lambda^{-3} + (16C_{n,\omega}^2 - 4(1 - \beta^2))\lambda^{-2} + (24i\beta C_{n,\omega})\lambda^{-1} \\
& \quad + (-32C_{n,\omega}^2 + 6(1 - \beta^2)) + (-24i\beta C_{n,\omega})\lambda + (16C_{n,\omega}^2 - 4(1 - \beta^2))\lambda^2 \\
& \quad \quad \quad + 8i\beta C_{n,\omega}\lambda^3 + (1 - \beta^2)\lambda^4 = 0. \tag{5.29}
\end{aligned}$$

When $\beta \neq 1$ the condition has eight solutions for λ :

$$\lambda = \pm 1, \pm 1, \frac{-2iC_{n,\omega} \pm \sqrt{(1 - \beta)^2 - 4C_{n,\omega}^2}}{1 - \beta}, \frac{2iC_{n,\omega} \pm \sqrt{(1 + \beta)^2 - 4C_{n,\omega}^2}}{(\beta + 1)}.$$

All of those values are in modulus equal to 1 if and only if $n/\omega = C_{n,\omega} \leq \min\{|1 + \beta|/2, |1 - \beta|/2\}$, so we expect algebraic behaviour in this regime and the onset of super-algebraic growth to occur when $n \approx \omega \min\{|1 + \beta|/2, |1 - \beta|/2\}$. This behaviour is confirmed in Fig. 5.3b. When $n/\omega = C_{n,\omega} \geq \min\{|1 + \beta|/2, |1 - \beta|/2\}$ at most two of the values for λ have modulus greater than 1 and we thus expect Oliver's algorithm (Oliver, 1968) with six initial and two endpoint values to provide a stable way of computing the remainder of the

moments. As we explained in §5.4.1, in practical applications of Filon methods it is less important to compute moments when $n \gtrsim \omega$ since at that point classical quadrature is just as expensive as the Filon method. In the interest of brevity we therefore omit a discussion of the application of Oliver’s algorithm, but note that we performed initial numerical experiments which suggest that this indeed provides a satisfactory way for computing the remaining quadrature moments.

When $\beta = 1$, Eq. (5.29) has six solutions, $\lambda = \pm 1, \pm 1, iC_{n,\omega} \pm \sqrt{1 - C_{n,\omega}^2}$. These solutions have modulus equal to 1 whenever $n/\omega = C_{n,\omega} \leq 1$ which suggests the onset of super-algebraic growth lies around $n \approx \omega$. This is confirmed in Fig. 5.3a. Moreover, when $n/\omega > 1$ one of the solutions for λ has modulus greater than 1, which indicates that Oliver’s algorithm with five initial and one endpoint value can be used to compute the remaining moments in a stable way.

5.5.4 Numerical evidence of stable forward propagation

Similar to Eq. (5.21) we can write the recurrence Eq. (5.25) in the form

$$\mathbf{x}_{N+1}^{(j)} = \prod_{n=1}^N \mathbf{B}_n^{(j)} \mathbf{x}_1^{(j)}, \quad \forall N \in \mathbb{Z},$$

where $j = 1$ corresponds to the case $\beta = 1$ and $j = 2$ covers the case $\beta \neq 1$. Here $\mathbf{x}_n^{(1)} = (x_{n+2}, \dots, x_{n-3})^T$, $\mathbf{x}_n^{(2)} = (x_{n+3}, \dots, x_{n-4})^T$, and $\mathbf{B}_n^{(j)}$, $n \geq 0, j = 1, 2$, are 6×6 , and 8×8 matrices respectively whose entries are, analogously to Eq. (5.22), given by the coefficients of the recurrence Eq. (5.25) in the top row and the bottom left 5×5 and 7×7 entries are given by the identity matrix of respective size l_5, l_7 . This means that the matrices are of the shape

$$\mathbf{B}_n^{(1)} = \begin{pmatrix} \star & \star \\ l_5 & 0 \end{pmatrix}, \quad \mathbf{B}_n^{(2)} = \begin{pmatrix} \star & \star \\ l_7 & 0 \end{pmatrix},$$

where \star is a placeholder for the nonzero entries given by the coefficients of the recurrence Eq. (5.25) which are not repeated in the interest of brevity. Similar to §5.4.4 we have the upper bound $\|\mathbf{x}_N^{(j)}\| \leq \left\| \prod_{n=1}^N \mathbf{B}_n^{(j)} \right\| \|\mathbf{x}_1^{(j)}\|$, thus we can look at the norm of the matrix product to find an upper bound on the growth of solutions to the recurrence Eq. (5.25). In Fig. 5.3 we plot these norms and we see initial algebraic growth which transitions to super-algebraic growth roughly at the points predicted in §5.5.3: when $n \approx \omega$ in Fig. 5.3a and when $n \approx 0.4\omega$ in Fig. 5.3b.

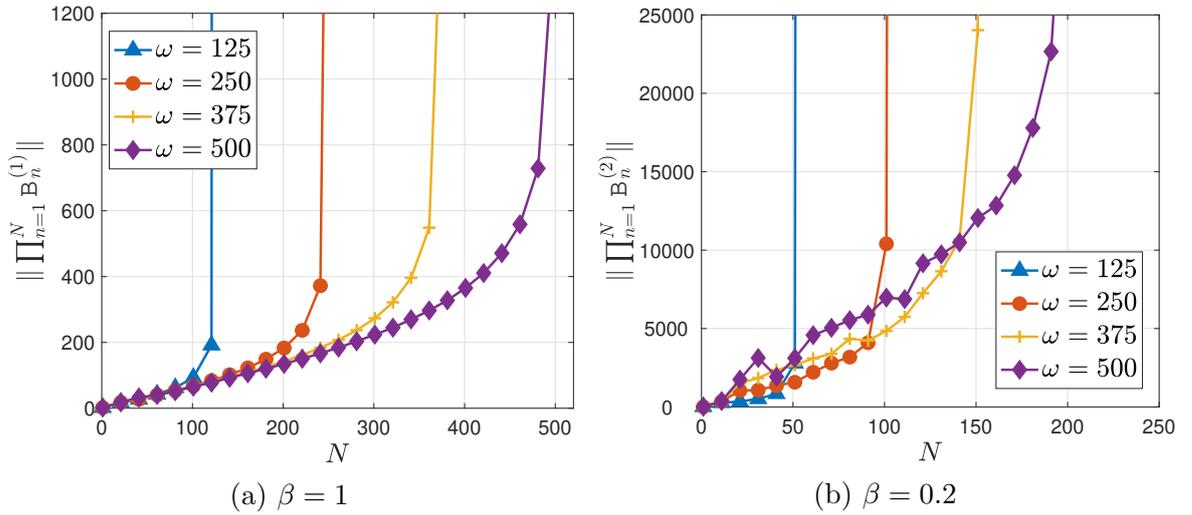


Figure 5.3: The growth of solutions to Eq. (5.18) as measured by $\left\| \prod_{n=1}^N \mathbf{B}_n^{(j)} \right\|$, $j = 1, 2$. The initial algebraic transitions to super-algebraic growth roughly at the points predicted heuristically in §5.5.3.

5.5.5 Wave scattering on a screen in two dimensions

Integrals of the form Eq. (5.23) appear in hybrid numerical-asymptotic collocation methods for high-frequency wave scattering on screens in two dimensions (see for instance Hewett et al. (2015) and Parolin (2015)). Closely related integrals also appear in the impedance boundary value problem for the Helmholtz equation in a half-plane with piecewise constant boundary data (Langdon and Chandler-Wilde, 2006). Recently, Gibbs et al. (2020) constructed a very efficient numerical steepest descent (NSD) method that can be used to assemble the matrix and right hand side in the corresponding collocation system at frequency-independent cost. In this example, we demonstrate that our direct Filon method with recursive moment computation can be applied to achieve the same goal.

As a conceptual difference, we highlight that our method relies on evaluations of the integrand strictly in the domain of integration, in contrast to NSD where a complex extension of the integrand is evaluated along steepest descent paths. This can be of advantage when the amplitude of the incident field is complicated, or not available in closed functional form. For instance, this arises in hybrid methods that employ ad hoc high frequency approximations to extend the Kirchhoff approximation to the shadow boundary on the scatterer, by exploiting knowledge from low-frequency scattering solutions (see Mavaleix-Marchessoux et al., 2020, §3.4). In the present set-up, one might imagine a screen placed in the shadow of a second object. In this case, an incident plane wave of unit amplitude would, to leading order in the small wavelength, result in an incident plane wave on the screen that has an amplitude varying from one in the illuminated region to zero in the shadow.

The aforementioned feature, of Filon methods evaluating the integrand only on the

domain of integration, is also useful when a complex extension of the highly oscillatory integrand is not readily available or difficult to evaluate due to the presence of branch cuts in the complex plane. Such a situation arises, for instance, when we consider the scattering of a two-dimensional highly oscillatory Gaussian beam by a finite screen Γ extending from $(-1, 0)$ to $(1, 0)$ in \mathbb{R}^2 . This scattering problem can be modelled by the Dirichlet problem of the Helmholtz equation on this boundary, and its solution is visualised in Fig. 5.4.

The solution of this problem can be interpreted physically as the z -component of the electric field arising from the scattering of a laser beam by a perfectly conducting screen in three-dimensional space extending along $\{(x, y, z) \mid -1 \leq x \leq 1, y = 0, z \in \mathbb{R}\}$. A more detailed description of the physical set-up can be found in the work by Kozaki (1982) who considered a similar scattering problem by an infinitely extending dielectric cylinder.

In the high-frequency regime the propagation of a laser beam can be described (at leading order in $(\omega a)^{-1} \ll 1$, where ω is the non-dimensional frequency of the beam and a is the non-dimensionalised width of the beam) by an incident field of the form

$$\psi_i(x, y) = \left(1 + \frac{i\tilde{y}}{\omega a^2}\right)^{-\frac{1}{2}} \exp\left(i\omega\tilde{y} - \frac{\tilde{x}^2}{2a^2} \left(1 + \frac{i\tilde{y}}{\omega a^2}\right)^{-\frac{1}{2}}\right),$$

where $\tilde{x} = x \sin \theta - y \cos \theta$, $\tilde{y} = x \cos \theta + y \sin \theta$ (cf. Kravtsov (1967, Eq. (55)), Keller and Streifer (1971, Eq. (17)), and Kozaki (1982, Eq. (7))). This describes a Gaussian beam focused at $(x, y) = (0, 0)$, which propagates in the direction $(\cos \theta, \sin \theta)$ at frequency ω and has width a in the plane $\{\tilde{y} = x \cos \theta + y \sin \theta = 0\}$. The incident field with $a = 0.25$, $\theta = \pi$, $\omega = 16$ can be seen in Fig. 5.4a.

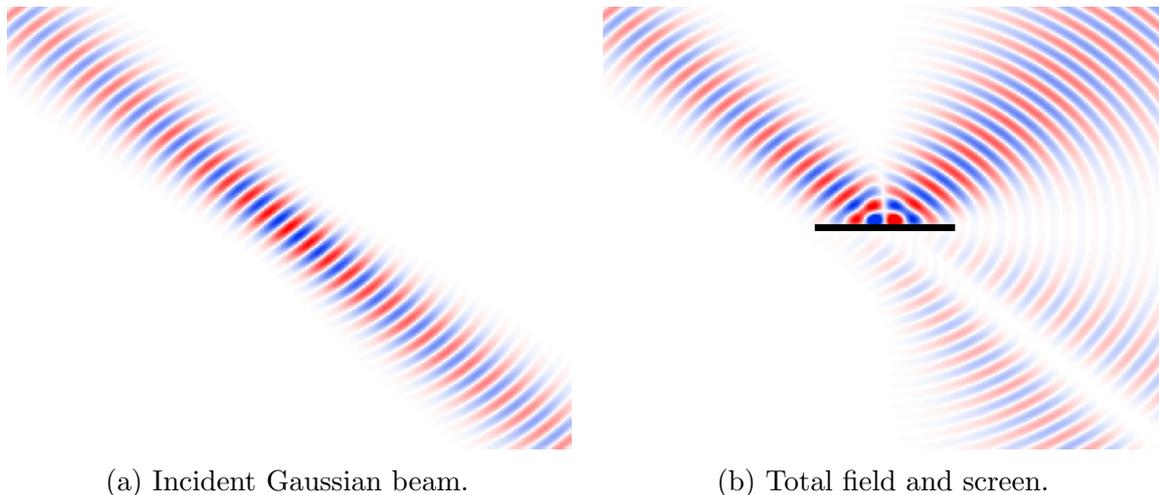


Figure 5.4: Scattering of a Gaussian beam on a perfectly conducting finite screen.

This scattering problem on the perfectly conducting screen (i.e. with Dirichlet boundary

conditions) can be written in the single layer formulation Gibbs et al. (2020, Eq. (7))

$$\psi_i(s_x) = \mathcal{S}(\partial_n \psi)(s_x) := \frac{i}{4} \int_{-1}^1 H_0^{(1)}(\omega|s_x - s_y|) \partial_n \psi(s_y) ds_y, \quad s_x \in [-1, 1],$$

where ψ is the unknown scattered field and s_x, s_y are coordinates in arclength along the screen.

We follow the hybrid Ansatz described by Hewett et al. (2015) where the unknown $\partial_n \psi$ is expanded in the form

$$\partial_n \psi(s_y) = V_0(s_y; \omega) + \sum_{l=1}^L V_l^+(s_y) e^{i\omega s_y} + V_l^-(s_y) e^{-i\omega s_y},$$

where $V_0(s_y; \omega) = 2\partial\psi_i(s_y)/\partial n$ is the geometrical optics approximation and V_l^\pm are piecewise polynomials of low degree defined on a mesh graded towards the endpoints of the screen. Thus the collocation system for the free parameters in V_l^\pm takes the form

$$\sum_{l=1}^L \mathcal{S}(V_l^+(\cdot) e^{i\omega \cdot} + V_l^-(\cdot) e^{-i\omega \cdot})(s_m) = \psi_i(s_m) - 2\mathcal{S}(\partial_n \psi_i)(s_m), \quad (5.30)$$

for some collocation points $s_m \in [-1, 1], m = 1, \dots, M$. In the present example we shall focus on the evaluation of the geometrical optics contribution, but we note that the integrals over the basis terms $V_l^\pm(s_y) \exp(\pm i\omega s_y)$, where V_l^\pm are piecewise polynomials, can also be approximated efficiently using expressions of the form Eq. (5.26). On the blade, the incident Gaussian beam takes the form

$$\begin{aligned} \partial_n \psi_i(s_y, 0) &= \left[\frac{-8ia^2\omega s_y \cos \theta + 3s_y^2 \cos(2\theta) + 5s_y^2}{8a^4\omega^2} \left(1 + \frac{is_y \cos \theta}{a^2\omega}\right)^{-2} \right. \\ &\quad \left. + \frac{-1}{2a^2\omega^2} \left(1 + \frac{is_y \cos \theta}{a^2\omega}\right)^{-\frac{3}{2}} + \left(1 + \frac{is_y \cos \theta}{a^2\omega}\right)^{-\frac{1}{2}} \right] \\ &\quad i\omega \sin \theta \exp \left(i\omega s_y \cos \theta - \frac{s_y^2 \sin^2 \theta}{2a^2} \left(1 + \frac{is_y \cos \theta}{\omega a^2}\right)^{-\frac{1}{2}} \right) \\ &=: A(s_y; \omega) \exp(i\omega s_y \cos \theta), \end{aligned}$$

where we have extracted the amplitude of $\partial_n \psi_i$ in A . Note the only ω dependency of A is via a constant multiplication out front and via the function $f(x) = (1 + ix \cos \theta/a)^{-1/2}$ in the form $f(x/\omega)$. We observe

$$\left| \frac{d^j}{dx^j} f(x) \right| = \frac{(2j-1)!!}{2^j} |\cos \theta|^j a^{-j} \left| 1 + \frac{ix \cos \theta}{\omega a^2} \right|^{-\frac{1}{2}-j} \leq C_j x^{-\frac{1}{2}-j}, \quad \forall j \geq 0,$$

where $(2j-1)!! = (2j-1)(2j-3)\cdots 1$. Thus tangential derivatives of $A(s_y; \omega)$ do not grow in ω , i.e. this is a smooth non-oscillatory function is well-approximable by polynomials on $[-1, 1]$ uniformly in $\omega \geq 1$ in the sense of Eq. (5.24). The geometrical optics approximation in a collocation method with the Ansatz Eq. (5.30) requires us to compute, for each collocation point $s_m \in [-1, 1]$,

$$\begin{aligned} & 2 \int_{-1}^1 \mathrm{H}_0^{(1)}(\omega|s - s_m|) \partial_n \psi_i(s, 0) \, ds \\ &= 2 \int_{-1}^{s_0} \mathrm{H}_0^{(1)}(\omega|s - s_m|) A(s) e^{i\omega s \cos \theta} \, ds + 2 \int_{s_0}^1 \mathrm{H}_0^{(1)}(\omega|s - s_m|) A(s) e^{i\omega s \cos \theta} \, ds \\ &= 2(1 + s_m) e^{i\omega s_0 \cos \theta} \int_0^1 \mathrm{H}_0^{(1)}(\omega(1 + s_m)t) A(s_m - (1 + s_m)t) e^{-i\omega(1+s_m)t \cos \theta} \, dt \\ &\quad + 2e^{i\omega s_m \cos \theta} \int_0^{s_m} \mathrm{H}_0^{(1)}(\omega t) A(t + s_m) e^{i\omega t \cos \theta} \, dt. \end{aligned}$$

We can write these integrals in the form

$$e^{i\omega s_m \cos \theta} \left((1 + s_m) I_{\omega(1+s_m), -\cos \theta}^{(3)}[A_1] + (1 - s_m) I_{\omega(1-s_m), \cos \theta}^{(3)}[A_2] \right), \quad (5.31)$$

where we took $A_1(x) = A((s_0 - 1)/2 - x(s_0 + 1)/2)$, $A_2(x) = A((s_0 + 1)/2 + x(1 - s_0)/2)$. Therefore, we can consider the approximation of Eq. (5.31) using the direct Filon method with recursive moment computation as described in §5.5.1. The performance of this method is demonstrated by the results in Fig. 5.5. Here we choose $a = 0.25, \theta = \pi/4, s_m = 0$. In Fig. 5.5a we see the behaviour of the relative error of the direct Filon method as a function of ω with $\nu = 6$ fixed. We recall from Prop. 5.5.2 that the direct Filon method has asymptotic error $\mathcal{O}(\omega^{-s-2} \log \omega)$ and a similar argument shows that $2\mathcal{S}(\partial_n \psi_i)(s_n)$ has asymptotic behaviour $\mathcal{O}(\omega^{-1} \log \omega)$. Thus we expect the relative error to behave like $\mathcal{O}(\omega^{-s-1})$ which is confirmed in Fig. 5.5a. This means that the direct Filon method can approximate the integral to a fixed relative error at uniform cost in ω .

In Fig. 5.5b we consider the convergence properties of the method for a fixed $\omega = 100$ as ν increases. Since A_1, A_2 are smooth we expect, by Corollary 5.5.3, to find spectral convergence in ν for any fixed value of ω . This is indeed confirmed in Fig. 5.5b. In both numerical examples the reference solution for the true integral was computed with a graded Clenshaw–Curtis method (as described in §5.4.4) with $M = 6000, \nu = 10, r = 40$.

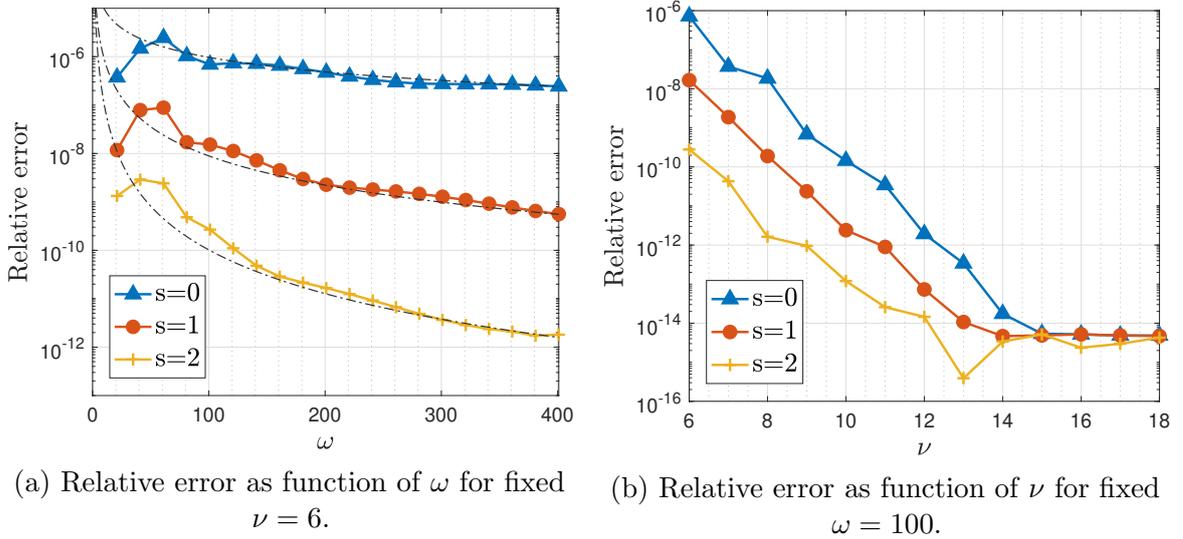


Figure 5.5: Relative error of the direct Filon method for evaluating $\mathcal{S}(\partial_n \psi_i)(s_m)$ with $a = 0.25, \theta = \pi/4, s_m = 0$. The black dash-dotted lines in (a) mark $\mathcal{O}(\omega^{-s-1})$ for $s = 0, 1, 2$ respectively.

5.6 Conclusions

In this chapter we sought to address the ‘moment-problem’ for Filon methods by providing a general framework for constructing recurrences satisfied by the Filon quadrature moments. This framework is based on the observation that many physically relevant oscillatory kernels are in the null space of certain differential operators whose action on the interpolation basis is represented by a banded matrix. The recursive moment computation allowed us to construct direct Filon methods for several examples of interest, two of which we studied in further detail: integrals with algebraic singularities and stationary points and integrals involving a Hankel function. For the former we proved rigorous stability results guaranteeing that the initial moments can be computed with at worst linear error growth. We also demonstrated the advantageous properties of the direct Filon method which perfectly matches the asymptotic behaviour of the integral up to certain order in the frequency ω . The second type of integrals are relevant in evaluating the geometrical optics approximation in high-frequency wave scattering. Based on numerical evidence we found that the recurrences incur only weak (algebraic) error growth as long as $N \lesssim \omega$, meaning the recurrences are a suitable means for computing the quadrature moments for most practical purposes. We provided rigorous convergence results that facilitate the understanding of both ν - and ω -dependency of the quadrature error and showed an application to high-frequency wave scattering of a Gaussian beam on a finite screen in two dimensions.

List of symbols

Symbol	Description
$(\cdot, \cdot)_{L^2}$	Inner product on $L^2([a, b], W)$.
$I_\omega[f]$	Oscillatory integral with frequency ω and amplitude function f .
J_n, Y_n	Bessel functions of the first and second kind of order n .
T_n	Chebyshev polynomial of the first kind of degree n .
$H_0^{(1)}$	Hankel function of the first kind of order zero.
L_n	Laguerre polynomial of degree n .
P_n	Legendre polynomial of degree n .
α	Parameter determining the strength of the algebraic singularity in $I_\omega^{(2)}[f]$.
β	Parameter determining the phase of the oscillatory kernel in $I_{\omega, \beta}^{(3)}[f]$.
$\gamma(\cdot, \cdot)$	Lower incomplete Gamma function.
$\langle g, f \rangle$	Action of $g \in H^{-s}([a, b], W)$ on $f \in H^s([a, b], W)$.
\mathcal{L}_ω^*	Adjoint of the differential operator \mathcal{L}_ω .
\mathcal{L}_ω	Differential operator in mapping h_ω to zero.
\mathcal{Q}	A quadrature technique (different versions depending on the context).
\mathbf{B}	Banded infinite matrix representation of the action of \mathcal{L}_ω on $\{\overline{\phi_n}\}_{n \in \mathcal{I}}$.
$\nu + 1$	Number of quadrature nodes (for function values) in the Filon method.
ω	Frequency in the integral.
\bar{z}	Complex conjugate of $z \in \mathbb{C}$.
σ_n	Filon quadrature moments.
$\{\phi_n\}_{n \in \mathcal{I}}$	Set of interpolation basis functions.
f	Non-oscillatory amplitude function in the integral.
h_ω	Oscillatory kernel function in the integral.
j_n, y_n	Spherical Bessel functions of the first and second kind.
p, q	Polynomials interpolating the amplitude function f .
s	Number of interpolated derivative values of f in the Filon method.

5.A Fast interpolation at Clenshaw–Curtis points, mid- and endpoint derivatives

Here we provide a some more detail on how the interpolation problem Eq. (5.3) can be solved at cost $\mathcal{O}(\nu \log \nu + s\nu + s^3)$ as mentioned in §5.2.2. Recall that we wish to solve the following interpolation problem:

$$q^{(j)}(0) = f^{(j)}(0), \quad q^{(j)}(\pm 1) = f^{(j)}(\pm 1), \quad j = 0, \dots, s, \quad \text{and} \quad q(c_n) = f(c_n), \quad n = 1, \dots, \nu,$$

where ν is odd, using an expansion in Chebyshev polynomials $q(x) = \sum_{n=0}^{\nu+3s+1} q_n T_n(x)$. Let us adopt the notation used by Gao and Iserles (2017a) and define

$$\begin{aligned} \hat{q}_0 &= 2q_0, \quad \hat{q}_k = q_k, \quad k = 1, \dots, \nu, \quad \hat{q}_{\nu+1} = 2q_{\nu+1}, \\ h_j &= f_j - \sum_{m=\nu+2}^{\nu+3s} q_m \cos\left(\frac{jm\pi}{\nu+1}\right), \quad f_j = f\left(\cos\frac{j\pi}{\nu+1}\right), \quad j = 0, \dots, \nu+1. \end{aligned}$$

Then the interpolation conditions $q(c_n) = f(c_n), n = 0, \dots, \nu+1$, are equivalent to saying that $\mathcal{C}_{\nu+1}\hat{\mathbf{q}} = \mathbf{h}$, where $\mathcal{C}_{\nu+1}$ is the discrete cosine transform DCT-I. The inverse is

$$\hat{q}_m = (\mathcal{C}_{\nu+1}^{-1}\mathbf{h})_m = \frac{2}{\nu+1} \sum_{j=0}^{\nu+1}{}'' h_j \cos\left(\frac{jm\pi}{\nu+1}\right) \quad \text{for } m = 0, \dots, \nu+1, \quad (5.32)$$

where $\sum_{j=0}^{\nu+1}{}''$ means that for $j = 0$ and $j = \nu+1$ the terms are halved. We can simplify the expressions for h_j as follows:

$$\begin{aligned} h_j &= f_j - \sum_{m=\nu+2}^{\nu+3s+1} \cos\left(\frac{jm\pi}{\nu+1}\right) q_m \\ &= f_j - \sum_{m=1}^{3s} (-1)^j \cos\left(\frac{jm\pi}{\nu+1}\right) q_{\nu+1+m}, \quad j = 0, \dots, \nu+1. \end{aligned} \quad (5.33)$$

Using Eq. (5.33) in Eq. (5.32) we find for $m = 0, \dots, \nu+1$ (and with $\check{\mathbf{q}} = \mathcal{C}_{\nu+1}^{-1}\mathbf{f}$):

$$\begin{aligned} \hat{q}_m &= \check{q}_m - \frac{2}{\nu+1} \sum_{n=1}^{3s} q_{\nu+1+n} \left[\sum_{j=0}^{\nu+1}{}'' (-1)^j \cos\left(\frac{jn\pi}{\nu+1}\right) \cos\left(\frac{jm\pi}{\nu+1}\right) \right] \\ &= \check{q}_m - \frac{1}{\nu+1} \sum_{n=1}^{3s} q_{\nu+1+n} \left[\sum_{j=0}^{\nu+1}{}'' (-1)^j \cos\left(\frac{j(n+m)\pi}{\nu+1}\right) \right. \\ &\quad \left. + \sum_{j=0}^{\nu+1}{}'' (-1)^j \cos\left(\frac{j(n-m)\pi}{\nu+1}\right) \right]. \end{aligned}$$

Now, because ν is odd, one can quickly check using standard trigonometric identities that

$$\sum_{j=0}^{\nu+1} (-1)^j \cos\left(\frac{j(n+m)\pi}{\nu+1}\right) = \begin{cases} 0, & n+m \neq \nu+1, \\ \nu+1, & n+m = \nu+1. \end{cases}$$

Thus we find $\hat{q}_m = \check{q}_m - \sum_{n=1}^{3s} q_{\nu+1+n}(\delta_{n+m,\nu+1} + \delta_{n-m,\nu+1})$ which implies

$$\begin{aligned} q_n &= \frac{1}{2}\check{q}_n, & n = 0, \nu+1, \\ q_n &= \check{q}_n, & n = 1, \dots, \nu-3s, \\ q_n &= \check{q}_n - q_{2\nu-n+2}, & n = \nu-3s+1, \dots, \nu. \end{aligned}$$

The remaining interpolation conditions, $q^{(j)}(0) = f^{(j)}(0), q^{(j)}(\pm 1) = f^{(j)}(\pm 1)$, for $j = 1, \dots, s$, are equivalent to the following $3s \times 3s$ system allowing us to find $q_{\nu+2}, \dots, q_{\nu+3s+1}$:

$$\begin{aligned} \sum_{n=1}^{3s} q_{\nu+1+n} \left[\mathbb{T}_{\nu+1+n}^{(j)}(-1) - \mathbb{T}_{\nu+1-n}^{(j)}(-1) \right] &= f^{(j)}(-1) - \sum_{n=0}^{\nu+1} \check{q}_n \mathbb{T}_n^{(j)}(-1), \\ \sum_{n=1}^{3s} q_{\nu+1+n} \left[\mathbb{T}_{\nu+1+n}^{(j)}(0) - \mathbb{T}_{\nu+1-n}^{(j)}(0) \right] &= f^{(j)}(0) - \sum_{n=0}^{\nu+1} \check{q}_n \mathbb{T}_n^{(j)}(0), \\ \sum_{n=1}^{3s} q_{\nu+1+n} \left[\mathbb{T}_{\nu+1+n}^{(j)}(1) - \mathbb{T}_{\nu+1-n}^{(j)}(1) \right] &= f^{(j)}(1) - \sum_{n=0}^{\nu+1} \check{q}_n \mathbb{T}_n^{(j)}(1). \end{aligned}$$

Note that the coefficients in this linear system can be found explicitly:

$$\begin{aligned} \mathbb{T}_n^{(j)}(\pm 1) &= (\pm 1)^{n-j} \frac{2^j j! n(n+j-1)!}{(2j)!(n-j)!}, & \text{for } 0 \leq j \leq n \text{ and } n+j \geq 1, \\ \mathbb{T}_n^{(j)}(0) &= \begin{cases} (-1)^r \frac{n(n-r-1)!}{r!} 2^{j-1}, & \text{if } r = (n-j)/2 \in \mathbb{N} \cup \{0\}, \\ 0, & \text{otherwise,} \end{cases} \end{aligned} \quad (5.34)$$

where the former expression is proved in Gao and Iserles (2017a, Eq. (2.3)) and the latter follows from the expansion of \mathbb{T}_n in the usual monomial basis (Abramowitz and Stegun, 1965, Eq. 22.3.6).

5.B Proof of Theorem 5.4.1

We recall the statement of Thm. 5.4.1:

Theorem 5.B.1. *Suppose the moments $\check{\sigma}_n$ are computed using Eq. (5.19) with slightly perturbed initial conditions: $\check{\sigma}_0 = \tilde{\sigma}_0 + \epsilon_0, \check{\sigma}_2 = \tilde{\sigma}_2 + \epsilon_2$, for some $|\epsilon_0|, |\epsilon_2| < \epsilon$. Then, for*

any n with $2n + 1 < \omega$,

$$|\check{\sigma}_{2n} - \tilde{\sigma}_{2n}| < \frac{8n\omega^{\frac{1}{2}}}{3(\omega^2 - (2n + 1)^2)^{\frac{1}{4}}} \left(2 + \frac{1}{\omega}\right) \epsilon.$$

Proof. Let $x_n = \check{\sigma}_n - \tilde{\sigma}_n$. By linearity it suffices to solve the recurrence Eq. (5.19) for x_n with initial conditions $x_0 = \epsilon_0$, $x_2 = \epsilon_2$, and $x_{2n+1} = 0$. Substitute $x_n = \rho_{n+1} - \rho_{n-1}$, $n \geq 0$, with $\rho_{-1} := -\epsilon_0/2$ and let $\gamma_n = \rho_{n+2} + \frac{2n}{i\omega}\rho_n - \rho_{n-2}$ for $n \geq 1$. We also note that $x_{2n+1} = 0$ so we may, without loss of generality, choose $\rho_{2n} = 0$, $n \geq 0$. Then the recurrence Eq. (5.19) is equivalent to solving

$$\gamma_{n+2} - 2\gamma_n + \gamma_{n-2} = 0, \quad n \geq 3,$$

with the initial conditions $\gamma_1 = \gamma_3 = \rho_3 + \frac{2}{i\omega}\rho_1 - \rho_{-1} = \epsilon_2 + \epsilon_0(1 + \frac{1}{i\omega})$. Here equality of $\gamma_1 = \gamma_3$ was achieved by setting $\rho_{-1} = -\epsilon_0/2$ and by using Eq. (5.19) for $n = 1$. Hence we have $\gamma_{2n+1} = \gamma_1$, $\forall n \geq 1$. Therefore, the problem of finding x_{2n} from given initial conditions is equivalent to

$$\rho_{n+2} + \frac{2n}{i\omega}\rho_n - \rho_{n-2} = \gamma_1, \quad n \geq 1, \quad (5.35)$$

subject to $\rho_1 = \epsilon_0/2$, $\rho_3 = \epsilon_2 + \epsilon_0/2$ and $\rho_{2n} = 0$, where $\gamma_1 = \epsilon_2 + \epsilon_0(1 + \frac{1}{i\omega})$. To solve this let us consider the homogeneous recurrence

$$a_{2n+3} + \frac{2(2n+1)}{i\omega}a_{2n+1} - a_{2n-1} = 0. \quad (5.36)$$

This has two linearly independent solutions that can be expressed in terms of spherical Bessel functions j_n, y_n (see Abramowitz and Stegun, 1965, §10.1) namely

$$a_{2n+1} = Ae^{i\frac{\pi}{2}n}j_n\left(\frac{\omega}{2}\right) + Be^{i\frac{\pi}{2}n}y_n\left(\frac{\omega}{2}\right),$$

for $A, B \in \mathbb{C}$. Let us write $\tilde{j}_n(z) := i^n j_n(z)$ and $\tilde{y}_n(z) := i^n y_n(z)$, then the solution to Eq. (5.36) with initial conditions a_1, a_3 is given by

$$\begin{pmatrix} a_{2n+3} \\ a_{2n+1} \end{pmatrix} = \begin{pmatrix} \tilde{j}_{n+1}(\omega/2) & \tilde{y}_{n+1}(\omega/2) \\ \tilde{j}_n(\omega/2) & \tilde{y}_n(\omega/2) \end{pmatrix} \begin{pmatrix} \tilde{j}_1(\omega/2) & \tilde{y}_1(\omega/2) \\ \tilde{j}_0(\omega/2) & \tilde{y}_0(\omega/2) \end{pmatrix}^{-1} \begin{pmatrix} a_3 \\ a_1 \end{pmatrix}, \quad n \geq 1.$$

Now we have the following useful identity (Abramowitz and Stegun, 1965, Eq. 10.1.31):

$$\tilde{j}_n(z)\tilde{y}_{n-1}(z) - \tilde{j}_{n-1}(z)\tilde{y}_n(z) = (-1)^{n+1}iz^{-2}, \quad n \geq 1.$$

Hence we have

$$\det \begin{pmatrix} \tilde{j}_{n+1}(\omega/2) & \tilde{y}_{n+1}(\omega/2) \\ \tilde{j}_n(\omega/2) & \tilde{y}_n(\omega/2) \end{pmatrix} = (-1)^{n+1} i \left(\frac{2}{\omega} \right)^2.$$

Thus, we can write the solution to Eq. (5.35), by discrete variation of constants, as

$$\begin{pmatrix} \rho_{2n+3} \\ \rho_{2n+1} \end{pmatrix} = i \left(\frac{\omega}{2} \right)^2 \begin{pmatrix} \tilde{j}_{n+1}(\omega/2) & \tilde{y}_{n+1}(\omega/2) \\ \tilde{j}_n(\omega/2) & \tilde{y}_n(\omega/2) \end{pmatrix} \left[\sum_{k=1}^n (-1)^{k+1} \begin{pmatrix} \tilde{y}_k(\omega/2) & -\tilde{y}_{k+1}(\omega/2) \\ -\tilde{j}_k(\omega/2) & \tilde{j}_{k+1}(\omega/2) \end{pmatrix} \begin{pmatrix} \gamma_1 \\ 0 \end{pmatrix} \right. \\ \left. + (-1) \begin{pmatrix} \tilde{y}_0(\omega/2) & -\tilde{y}_1(\omega/2) \\ -\tilde{j}_0(\omega/2) & \tilde{j}_1(\omega/2) \end{pmatrix} \begin{pmatrix} \rho_3 \\ \rho_1 \end{pmatrix} \right]. \quad (5.37)$$

According to Watson (1995, Eq. (1) in §13.74) we have, for $z \geq \nu + 1/2 \geq 1$,

$$|j_\nu(z)|^2 + |y_\nu(z)|^2 < \frac{1}{|z| \sqrt{z^2 - (\nu + 1/2)^2}}.$$

Thus, we can apply Cauchy–Schwarz to Eq. (5.37) and find

$$\begin{aligned} |\rho_{2n+3}| &\leq \left(\frac{\omega}{2} \right)^2 \left(\frac{1}{|\frac{\omega}{2}| \sqrt{(\frac{\omega}{2})^2 - (n + 3/2)^2}} \right)^{\frac{1}{2}} \left[\sum_{k=1}^n \left(\frac{1}{|\frac{\omega}{2}| \sqrt{(\frac{\omega}{2})^2 - (k + 1/2)^2}} \right)^{\frac{1}{2}} |\gamma_1| \right. \\ &\quad \left. + \left(\frac{1}{|\frac{\omega}{2}| \sqrt{(\frac{\omega}{2})^2 - (3/2)^2}} \right)^{\frac{1}{2}} |\rho_1| + \left(\frac{1}{|\frac{\omega}{2}| \sqrt{(\frac{\omega}{2})^2 - (1/2)^2}} \right)^{\frac{1}{2}} |\rho_3| \right] \\ &\leq \frac{\omega}{(\omega^2 - (2n + 3)^2)^{\frac{1}{4}}} \left[\sum_{k=1}^n \frac{|\gamma_1|}{(\omega^2 - (2k + 1)^2)^{\frac{1}{4}}} + \frac{|\rho_1|}{(\omega^2 - 3^2)^{\frac{1}{4}}} + \frac{|\rho_3|}{(\omega^2 - 1)^{\frac{1}{4}}} \right]. \end{aligned}$$

Finally, we notice by the integral test for $2n + 1 < \omega$:

$$\begin{aligned} \sum_{k=1}^n \frac{1}{(\omega^2 - (2k + 1)^2)^{\frac{1}{4}}} &\leq \int_0^n \frac{1}{(\omega + 2x + 1)^{\frac{1}{4}} (\omega - 2x - 1)^{\frac{1}{4}}} dx \leq \omega^{-\frac{1}{4}} \int_0^n \frac{dx}{(\omega - 2x - 1)^{\frac{1}{4}}} \\ &\leq \frac{2}{3} \omega^{-\frac{1}{4}} \left((\omega - 1)^{\frac{3}{4}} - (\omega - 2n - 1)^{\frac{3}{4}} \right) \\ &\leq \frac{2}{3} \omega^{-\frac{1}{2}} (\omega - (\omega - (2n + 1))) = \omega^{-\frac{1}{2}} \frac{2(2n + 1)}{3}. \end{aligned}$$

Thus, we have overall

$$|\rho_{2n+3}| \leq \frac{\omega^{\frac{1}{2}}}{(\omega^2 - (2n + 3)^2)^{\frac{1}{4}}} \left[\frac{2(2n + 1)}{3} \left(2 + \frac{1}{\omega} \right) + \frac{1}{2 \left(1 - \frac{3^2}{\omega^2} \right)^{\frac{1}{4}}} + \frac{3}{2 \left(1 - \frac{1^2}{\omega^2} \right)^{\frac{1}{4}}} \right] \epsilon.$$

A similar estimate holds for ρ_{2n+1} and hence the result follows, since $x_{2n+2} = \rho_{2n+3} - \rho_{2n+1}$. \square

5.C Proof of Theorem 5.4.2

We recall the statement of Thm. 5.4.2:

Theorem 5.C.1. *Suppose the moments $\check{\sigma}_n$ are computed using Eq. (5.18) with the perturbed initial conditions $\check{\sigma}_0 = \tilde{\sigma} + \epsilon_0$, $\check{\sigma}_1 = \check{\sigma}_{-1} = \tilde{\sigma}_1 + \epsilon_1$, $\check{\sigma}_2 = \check{\sigma}_{-2} = \tilde{\sigma}_2 + \epsilon_2$, $|\epsilon_j| < \epsilon$ for some $\epsilon > 0$, and assume $\check{\sigma}_3 = \check{\sigma}_{-3}$. Then, whenever $n + 1 < \min\{C\sqrt{\omega}, \omega\}$ for a given $C > 0$, we have*

$$|\check{\sigma}_n - \tilde{\sigma}_n| \leq \epsilon \frac{(K_0 + nK_1)}{2} \left(\frac{K_2\omega^{\frac{1}{2}}}{K_2\omega^{\frac{1}{2}} - 1} \exp\left(\frac{C}{K_2 - \omega^{-\frac{1}{2}}}\right) + 1 \right),$$

where the constants K_0, K_1, K_2 are independent of n and are given by

$$K_0 = \frac{2\sqrt{\omega}}{\sqrt{\omega - C^2}}, \quad K_1 = \frac{\omega + 2 + |\alpha|}{\sqrt{\omega^2 - C^2\omega}}, \quad K_2 = \frac{(\omega - C^2)^{\frac{1}{4}}}{\omega^{\frac{1}{4}}\sqrt{2|\alpha| + 2}}.$$

Proof. Define $x_n := \check{\sigma}_n - \tilde{\sigma}_n$. Then, by linearity, it suffices to solve for x_n which satisfies Eq. (5.18) subject to $x_0 = \epsilon_0, x_1 = x_{-1} = \epsilon_1, x_2 = x_{-2} = \epsilon_2, x_3 = x_{-3}$. We can formulate the recurrence equivalently in the form

$$\begin{aligned} x_{n-3} - \frac{2(n-2)}{i\omega}x_{n-2} - x_{n-1} + \frac{4}{i\omega}x_n - x_{n+1} + \frac{2(n+2)}{i\omega}x_{n+2} + x_{n+3} \\ = -\frac{2\alpha}{i\omega}(x_{n-2} - 2x_n + x_{n+2}) - \frac{2}{i\omega}(x_{n-2} + 2x_n + x_{n+2}) \end{aligned}$$

for $n \geq 0$. We can solve the homogeneous difference equation corresponding to the left hand side exactly, and we view the right hand side as a perturbation of the recurrence in the following sense: Let

$$x_n = \sum_{j=0}^{n-2} \omega^{-j} x_n^{(j)}, \quad n \geq 2.$$

Then x_n is the unique solution of the recurrence Eq. (5.18) with the specified initial conditions if we define

$$x_{n-3}^{(0)} - \frac{2(n-2)}{i\omega}x_{n-2}^{(0)} - x_{n-1}^{(0)} - x_{n+1}^{(0)} + \frac{2(n+2)}{i\omega}x_{n+2}^{(0)} + x_{n+3}^{(0)} = 0, \quad n \geq 0, \quad (5.38)$$

with the initial conditions $x_n^{(0)} = \epsilon_0, x_{-1}^{(0)} = x_1^{(0)} = \epsilon_1, x_{-2}^{(0)} = x_2^{(0)} = \epsilon_2$, and

$$x_{-3}^{(0)} = x_3^{(0)} = \frac{2\alpha - 2}{i\omega}\epsilon_0 + \epsilon_1 - \frac{2(3 + \alpha)}{i\omega}\epsilon_2,$$

and if we further choose

$$\begin{aligned} x_{n-3}^{(j+1)} - \frac{2(n-2)}{i\omega}x_{n-2}^{(j+1)} - x_{n-1}^{(j+1)} - x_{n+1}^{(j+1)} + \frac{2(n+2)}{i\omega}x_{n+2}^{(j+1)} + x_{n+3}^{(j+1)} \\ = 2i\alpha(x_{n-2}^{(j)} - 2x_n^{(j)} + x_{n+2}^{(j)}) + 2i(x_{n-2}^{(j)} + 2x_n^{(j)} + x_{n+2}^{(j)}) \end{aligned} \quad (5.39)$$

for $j \geq 0$ and $n \geq j$, under the extra symmetry condition $x_n^{(j)} = x_{-n}^{(j)}$ and with the initial conditions $x_{j+2}^{(j+1)}, x_{j+1}^{(j+1)}, x_j^{(j+1)}, x_{j-1}^{(j+1)} = 0$.

Let us firstly solve Eq. (5.38): We let $\gamma_n^{(0)} = x_{n+1}^{(0)} + \frac{2n}{i\omega}x_n^{(0)} - x_{n-1}^{(0)}$, $n \geq -2$, which ensures that Eq. (5.38) is equivalent to

$$\begin{aligned} \gamma_{n+2}^{(j+1)} - \gamma_{n-2}^{(j+1)} = 0, \quad n \geq 0, \\ \gamma_{-2}^{(0)} = \frac{2-2\alpha}{i\omega}\epsilon_0 + \frac{2+2\alpha}{i\omega}\epsilon_2, \quad \gamma_{-1}^{(0)} = \epsilon_0 - \frac{2}{i\omega}\epsilon_1 - \epsilon_2, \quad \gamma_0^{(0)} = 0, \quad \gamma_1^{(0)} = \epsilon_2 + \frac{2}{i\omega}\epsilon_1 - \epsilon_0. \end{aligned}$$

Thus $\gamma_{4n+j}^{(0)} = \gamma_j^{(0)}$ for $j = -2, \dots, 1$ and $n \geq 0$. Hence, it remains to solve

$$x_{n+1}^{(0)} + \frac{2n}{i\omega}x_n^{(0)} - x_{n-1}^{(0)} = \gamma_n^{(0)}, \quad n \geq 1, \quad (5.40)$$

with initial conditions $x_0^{(0)} = \epsilon_0, x_1^{(0)} = \epsilon_1$. As described by Domínguez et al. (2011) the homogeneous solutions of this recurrence can be expressed in terms of Bessel functions, where it will be convenient to express the solutions in terms of the functions $\tilde{J}_n(\omega) := i^n J_n(\omega)$, $\tilde{Y}_n(\omega) := i^n Y_n(\omega)$, where $J_n(\omega), Y_n(\omega)$ are the standard Bessel functions of the first and second kind, as defined for instance in Abramowitz and Stegun (1965). The solution to Eq. (5.40) can be written using discrete variation of constants as

$$\begin{aligned} \begin{pmatrix} x_{n+1}^{(0)} \\ x_n^{(0)} \end{pmatrix} = \frac{i\pi\omega}{2} \begin{pmatrix} \tilde{J}_{n+1}(\omega) & \tilde{Y}_{n+1}(\omega) \\ \tilde{J}_n(\omega) & \tilde{Y}_n(\omega) \end{pmatrix} \left(\sum_{k=1}^n (-1)^{k+1} \begin{pmatrix} \tilde{Y}_k(\omega) & -\tilde{Y}_{k+1}(\omega) \\ -\tilde{J}_k(\omega) & \tilde{J}_{k+1}(\omega) \end{pmatrix} \begin{pmatrix} \gamma_k^{(0)} \\ 0 \end{pmatrix} \right. \\ \left. + (-1) \begin{pmatrix} \tilde{Y}_0(\omega) & -\tilde{Y}_1(\omega) \\ -\tilde{J}_0(\omega) & \tilde{J}_1(\omega) \end{pmatrix} \begin{pmatrix} \epsilon_1 \\ \epsilon_0 \end{pmatrix} \right), \end{aligned} \quad (5.41)$$

for $n \geq 1$. Here, analogously to Domínguez et al. (2011), we used the identity (Abramowitz and Stegun, 1965, Eq. (9.1.16))

$$\det \begin{pmatrix} \tilde{J}_{n+1}(\omega) & \tilde{Y}_{n+1}(\omega) \\ \tilde{J}_n(\omega) & \tilde{Y}_n(\omega) \end{pmatrix} = (-1)^{n+1} \frac{-2i}{\pi\omega}.$$

We can now perform a similar estimate to Domínguez et al. (2011, p. 1271) on Eq. (5.41): Note the upper bound given by Watson (1995, §13.74)

$$|\mathbf{J}_n(\omega)|^2 + |\mathbf{Y}_n(\omega)|^2 \leq \frac{2}{\pi} \frac{1}{\sqrt{\omega^2 - n^2}}, \quad \text{for } \omega > n > 1/2. \quad (5.42)$$

Combining this with Cauchy–Schwarz on Eq. (5.41) yields, for $n \geq 2$,

$$|x_n^{(0)}| \leq \frac{\omega}{(\omega^2 - n^2)^{\frac{1}{4}}} \left(\sum_{k=1}^{n-1} \frac{1}{(\omega^2 - k^2)^{\frac{1}{4}}} |\gamma_k^{(0)}| + \epsilon \sum_{j=0}^1 \frac{1}{(\omega^2 - j^2)^{\frac{1}{4}}} \right).$$

Thus, summing these contributions, we obtain the following estimate when $n < \min\{C\sqrt{\omega}, \omega\}$:

$$|x_n^{(0)}| \leq \frac{\epsilon}{\sqrt{1 - \frac{C^2}{\omega}}} \left(n \left(1 + \frac{2 + |\alpha|}{\omega} \right) + 2 \right). \quad (5.43)$$

We now consider the perturbed recurrence Eq. (5.39) order by order. To do so, let us write

$$f_n^{(j)} = 2i(1 + \alpha) \left(x_{n-2}^{(j)} + x_{n+2}^{(j)} \right) + 2i(1 - \alpha)x_n^{(j)}, \quad n \geq j.$$

Thus, for $j \geq 0$, we need to solve the recurrence

$$x_{n-3}^{(j+1)} - \frac{2(n-2)}{i\omega} x_{n-2}^{(j+1)} - x_{n-1}^{(j+1)} - x_{n+1}^{(j+1)} + \frac{2(n+2)}{i\omega} x_{n+2}^{(j+1)} + x_{n+3}^{(j+1)} = f_n^{(j)}, \quad n \geq j$$

subject to the initial conditions $x_{j+2}^{(j+1)}, x_{j+1}^{(j+1)}, x_j^{(j+1)}, x_{j-1}^{(j+1)} = 0$. We again substitute $\gamma_n^{(j)} = x_{n+1}^{(j)} + \frac{2n}{i\omega} x_n^{(j)} - x_{n-1}^{(j)}$, $n \geq j - 2$, which ensures that this recurrence is equivalent to

$$\gamma_{n+2}^{(j+1)} - \gamma_{n-2}^{(j+1)} = f_n^{(j)}, \quad n \geq j - 2, \quad \gamma_{j-2}^{(j+1)}, \gamma_{j-1}^{(j+1)}, \gamma_j^{(j+1)}, \gamma_{j+1}^{(j+1)} = 0$$

Therefore, we easily find

$$\gamma_{4n+j}^{(j+1)} = \sum_{m=0}^{n-1} f_{4m+j+2}^{(j)}, \quad \gamma_{4n+j+1}^{(j+1)} = \sum_{m=0}^{n-1} f_{4m+j+3}^{(j)}, \quad \gamma_{4n+j-1}^{(j+1)} = \sum_{m=0}^{n-1} f_{4m+j+1}^{(j)}, \quad \gamma_{4n+j-2}^{(j+1)} = \sum_{m=0}^{n-1} f_{4m+j}^{(j)}.$$

Now it remains to solve

$$\gamma_n^{(j)} = x_{n+1}^{(j)} + \frac{2n}{i\omega} x_n^{(j)} - x_{n-1}^{(j)}, \quad n \geq j + 2$$

with the initial conditions $x_{j+2}^{(j+1)}, x_{j+1}^{(j+1)} = 0$. Similarly to the case for $x_n^{(0)}$ we can write the solution in terms of $\tilde{\mathbf{J}}_n(\omega), \tilde{\mathbf{Y}}_n(\omega)$, which yields

$$\begin{pmatrix} x_{n+1}^{(j+1)} \\ x_n^{(j+1)} \end{pmatrix} = \frac{i\pi\omega}{2} \begin{pmatrix} \tilde{J}_{n+1}(\omega) & \tilde{Y}_{n+1}(\omega) \\ \tilde{J}_n(\omega) & \tilde{Y}_n(\omega) \end{pmatrix} \sum_{k=j+2}^n (-1)^{k+1} \begin{pmatrix} \tilde{Y}_k(\omega) & -\tilde{Y}_{k+1}(\omega) \\ -\tilde{J}_k(\omega) & \tilde{J}_{k+1}(\omega) \end{pmatrix} \begin{pmatrix} \gamma_k^{(j+1)} \\ 0 \end{pmatrix}.$$

Therefore, we can use Cauchy–Schwarz and Eq. (5.42), as we did previously, to estimate

$$|x_n^{(j+1)}| \leq \frac{\pi\omega}{2} \frac{2}{\pi} \frac{1}{(\omega^2 - n^2)^{\frac{1}{4}}} \sum_{k=j+2}^{n-1} \frac{1}{(\omega^2 - k^2)^{\frac{1}{4}}} |\gamma_k^{(j+1)}|. \quad (5.44)$$

Now we recall $f_n^{(j)} = 2i\alpha(x_{n-2}^{(j)} - 2x_n^{(j)} + x_{n+2}^{(j)}) + 2i(x_{n-2}^{(j)} + 2x_n^{(j)} + x_{n+2}^{(j)})$, which means that

$$\begin{aligned} \gamma_{4n+j}^{(j+1)} &= \sum_{m=0}^{n-1} f_{4m+j+2}^{(j)} = 2i\alpha \left(x_{4n+j}^{(j)} + 2 \sum_{l=1}^{2n-1} (-1)^l x_{j+2l}^{(j)} \right) + 2i \left(x_{4n+j}^{(j)} + 2 \sum_{l=1}^{2n-1} x_{j+2l}^{(j)} \right), \\ \therefore |\gamma_{4n+j}^{(j+1)}| &\leq (2\alpha + 2) \left(|x_{4n+j}^{(j)}| + 2 \sum_{l=1}^{2n-1} |x_{j+2l}^{(j)}| \right). \end{aligned}$$

Analogously we find for $k = -2, -1, 1$:

$$|\gamma_{4n+k+j}^{(j+1)}| \leq (2\alpha + 2) \left(|x_{4n+k+j}^{(j)}| + 2 \sum_{l=1}^{2n-1} |x_{j+2l+k}^{(j)}| \right),$$

where, of course, $x_j^{(j)}, x_{j+1}^{(j)} = 0$. To complete a total estimate on the size of $x_n^{(j+1)}$ we proceed as follows:

Claim 5.C.2. *If $|x_{n+j}^{(j)}| \leq n^b$, for $b \geq 0$ and all $n + j + 1 < C\sqrt{\omega}$, then*

$$|x_{n+j+1}^{(j+1)}| \leq \frac{2 + 2|\alpha|}{\sqrt{1 - \frac{C^2}{\omega}}} \left(\frac{n^{b+2}}{(b+2)(b+1)} + 2\frac{n^{b+1}}{b+1} + n^b \right).$$

Proof of Claim.

$$\begin{aligned} |\gamma_{4n+j}^{(j+1)}| &\leq (2|\alpha| + 2) \left(|x_{4n+j}^{(j)}| + 2 \sum_{l=1}^{2n-1} |x_{j+2l}^{(j)}| \right) \leq C(2|\alpha| + 2) \left((4n)^b + 2 \sum_{l=1}^{2n-1} (2l)^b \right) \\ &\leq C(2|\alpha| + 2) \left((4n)^b + 2 \int_0^{2n} (2x)^b dx \right) = C(2|\alpha| + 2) \left((4n)^b + \frac{1}{b+1} (4n)^{b+1} \right), \end{aligned}$$

where in the final step we used the integral test to find an upper bound. Analogously, we find in general,

$$|\gamma_{k+j}^{(j+1)}| \leq C(2|\alpha| + 2) \left(\frac{1}{b+1} k^{b+1} + k^b \right), \quad k \geq 2.$$

Thus we have, based on Eq. (5.44),

$$\begin{aligned} |x_{n+j+1}^{(j+1)}| &\leq \frac{\omega}{(\omega^2 - (n+j+1)^2)^{\frac{1}{4}}} \sum_{k=j+2}^{n+j} \frac{1}{(\omega^2 - k^2)^{\frac{1}{4}}} |\gamma_k^{(j+1)}| \\ &\leq \frac{C(2|\alpha|+2)\omega}{(\omega^2 - (n+j+1)^2)^{\frac{1}{4}}} \sum_{k=2}^n \frac{1}{(\omega^2 - (j+k)^2)^{\frac{1}{4}}} \left(\frac{1}{b+1} k^{b+1} + k^b \right). \end{aligned}$$

Thus, if $n+1+j < C\sqrt{\omega}$, we can simplify the above estimate to complete the proof of the claim:

$$|x_{n+j+1}^{(j+1)}| \leq \frac{C(2|\alpha|+2)}{\sqrt{1-\frac{C^2}{\omega}}} \sum_{k=2}^n \frac{1}{b+1} k^{b+1} + k^b \leq \frac{C(2|\alpha|+2)}{\sqrt{1-\frac{C^2}{\omega}}} \left[\frac{n^{b+2}}{(b+2)(b+1)} + 2\frac{n^{b+1}}{(b+1)} + n^b \right].$$

■

We have shown in Eq. (5.43) that

$$|x_n^{(0)}| \leq n\epsilon \underbrace{\frac{\left(1 + \frac{2+|\alpha|}{\omega}\right)}{\sqrt{1-\frac{C^2}{\omega}}}}_{K_1} + \epsilon \underbrace{\frac{2}{\sqrt{1-\frac{C^2}{\omega}}}}_{K_0}.$$

Thus we have, by linearity, for $n \geq 2$

$$\begin{aligned} |x_n| &\leq \sum_{j=0}^{n-2} \omega^{-j} |x_n^{(j)}| \leq \sum_{j=0}^{n-2} \left(\frac{(2|\alpha|+2)}{\omega\sqrt{1-\frac{C^2}{\omega}}} \right)^j \left[\sum_{l=0}^{2j} \binom{2j}{l} \frac{(n-j)^l}{l!} \left(K_0\epsilon + K_1\epsilon \frac{n-j}{l+1} \right) \right] \\ &\leq (K_0 + nK_1)\epsilon \sum_{j=0}^{n-2} \left(\frac{(2|\alpha|+2)}{\omega\sqrt{1-\frac{C^2}{\omega}}} \right)^j L_{2j}(-n+j), \end{aligned} \quad (5.45)$$

where L_{2j} are Laguerre polynomials and the final line follows from the explicit expansion of L_{2j} in the usual monomial basis (Abramowitz and Stegun, 1965, Eq. 22.3.9). Thus we seek to find an upper bound for the function $f(z, n) = \sum_{j=0}^{n-2} z^{-j} L_{2j}(-n+j)$. Note to begin with that Laguerre polynomials are strictly monotonically decreasing for negative arguments, i.e. for $x < y < 0$ we have for any $n \geq 1$

$$L_n(x) > L_n(y) > L_n(0) = 1, \text{ and } L_0 \equiv 1.$$

This follows by induction from the following identity (Abramowitz and Stegun, 1965, Eqs. 22.5.17 & 22.7.30)

$$\frac{d}{dx}L_{n+1} = \frac{d}{dx}L_n - L_n, \quad n \geq 0,$$

since we have $L_n(x) > 0$ for any $x < 0, n \geq 0$ (because the zeros of the Laguerre polynomials are located in $[0, \infty)$ and $L_n(0) = 1$). Thus, we may estimate

$$f(z, n) \leq \sum_{j=0}^{n-2} \left(z^{\frac{1}{2}}\right)^{-2j} L_{2j}(-n) \leq \frac{1}{2} \sum_{j=0}^{n-2} \left(z^{\frac{1}{2}}\right)^{-2j} L_{2j}(-n) + \frac{1}{2} \sum_{j=0}^{n-2} \left(-z^{\frac{1}{2}}\right)^{-2j} L_{2j}(-n).$$

By the three-term recurrence for Laguerre polynomials (Abramowitz and Stegun, 1965, 22.7.12),

$$L_{n+1}(x) = L_n(x) - \frac{x}{n+1}L_n + \frac{n}{n+1}(L_n(x) - L_{n-1}(x)),$$

and by induction we have $L_{n+1}(x) > L_n(x)$ for any $x < 0, n \geq 0$. Therefore,

$$\begin{aligned} f(z, n) &= f(z, n) + z^{-\frac{1}{2}}f(z, n) - z^{-\frac{1}{2}}f(z, n) \\ &\leq \frac{1}{2} \sum_{j=0}^{\infty} z^{-\frac{j}{2}}L_j(-n) + \frac{1}{2} \sum_{j=0}^{\infty} \left(-z^{\frac{1}{2}}\right)^j L_j(-n). \end{aligned} \quad (5.46)$$

The generating function of the Laguerre polynomials (Abramowitz and Stegun, 1965, Eq. 22.9.15) is

$$\sum_{j=0}^{\infty} a^j L_j(x) = \frac{1}{1-a} \exp\left(\frac{xa}{a-1}\right), \quad \forall x \in \mathbb{R}, |a| < 1,$$

which allows us to simplify the estimate Eq. (5.46) to

$$f(z, n) \leq \frac{1}{2} \frac{z^{\frac{1}{2}}}{z^{\frac{1}{2}} - 1} \exp\left(\frac{n}{z^{\frac{1}{2}} - 1}\right) + \frac{1}{2} \frac{1}{1 + z^{\frac{1}{2}}} \exp\left(-\frac{n}{z^{\frac{1}{2}} + 1}\right).$$

Therefore we conclude, by Eq. (5.45),

$$|x_n| \leq \frac{(K_0 + nK_1)}{2} \epsilon \left(\frac{K_2 \omega^{\frac{1}{2}}}{K_2 \omega^{\frac{1}{2}} - 1} \exp\left(\frac{n}{\omega^{\frac{1}{2}} K_2 - 1}\right) + \frac{1}{K_2 \omega^{\frac{1}{2}} + 1} \exp\left(-\frac{n}{\omega^{\frac{1}{2}} K_2 + 1}\right) \right),$$

where $K_2 = (2|\alpha| + 2)^{-1/2}(1 - C^2/\omega)^{1/4}$ and the result follows. \square

5.D Expression for initial moments in §5.5.2

Let us define the standard moments by $\rho_n := I_{\omega, \beta}^{(3)}[x^n]$, then we have the expressions

$$\tilde{\sigma}_0 = \rho_0, \quad \tilde{\sigma}_1 = 2\rho_1 - \rho_0, \quad \tilde{\sigma}_2 = 8\rho_2 - 8\rho_1 + \rho_0, \quad \tilde{\sigma}_3 = 32\rho_3 - 48\rho_2 + 18\rho_1 - \rho_0.$$

Differentiating Eq. (5.26) with respect to β yields:

$$\begin{aligned} i\omega\rho_1 &= -\frac{2i}{\pi}\omega^{-\frac{3}{2}}e^{i(\beta+1)\omega} \int_0^\infty \left[i\omega g_\beta \left(\frac{t}{\omega} \right) + \left(\partial_\beta g_\beta \left(\frac{t}{\omega} \right) \right) \right] \frac{1}{\sqrt{t}} e^{-t} dt \\ &\quad + \begin{cases} \frac{1}{1-\beta^2} + \frac{2\beta \tanh^{-1}\left(\sqrt{\frac{\beta-1}{\beta+1}}\right)}{(\beta^2-1)^{3/2}}, & \beta > -1, \beta \neq 1, \\ 1/3, & \beta = 1, \\ \frac{1}{1-\beta^2} - \frac{2\beta \tanh^{-1}\left(\sqrt{\frac{\beta-1}{\beta+1}}\right)}{(\beta^2-1)^{3/2}}, & \beta < -1, \end{cases} \\ -\omega^2\rho_2 &= -\frac{2i}{\pi}\omega^{-\frac{3}{2}} \int_0^\infty \left[-\omega^2 g_\beta \left(\frac{t}{\omega} \right) + 2i\omega \left(\partial_\beta g_\beta \left(\frac{t}{\omega} \right) \right) + \left(\partial_\beta^2 g_\beta \left(\frac{t}{\omega} \right) \right) \right] \frac{1}{\sqrt{t}} e^{-t} dt \\ &\quad + \begin{cases} \frac{3\beta}{(\beta^2-1)^2} - \frac{(4\beta^2+2) \tanh^{-1}\left(\sqrt{\frac{\beta-1}{\beta+1}}\right)}{(\beta^2-1)^{5/2}}, & \beta > -1, \beta \neq 1, \\ -4/15, & \beta = 1, \\ \frac{3\beta}{(\beta^2-1)^2} + \frac{(4\beta^2+2) \tanh^{-1}\left(\sqrt{\frac{\beta-1}{\beta+1}}\right)}{(\beta^2-1)^{5/2}}, & \beta < -1, \end{cases} \end{aligned}$$

and, when $\beta \neq 1$, we find additionally:

$$\begin{aligned} -i\omega^3\rho_3 &= -\frac{2i}{\pi}\omega^{-\frac{3}{2}}e^{i(\beta+1)\omega} \int_0^\infty \left[-i\omega^3 g_\beta \left(\frac{t}{\omega} \right) - 3\omega^2 \left(\partial_\beta g_\beta \left(\frac{t}{\omega} \right) \right) \right. \\ &\quad \left. + 3i\omega \left(\partial_\beta^2 g_\beta \left(\frac{t}{\omega} \right) \right) + \left(\partial_\beta^3 g_\beta \left(\frac{t}{\omega} \right) \right) \right] \frac{1}{\sqrt{t}} e^{-t} dt \\ &\quad + \frac{2}{\pi}\omega^{-1} \begin{cases} \frac{11\beta^2+4}{(\beta^2-1)^3} - \frac{6\beta(2\beta^2+3) \tanh^{-1}\left(\sqrt{\frac{\beta-1}{\beta+1}}\right)}{(\beta-1)^{7/2}(\beta+1)^{7/2}}, & \beta > -1, \beta \neq 1, \\ \frac{11\beta^2+4}{(\beta^2-1)^3} + \frac{6\beta(2\beta^2+3) \tanh^{-1}\left(\sqrt{\frac{\beta-1}{\beta+1}}\right)}{(\beta-1)^{7/2}(\beta+1)^{7/2}}, & \beta < -1. \end{cases} \end{aligned}$$

5.E Proof of Proposition 5.5.2

We recall the statement of Prop. 5.5.2:

Proposition 5.E.1 (Filon paradigm for $I_{\omega, \beta}^{(3)}$). *For any $k \in \mathbb{N}$ there is a constant $C_k > 0$ such that for all $\beta \in \mathbb{R}, \beta \neq -1, \omega \geq 1$ and any function $\tilde{f} \in C^{k+2}[0, 1]$ with $\tilde{f}^{(j)}(\pm 1) = 0$ for $j = 0, \dots, k$:*

$$\left| I_{\omega, \beta}^{(3)}[\tilde{f}] \right| \leq C_k \left(\omega^{-(k+2)} \|\tilde{f}^{(k+1)}\|_\infty \frac{|\beta+1|^{k+3} - 1}{|\beta+1| - 1} + \omega^{-(k+2)} \log \omega \|\tilde{f}^{(k+2)}\|_\infty |\beta+1|^{-(k+2)} \right).$$

Proof. We write

$$I_{\omega,\beta}^{(3)}[\tilde{f}] = \int_0^{\omega^{-1}} \mathbf{H}_0^{(1)}(\omega x) \tilde{f}(2x-1) e^{i\omega\beta x} dx + \int_{\omega^{-1}}^1 \mathbf{H}_0^{(1)}(\omega x) \tilde{f}(2x-1) e^{i\omega\beta x} dx. \quad (5.47)$$

By Taylor's theorem we have $|\tilde{f}^{(j)}(2x-1)| \leq \tilde{C}_k x^{k+1-j} \|\tilde{f}^{(k+1)}\|_\infty$, $j = 0, \dots, k+1$, for all $x \in [-1, 1]$ and for some constant $\tilde{C}_k > 0$ independent of x . To bound the first integral note, by Lemma 5.5.1 with $n = 0$, that

$$\left| \mathbf{H}_0^{(1)}(\omega x) \right| \leq 2C_0(\omega x)^{-1/2}, \quad \forall x > 0,$$

since $1 + \log(1/z) \leq 2z^{-1/2}$, $|h_0(z)| \leq 2C_0 z^{-1/2}$ when $z \leq 1$. Thus we have

$$\begin{aligned} \left| \int_0^{\omega^{-1}} \mathbf{H}_0^{(1)}(\omega x) \tilde{f}(2x-1) e^{i\omega\beta x} dx \right| &\leq 2C_0 \tilde{C}_k \omega^{-\frac{1}{2}} \|\tilde{f}^{(k+1)}\|_\infty \int_0^{\omega^{-1}} x^{k+\frac{1}{2}} dx \\ &\lesssim \omega^{-k-2} \|\tilde{f}^{(k+1)}\|_\infty, \end{aligned} \quad (5.48)$$

where $A(\omega) \lesssim B(\omega)$ means $A(\omega) \leq KB(\omega)$ for a constant $K > 0$ independent of ω . Moreover, by integration by parts and noting that h_0 is non-singular on $(0, 1]$, and $\tilde{f}^{(j)}(1) = 0$, $j = 0, \dots, k$, we have

$$\begin{aligned} \int_{\omega^{-1}}^1 \mathbf{H}_0^{(1)}(\omega x) \tilde{f}(2x-1) e^{i\omega\beta x} dx &= \int_{\omega^{-1}}^1 h_0(\omega x) \tilde{f}(2x-1) e^{i\omega(\beta+1)x} dx \\ &= - \sum_{j=0}^{k+1} \left(\frac{-1}{i\omega(\beta+1)} \right)^{j+1} \left[e^{i\omega(\beta+1)x} \frac{d^j}{dx^j} \left(h_0(\omega x) \tilde{f}(2x-1) \right) \right]_{x=\omega^{-1}} \\ &\quad + \left(\frac{-1}{i\omega(\beta+1)} \right)^{k+2} \left[e^{i\omega(\beta+1)x} \frac{d^{k+1}}{dx^{k+1}} \left(h_0(\omega x) \tilde{f}(2x-1) \right) \right]_{x=1} \\ &\quad + \left(\frac{-1}{i\omega(\beta+1)} \right)^{k+2} \int_{\omega^{-1}}^1 e^{i\omega(\beta+1)x} \frac{d^{k+2}}{dx^{k+2}} \left(h_0(\omega x) \tilde{f}(2x-1) \right) dx. \end{aligned} \quad (5.49)$$

We bound each term in turn using the Leibniz rule for the derivatives of a product:

$$\begin{aligned} \left| \left[\frac{d^j}{dx^j} \left(h_0(\omega x) \tilde{f}(2x-1) \right) \right]_{x=\omega^{-1}} \right| &\lesssim \sum_{l=0}^j \omega^l \left| \left[\frac{d^l h_0}{dx^l} \right]_{x=1} \right| \left| \tilde{f}^{(j-l)}(2\omega^{-1}-1) \right| \\ &\lesssim \omega^{-k-1+j} \|\tilde{f}^{(k+1)}\|_{k+1}, \end{aligned} \quad (5.50)$$

where we used $\left| \tilde{f}^{(l)}(-1+2\omega^{-1}) \right| \leq \tilde{C}_k \omega^{k+1-l} \|\tilde{f}^{(k+1)}\|_{k+1}$, $l = 0, \dots, k+1$. Similarly we find

$$\left| \left[\frac{d^{k+1}}{dx^{k+1}} \left(h_0(\omega x) \tilde{f}(2x-1) \right) \right]_{x=1} \right| = 2^{k+1} \left| \left(h_0(\omega) \tilde{f}^{(k+1)}(1) \right) \right| \lesssim \omega^{-1/2} \|\tilde{f}^{(k+1)}\|_\infty, \quad (5.51)$$

where the first equality holds because $\tilde{f}^{(j)}(1) = 0, j = 0, \dots, k$. Finally, we have

$$\begin{aligned} \left| \frac{d^{k+2}}{dx^{k+2}} \left(h_0(\omega x) \tilde{f}(2x-1) \right) \right| &\leq \sum_{l=0}^{k+2} \binom{k+2}{l} \left| \frac{d^l}{dx^l} h_0(\omega x) \right| \left| 2^{k+2-l} \tilde{f}^{(k+2-l)}(2x-1) \right| \\ &\lesssim \omega^{-1/2} x^{-1/2} \left| \tilde{f}^{(k+2)}(2x-1) \right| + \sum_{l=1}^{k+2} \omega^{-1/2} x^{-l-1/2} \left| \tilde{f}^{(k+2-l)}(2x-1) \right| \\ &\lesssim \omega^{-1/2} x^{-1/2} \|\tilde{f}^{(k+2)}\|_\infty + \sum_{l=1}^{k+2} \omega^{-1/2} x^{-l-1/2} x^{l-1} \|\tilde{f}^{(k+1)}\|_\infty \lesssim x^{-1} \|\tilde{f}^{(k+2)}\|_\infty, \end{aligned}$$

where the final estimate holds uniformly in $\omega^{-1} \leq x \leq 1$, since in that region $\omega^{-1/2} x^{-1/2} \leq 1$. Therefore,

$$\begin{aligned} \left| \int_{\omega^{-1}}^1 e^{i\omega(\beta+1)x} \frac{d^{k+2}}{dx^{k+2}} \left(h_0(\omega x) \tilde{f}(2x-1) \right) dx \right| &\lesssim \|\tilde{f}^{(k+2)}\|_\infty \int_{\omega^{-1}}^1 x^{-1} dx \\ &\lesssim \|\tilde{f}^{(k+2)}\|_\infty \log \omega. \end{aligned} \quad (5.52)$$

Thus, combining Eqs. (5.47)–(5.52) yields the estimate

$$\left| I_{\omega, \beta}^{(3)}[\tilde{f}] \right| \lesssim \omega^{-(k+2)} \|\tilde{f}^{(k+1)}\|_\infty \sum_{j=0}^{k+2} |\beta+1|^{-j} + \omega^{-(k+2)} \log \omega |\beta+1|^{-(k+2)} \|\tilde{f}^{(k+2)}\|_\infty,$$

which completes the proof. \square

Chapter 6

An efficient Levin–Clenshaw–Curtis method for a class of highly oscillatory integrals

6.1 Introduction

In this chapter we will explore a further application of the ideas from Thm. 5.3.2 which we proved in Chapter 5. In Thm. 5.3.2 we showed that a function which lies in the nullspace of a certain differential operator has Chebyshev coefficients that satisfy a linear recurrence involving a finite number of terms. In the previous chapter this insight was used to construct a very efficient way of computing the Filon quadrature moments in certain cases. Here we will use this observation to construct an efficient Levin–Clenshaw–Curtis method for the computation of highly oscillatory integrals similar to the ones considered in the previous chapter. Originally introduced in two seminal papers by Levin (1982, 1996), the Levin method provides an alternative to numerical steepest descent (NSD) and Filon methods. The central idea, which we will outline in some more detail in §6.2, is to replace the computation of a highly oscillatory integral with the solution of a non-oscillatory ordinary differential equation (ODE). One can then proceed to solve this ODE with a collocation method, and, if the collocation points are chosen carefully, the solution of the ODE provides a way of approximating the original integral at frequency independent cost. The Levin method has the advantage that, unlike NSD, it relies on evaluations of the amplitude function (f as defined in Eq. (6.1)) strictly on the domain of integration, and, unlike the Filon method, it does not require the computation of quadrature moments. However, the Levin method cannot be constructed in the presence of singularities and stationary points (Deaño et al., 2017), and therefore its practical use is more limited than that of the other two methods. Nevertheless, in those cases when the Levin method can be constructed, it provides a fast and uniformly accurate alternative to approximate

highly-oscillatory integrals.

The collocation part in the Levin method typically leads to a dense $(\nu + 2) \times (\nu + 2)$ linear system, where $\nu + 2$ is the number of collocation points, and thus, in general, the Levin method costs $\mathcal{O}(\nu^3)$ operations to evaluate. In a similar spirit as the speed-up of the Filon method achieved by the Filon–Clenshaw–Curtis construction (see Domínguez et al. (2011) and Gao and Iserles (2017a)), we are able to exploit the use of a Chebyshev basis and the recursion for the spectral coefficients of the oscillatory kernel to achieve a significant acceleration, solving the collocation equations in just $\mathcal{O}(\nu \log \nu)$ operations.

This chapter is structured as follows: In §6.2 we begin with a brief introduction to Levin methods as can be found in the relevant literature. In §6.3 we introduce an efficient Levin–Clenshaw–Curtis method, which is able to compute the full Levin approximation in just $\mathcal{O}(\nu \log \nu)$ operations. The algorithm is summarised in §6.3.1, and we provide two numerical examples in §6.4 that show how the method performs in practice, highlighting both the uniform convergence across all frequencies $\omega \geq 0$ and the significant speed-up achieved when compared to a direct solution of the collocation equations. We conclude the chapter in §6.5 with a summary of our results.

6.2 The Levin method

In this chapter we are interested in constructing efficient approximations to integrals of the form

$$I_\omega[f] := \int_{-1}^1 f(x) e^{i\omega g(x)} dx, \quad (6.1)$$

where $g(x)$ is a polynomial of finite degree $d \in \mathbb{N}$, with $g'(x) \neq 0, \forall x \in [-1, 1]$. Our goal is to build on the method introduced by Levin (1982, 1996) and to demonstrate how the use of a Chebyshev polynomial basis and the discrete computation of Chebyshev spectral coefficients can lead to a significant speed-up of the Levin method in the setting of Eq. (6.1). Let us begin with a quick overview of how a general Levin method for $I_\omega[f]$ may be constructed, and refer the interested reader to Deaño et al. (2017, §3.3) and the original papers by Levin (1982, 1996) for a comprehensive introduction to the topic.

The central idea of the Levin method is to relate the computation of $I_\omega[f]$ to the solution of an ordinary differential equation of non-oscillatory nature. As a first step consider a function $u_1(x)$ which satisfies

$$u_1'(x) + i\omega g'(x)u_1(x) = f(x), \quad x \in [-1, 1]. \quad (6.2)$$

We can use the function u_1 to write the integrand of $I_\omega[f]$ as an exact differential

$$f(x)e^{i\omega g(x)} dx = \frac{d}{dx} (u_1(x)e^{i\omega g(x)}) dx = d(u_1(x)e^{i\omega g(x)}),$$

and therefore to compute $I_\omega[f]$ by evaluating u_1 :

$$I_\omega[f] = u_1(1)e^{i\omega g(1)} - u_1(-1)e^{i\omega g(-1)}.$$

The idea is that in solving for a non-oscillatory solution of Eq. (6.2) the numerical method requires significantly fewer degrees of freedom than would be required to approximate the oscillatory integral $I_\omega[f]$ directly. This idea is supported by the following observation:

Theorem 6.2.1 (See Levin (1996, Appendix)). *Suppose $f \in C^\infty([-1, 1])$, and that $f(g^{-1}(\xi))/g'(g^{-1}(\xi))$ is slowly oscillatory in the sense that its spectrum is bounded, i.e. there is a smooth function $H \in C^\infty([-1, 1])$ and $w_0 > 0$ such that $w_0 \ll \omega$ and*

$$\frac{f(g^{-1}(\xi))}{g'(g^{-1}(\xi))} = \int_{-w_0}^{w_0} H(t)e^{i\xi t} dt.$$

Then Eq. (6.2) has a slowly varying solution $u_1^{(\omega)} \in C^\infty([-1, 1])$ with spectrum in $[-w_0, w_0]$, i.e. there is a smooth $\tilde{H} \in C^\infty([-1, 1])$ such that

$$u_1^{(\omega)}(g^{-1}(\xi)) = \int_{-w_0}^{w_0} \tilde{H}(t)e^{i\xi t} dt.$$

The Levin method then solves Eq. (6.2) using collocation, i.e. we seek an approximation $u_1^{[\nu]} = \sum_{j=0}^{\nu+1} \alpha_j \psi_j(x)$ for some $\nu \in \mathbb{N}$ and some basis functions $\{\psi_j\}_{j=0}^{\nu+1}$, by solving the following linear system:

$$\sum_{j=0}^{\nu+1} \alpha_j (\psi_j'(c_m) + i\omega g'(c_m) \psi_j(c_m)) = f(c_m), \quad m = 0, \dots, \nu + 1, \quad (6.3)$$

where $-1 = c_0 < c_1 < \dots < c_{\nu+1} = 1$ are the collocation points. A useful class of basis functions are given by so-called Chebyshev sets. They can be defined through the following equivalent properties (cf. Powell, 1981, §7.3):

Definition 6.2.2. *Let $N \in \mathbb{N}$. A set of N linearly independent functions $\mathcal{A} = \{\phi_n\}_{n=1}^N \subset C([-1, 1])$, defined on the interval $[-1, 1]$, is called a Chebyshev set if it satisfies any of the following two equivalent properties:*

1. *If a linear combination of functions in \mathcal{A} has more than $N - 1$ zeros then it must be identically zero.*

2. Let $\{\xi_j\}_{j=1}^N$ be any set of distinct points in $[-1, 1]$ then the $N \times N$ matrix with entries $\{\phi_n(\xi_j)\}_{n,j=1}^N$ is non-singular.

It can be shown that if $\{\psi_j\}_{j=0}^{\nu+2}$ is a Chebyshev set (see Def. 6.2.2) then the solution of Eq. (6.3) approximates a slowly varying solution of Eq. (6.2):

Proposition 6.2.3 (Prop. 3.4 in Deaño et al. (2017)). *Let $\{\psi_j\}_{j=0}^{\nu+2}$ be a Chebyshev set and let α be the solution of Eq. (6.3). Then, to leading order in $\omega \gg 1$, each α_n is a rational function of ω (more specifically a ratio of a $(\nu + 1)$ -degree polynomial over a $(\nu + 2)$ -degree polynomial). Thus $u_1^{[\nu]}$ varies slowly in ω .*

The result follows essentially from observing that for $\omega \gg 1$ the dominant part of the collocation equations Eq. (6.3) is

$$\sum_{n=0}^{\nu+1} \alpha_n \psi_n(c_m) = \frac{f(c_m)}{i\omega g'(c_m)} + \mathcal{O}(\omega^{-2}), \quad m = 0, \dots, \nu + 1. \quad (6.4)$$

Therefore, if the interpolation basis forms a Chebyshev set the system Eq. (6.4) must have a solution for any distribution of collocation points $-1 = c_0 < c_1 < \dots < c_{\nu+1} = 1$ provided ω is sufficiently large. Following the solution of the linear system Eq. (6.3) the Levin quadrature rule is then given by

$$\mathcal{Q}_\omega^{L, [\nu]}[f] = u_1^{[\nu]}(1)e^{i\omega g(1)} - u_1^{[\nu]}(-1)e^{i\omega g(-1)}.$$

So long as we ensure $c_0 = -1, c_{\nu+1} = 1$, the Levin method is well-suited for approximating $I_\omega[f]$ at frequency independent cost, because the quadrature method $\mathcal{Q}_\omega^{L, [\nu]}[f]$ is able to match the first term in the Poincaré series of $I_\omega[f]$ as the following result demonstrates.

Theorem 6.2.4 (Thm. 3.5 in Deaño et al. (2017)). *If $f \in C^2([-1, 1])$ and $g'(c_n) \neq 0$ for all $n = 0, \dots, \nu + 1$, the Levin method has an asymptotic error of the form*

$$|\mathcal{Q}_\omega^{L, [\nu]}[f] - I_\omega[f]| = \mathcal{O}(\omega^{-2}), \quad \text{as } \omega \rightarrow \infty.$$

However, the linear collocation system Eq. (6.3) is, in general, a dense $(\nu + 2) \times (\nu + 2)$ -linear system and hence requires $\mathcal{O}(\nu^3)$ operations to solve. This is the point where banded spectral computations become useful as we shall see in the following section. As a final remark concerning the properties of Levin methods in general, we highlight that the conditions of Thm. 6.2.4 exhibit a central limitation of the Levin method, which is that it cannot be constructed in the presence of stationary points: Similar to the Filon method (cf. §1.2.3) we would require the stationary point to be amongst the collocation points in order to match the asymptotic behaviour of the approximated integral, but this is prevented by the requirement $g'(c_m) \neq 0, m = 0, \dots, \nu + 1$ (Deaño et al., 2017, p. 46).

6.3 Accelerating the Levin method using banded matrix computations of Chebyshev coefficients

We shall now discuss how one may use a Chebyshev polynomial basis for the approximate solution of Eq. (6.2) in order to achieve a significant acceleration in finding the coefficients α as defined in Eq. (6.3). Ultimately, we will see that our method is able to compute the full approximation to the integral $I_\omega[f]$ in just $\mathcal{O}(\nu \log \nu + d^2 \nu)$ operations, where we recall that d is the polynomial degree of the phase function, $d = \deg g$. As is explained in sequence, we require ν to be even and we assume $\nu > d$. The latter assumption is made since we shall rely on the solution of a banded matrix system, with bandwidth $2d + 3$, so in the case $d \geq \nu$ the matrix would be dense and yield no significant speed up over the direct solution of Eq. (6.5). Our Ansatz is, therefore,

$$u_1(x) = \sum_{l=0}^{\nu+1} \alpha_l T_l(x),$$

and we use Clenshaw–Curtis collocation points $c_m = \cos(m\pi/(\nu + 1))$, $0 \leq m \leq \nu + 1$. Thus, we wish to find the coefficients α from the following set of interpolation conditions

$$\mathbf{A}\alpha = \mathbf{f}, \tag{6.5}$$

where $A_{mn} = \mathcal{L}_\omega T_n(c_m)$, $f_m = f(c_m)$, $m, n = 0, \dots, \nu + 1$, and here we denoted

$$\mathcal{L}_\omega = \frac{d}{dx} + i\omega g'(x).$$

As a first observation, we note that once α is found, we can evaluate the Levin approximation to $I_\omega[f]$ using just $\mathcal{O}(\nu)$ additions. Indeed, by the expressions Eq. (5.34) given in Appendix 5.A of the previous chapter we have $T_n(\pm 1) = (\pm 1)^n$, and hence

$$u_1^{[\nu]}(\pm 1) = \sum_{j=0}^{\nu+1} (\pm 1)^j \alpha_j.$$

To speed up the solution of Eq. (6.2) we will show that ν of the $\nu + 2$ degrees of freedom in the equation can be found through the solution of a banded matrix system of bandwidth $2d + 3$.

In order to do so we begin by transforming \mathcal{L}_ω to a form that has banded action on the Chebyshev polynomial basis. We recall Lemma 5.3.5:

Lemma 6.3.1 (Eqs. 22.7.4 & 22.8.3 in Abramowitz and Stegun (1965)). *For all $n \geq \mathbb{Z}$:*

$$xT_n(x) = \frac{1}{2}T_{n-1}(x) + \frac{1}{2}T_{n+1}(x), \text{ and } (1-x^2)T'_n(x) = \frac{n}{2}T_{n-1}(x) - \frac{n}{2}T_{n+1}(x).$$

In particular, the actions of $x, (1-x^2)d/dx$ on $\{T_n\}_{n=0}^\infty$ are both of bandwidth 3.

Thus let us consider the differential operator

$$\tilde{\mathcal{L}}_\omega = (1-x^2)\mathcal{L}_\omega = (1-x^2)\frac{d}{dx} + i\omega(1-x^2)g'(x).$$

The effect of replacing \mathcal{L}_ω by $\tilde{\mathcal{L}}_\omega$ in the collocation equations Eq. (6.5) is simply to multiply every row of the linear system by a factor of $(1-c_m^2)$, i.e. $\tilde{\mathbf{A}}_{mn} = (1-c_m^2)\mathbf{A}_{mn}$, $m, n = 0, \dots, \nu+1$, where $\tilde{\mathbf{A}}_{mn} = \tilde{\mathcal{L}}_\omega T_n(c_m)$. Thus the solution $\boldsymbol{\alpha}$ of Eq. (6.5) must also satisfy the $(\nu+2) \times (\nu+2)$ matrix system

$$\tilde{\mathbf{A}}\boldsymbol{\alpha} = \tilde{\mathbf{f}}. \tag{6.6}$$

By construction (since $c_0 = 1, c_{\nu+1} = -1$) we have $f_0 = f_{\nu+1} = 0$ and $\tilde{\mathbf{A}}_{0n} = \tilde{\mathbf{A}}_{(\nu+1)n} = 0$ for all $0 \leq n \leq \nu+1$. Hence the linear system Eq. (6.6) is ill-posed. We can nevertheless work with the system in the following way: Let \mathbf{P}_ν be projection onto the middle ν coordinates, i.e. $\mathbf{P}_\nu : (x_0, \dots, x_{\nu+1}) \mapsto (0, x_1, \dots, x_\nu, 0)$. Consider the linear system

$$\mathbf{P}_\nu \tilde{\mathbf{A}} \mathbf{P}_\nu \boldsymbol{\alpha}_0 = \mathbf{P}_\nu \tilde{\mathbf{f}}. \tag{6.7}$$

This corresponds (up to multiplication of each row by a nonzero constant) to the interpolation conditions

$$\sum_{n=1}^{\nu} (\boldsymbol{\alpha}_0)_n \mathcal{L}_\omega T_n(c_m) = f(c_m), \quad m = 1, \dots, \nu.$$

Our central observation is that we can solve Eq. (6.7) very efficiently:

Lemma 6.3.2. *For ω sufficiently large, the system Eq. (6.7) has a unique solution $\boldsymbol{\alpha}_0 \in \{\mathbf{x} \in \mathbb{C}^{\nu+2} \mid x_0 = x_{\nu+1} = 0\}$. Moreover, whenever such a solution exists, we can solve Eq. (6.7) using $\mathcal{O}(\nu \log \nu + d^2 \nu)$ operations, by the application of one discrete cosine transform (DCT-I) and the solution of a banded matrix system of bandwidth $2d+3$.*

Proof. We recall $\tilde{\mathbf{A}}_{mn} = \tilde{\mathcal{L}}_\omega T_n(c_m)$. By Lemma 6.3.1 the action of $\tilde{\mathcal{L}}_\omega$ on $\{T_n\}_{n=0}^\infty$ is represented by a banded (infinite) matrix \mathbf{B} with bandwidth $2d+3$ such that

$$\tilde{\mathcal{L}}_\omega T_n(x) = \sum_{k=\max\{0, n-(d+1)\}}^{n+d+1} \mathbf{B}_{kn} T_k(x). \tag{6.8}$$

Therefore we can express the matrix $\tilde{\mathbf{A}}$ in the form

$$\tilde{\mathbf{A}}_{mn} = \sum_{k=\max\{0, n-(d+1)\}}^{n+d+1} \mathbf{B}_{kn} \mathbf{T}_k(c_m), \quad m, n = 0, \dots, \nu + 1.$$

To simplify this further note that for $0 \leq l, m \leq \nu + 1$ we have

$$\begin{aligned} \mathbf{T}_{\nu+1+l}(c_m) &= \cos\left(\frac{(\nu+1+l)m}{\nu+1}\pi\right) = (-1)^m \cos\left(\frac{lm}{\nu+1}\pi\right) \\ &= (-1)^m \cos\left(\frac{(-l)m}{\nu+1}\pi\right) = \mathbf{T}_{\nu+1-l}(c_m). \end{aligned}$$

Thus let us define $\tilde{\mathbf{B}}$ as a $(\nu+2) \times (\nu+2)$ matrix by

$$\tilde{\mathbf{B}}_{nm} = \begin{cases} \mathbf{B}_{nm}, & 0 \leq n < \nu - d, n = \nu + 1, \\ \mathbf{B}_{nm} + \mathbf{B}_{(2\nu+2-n)m}, & \nu - d \leq n \leq \nu. \end{cases}$$

Then we have, since $\nu > d$,

$$\tilde{\mathbf{A}}_{mn} = \sum_{k=\max\{0, n-(d+1)\}}^n \mathbf{C}_{mk} \tilde{\mathbf{B}}_{kn},$$

where $\mathbf{C}_{mk} = \mathbf{T}_k(c_m)$. We now observe that the action of \mathbf{C} essentially represents a discrete cosine transform (DCT-I): Let $\mathbf{x} \in \mathbb{C}^{\nu+2}$ then

$$(\mathbf{C}\mathbf{x})_m = \sum_{n=0}^{\nu+1} \mathbf{C}_{mn} x_n = \sum_{n=0}^{\nu+1} \cos\left(\frac{mn\pi}{\nu+1}\right) x_n = \sum_{n=0}^{\nu+1}{}'' \cos\left(\frac{mn\pi}{\nu+1}\right) \tilde{x}_n =: (\mathcal{C}_{\nu+1}\tilde{\mathbf{x}}),$$

where we denoted by $\mathcal{C}_{\nu+1}$ the DCT-I on the space $\mathbb{C}^{\nu+2}$. Here $\sum_{n=0}{}''^{\nu+1}$ means that for $j = 0$ and $j = \nu + 1$ the terms in the sum are halved, and we defined

$$\tilde{x}_n = \begin{cases} x_n, & 1 \leq n \leq \nu, \\ 2x_n, & n = 0, \nu + 1. \end{cases}$$

The inverse of $\mathcal{C}_{\nu+1}$ is again a DCT-I, $\mathcal{C}_{\nu+1}^{-1} = 2/(\nu+1)\mathcal{C}_{\nu+1}$. It is well-known that the application of $\mathcal{C}_{\nu+1}$ can be computed efficiently in $\mathcal{O}(\nu \log \nu)$ operations (Trefethen, 2008), so we see that the action of \mathbf{C} and \mathbf{C}^{-1} can both be computed in $\mathcal{O}(\nu \log \nu)$ operations. Moreover, since we assumed ν to be even, we have the following convenient fact.

Claim 6.3.3. *If $\mathbf{x}, \mathbf{y} \in \mathbb{C}^{\nu+2}$ are such that $(\mathbf{I} - \mathbf{P}_\nu)\mathbf{C}\mathbf{x} = (\mathbf{I} - \mathbf{P}_\nu)\mathbf{C}\mathbf{y} = \mathbf{0}$, where \mathbf{I} is the*

$(\nu + 2) \times (\nu + 2)$ identity matrix, then

$$\mathbf{P}_\nu \mathbf{x} = \mathbf{P}_\nu \mathbf{y} \iff \mathbf{P}_\nu \mathbf{C} \mathbf{x} = \mathbf{P}_\nu \mathbf{C} \mathbf{y}.$$

Proof of Claim 6.3.3. We know that the first and final rows of \mathbf{C} take the forms $C_{0,n} = 1, C_{\nu+1,n} = (-1)^n, n = 0, \dots, \nu + 1$. Thus if $(\mathbf{I} - \mathbf{P}_\nu) \mathbf{C} \mathbf{x} = \mathbf{0}$, then

$$\begin{aligned} x_0 + x_{\nu+1} &= - \sum_{n=1}^{\nu} x_n, \\ x_0 - x_{\nu+1} &= - \sum_{n=1}^{\nu} (-1)^n x_n, \end{aligned}$$

and so $x_0, x_{\nu+1}$ are uniquely determined by the remaining entries of \mathbf{x} . Therefore we have

$$\mathbf{P}_\nu \mathbf{x} = \mathbf{P}_\nu \mathbf{y} \implies \mathbf{x} = \mathbf{y} \implies \mathbf{P}_\nu \mathbf{C} \mathbf{x} = \mathbf{P}_\nu \mathbf{C} \mathbf{y}.$$

The implication in the other direction follows, by recalling that $(\mathbf{I} - \mathbf{P}_\nu) \mathbf{C} \mathbf{x} = \mathbf{0}$, thus:

$$\mathbf{P}_\nu \mathbf{C} \mathbf{x} = \mathbf{P}_\nu \mathbf{C} \mathbf{y} \implies \mathbf{C} \mathbf{x} = \mathbf{P}_\nu \mathbf{C} \mathbf{x} = \mathbf{P}_\nu \mathbf{C} \mathbf{y} = \mathbf{C} \mathbf{y} \implies \mathbf{x} = \mathbf{y} \implies \mathbf{P}_\nu \mathbf{x} = \mathbf{P}_\nu \mathbf{y},$$

where the penultimate implication follows since \mathbf{C} is invertible. ■

First of all, by Claim 6.3.3, Eq. (6.7) has a solution for ω sufficiently large, because the dominant part of the linear system (similar to Eq. (6.4)) can be written as

$$\sum_{n=1}^{\nu} \alpha_n \mathbf{T}_n(c_m) = \frac{f(c_m)}{i\omega g'(c_m)} + \mathcal{O}(\omega^{-2}), \quad m = 0, \dots, \nu + 1,$$

i.e. $\mathbf{P}_\nu \mathbf{C} \mathbf{P}_\nu \boldsymbol{\alpha}_0 = \mathbf{P}_\nu \mathbf{h} + \mathcal{O}(\omega^{-2})$. Thus it is sufficient to prove that the equation

$$\mathbf{P}_\nu \mathbf{C} \mathbf{y} = \mathbf{P}_\nu \mathbf{z}, \tag{6.9}$$

has a solution $\mathbf{y} \in \{\mathbf{x} \in \mathbb{C}^{\nu+2} \mid x_0 = x_{\nu+1} = 0\}$ for all $\mathbf{z} \in \mathbb{C}^{\nu+2}$. We recall that \mathbf{C} is (up to rescaling of the first and final coordinates) a DCT-I, thus one can easily check that

$$\mathbf{y}_1 = (1/2, 1, 1, 1, \dots, 1, 1, 1/2), \tag{6.10}$$

$$\mathbf{y}_2 = (1/2, -1, 1, -1, \dots, -1, 1, -1/2), \tag{6.11}$$

are such that $\mathbf{P}_\nu \mathbf{C} \mathbf{y}_j = \mathbf{0}, j = 1, 2$. The final coordinate of \mathbf{y}_2 has negative sign because ν is even. Therefore, whenever a solution $\mathbf{y}_0 \in \mathbb{C}^{\nu+2}$ of Eq. (6.9) exists, we can find a linear combination of $\mathbf{y}_0, \mathbf{y}_1, \mathbf{y}_2$ which is in $\mathbf{z} \in \{\mathbf{x} \in \mathbb{C}^{\nu+2} \mid x_0 = x_{\nu+1} = 0\}$ and also solves Eq. (6.9). By Claim 6.3.3 one such solution is $\mathbf{y}_0 = \mathbf{C}^{-1} \mathbf{P}_\nu \mathbf{z}$, thus completing the proof of

the first part of the lemma.

Now let us show how we can solve Eq. (6.7) efficiently whenever a solution exists. By Claim 6.3.3 the system Eq. (6.7) is equivalent to

$$\mathbf{P}_\nu \tilde{\mathbf{B}} \mathbf{P}_\nu \boldsymbol{\alpha}_0 = \mathbf{P}_\nu \mathbf{C}^{-1} \tilde{\mathbf{f}}. \quad (6.12)$$

Now we note $\mathbf{P}_\nu \tilde{\mathbf{B}} \mathbf{P}_\nu \boldsymbol{\alpha}$ is a banded matrix of bandwidth $2d + 3$, so a solution of Eq. (6.12) can be found, using Gaussian elimination, in $\mathcal{O}(d^2\nu)$ operations. Thus to find $\boldsymbol{\alpha}_0$, we need to apply a single DCT-I and then solve Eq. (6.12), i.e. we incur an overall cost of $\mathcal{O}(\nu \log \nu + d^2\nu)$ operations. \square

Since $\tilde{A}_{mn} = (1 - c_m^2)A_{mn}$, $m, n = 0, \dots, \nu + 1$, we therefore found $\boldsymbol{\alpha}_0$ such that

$$\mathbf{P}_\nu \mathbf{A} \boldsymbol{\alpha}_0 = \mathbf{P}_\nu \mathbf{f}.$$

By construction $\text{Nullity}(\tilde{\mathbf{A}}) = \text{Nullity}(\mathbf{A}) + 2$, therefore, provided the collocation problem Eq. (6.5) is soluble for all f , we know that there are two linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{C}^{\nu+2}$ such that

$$\mathbf{P}_\nu \mathbf{A} \mathbf{v}_j = \mathbf{0}, \quad j = 1, 2. \quad (6.13)$$

We can construct those as follows: Let $\mathbf{e}_0 = (1, 0, \dots, 0)^T$ and $\mathbf{e}_{\nu+1} = (0, \dots, 0, 1)^T$ then by the same process as in Lemma 6.3.2 we can find $\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2 \in \{\mathbf{x} \in \mathbb{C}^{\nu+2} \mid x_0 = x_{\nu+1} = 0\}$ with

$$\mathbf{P}_\nu \mathbf{A} \mathbf{P}_\nu \tilde{\mathbf{v}}_1 = -\mathbf{P}_\nu \mathbf{A} \mathbf{e}_0, \quad \mathbf{P}_\nu \mathbf{A} \mathbf{P}_\nu \tilde{\mathbf{v}}_2 = -\mathbf{P}_\nu \mathbf{A} \mathbf{e}_{\nu+1}. \quad (6.14)$$

We note that the first and last columns of \mathbf{A} have at least one nonzero entry in a row with index $1 \leq m \leq \nu$, so $\mathbf{P}_\nu \mathbf{A} \mathbf{e}_0, \mathbf{P}_\nu \mathbf{A} \mathbf{e}_{\nu+1} \neq \mathbf{0}$. The vectors $\mathbf{v}_1 = \mathbf{e}_0 + \tilde{\mathbf{v}}_1, \mathbf{v}_2 = \mathbf{e}_{\nu+1} + \tilde{\mathbf{v}}_2$ are then clearly linearly independent (since $\mathbf{e}_0^T \mathbf{v}_2 = 0 \neq \mathbf{e}_0^T \mathbf{v}_1$), and they satisfy Eq. (6.13). The computation of the vectors takes again $\mathcal{O}(\nu \log \nu + d^2\nu)$ operations, as per Lemma 6.3.2. We can finally find $\boldsymbol{\alpha}$, the solution to Eq. (6.5), by solving the remaining linear system:

$$\delta_1 \mathbf{A} \mathbf{v}_1 + \delta_2 \mathbf{A} \mathbf{v}_2 = \mathbf{f} - \mathbf{A} \boldsymbol{\alpha}_0. \quad (6.15)$$

Recall that by construction of $\mathbf{v}_1, \mathbf{v}_2, \boldsymbol{\alpha}_0$, the rows $m = 1, \dots, \nu$ in Eq. (6.15) are trivially satisfied ($0 = 0$), and so this reduces to a 2×2 linear system for the coefficients δ_1, δ_2 :

$$\begin{pmatrix} (\mathbf{A} \mathbf{v}_1)_0 & (\mathbf{A} \mathbf{v}_2)_0 \\ (\mathbf{A} \mathbf{v}_1)_{\nu+1} & (\mathbf{A} \mathbf{v}_2)_{\nu+1} \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = \begin{pmatrix} f_0 - (\mathbf{A} \boldsymbol{\alpha}_0)_0 \\ f_{\nu+1} - (\mathbf{A} \boldsymbol{\alpha}_0)_{\nu+1} \end{pmatrix}. \quad (6.16)$$

The coefficients in this system can be found in $\mathcal{O}(\nu)$ operations (since for instance $(\mathbf{A}\mathbf{v}_1)_1 = \sum_{n=0}^{\nu+1} \mathbf{A}_{0n}(\mathbf{v}_1)_n$). Furthermore, because \mathbf{A} is invertible, Eq. (6.15) must have a unique solution for δ_1, δ_2 . This completes our construction of $\boldsymbol{\alpha}$, and hence the computation of $\mathcal{Q}_\omega^{\text{L},[\nu]}[f]$.

6.3.1 Algorithm for the efficient construction of the Levin method

Before moving on to provide some numerical examples let us briefly summarise the above algorithm for constructing the Levin method in $\mathcal{O}(\nu \log \nu + d^2\nu)$ operations:

1. Solve Eq. (6.7) using the method described in Lemma 6.3.2, to find $\boldsymbol{\alpha}_0 \in \{\mathbf{x} \in \mathbb{C}^{\nu+2} \mid x_0 = x_{\nu+1} = 0\}$ with

$$\mathbf{P}_\nu \mathbf{A} \boldsymbol{\alpha}_0 = \mathbf{P}_\nu \mathbf{f}.$$

2. Solve Eq. (6.14) using the method described in Lemma 6.3.2, to find two linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2$ with

$$\mathbf{P}_\nu \mathbf{A} \mathbf{v}_j = \mathbf{0}, \quad j = 1, 2.$$

3. Compute the coefficients in Eq. (6.16) and solve the resulting 2×2 linear system for δ_1, δ_2 .
4. Finally, let $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \delta_1 \mathbf{v}_1 + \delta_2 \mathbf{v}_2$, and compute

$$\mathcal{Q}_\omega^{\text{L},[\nu]}[f] = \sum_{n=0}^{\nu+1} \alpha_n e^{i\omega g(1)} - \sum_{n=0}^{\nu+1} (-1)^n \alpha_n e^{i\omega g(-1)}. \quad (6.17)$$

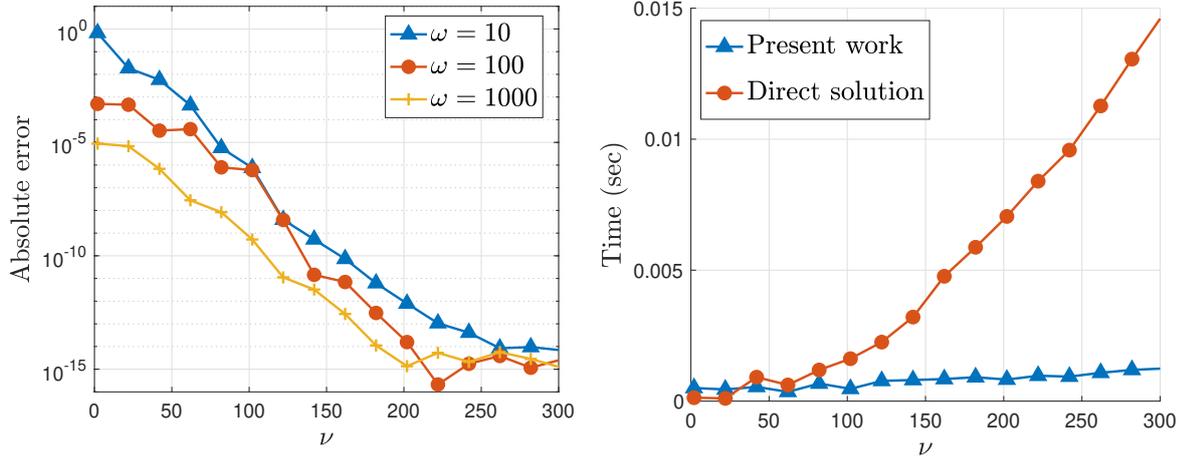
6.4 Numerical examples

To illustrate the performance of this method in practice let us consider two numerical examples. In both cases the reference solution for the exact integral was computed using Clenshaw–Curtis quadrature on the full oscillatory integrand using 10^7 points, which is a very time-consuming but accurate way for computing the integrals numerically.

6.4.1 Linear phase function

In the first example, we consider the case $g(x) = x$, and $f(x) = x/(x^2 + 0.02)$, i.e. we wish to approximate the integral

$$I_\omega^{(1)}[f] = \int_{-1}^1 \frac{x}{x^2 + 0.02} e^{i\omega x} dx.$$



(a) Absolute error, $|\mathcal{Q}_\omega^{L, [\nu]}[f] - I_\omega^{(1)}[f]|$. (b) Average time to solve Eq. (6.5), $\omega = 100$.

Figure 6.2: The error and timing of the Levin–Clenshaw–Curtis method as a function of the number of collocation points ν for fixed frequency ω .

6.4.2 Quadratic phase function

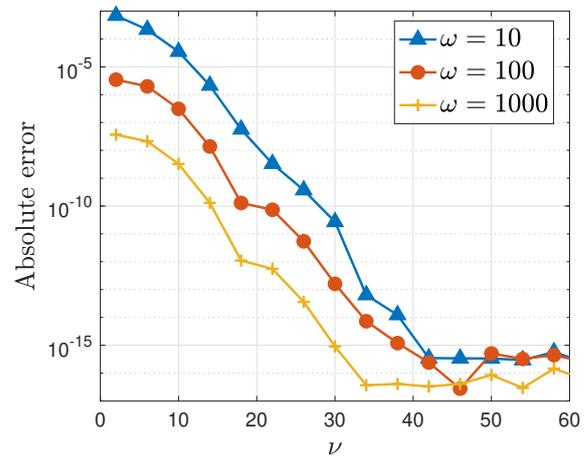
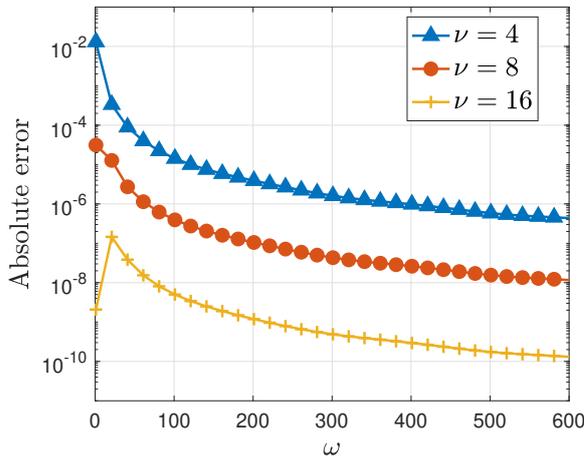
In our second example we consider a phase function of the form $g(x) = 3x - x^2$ and an amplitude of the form $f(x) = 1/(1 + x^4)$, i.e we aim to approximate the following integral

$$I_\omega^{(2)}[f] = \int_{-1}^1 \frac{1}{1 + x^4} e^{i\omega(3x - x^2)} dx.$$

In this case \mathcal{L}_ω takes the form $\mathcal{L}_\omega = \frac{d}{dx} + i\omega(3 - 2x)$ and one can check that $\tilde{\mathcal{L}}_\omega$ has the following banded matrix representation when acting on $\{\mathbb{T}_n\}_{n=0}^\infty$:

$$\mathbf{B} = \begin{pmatrix} 3i\omega/2 & 1/2 - i\omega/4 & -3i\omega/4 & i\omega/4 & 0 & & & & & \\ -i\omega/2 & 3i\omega/4 & 1 & -3i\omega/4 & i\omega/4 & 0 & & & & \\ -3i\omega/2 & -1/2 & 3i\omega/2 & 3/2 - i\omega/4 & -3i\omega/4 & i\omega/4 & 0 & & & \\ i\omega/2 & -3i\omega/4 & -1 - i\omega/4 & 3i\omega/2 & 2 - i\omega/4 & -3i\omega/4 & i\omega/4 & 0 & & \\ 0 & i\omega/4 & -3i\omega/4 & -3/2 - i\omega/4 & 3i\omega/2 & 5/2 - i\omega/4 & -3i\omega/4 & i\omega/4 & 0 & \\ & & \ddots & \\ & & & & & & & & & \ddots \end{pmatrix}.$$

In Fig. 6.3 we show the error of the Levin–Clenshaw–Curtis method applied to $I_\omega^{(2)}[f]$. As in the previous example, we observe uniform convergence of the method across all values $\omega \geq 0$ and spectral convergence with respect to the number of collocation points ν . The timing of the method is comparable to the previous case and so we shall not repeat the experiment in the interest of brevity.



(a) Absolute error as a function of ω for fixed ν . (b) Absolute error as a function of ν for fixed ω .

Figure 6.3: The error in the Levin–Clenshaw–Curtis method for $I_\omega^{(2)}[f]$.

6.5 Conclusions

In this chapter, we have seen how recursions for Chebyshev moments can be exploited to construct a fast Levin–Clenshaw–Curtis method which computes a Levin approximation to certain highly oscillatory integrals in $\mathcal{O}(\nu \log \nu)$ operations. This presents a significant speed-up compared to the $\mathcal{O}(\nu^3)$ operations required to solve the Levin collocation system by direct Gaussian elimination. We provided a detailed description of this algorithm as well as two numerical examples that demonstrate the favourable properties of the method in practice. Indeed, the experiments confirmed that the method delivers the promised speed, and, like Levin methods generally, converges uniformly for all values of ω as we increase the number of collocation points ν . In fact, in our examples with locally analytic amplitude functions, we observed spectral convergence with respect to ν .

List of symbols

Symbol	Description
$I_\omega[f]$	Highly oscillatory integral with amplitude function f .
T_n	Chebyshev polynomial of the first kind of degree n .
α	Unknown coefficients in the Levin collocation equations.
$\mathcal{C}_{\nu+1}$	Discrete cosine transform DCT-I on the space $\mathcal{C}^{\nu+2}$.
\mathcal{L}_ω	Levin differential operator.
$Q_\omega^{L, [\nu]}$	Levin quadrature rule using $\nu + 2$ collocation points.
A	Discretization matrix in the Levin collocation equations.
B	Banded matrix representation of the action of $\tilde{\mathcal{L}}_\omega$ on Chebyshev polynomials.
P_ν	Projection onto the middle ν coordinates, $\mathbf{x} \mapsto (0, x_1, \dots, x_\nu, 0)^T$.
$\nu + 2$	Number of collocation points in the Levin method.
ω	Frequency in the highly oscillatory integral.
$\tilde{\mathcal{L}}_\omega$	Modified Levin differential operator, $\tilde{\mathcal{L}}_\omega = (1 - x^2)\mathcal{L}_\omega$.
\tilde{A}	Discretization matrix in the modified Levin collocation equations.
$\{c_m\}_{m=0}^{\nu+1}$	Levin collocation points.
d	Polynomial degree of the phase function $g(x)$.
$f(x)$	Amplitude function in the highly oscillatory integral.
$g(x)$	Phase function in the highly oscillatory integral.
u_1	Non-oscillatory solution to the adjoint equation.

Chapter 7

Concluding remarks

In this thesis, we described a number of new results concerning a range of aspects in wave scattering problems: Starting from analytical treatment using the Wiener–Hopf technique, we covered some of the physics of acoustics in the presence of mean flow and vorticity, looked more closely at numerical techniques for boundary element methods, and, finally, examined highly oscillatory quadrature.

In a few summarising words our main contributions are the following: in Chapter 2 we provided a new Wiener–Hopf solution to the scattering problem of vortical and acoustic waves by an infinite arrangement of finite-length flat blades in uniform mean flow. Our new approach removes the requirement of overlap in consecutive blades which was a crucial limitation in previous solutions based on the Wiener–Hopf method. We used this solution to study the balance of incoming and outgoing acoustic energy flux for acoustic waves scattering by the cascade in Chapter 3. Based on this study we provided new analytical insights, proving the existence of symmetries and points of zero acoustic reflections for certain angles of propagation of the incident acoustic waves, as well as comprehensive numerical results outlining the effects of mean flow on the scattered field. This was followed by the study of numerical techniques relevant to computational wave scattering: firstly, in Chapter 4 we provided a rigorous analysis of an oversampled collocation method for Fredholm integral equations and showed that oversampling can be used to improve robustness and convergence properties of the method. Motivated by the collocation method and hybrid numerical-asymptotic bases for high-frequency wave scattering, we then considered Filon methods for the efficient approximation of highly oscillatory integrals. In Chapter 5 we addressed the moment-problem of Filon methods by providing a general framework for constructing moment-recurrences that can be used to efficiently compute the Filon moments in a number of relevant cases. We studied this methodology more closely in several cases of practical interest, and specifically applied the resulting direct Filon method in the context of high-frequency wave scattering on a finite screen in two dimensions. Finally, in Chapter 6 we considered an alternative methodology for the approximation

of highly oscillatory integrals, so-called Levin methods, and we provided a structured design of such a method that can lead to significant speed-up in the computation of the approximation. In fact, we showed that our Levin–Clenshaw–Curtis method is able to compute the approximation to the integral in just $\mathcal{O}(\nu \log \nu)$ rather than the usual $\mathcal{O}(\nu^3)$ operations, where ν is the Levin analogue of quadrature points.

In conclusion, although wave scattering problems have been tackled by mathematicians for centuries, we found that the field still possesses many challenging and open questions. Perhaps one of the main messages we take away from this thesis is that, in the modern treatment of wave scattering, numerical and analytical techniques often go hand-in-hand to achieve optimal results. For instance, we saw how the numerical finite section method provided the successful final step in the Wiener–Hopf treatment of waves scattering by a cascade of blades (cf. Chapter 2), and how the detailed understanding of the asymptotic behaviour, specifically the Poincaré series, of oscillatory integrals could be exploited to construct very efficient quadrature techniques (cf. Chapters 5 & 6). We, therefore, expect that the combination of computational and analytical research will continue to deliver fruitful insights in this field in the near and distant future.

7.1 Future work by chapter

“Die Wissenschaft fängt eigentlich erst da an, interessant zu werden, wo sie aufhört.”¹
 –Justus von Liebig

One beautiful aspect of scientific research is that every solution to a problem raises a whole new set of interesting questions. With definite excitement and the sincere hope to tackle some of the following in times to come, we thus wish to conclude this thesis by offering a few indications of how the ideas presented may be taken forward in future research.

Chapter 2

As we described in §2.6, the Wiener–Hopf solution presented in Chapter 2 is equivalent to previous work by Peake (1993) and Glegg (1999) in the case of overlapping blades. These previous analytical solutions were used as a basis to include more realistic effects to model the features of fan components in turbomachinery in further detail, such as the inclusion of blade camber and thickness (Baddoo and Ayton, 2020*b*) and the consideration of full three-dimensional annular geometries (Posson et al., 2010). We expect that our solution will be similarly suited for such extensions. In particular, the work by (Baddoo and Ayton, 2020*b*) is built on the Wiener–Hopf solution for overlapping cascades given by Glegg (1999) and which is very similar in nature to our approach, hence suggesting that the velocity

¹“Science really only begins to become interesting where it ends.”

potential and streamfunction coordinate system exploited by Baddoo and Ayton (2020*b*) could be used together with the Wiener–Hopf solution in Chapter 2 to account for realistic airfoil shapes in the non-overlapping case.

A further assumption of our model that could potentially be weakened is the zero angle of attack of the incident mean flow. Indeed Myers and Kerschen (1995) provided an asymptotic analysis (for high frequencies and small incidence angles) which allows the study of such effects for single blades. Peake and Kerschen (1997, 2004) pursued a successful program to apply this type of asymptotic approach to cascades with mean loading, albeit overlapping cascades are implicitly assumed in their analysis (by studying duct modes present in the regions between consecutive cascade blades). Since the leading order term of their asymptotic expansion is simply the case of uniform mean-flow, we expect that it might be possible to study these perturbations in a similar way for a cascade of non-overlapping blades based on our solution from Chapter 2 – the essential step in the approach being again a change of coordinates to a velocity potential and streamfunction coordinate system.

Finally, although our analytical solution has guaranteed convergence, as we proved in Appendix 2.D, the convergence occurs at an algebraic rate in the number of modes retained in the linear system, as opposed to the spectral convergence observed in previous formulations for overlapping cascades. Therefore, it would be interesting to compare our approach to alternative numerical solutions of the scattering problem and to understand whether a similar slow-down in convergence rates occurs as the blade spacing is increased. One possibility to do so is to construct a direct collocation method based on the integral equation Eq. (3.4) similar to the work by Porter (2021).

Chapter 3

As discussed for Chapter 2 above, the effect of nonzero angles of attack of the mean flow might also be studied in the context of the acoustic energy balance for waves scattering by the cascade of blades. Indeed, we would expect to find a similar balance of incoming and outgoing energy fluxes as described in Chapter 3, noting in particular that the mean flow will become uniform far upstream and downstream of the cascade blades. Thus, in those regions, the expressions for the acoustic energy flux will remain similar to the case considered in Chapter 3. However, close to the trailing edges, the flow will take a much more complicated shape. One might therefore use the balance of incoming acoustic power against acoustic and hydrodynamic power radiated to the far-field, in order to determine the sound power absorbed by the wake without having to know the exact form of the field in the proximity of the cascade. This might provide useful insights into the process of vortex shedding and the role of the Kutta condition in this case.

A further research direction is the study of the effect of including a small, nonzero

amount of viscosity in the body of the fluid (together with appropriately accounting for viscous effects at the boundary of the blades), especially in relation to the stability of the outgoing waves and the Kutta condition. Indeed, due to the possibility of acoustic energy emission from the wake, which we observed in §3.4.3, the ‘activation energy’ (a term coined by Benjamin (1963); cf. also later work by Cairns (1979)) of an acoustic mode relative to the background flow and vortex sheets may be negative. Thus the stability of acoustic and hydrodynamic modes remains a challenging problem in the presence of mean flow and could provide the foundation for exciting future research.

Finally, we highlight that in our study we used the isentropic assumption for the fluid, meaning we ignored the effects of temperature. It is possible to find an energy corollary involving only quantities up to first order for perturbations to general steady flows as well (cf. Myers, 1991), and it would be interesting to study whether the effects of sound generation by cascades are cancelled if the effects of temperature are taken into account.

Chapter 4

In this chapter we saw that on a periodic domain with spline basis functions an optimal choice of collocation points is readily available and given as equispaced points that match the mesh of the basis functions. However, the relative simplicity of collocation methods is especially useful in more complicated settings such as three-dimensional scattering problems. In such settings, an optimal choice of collocation points may no longer be obvious and we can regard our results in Chapter 4 as an initial analysis of the underlying mechanism with the expectation that some properties will translate to more complicated settings.

Thus we expect that future research on this topic will focus on exploring the properties of oversampling for collocation methods in more general settings. This includes the extension of the present results to three-dimensional boundary integral equations and more complicated boundary conditions such as impedance or penetrable boundaries. Integral equations for these boundaries can be constructed as combinations of single layer, double layer, adjoint double layer and hypersingular operators (cf. Chandler-Wilde, Langdon and Mokgolele, 2012, Groth et al., 2015) and the resulting integral equations can be solved using oversampled collocation methods in precisely the same way as introduced in Chapter 4. Therefore it would be interesting to study the convergence properties of the oversampled collocation method applied in these settings. It would also be interesting to study the effects of oversampling when there is redundancy in the trial spaces. Based on recent results by Adcock and Huybrechs (2020) we expect that oversampling together with suitable regularisation can act as a stabiliser towards redundancies in the basis spaces, which might provide a framework for rigorous analysis of collocation methods involving more general approximation spaces such as hybrid numerical-asymptotic basis functions

as described by Chandler-Wilde, Graham, Langdon and Spence (2012) and Gibbs et al. (2020).

We highlight that the results concerning optimal convergence with quadratic rates of oversampling in Thms. 4.4.10, 4.F.1 & 4.G.1 were proved for integral operators arising from the single and double layer formulations of Laplace's equation on general smooth domains and the single layer formulation of the Helmholtz equation on such domains. We expect that a similar analysis can be performed for the double layer formulation of the Helmholtz equation on smooth domains, but we note that further work is necessary to accommodate this in the proof of Thm. 4.G.1. This work can be started by considering the explicit expressions and properties of the double layer integral operator formulated in terms of a parametrisation as can be found for instance in (Kress, 1991, Meier et al., 2000).

A further direction for future research is the investigation of the merits of relating a discrete sampling process to its continuous limit, and potentially the use of oversampling, to understand the convergence rates of fully discrete schemes such as Nystöm methods (Bremer and Gimbutas, 2012, Bruno et al., 2013, Hao et al., 2014).

Chapter 5

In the application to wave scattering problems (see §5.5) we found that, even when initial Filon moments cannot be expressed explicitly in terms of simple functions, one may still be able to provide an expression that is easy to evaluate numerically (for instance an exponentially decaying integral). One may then compute the remaining moments efficiently using the recurrences provided by Thm. 5.3.2. This idea of computing the initial conditions through exponentially decaying integrals is closely related to the paradigm of numerical steepest descent (NSD), where oscillatory integrals are written in terms of exponentially decaying integrals by moving to the complex plane. However, a direct application of NSD requires analyticity of the integrand at least in a neighbourhood of the domain of integration, which may not be guaranteed in general. This is in contrast to Filon methods which only require the weaker condition that the amplitude is well-approximable by polynomials. Therefore, we believe future research could focus on combining the two approaches, by providing a polynomial approximation to the amplitude function in an oscillatory integral followed by the use of NSD to help evaluate the oscillatory integral over the polynomial which is an entire function.

A further aspect that remains largely unstudied is a rigorous analysis of the stability of the moment recurrences in regimes beyond $N \propto \omega$, where N is the number of moments computed with the recurrence and ω is the frequency of the oscillatory integrand. As we mentioned in Chapter 5, we performed a range of numerical experiments which indicated that Oliver's algorithm (Oliver, 1968) provides a satisfactory way to compute the tail in

this recurrence, but rigorous analysis would have to focus on understanding the extremal singular values of variable-coefficient Toeplitz matrices. This is a fairly challenging topic: although the distribution of the spectrum of such methods is well-studied (Garoni and Serra-Capizzano, 2017), the extremal values for the variable Toeplitz case have so far escaped rigorous treatment. There are some results available in the literature for constant-coefficient Toeplitz matrices (Böttcher and Grudsky, 1998, Garoni, 2013), though these do not easily generalise to the cases of relevance for recursive moment computations.

Chapter 6

In a similar way to the matrices appearing in Oliver’s algorithm mentioned above, the matrices \mathbf{B} in Chapter 6 also take the form of variable-coefficient Toeplitz matrices (with low-rank degeneracies). Thus, to understand the stability of the Levin–Clenshaw–Curtis method, especially for a large number of Levin collocation points ν , a study of the singular values of variable coefficient Toeplitz matrices would be required.

Additionally, the ideas presented in Chapter 6 might be extended to more general settings, for instance allowing for more complicated oscillatory integral kernels, which may satisfy higher-order ordinary differential equations with polynomial coefficients, such as Bessel or Hankel functions. The original work by Levin (1996) is able to allow for those cases, essentially by presenting a vectorised form of the Levin method, and we believe that an acceleration using our structured Levin–Clenshaw–Curtis approach might be applied to this setting as well. A final possible extension involves the inclusion of derivative values in the collocation equations, which would allow the Levin quadrature rule to match more terms in the Poincaré series of the integral, thus leading to smaller asymptotic error. This was done for general Levin methods by Olver (2008) and might be incorporated in the fast Levin–Clenshaw–Curtis setting as well.

References

- Abrahams, I. D. (1997), ‘On the solution of Wiener–Hopf problems involving noncommutative matrix kernel decompositions’, *SIAM Journal on Applied Mathematics* **57**(2), 541–567. [Cited on page 16.]
- Abrahams, I. D. and Aitken, M. G. (2019), ‘A Brief and Gentle Introduction to the Wiener–Hopf Technique’. Lecture notes, Summer School “Bringing pure and applied analysis together via the Wiener–Hopf technique”. [Cited on page 12.]
- Abramowitz, M. and Stegun, I. (1965), *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, Dover Publications. [Cited on pages 22, 28, 196, 197, 198, 208, 210, 223, 224, 227, 230, 231, and 240.]
- Adcock, B. and Hansen, A. C. (2012), ‘Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon’, *Applied and Computational Harmonic Analysis* **32**(3), 357–388. [Cited on pages 113 and 125.]
- Adcock, B. and Huybrechs, D. (2019), ‘Frames and Numerical Approximation’, *SIAM Review* **61**(3), 443–473. [Cited on page 119.]
- Adcock, B. and Huybrechs, D. (2020), ‘Frames and Numerical Approximation II: Generalized Sampling’, *Journal of Fourier Analysis and Applications* **26**(6), 87. [Cited on pages 113, 119, and 254.]
- Adcock, B., Huybrechs, D. and Martín-Vaquero, J. (2014), ‘On the Numerical Stability of Fourier Extensions’, *Foundations of Computational Mathematics* **14**(4), 635–687. [Cited on page 113.]
- Al Azah, M. (2017), Fast evaluation of special functions by the modified trapezium rule, PhD thesis, University of Reading. [Cited on page 212.]
- Al Azah, M. and Chandler-Wilde, S. N. (2021), ‘Computation of the complex error function using modified trapezoidal rules’, *SIAM Journal on Numerical Analysis* **59**(5), 2346–2367. [Cited on page 212.]

- Andrews, D. R. (2003), Ultrasonics and Acoustics, *in* R. A. Meyers, ed., ‘Encyclopedia of Physical Science and Technology (Third Edition)’, third edn, Academic Press, New York, pp. 269–287. [Cited on page 7.]
- Anselone, P. M. (1971), *Collectively Compact Operator Approximation Theory and Applications to Integral Equations*, Prentice-Hall, Inc. [Cited on page 115.]
- Arens, T. and Rösch, T. (2016), ‘A collocation method for weakly singular integral equations with super-algebraic convergence rate’, *Numerische Mathematik* **134**(3), 441–472. [Cited on page 114.]
- Arnold, D. N. and Wendland, W. L. (1983), ‘On the asymptotic convergence of collocation methods’, *Mathematics of Computation* **41**(164), 349–381. [Cited on pages 9, 11, 114, 123, 146, and 152.]
- Arnold, D. N. and Wendland, W. L. (1985), ‘The convergence of spline collocation for strongly elliptic equations on curves’, *Numerische Mathematik* **47**(3), 317–341. [Cited on pages 114, 136, 152, 164, and 165.]
- Arzoumanian, S. (2011), *Stability of Fluid-Loaded Structures*, PhD thesis, University of Cambridge. [Cited on pages 84, 95, and 96.]
- Atkinson, K. E. (1997), *The Numerical Solution of Integral Equations of the Second Kind*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press. [Cited on page 115.]
- Ayton, L. J. and Peake, N. (2013), ‘On high-frequency noise scattering by aerofoils in flow’, *Journal of Fluid Mechanics* **734**, 144–182. [Cited on page 36.]
- Babuška, I. and Aziz, A. K. (1972), Part I: Survey Lectures on the Mathematical Foundations of the Finite Element Method, *in* A. K. Aziz, ed., ‘The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations’, Academic Press. [Cited on pages 122 and 152.]
- Baddoo, P. J. (2020), *Analytic Solutions for Flows Through Cascades*, PhD thesis, University of Cambridge. [Cited on pages 4 and 5.]
- Baddoo, P. J. and Ayton, L. J. (2020*a*), ‘Acoustic scattering by cascades with complex boundary conditions: compliance, porosity and impedance’, *Journal of Fluid Mechanics* **898**, A16. [Cited on page 36.]
- Baddoo, P. J. and Ayton, L. J. (2020*b*), ‘An analytic solution for gust–cascade interaction noise including effects of realistic aerofoil geometry’, *Journal of Fluid Mechanics* **886**, A1. [Cited on pages 5, 36, 252, and 253.]

- Banaugh, R. P. and Goldsmith, W. (1963), ‘Diffraction of steady acoustic waves by surfaces of arbitrary shape’, *The Journal of the Acoustical Society of America* **35**(10), 1590–1601. [Cited on page 7.]
- Barnett, A. H. and Betcke, T. (2018), ‘Stability and convergence of the Method of Fundamental Solutions for Helmholtz problems on analytic domains’, *J. Comput. Phys.* **227**, 7003–7026. [Cited on page 115.]
- Bechert, D., Michel, U. and Pfizenmaier, E. (1977), ‘Experiments on the transmission of sound through jets’, *4th AIAA Aeroacoustics Conference, Atlanta GA, October 3-5*. [Cited on pages 6, 84, and 106.]
- Ben-Artzi, J., Colbrook, M. J., Hansen, A. C., Nevanlinna, O. and Seidel, M. (2020), ‘Computing Spectra – On the Solvability Complexity Index Hierarchy and Towers of Algorithms’, *arXiv:1508.03280v5*. [Cited on page 76.]
- Bender, C. M. and Orszag, S. A. (2013), *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory*, Springer, New York. [Cited on page 201.]
- Benjamin, T. B. (1963), ‘The threefold classification of unstable disturbances in flexible surfaces bounding inviscid flows’, *Journal of Fluid Mechanics* **16**(3), 436–450. [Cited on page 254.]
- Berenger, J.-P. (1994), ‘A perfectly matched layer for the absorption of electromagnetic waves’, *Journal of Computational Physics* **114**(2), 185–200. [Cited on page 9.]
- Bonnet-Ben Dhia, A.-S., Carvalho, C., Chesnel, L. and Ciarlet, P. (2016), ‘On the use of Perfectly Matched Layers at corners for scattering problems with sign-changing coefficients’, *Journal of Computational Physics* **322**, 224–247. [Cited on page 10.]
- Böttcher, A. and Grudsky, S. M. (1998), ‘On the condition numbers of large semidefinite Toeplitz matrices’, *Linear algebra and its applications* **279**(1-3), 285–301. [Cited on page 256.]
- Bremer, J. and Gimbutas, Z. (2012), ‘A Nyström method for weakly singular integral operators on surfaces’, *Journal of Computational Physics* **231**(14), 4885–4903. [Cited on pages 115 and 255.]
- Bruno, O. P., Domínguez, V. and Sayas, F.-J. (2013), ‘Convergence analysis of a high-order Nyström integral-equation method for surface scattering problems’, *Numerische Mathematik* **124**(4), 603–645. [Cited on pages 115 and 255.]

- Burnside, W., Yu, C. and Marhefka, R. (1975), ‘A technique to combine the geometrical theory of diffraction and the moment method’, *IEEE Transactions on Antennas and Propagation* **23**(4), 551–558. [Cited on page 11.]
- Cairns, R. A. (1979), ‘The role of negative energy waves in some instabilities of parallel flows’, *Journal of Fluid Mechanics* **92**(1), 1–14. [Cited on page 254.]
- Cantrell, R. H. and Hart, R. W. (1964), ‘Interaction between Sound and Flow in Acoustic Cavities: Mass, Momentum, and Energy Considerations’, *The Journal of the Acoustical Society of America* **36**(4), 697–706. [Cited on page 84.]
- Carlson, J. and Heins, A. (1947), ‘The reflection of an electromagnetic plane wave by an infinite set of plates, I’, *Quarterly of Applied Mathematics* **4**(4), 313–329. [Cited on pages 35 and 47.]
- Chandler, G. A. and Sloan, I. H. (1990), ‘Spline quadrature methods for boundary integral equations’, *Numerische Mathematik* **58**(1), 537–567. [Cited on pages 114, 117, 134, 136, 138, and 158.]
- Chandler-Wilde, S. N., Graham, I. G., Langdon, S. and Spence, E. A. (2012), ‘Numerical-asymptotic boundary integral methods in high-frequency acoustic scattering’, *Acta numerica* **21**, 89–305. [Cited on pages 7, 10, 12, 18, 19, 21, 22, 23, 24, 119, 142, 187, 189, 208, and 255.]
- Chandler-Wilde, S. N. and Langdon, S. (2007), ‘A Galerkin boundary element method for high frequency scattering by convex polygons’, *SIAM Journal on Numerical Analysis* **45**(2), 610–640. [Cited on page 143.]
- Chandler-Wilde, S. N., Langdon, S. and Mokgolele, M. (2012), ‘A high frequency boundary element method for scattering by convex polygons with impedance boundary conditions’, *Communications in Computational Physics* **11**(2), 573–593. [Cited on page 254.]
- Cheng, A. H.-D. and Cheng, D. T. (2005), ‘Heritage and early history of the boundary element method’, *Engineering Analysis with Boundary Elements* **29**(3), 268–302. [Cited on pages 10 and 11.]
- Ciarlet, P. (2002), *The Finite Element Method for Elliptic Problems*, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics. [Cited on page 127.]
- Cody, W. J. (1976), An overview of software development for special functions, *in* G. A. Watson, ed., ‘Numerical Analysis’, Springer, pp. 38–48. [Cited on page 197.]

- Colton, D. and Kress, R. (1998), *Integral Equation Methods in Scattering Theory*, Classics in Applied Mathematics, second edn, Society for Industrial and Applied Mathematics. [Cited on page 115.]
- Condon, M., Deaño, A. and Iserles, A. (2009), ‘On highly oscillatory problems arising in electronic engineering’, *ESAIM: Mathematical Modelling and Numerical Analysis* **43**(4), 785–804. [Cited on page 187.]
- Conway, J. B. (1978), *Functions of One Complex Variable I*, second edn, Springer, New York. [Cited on pages 46 and 70.]
- Conway, J. B. (2010), *A Course in Functional Analysis*, Graduate Texts in Mathematics, second edn, Springer New York. [Cited on page 77.]
- Costabel, M. (1988), ‘Boundary Integral Operators on Lipschitz Domains: Elementary Results’, *SIAM journal on Mathematical Analysis* **19**(3), 613–626. [Cited on page 21.]
- Crighton, D. G. (1972), ‘Radiation from vortex filament motion near a half plane’, *Journal of Fluid Mechanics* **51**(2), 357–362. [Cited on page 84.]
- Crighton, D. G. (1985), ‘The Kutta Condition in Unsteady Flow’, *Annual Review of Fluid Mechanics* **17**(1), 411–445. [Cited on pages 5, 39, and 40.]
- Crighton, D. G. and Oswell, J. (1991), ‘Fluid loading with mean flow. I. Response of an elastic plate localized excitation’, *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences* **335**(1639), 557–592. [Cited on pages 95 and 96.]
- Daniele, V. (1978), ‘On the factorization of Wiener–Hopf matrices in problems solvable with Hurd’s method’, *IEEE Transactions on Antennas and Propagation* **26**(4), 614–616. [Cited on page 16.]
- Deaño, A., Huybrechs, D. and Iserles, A. (2017), *Computing Highly Oscillatory Integrals*, Other Titles in Applied Mathematics, Society for Industrial and Applied Mathematics. [Cited on pages 12, 24, 26, 30, 114, 188, 191, 192, 196, 235, 236, and 238.]
- DLMF (2021), ‘NIST Digital Library of Mathematical Functions’. Release 1.1.1 of 2021-03-15, Olver, F. W. J., Olde Daalhuis, A. B., Lozier, D. W., Schneider, B. I., Boisvert, R. F., Clark, C. W., Miller, B. R., Saunders, B. V., Cohl, H. S. and McClain M. A., eds. **URL:** <http://dlmf.nist.gov/> [Cited on page 211.]
- Domínguez, V. (2014), ‘Filon–Clenshaw–Curtis rules for a class of highly-oscillatory integrals with logarithmic singularities’, *Journal of Computational and Applied Mathematics* **261**, 299–319. [Cited on page 188.]

- Domínguez, V., Graham, I. G. and Kim, T. (2013), ‘Filon–Clenshaw–Curtis rules for highly oscillatory integrals with algebraic singularities and stationary points’, *SIAM Journal on Numerical Analysis* **51**(3), 1542–1566. [Cited on pages 189, 200, 201, 206, and 207.]
- Domínguez, V., Graham, I. G. and Smyshlyaev, V. P. (2011), ‘Stability and error estimates for Filon–Clenshaw–Curtis rules for highly oscillatory integrals’, *IMA Journal of Numerical Analysis* **31**(4), 1253–1280. [Cited on pages 27, 188, 192, 202, 203, 205, 206, 209, 227, 228, and 236.]
- Dowling, A., Greitzer, E., Hynes, T., Hileman, J. and Spakovszky, Z. (2007), ‘The Silent Aircraft’, *AeroAstro* **4**. Accessed: 09-05-2021.
URL: <http://web.mit.edu/aeroastro/news/magazine/aeroastro-no4/silentaircraft.html>
 [Cited on page 3.]
- Edmunds, D. and Evans, D. (2018), *Spectral Theory and Differential Operators*, Oxford Mathematical Monographs, second edn, OUP Oxford.
URL: <https://books.google.fr/books?id=C4FZDwAAQBAJ> [Cited on page 23.]
- Engquist, B., Fokas, A., Hairer, E. and Iserles, A., eds (2009), *Highly Oscillatory Problems*, London Mathematical Society Lecture Note Series, Cambridge University Press. [Cited on page 7.]
- ENOVAL Consortium (2019), ‘Final report (FP7-604999)’. Accessed: 09-05-2021.
URL: <https://cordis.europa.eu/docs/results/604/604999/final1-enoval-final-reporting.pdf> [Cited on pages 3 and 4.]
- European Commission (2012), ‘Flightpath 2050 - Europe’s Vision for Aviation’. Accessed: 09-05-2021.
URL: <https://op.europa.eu/s/o9UU> [Cited on page 3.]
- European Environment Agency (2020), ‘Environmental noise in Europe – 2020 (EEA Report No 22/2019)’. Accessed: 26-04-2021.
URL: <https://www.eea.europa.eu/ds-resolveuid/6f3ccabbe95b4ee38fe521b4bc47a5b0>
 [Cited on page 1.]
- Evans, L. (2010), *Partial Differential Equations*, Graduate studies in mathematics, second edn, American Mathematical Society. [Cited on page 23.]
- Evers, I. and Peake, N. (2002), ‘On sound generation by the interaction between turbulence and a cascade of airfoils with non-uniform mean flow’, *Journal of Fluid Mechanics* **463**, 25–52. [Cited on page 36.]

- Ffowcs Williams, J. E. and Hall, L. H. (1970), ‘Aerodynamic sound generation by turbulent flow in the vicinity of a scattering half plane’, *Journal of Fluid Mechanics* **40**(4), 657–670. [Cited on page 84.]
- Filon, L. N. G. (1930), ‘On a quadrature formula for trigonometric integrals’, *Proceedings of the Royal Society of Edinburgh* **49**, 38–47. [Cited on pages 26 and 188.]
- Fokas, A., Iserles, A. and Smitheman, S. (2013), ‘The unified method in polygonal domains via the explicit fourier transform of legendre polynomials’. [Cited on page 200.]
- Fredholm, I. (1903), ‘Sur une classe d’équations fonctionnelles’, *Acta Mathematica* **27**(1), 365–390. [Cited on page 10.]
- Ganesh, M. and Graham, I. (2004), ‘A high-order algorithm for obstacle scattering in three dimensions’, *Journal of Computational Physics* **198**(1), 211–242. [Cited on page 115.]
- Gao, J. and Iserles, A. (2017*a*), ‘A generalization of Filon-Clenshaw-Curtis quadrature for highly oscillatory integrals’, *BIT Numerical Mathematics* **57**(4), 943–961. [Cited on pages 188, 189, 190, 192, 222, 223, and 236.]
- Gao, J. and Iserles, A. (2017*b*), ‘Error analysis of the extended Filon-type method for highly oscillatory integrals’, *Research in the Mathematical Sciences* **4**(21). [Cited on pages 188, 189, and 190.]
- Garoni, C. (2013), ‘Estimates for the minimum eigenvalue and the condition number of Hermitian (block) Toeplitz matrices’, *Linear Algebra and its Applications* **439**(3), 707–728. [Cited on page 256.]
- Garoni, C. and Serra-Capizzano, S. (2017), *Generalized Locally Toeplitz sequences: theory and applications*, Vol. 1, Springer. [Cited on page 256.]
- Gautschi, W. (1979), ‘A computational procedure for incomplete gamma functions’, *ACM Transactions on Mathematical Software (TOMS)* **5**(4), 466–481. [Cited on page 197.]
- Gibbs, A. (2020*a*), ‘HNABEMLAB’. Accessed: 15-06-2021.
URL: <https://github.com/AndrewGibbs/HNABEMLAB> [Cited on page 189.]
- Gibbs, A. (2020*b*), ‘PathFinder’. Accessed: 15-06-2021.
URL: <https://github.com/AndrewGibbs/PathFinder> [Cited on page 189.]
- Gibbs, A., Hewett, D. P., Huybrechs, D. and Parolin, E. (2020), ‘Fast hybrid numerical-asymptotic boundary element methods for high frequency screen and aperture problems based on least-squares collocation’, *SN Partial Differential Equations and Applications* **1**(4), 21. [Cited on pages 11, 114, 115, 189, 215, 217, and 255.]

- Glegg, S. A. (1999), ‘The response of a swept blade row to a three-dimensional gust’, *Journal of Sound and Vibration* **227**(1), 29–64. [Cited on pages 5, 36, 45, 47, 50, 52, 56, 62, 70, 85, 106, and 252.]
- Goldstein, M. E. (1976), *Aeroacoustics*, McGraw-Hill, New York. [Cited on pages 4, 38, 84, 85, and 90.]
- Gradshteyn, I. S. and Ryzhik, I. M. (2000), *Table of Integrals, Series, and Products*, sixth edn, Academic press. [Cited on page 211.]
- Green, G. (1828), *An essay on the application of mathematical analysis to the theories of electricity and magnetism*, Nottingham. [Cited on page 10.]
- Groth, S. P., Hewett, D. P. and Langdon, S. (2015), ‘Hybrid numerical–asymptotic approximation for high-frequency scattering by penetrable convex polygons’, *IMA Journal of Applied Mathematics* **80**(2), 324–353. [Cited on page 254.]
- Guo, Y. P. (1991), ‘Sound diffraction and dissipation at a sharp trailing edge in a supersonic flow’, *Journal of Sound and Vibration* **145**(2), 179–193. [Cited on page 94.]
- Hao, S., Barnett, A. H., Martinsson, P.-G. and Young, P. (2014), ‘High-order accurate methods for Nyström discretization of integral equations on smooth curves in the plane’, *Advances in Computational Mathematics* **40**(1), 245–272. [Cited on pages 115 and 255.]
- Heins, A. E. (1950), ‘The reflection of an electromagnetic plane wave by an infinite set of plates, III’, *Quarterly of Applied Mathematics* **8**(3), 281–291. [Cited on page 35.]
- Heins, A. E. and Carlson, J. (1947), ‘The reflection of an electromagnetic plane wave by an infinite set of plates, II’, *Quarterly of Applied Mathematics* **5**(1), 82–88. [Cited on page 35.]
- Hewett, D. P., Langdon, S. and Chandler-Wilde, S. N. (2015), ‘A frequency-independent boundary element method for scattering by two-dimensional screens and apertures’, *IMA Journal of Numerical Analysis* **35**(4), 1698–1728. [Cited on pages 215 and 217.]
- Holst, L. (1980), ‘On the Lengths of the Pieces of a Stick Broken at Random’, *Journal of Applied Probability* **17**(3), 623–634. [Cited on page 149.]
- Howe, M. S. (1979), ‘Attenuation of sound in a low Mach Number nozzle flow’, *Journal of Fluid Mechanics* **91**(2), 209–229. [Cited on pages 6 and 84.]
- Howe, M. S. (1980), ‘The dissipation of sound at an edge’, *Journal of Sound and Vibration* **70**(3), 407–411. [Cited on pages 84 and 94.]

- Howe, M. S. (1984), ‘On the Absorption of Sound by Turbulence and Other Hydrodynamic Flows’, *IMA Journal of Applied Mathematics* **32**(1-3), 187–209. [Cited on page 95.]
- Howe, M. S. (1998), *Acoustics of Fluid-Structure Interactions*, Cambridge Monographs on Mechanics, Cambridge University Press, Cambridge. [Cited on page 94.]
- Hsiao, G. C., Steinbach, O. and Wendland, W. L. (2017), Boundary Element Methods: Foundation and Error Analysis, in E. Stein, R. de Borst and T. J. R. Hughes, eds, ‘Encyclopedia of Computational Mechanics Second Edition, Volume 2’, Wiley, pp. 841–902. [Cited on pages 11, 18, 122, 132, 146, and 150.]
- Hsiao, G. C. and Wendland, W. L. (1981), ‘The Aubin–Nitsche Lemma for Integral Equations’, *Journal of Integral Equations* **3**(4), 299–315. [Cited on page 132.]
- Huybrechs, D. (2006), Multiscale and hybrid methods for the solution of oscillatory integral equations, PhD thesis, Katholieke Universiteit Leuven. [Cited on page 21.]
- Huybrechs, D. (2021), ‘IntegralEquations2D.jl’. Julia package, accessed: 20-12-2021. **URL:** <https://github.com/daanhb/IntegralEquations2D.jl> [Cited on pages iii and 8.]
- Huybrechs, D. and Olteanu, A.-E. (2019), ‘An oversampled collocation approach of the Wave Based Method for Helmholtz problems’, *Wave Motion* **87**, 92–105. [Cited on pages 11 and 115.]
- Huybrechs, D. and Vandewalle, S. (2006), ‘On the evaluation of highly oscillatory integrals by analytic continuation’, *SIAM Journal on Numerical Analysis* **44**(3), 1026–1048. [Cited on pages 30 and 189.]
- Innovate UK (2019), ‘Electech sector: a roadmap for the UK’. Accessed: 03-06-2021. **URL:** <https://www.gov.uk/government/publications/electech-sector-a-roadmap-for-the-uk> [Cited on page 1.]
- Iserles, A. (2004), ‘On the numerical quadrature of highly-oscillating integrals I: Fourier transforms’, *IMA Journal of Numerical Analysis* **24**(3), 365–391. [Cited on pages 26 and 188.]
- Iserles, A. (2005), ‘On the numerical quadrature of highly-oscillating integrals II: Irregular oscillators’, *IMA Journal of Numerical Analysis* **25**(1), 25–44. [Cited on pages 26 and 188.]
- Iserles, A., Kropielnicka, K. and Singh, P. (2019a), ‘Compact schemes for laser–matter interaction in Schrödinger equation based on effective splittings of Magnus expansion’, *Computer Physics Communications* **234**, 195–201. [Cited on page 187.]

- Iserles, A., Kropielnicka, K. and Singh, P. (2019*b*), ‘Solving Schrödinger equation in semiclassical regime with highly oscillatory time-dependent potentials’, *Journal of Computational Physics* **376**, 564–584. [Cited on page 187.]
- Iserles, A. and Nørsett, S. P. (2004), ‘On quadrature methods for highly oscillatory integrals and their implementation’, *BIT Numerical Mathematics* **44**(4), 755–772. [Cited on pages 26 and 188.]
- Iserles, A. and Nørsett, S. P. (2005), ‘Efficient quadrature of highly oscillatory integrals using derivatives’, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **461**(2057), 1383–1399. [Cited on pages 29, 188, and 190.]
- Jaswon, M. A. (1963), ‘Integral equation methods in potential theory. I’, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **275**(1360), 23–32. [Cited on page 11.]
- Jones, D. S. (1972), ‘Aerodynamic Sound Due to a Source Near a Half-Plane’, *IMA Journal of Applied Mathematics* **9**(1), 114–122. [Cited on page 96.]
- Kaji, S. and Okazaki, T. (1970), ‘Propagation of sound waves through a blade row: II. Analysis based on the acceleration potential method’, *Journal of Sound and Vibration* **11**(3), 355–375. [Cited on pages 5, 35, and 62.]
- Keller, J. B. and Streifer, W. (1971), ‘Complex Rays with an Application to Gaussian Beams’, *Journal of the Optical Society of America* **61**(1), 40–43. [Cited on page 216.]
- Keller, P. (1999), ‘Indefinite integration of oscillatory functions’, *Applicaciones Mathematicae* **25**(3), 301–311. [Cited on page 188.]
- Keller, P. (2007), ‘A method for indefinite integration of oscillatory and singular functions’, *Numerical Algorithms* **46**(3), 219–251. [Cited on pages 188 and 193.]
- Khrapkov, A. (1971), ‘Certain cases of the elastic equilibrium of an infinite wedge with a nonsymmetric notch at the vertex, subjected to concentrated forces: PMM vol. 35, no. 4, 1971, pp. 677–689’, *Journal of Applied Mathematics and Mechanics* **35**(4), 625–637. [Cited on page 16.]
- Kim, T. (2012), Asymptotic and numerical methods for high-frequency scattering problems, PhD thesis, University of Bath. [Cited on page 189.]
- Kisil, A. (2015), Approximate Wiener–Hopf Factorisation with Stability Analysis, PhD thesis, University of Cambridge. [Cited on page 12.]

- Kisil, A. V. (2018), ‘An iterative Wiener–Hopf method for triangular matrix functions with exponential factors’, *SIAM Journal on Applied Mathematics* **78**(1), 45–62. [Cited on page 16.]
- Koch, W. (1971), ‘On the transmission of sound waves through a blade row’, *Journal of Sound and Vibration* **18**(1), 111–128. [Cited on pages 5, 35, 62, 65, 66, 70, 96, and 99.]
- Koch, W. (2009), ‘Acoustic resonances and trapped modes in annular plate cascades’, *Journal of Fluid Mechanics* **628**, 155–180. [Cited on page 77.]
- Kozaki, S. (1982), ‘Scattering of a Gaussian beam by a homogeneous dielectric cylinder’, *Journal of Applied Physics* **53**(11), 7195–7200. [Cited on page 216.]
- Kravtsov, Y. A. (1967), ‘Complex rays and complex caustics’, *Radiophysics and Quantum Electronics* **10**, 719–730. [Cited on page 216.]
- Kress, R. (1991), ‘Boundary integral equations in time-harmonic acoustic scattering’, *Mathematical and Computer Modelling* **15**(3-5), 229–243. [Cited on page 255.]
- Kress, R. (1999), *Linear Integral Equations*, Applied Mathematical Sciences, second edn, Springer-Verlag New York. [Cited on page 115.]
- Kress, R. and Sloan, I. H. (1993), ‘On the numerical solution of a logarithmic integral equation of the first kind for the Helmholtz equation’, *Numerische Mathematik* **66**(1), 199–214. [Cited on pages 135, 161, 162, and 163.]
- Kupradze, V. and Aleksidze, M. (1964), ‘The method of functional equations for the approximate solution of certain boundary value problems’, *USSR Computational Mathematics and Mathematical Physics* **4**(4), 82–126. [Cited on page 11.]
- Langdon, S. and Chandler-Wilde, S. N. (2006), ‘A wavenumber independent boundary element method for an acoustic scattering problem’, *SIAM Journal on Numerical Analysis* **43**(6), 2450–2477. [Cited on page 215.]
- Lapworth, L. (2019), ‘Large scale industrial simulations using HPC’. Keynote talk at the HiPEAC Computer Systems Week Spring 2019. Accessed: 09-05-2021.
URL: <https://www.youtube.com/watch?v=VldCGjirZsw&t=146s> [Cited on page 7.]
- Lele, S. K. and Nichols, J. W. (2014), ‘A second golden age of aeroacoustics?’, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **372**(2022), 20130321. [Cited on page 7.]
- Levin, D. (1982), ‘Procedures for computing one-and two-dimensional integrals of functions with rapid irregular oscillations’, *Mathematics of Computation* **38**(158), 531–538. [Cited on pages 30, 235, and 236.]

- Levin, D. (1996), ‘Fast integration of rapidly oscillatory functions’, *Journal of Computational and Applied Mathematics* **67**(1), 95–101. [Cited on pages 30, 235, 236, 237, and 256.]
- Lewanowicz, S. (1991), ‘A new approach to the problem of constructing recurrence relations for the Jacobi coefficients’, *Applicationes Mathematicae* **21**(2), 303–326. [Cited on pages 189 and 193.]
- Lighthill, J. (1993), The Final Panel Discussion, in J. C. Hardin and M. Y. Hussaini, eds, ‘Computational Aeroacoustics’, Springer New York, pp. 499–513. [Cited on page 7.]
- Lighthill, M. J. (1952), ‘On sound generated aerodynamically I. General theory’, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **211**(1107), 564–587. [Cited on pages 6 and 84.]
- Maierhofer, G. and Huybrechs, D. (2022a), ‘An analysis of least-squares oversampled collocation methods for compactly perturbed boundary integral equations in two dimensions’, *Under review, arXiv preprint arXiv:2201.11683* . [Cited on page iii.]
- Maierhofer, G. and Huybrechs, D. (2022b), ‘Convergence analysis of oversampled collocation boundary element methods in 2D’, *Advances in Computational Mathematics (in press)* . [Cited on page iii.]
- Maierhofer, G., Iserles, A. and Peake, N. (2021), ‘Recursive moment computation in Filon methods and application to high-frequency wave scattering in two dimensions’, *Under review* . [Cited on page iii.]
- Maierhofer, G. and Peake, N. (2020), ‘Wave scattering by an infinite cascade of non-overlapping blades’, *Journal of Sound and Vibration* **481**, 115418. [Cited on page iii.]
- Maierhofer, G. and Peake, N. (2022), ‘Acoustic and hydrodynamic power of wave scattering by an infinite cascade of plates in mean flow’, *Journal of Sound and Vibration* **520**, 116564. [Cited on page iii.]
- Mani, R. and Horvay, G. (1970), ‘Sound transmission through blade rows’, *Journal of Sound and Vibration* **12**(1), 59–83. [Cited on page 47.]
- Mavaleix-Marchessoux, D., Bonnet, M., Chaillat, S. and Leblé, B. (2020), ‘A fast BEM procedure using the Z-transform and high-frequency approximations for large-scale 3D transient wave problems’, *International Journal for Numerical Methods in Engineering* **121**, 4734–4767. [Cited on page 215.]
- McLean, W. (2000), *Strongly Elliptic Systems and Boundary Integral Equations*, Cambridge University Press. [Cited on page 18.]

- Meier, A., Arens, T., Chandler-Wilde, S. N. and Kirsch, A. (2000), ‘A nyström method for a class of integral equations on the real line with applications to scattering by diffraction gratings and rough surfaces’, *The Journal of Integral Equations and Applications* pp. 281–321. [Cited on pages 115 and 255.]
- Melenk, J. M. (2010), ‘On the convergence of Filon quadrature’, *Journal of Computational and Applied Mathematics* **234**(6), 1692–1701. [Cited on pages 26, 27, 29, 188, 201, and 209.]
- Morfe, C. (1971), ‘Acoustic energy in non-uniform flows’, *Journal of Sound and Vibration* **14**(2), 159–170. [Cited on page 84.]
- Morse, P. and Feshbach, H. (1953), *Methods of Theoretical Physics. Part I*, International series in pure and applied physics, McGraw-Hill. [Cited on page 144.]
- Myers, M. K. (1986), ‘An exact energy corollary for homentropic flow’, *Journal of Sound and Vibration* **109**(2), 277–284. [Cited on pages 6, 83, and 84.]
- Myers, M. K. (1991), ‘Transport of energy by disturbances in arbitrary steady flows’, *Journal of Fluid Mechanics* **226**, 383–400. [Cited on page 254.]
- Myers, M. R. and Kerschen, E. J. (1995), ‘Influence of incidence angle on sound generation by airfoils interacting with high-frequency gusts’, *Journal of Fluid Mechanics* **292**, 271–304. [Cited on page 253.]
- Möhring, W. (1970), ‘Zum Energiesatz bei Schallausbreitung in stationär strömenden Medien’, *Zeitschrift für Angewandte Mathematik und Mechanik* **50**, 196–198. [Cited on page 84.]
- Möhring, W. (1973), ‘On energy, group velocity and small damping of sound waves in ducts with shear flow’, *Journal of Sound and Vibration* **29**(1), 93–101. [Cited on page 84.]
- Noble, B. (1958), *Methods based on the Wiener-Hopf Technique for the Solution of Partial Differential Equations*, Pergamon Press, London. [Cited on pages 12, 15, 16, 17, 46, 49, 52, 70, and 73.]
- Ofcom and HM Government (2020), ‘5G mobile technology: a guide’. Accessed: 03-06-2021.
URL: <https://www.gov.uk/government/publications/5g-mobile-technology-a-guide>
 [Cited on page 8.]
- Oliver, J. (1968), ‘The numerical solution of linear recurrence relations’, *Numerische Mathematik* **11**(4), 349–360. [Cited on pages 189, 202, 205, 213, and 255.]

- Olver, S. (2006), ‘Moment-free numerical integration of highly oscillatory functions’, *IMA Journal of Numerical Analysis* **26**(2), 213–227. [Cited on pages 188 and 200.]
- Olver, S. (2007), ‘Moment-free numerical approximation of highly oscillatory integrals with stationary points’, *European Journal of Applied Mathematics* **18**(4), 435–447. [Cited on pages 188 and 200.]
- Olver, S. (2008), Numerical Approximation of Highly Oscillatory Integrals, PhD thesis, University of Cambridge. [Cited on page 256.]
- Parker, R. (1966), ‘Resonance effects in wake shedding from parallel plates: Some experimental observations’, *Journal of Sound and Vibration* **4**(1), 62–72. [Cited on page 77.]
- Parolin, E. (2015), A hybrid numerical-asymptotic boundary element method for high-frequency wave scattering, Master’s thesis, University of Oxford. [Cited on pages 189 and 215.]
- Peake, N. (1992), ‘The interaction between a high-frequency gust and a blade row’, *Journal of Fluid Mechanics* **241**, 261–289. [Cited on pages 5, 36, 39, 40, 43, 47, 50, 52, 55, 70, and 73.]
- Peake, N. (1993), ‘The scattering of vorticity waves by an infinite cascade of flat plates in subsonic flow’, *Wave Motion* **18**(3), 255–271. [Cited on pages 5, 36, 39, 44, 56, 62, 63, and 252.]
- Peake, N. and Kerschen, E. J. (1995), ‘A uniform asymptotic approximation for high-frequency unsteady cascade flow’, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **449**(1935), 177–186. [Cited on page 36.]
- Peake, N. and Kerschen, E. J. (1997), ‘Influence of mean loading on noise generated by the interaction of gusts with a flat-plate cascade: upstream radiation’, *Journal of Fluid Mechanics* **347**, 315–346. [Cited on pages 36, 37, and 253.]
- Peake, N. and Kerschen, E. J. (2004), ‘Influence of mean loading on noise generated by the interaction of gusts with a cascade: downstream radiation’, *Journal of Fluid Mechanics* **515**, 99–133. [Cited on pages 36, 37, 39, and 253.]
- Peake, N. and Parry, A. B. (2012), ‘Modern challenges facing turbomachinery aeroacoustics’, *Annual Review of Fluid Mechanics* **44**(1), 227–248. [Cited on page 4.]
- Piessens, R. and Branders, M. (1983), ‘Modified Clenshaw–Curtis method for the computation of Bessel function integrals’, *BIT Numerical Mathematics* **23**(3), 370–381. [Cited on pages 188 and 202.]

- Porter, R. (2021), ‘Plate arrays as a perfectly-transmitting negative-refraction metamaterial’, *Wave Motion* **100**, 102673. [Cited on pages 6, 85, 99, and 253.]
- Posson, H., Moreau, S. and Roger, M. (2010), ‘On the use of a uniformly valid analytical cascade response function for fan broadband noise predictions’, *Journal of Sound and Vibration* **329**(18), 3721–3743. [Cited on pages 4, 36, and 252.]
- Powell, M. J. D. (1981), *Approximation theory and methods*, CUP, Cambridge. [Cited on page 237.]
- Prandtl, L. (1918), ‘Tragflügeltheorie. I. Mitteilung’, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* pp. 451–477. [Cited on page 94.]
- Priddin, M. J., Kisil, A. V. and Ayton, L. J. (2020), ‘Applying an iterative method numerically to solve $n \times n$ matrix Wiener–Hopf equations with exponential factors’, *Philosophical Transactions of the Royal Society A* **378**(2162), 20190241. [Cited on page 16.]
- Rienstra, S. W. (1981), ‘Sound diffraction at a trailing edge’, *Journal of Fluid Mechanics* **108**, 443–460. [Cited on pages 6, 84, 85, 94, 105, and 107.]
- Rienstra, S. W. (1984), ‘Acoustic radiation from a semi-infinite annular duct in a uniform subsonic mean flow’, *Journal of Sound and Vibration* **94**(2), 267–288. [Cited on page 84.]
- Ritz, W. (1909), ‘Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik.’, *Journal für die reine und angewandte Mathematik* pp. 1–61. [Cited on page 10.]
- Rizzo, F. J. (1967), ‘An integral equation approach to boundary value problems of classical elastostatics’, *Quarterly of applied mathematics* **25**(1), 83–95. [Cited on page 11.]
- Rokhlin, V. (1985), ‘Rapid solution of integral equations of classical potential theory’, *Journal of Computational Physics* **60**(2), 187–207. [Cited on page 10.]
- Saffman, P. G. (1993), *Vortex Dynamics*, Cambridge Monographs on Mechanics, Cambridge University Press, Cambridge. [Cited on page 94.]
- Saranen, J. and Vainikko, G. (2013), *Periodic Integral and Pseudodifferential Equations with Numerical Approximation*, Springer Monographs in Mathematics, Springer. [Cited on pages 123 and 138.]
- Saranen, J. and Wendland, W. L. (1985), ‘On the asymptotic convergence of collocation methods with spline functions of even degree’, *Mathematics of Computation* **45**(171), 91–108. [Cited on page 114.]

- Schwab, C. and Wendland, W. L. (1992), ‘On numerical cubatures of singular surface integrals in boundary element methods’, *Numerische Mathematik* **62**(1), 343–369. [Cited on page 10.]
- Schwarzschild, K. (1901), ‘Die Beugung und Polarisation des Lichts durch einen Spalt. I’, *Mathematische Annalen* **55**(2). [Cited on page 36.]
- Shadrin, A. (1995), ‘Error Bounds for Lagrange Interpolation’, *Journal of Approximation Theory* **80**(1), 25–49. [Cited on pages 27 and 209.]
- Sloan, I. H. (1988), ‘A quadrature-based approach to improving the collocation method’, *Numerische Mathematik* **54**(1), 41–56. [Cited on pages 114, 117, and 134.]
- Sloan, I. H. (1992), ‘Error analysis of boundary integral methods’, *Acta Numerica* **1**, 287–339. [Cited on pages 11, 134, 135, 136, 137, 138, and 149.]
- Sloan, I. H. and Wendland, W. L. (1989), ‘A quadrature-based approach to improving the collocation method for splines of even degree’, *Zeitschrift für Analysis und ihre Anwendungen* **8**(4), 361–376. [Cited on pages 114 and 117.]
- Symm, G. T. (1963), ‘Integral equation methods in potential theory. II’, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **275**(1360), 33–46. [Cited on page 11.]
- Torres, R. H. and Welland, G. V. (1993), ‘The Helmholtz Equation and Transmission Problems with Lipschitz Interfaces’, *Indiana University Mathematics Journal* **42**(4), 1457–1485. [Cited on page 142.]
- Trefethen, L. (2019), *Approximation Theory and Approximation Practice*, Other Titles in Applied Mathematics, extended edn, SIAM, Society for Industrial and Applied Mathematics. [Cited on page 245.]
- Trefethen, L. N. (2008), ‘Is Gauss quadrature better than Clenshaw–Curtis?’, *SIAM review* **50**(1), 67–87. [Cited on pages 198 and 241.]
- UK Acoustics Network (2019), ‘UK Acoustics: Sound Economics’. Accessed: 26-04-2021. **URL:** <https://acoustics.ac.uk/wp-content/uploads/2019/05/Acoustics-sound-economy-report-May-2019.pdf> [Cited on page 1.]
- Uncles, R. J. (1976), ‘Numerical solution of acoustic scattering problems at intermediate frequencies’, *The Journal of the Acoustical Society of America* **60**(1), 266–267. [Cited on page 11.]

- Vér, I. L. (1982), Perforated baffles prevent flow-induced acoustic resonances in heat exchangers, in 'Fortschritte der Akustik', FASE/DAGA '82, pp. 531–534. [Cited on page 6.]
- Watson, G. (1995), *A Treatise on the Theory of Bessel Functions*, second edn, Cambridge University Press. [Cited on pages 225 and 228.]
- Weisstein, E. W. (2002), *CRC Concise Encyclopedia of Mathematics*, second edn, CRC Press, London. [Cited on pages 17 and 144.]
- Whitehead, D. S. (1970), 'Vibration and Sound Generation in a Cascade of Flat Plates in Subsonic Flow', *University of Cambridge Department of Engineering Reports and Memoranda No. 3685*. [Cited on pages 5 and 35.]
- Wiener, N. and Hopf, E. (1931), 'Über eine Klasse singulärer Integralgleichungen', *Sitzungsberichte der Preuß. Akademie der Wissenschaften, Phys.-math. Klasse, 1931 XXXI*. [Cited on page 12.]

