Embe ing Metho s for the Numeri al Solution of Convolution Equations

James Atkinson

August 31, 2004

Abstract

In this report we study linear integral operators generated by a difference kernel. The embedding results of Porter [8] for such operators are extended to Toeplitz matrices which have analogous structure. This structure is preserved by a basic collocation discretisation of such integral equations in the sense that it results in a Toeplitz matrix equation. It follows that for this discretisation embedding can be employed at the computational level.

This application of the theory is used in the numerical solution of a simple two dimensional diffraction problem. When Green's method is used to reduce the dimension of the problem, embedding results can be applied to the resulting one dimensional equation. Our computations show how a thorough study can be completed numerically with embedding as a useful tool.

Acknowled ements

This project was funded by an EPSRC studentship.

The project was supervised by ave Porter to whom I am very grateful for the first hand account of an interesting subject and for the engaging supervision meetings. I would also lik

Contents

Chapter 1

Background

In this chapter we review some developments which provide context and motivation for this work. Our review is not comprehensive, we focus on the particular theory and application pertinent here.

1.1 Embeddin

In the Theory section of this report we are able to be quite precise about what we mean by embedding. In a wider context embedding has come to mean that the solution of a problem can be acquired directly from the solutions of other problems. Although vague this implies a structural result for the mapping in question. A class of equations which exhibit a lot of structure involve convolution operators which lead directly to embedding results.

1.2 Convolution Equations

Gokhberg and Feldman [2] studied extensively convolution structure in a broad sense, they derive results in both continuous and discrete systems. Here we are concerned with integral operators generated by a difference kernel on a finite interval which are referred to in [2] as truncated Wiener-Hopf integral operators. They prove an embedding result for such operators which is suggested by an analogous result for Toeplitz matrices (or truncated discrete Winer-Hopf operators). More extensive embedding results of Porter [8] for the same integral operators were motivated by a particular application which we now describe.

1.3 Wave iffraction

Many physical phenomena are described reasonably well by the scalar wave equation, for example surface waves on water, pressure waves in a fluid and electromagnetic waves. These correspond to increasing magnitudes of wave-number in ratio to the dimensions of the geometry. One can take advantage of approximations that can be made for large wave-numbers (short wavelengths) which are well established in the theory of electromagnetic wave propagation and referred to as Fresnel (second-order) and Fraunhofer (first-order) analysis. The use of exact, or near-field analysis, is necessary when the wave-number is smaller, like for water-wave and acoustic scattering problems.

A fundamental problem in this area, which can be considered in 2 , is the description of the diffracted field due to a plane wave incident on a thin barrier of infinite extent in which there is a barriers [13]. The class of equations for which embedding results are known to apply are expanded by these papers, to those involving integral operators with a matrix difference kernel and operators generated by sum-and-difference kernels.

1.5 Toeplit Matrices

The embedding results considered in this report are connected to those of Sakhnovich [9], which was shown by Porter in [8]. The results of Sakhnovich apply to integral operators however he also presents a discrete version applicable to Toeplitz matrices which is attributed to A.L.Sakhnovich [3]. This result bears the same relationship to the discrete case presented here as it does in the continuous case.

1.6 This Project

It can be beneficial to preserve structure when a continuous model is discretised for computation. For example one is often able to determine salient information about a solution, like symmetries or conserved quantities. We can place additional demands on our numerical method to ensure such features remain during computation. This project has that theme. The structure in this case is very rich, and as a result puts quite strict constraints on the discrete system.

In chapter 2, the theory section, we exhibit the structural components which lead to the embedding formula of Porter [8]. These are given discrete analogues so that we can derive an embedding result that applies to the analogous discrete system.

Chapter 3 concerns itself with the formwhic pictures of diffraction are included to illustrate this application of numerical methods.

Chapter 5 concludes the report and suggests avenues of investigation which it has thrown up.

Chapter 2

Theory

This chapter establishes what we mean by convolution equations and embedding. The development is abstract but we follow closely our motivational examples of the integral equation with a difference kernel and the Toeplitz matrix equation.

Initially for convenience and later through necessity we assume the vector space in which we work is over the field $\mathbb{C}$.

2.1 The enerator of Convolutions

At the heart of our discussion is a definition which associates with each vector in our space a linear transform on that space.

Definition. If i a linear pace then e refer to $: \longrightarrow \mathrm{Hom}(,)$ *a a Generator of Convolution hen, given any* $x, y \in$ *an* $\lambda, \mu \in \mathbb{C}$,

(i) (

4.

1. Suppose that A and B are of convolution type, then we get a straightforward embedding result from the fact that A and B commute. If $A\phi = y$ it follows that $AB\phi = By$ so that $B\phi$ is in the inverse image of By under A.

2. If y is such that $(y) = 1$ then $A\phi = y$ implies $A \quad (\phi) = \quad (y$

Theorem. If i injective an generate convolution on an if P i an invertible conjugate linear map on ,

$$P \quad {}^*(x) = \quad (x)P, \quad \forall x \in \qquad \Longleftrightarrow \qquad Q(x) = \quad {}^*(Px) \Rightarrow \begin{array}{l} Q \textit{ generate} \\ \textit{convolution on} \quad . \end{array}$$

Proof.

$\Rightarrow$) Axioms (i) and (iv) are trivial, it remains to verify axiom (ii) .

$$\begin{aligned} Q(x)y &= \quad {}^*(Px)y \\ &= P^{-1} \quad (Px)Py \\ &= P^{-1} \quad (Py)Px \\ &= \quad {}^*(Py)x \\ &= Q(y)x. \end{aligned}$$

$\Leftarrow$) Assuming Q is a generator of convolutions we can use (vi) to start the following string of implication

$$\begin{aligned} Q^*(Q(P^{-1}y)x)z &= Q^*(P^{-1}y)Q^*(x)z \quad \forall x,y,z \in \\ \Rightarrow \qquad (P \quad {}^*(y)x)z &= \quad (y) \quad (Px)z \\ \Rightarrow \qquad (z)P \quad {}^*(y)x &= \quad (z) \quad (y)Px \\ \Rightarrow \qquad (z)Hx &= \qquad \forall x,z \in \qquad \text{where } H = P \quad {}^*(y) - \quad (y)P \\ \Rightarrow \qquad (Hx)z &= \\ \Rightarrow \qquad (Hx) &= \qquad \forall x \in \\ \Rightarrow \qquad H &= \qquad \text{by assuming the injectivity of} \\ \Rightarrow \qquad P \quad {}^*(y) &= \quad (y)P \quad \forall y \in \quad . \end{aligned}$$

□

Our move now is in the direction developed by Porter in [8].

oy[htrwb)l2e44q2qqew4r8l6531244sdqR4S1q24qR4wS2Re
e &r)lisew1 q4q=9iw66AF

ii

yU

2.4 The Projection Property

The subset of   to which our embedding method applies is defined as follows. If there exists $\sigma \in \mathbb{R}$ so that $X = \sigma + \ (x) + \ ^*(x)$ is the projection onto x, i.e.

$$X\phi = \langle\, \phi \,|\, x \,\rangle x$$

for all $\phi \in$, then we say that x has the projection property. It turns out that we can find all such vectors for our core examples, although the arguments are slightly lengthy.

1. Suppose that f has the projection property, then for all $\phi \in L_2([\ ,1])$

$$\sigma\phi(x) + \int_0^x f(x-t)\phi(t)dt + \int_x^1 -$$

2. If v has the projection property then for all $\phi \in \mathbb{R}^n$

$$\sigma\phi_k + \sum_{j=0}^{k} v_{k-j}\phi_j + \sum_{j=k}^{n} \overline{v}_{j-k}\phi_j = v_k \sum_{j=0}^{n} \phi_j \overline{v}_j.$$

It is sufficient that this be true for the canonical basis so that choosing $\phi = e_l$ we find that v must satisfy

$$v_{k-l} = \overline{v}_l v_k, \qquad l < k, \tag{2.1}$$

$$\sigma + v_0 + \overline{v}_0 = \overline{v}_k v_k, \qquad l = k \quad \text{and} \tag{2.2}$$

$$\overline{v}_{l-k} = \overline{v}_l v_k, \qquad l > k, \tag{2.3}$$

for $k \in \{\ ,1 \ldots n\}$. Clearly from 2.2 $|v_0|^2 = |v_1|^2 = \ldots = |v_n|^2$ so that $v_k = \ \Rightarrow v = \ $. Pursuing non-trivial v we can assume $v_k \neq \ $ for all k, hence choosing $l = \ $ in 2.1 we can write $v_k = \overline{v}_0 v_k$ which implies $v_0 = 1$ therefore $|v_0|^2 = |v_1|^2 = \ldots = |v_n|^2 = 1$. This in fact fixes σ because 2.2 now implies $\sigma + 1 + 1 = 1$ so that $\sigma = -1$. Choosing $l = 1$ in 2.1 and $l = k + 1$ in 2.3 gives the recurrence $v_k = v_1 v_{k-1}$ and $v_{k+1} = v_1 v_k$ which both imply that v necessarily satisfies $v_k = v_1^k$. In fact such a v is also sufficient to satisfy 2.1 and 2.2, which is easy to verify. Finally, v_1 can be any number with unit magnitude, we choose $v_1 = e^{-i\gamma/n}$ where $\gamma \in \mathbb{R}$ is arbitrary. Thus we have determined that

$$v_j = e^{-i\gamma j/n}, \quad j \in \{\ ,1 \ldots n\}.$$

We can establish some elementary facts about vectors which have the projection property. First we notice that X is self-adjoint so, from the end of the last section, it commutes with P and hence X and P share eigenvectors. Clearly then, because x is the sole eigenvector of X, it is also an eigenvector of P. Appendix A documents properties of such eigenvectors which we shall use freely in the following. Note that in both our examples we have contrived to write the projection type vectors so that their eigenvalue under P is $e^{i\gamma}$.

If v and

$^*(Pw)Pv = \lambda$

Now, $_0 A\psi_0 = \ _0 w_0 = v$ and collecting together multiples of $v = A\phi$ in the above and for convenience writing the constant as c_0 we have that

$$A(c_0\phi - \ _0\psi_0) = \alpha_0 \langle\, \psi_0 \,|\, v \,\rangle \ \ ^*(m)v.$$

Notice here that if $\langle\, \psi_0 \,|\, v \,\rangle =$ we arrive straightaway at an embedding result. More generally we can use the parallel result

$$A(c_1\phi - \ _1\psi_1) = \alpha_1 \langle\, \psi_1 \,|\, v \,\rangle \ \ ^*(m)v,$$

to eliminate $\ ^*(m)v$. oing this and again collecting multiples of ϕ, then using the injectivity of A we find

$$c\phi = \alpha_1 \langle\, \psi_1 \,|\, v \,\rangle \ \ _0\psi_0 - \alpha_0 \langle\, \psi_0 \,|\, v \,\rangle \ \ _1\psi_1$$

for some $c \in \mathbb{C}$. This is our main result. We have left c in the formula although it is found quite easily by the less direct route of taking the inner product of this with w_0 or w_1, so that

$$c\langle\, \phi \,|\, w_0 \,\rangle = \alpha_1 \langle\, \psi_1 \,|\, v \,\rangle \langle\, \psi_0 \,|\ \ _0^* w_0 \,\rangle - \alpha_0 \langle\, \psi_0 \,|\, v \,\rangle \langle\, \psi_1 \,|\ \ _1^* w_0 \,\rangle.$$

It is easy to verify that $\ _0^* w_0 = \lambda\overline{\mu}_0 v$ and $\overline{\alpha}_0 \ \ _1^* w_0 = (\overline{\alpha}_0 - \overline{\alpha}_1)w_0 + \overline{\alpha}_1\lambda\overline{\mu}_0 v$ and on substitution this reveals

$$c\langle\, \phi \,|\, w_0 \,\rangle = (\alpha_1 - \alpha_0)\langle\, \psi_0 \,|\, v \,\rangle \langle\, \psi_1 \,|\, w_0 \,\rangle.$$

Then $\langle\, \phi \,|\, w_0 \,\rangle = \mu_0\langle\, \phi \,|\, PA\psi_0 \,\rangle = \mu_0\langle\, \phi \,|\, A^*P\psi_0 \,\rangle = \mu_0\langle\, v \,|\, P\psi_0 \,\rangle = \mu_0\langle\, \psi_0 \,|\, Pv \,\rangle = \overline{\lambda}\mu_0\langle\, \psi_0 \,|\, v \,\rangle$ so that assuming v is such that $\langle\, \psi_0 \,|\, v \,\rangle \neq$ (or $\langle\, \psi_1 \,|\, v \,\rangle \neq$) we have found

$$c = \lambda\overline{\mu}_0\langle\, \psi_1 \,|\, w_0 \,\rangle(\alpha_1 - \alpha_0) = \lambda\overline{\mu}_1\langle\, \psi_0 \,|\, w_1 \,\rangle(\alpha_1 - \alpha_0).$$

The second result can be deduced directly because $\langle\, \psi_1 \,|\, w_0 \,\rangle = \mu_0\overline{\mu}_1\langle\, \psi_0 \,|\, w_1 \,\rangle$ using the same reasoning as above. We can now write the embedding formula in full, as

$$\lambda\overline{\mu}_0\langle\, \psi_1 \,|\, w_0 \,\rangle(\alpha_1 - \alpha_0)\phi = \alpha_1 \langle\, \psi_1 \,|\, v \,\rangle \ \ _0\psi_0 - \alpha_0 \langle\, \psi_0 \,|\, v \,\rangle \ \ _1\psi_1. \tag{2.6}$$

As a corollary to the main result, we also give an embedding formula for the inner products between vectors in S and their inverse image under A. We suppose additionally that $A\psi_2 = w_2 \in S$ and take the inner product of 2.6 with w_2. By reasoning exact

2.6 emarks

To summarise, we have introduced convolution as a set of commutative linear transforms in one-to-one correspondence with an underlying inner product space, . Our main supposition is the existence of P, a conjugate linear similarity transform between each convolution map and its adjoint. The idea of a generalised convolution operator, A, being a linear combination of the identity, a convolution map and the adjoint of a convolution map then becomes our focus. We define a set S of vectors with the projection property and suppose is such that S is non-empty. Finally we use convolution structure and projection property to construct an embedding formula for vectors in the inverse image of S under A.

It should be noted that the inverse image of vectors in the span of S are also embedded in the same sense, the calculation being a little more cumbersome.

Chapter

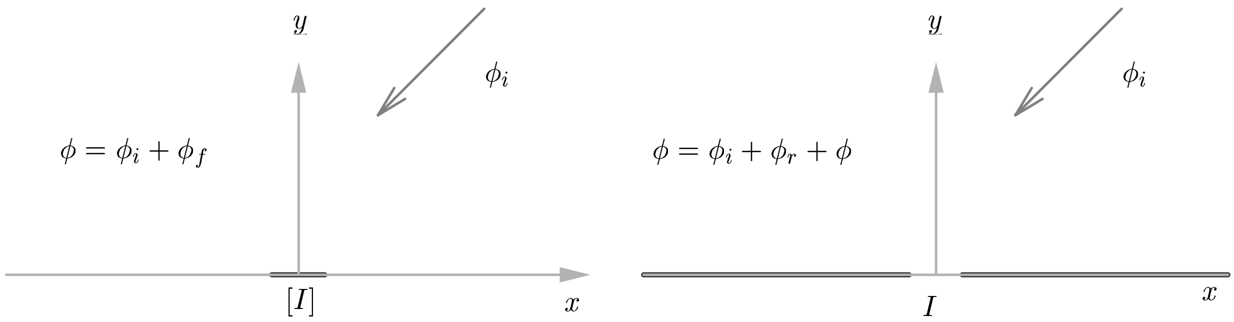


Figure 3.1: The complementary geometries.

3.2 The Solution in the Far-Field

In the far-field, I looks arbitrarily small. This gives physical plausibility to the important demand we place on the diffracted field, that it satisfy the Sommerfeld radiation condition,

$$\sqrt{r}\left(\frac{\partial}{\partial r} - ik\right)\phi_{f,} \;\to \qquad \text{as } r \to \infty.$$

Solutions of the Helmholtz equation in two dimensions which satisfy this condition are travelling toward the far-field and decaying like cylindrical waves (i.e. they look as if they originate from a point source). Such solutions are also unique which is a fact with immediate consequence for our problems.

It turns out that ϕ_f and ϕ are anti-symmetric through X. Although this is not obvious it reduces the problem so is introduced early for the sake of simplicity. It is straightforward to verify that the gradient of either diffracted field is continuous everywhere through X. If we then write it as the sum of odd and even parts, both of which must satisfy the Helmholtz equation (because the operator $\left(\nabla^2 + k^2\right)$ conserves such symmetry), we notice the even part has zero derivative everywhere through X. But the function which is everywhere zero also has zero derivative through X, satisfies the Helmholtz equation and Sommerfeld radiation condition, by the uniqueness of this solution we see that the even part of either diffracted field is zero.

3.3 The Solution Close to a Barrier Ed e

It is important later that we understand the behaviour of the diffracted field as we approach a barrier edge. Basic information is accessible to us by finding solutions of the Helmholtz equation separated in polar co-ordinates centred on the endpoint of a barrier in the limit $r \to$. Choosing the separation constant l^2, $l \in \mathbb{R}$, and writing $\phi_l = \rho(r)\sigma(\theta)$, ρ and σ satisfy

$$\left(r^2 \frac{\partial^2}{\partial r^2} + r\frac{\partial}{\partial r} + (kr)^2 - v \quad d\right.$$

3.5 eduction by reen's Method

We can establish an integral representation of the diffracted field using Green's method, for which we need a function G_0 which satisfies

$$\begin{aligned} \left(\nabla^2 + k^2\right) G_0 &= -\delta(x-x_0)\delta(y-y_0) \quad \text{in} \quad \Omega, \\ \sqrt{R}\left(\frac{\partial}{\partial R} - ik\right) G_0 & \end{aligned}$$

which is our desired integral representation of ϕ_f. sing G and ϕ in Green's second identity we find by similar reasoning that

$$\phi \ = -\int_I G \frac{\partial \phi}{\partial y} \ dx \quad \text{in } \Omega.$$

Of course, we do not know ϕ_f or $\frac{\partial \phi_g}{\partial y}$ on I; in fact we have reduced the total problem to one of finding these.

Note that the analysis of section 3.3 reveals its significance, informing us that $\frac{\partial \phi_g}{\partial y}$ is potentially singular close to the barrier edge which here means the endpoints of I.

3.6 The educed Equations

Let us defer ϕ_f for the moment; now we know ϕ on the

Asymptotically when r is much greater than $\sqrt{x^2+y^2} \quad 1$

$$r \quad \infty \quad \Rightarrow \quad \begin{array}{rcl} R & = & r - x\cos\theta - y\sin\theta + \quad (r^{-1}) \\ \widetilde{R} & = & r - x\cos\theta + y\sin\theta + \quad (r^{-1}). \end{array}$$

The asymptotic expansion of $H_0^{(1)}$ for large arguments (e.g. from [1]) leads directly to

$$R \quad \infty \quad \Rightarrow \quad G_0(R) = \frac{1}{2\sqrt{2 \ \ kR}} e^{i(kR+\frac{\pi}{4})} + \quad (R^{-\frac{3}{2}}).$$

We are now in a position to expand G and $\frac{\partial \hat{G}}{\partial y}$ asymptotically for $r \quad \infty$ and hence derive the far-field approximations to ϕ_f and ϕ , sparing the reader the details we find

$$r \quad \infty \quad \Rightarrow \quad \begin{array}{rcl} \phi_f & = & \dfrac{1}{\sqrt{2 \ \ kr}} e^{i(kr-\frac{3\pi}{4})} \left\{ ik\sin\theta \displaystyle\int_I \psi_\alpha \, e^{-ikx\cos\theta} \, dx \right\} + \quad (r^{-\frac{3}{2}}) \\ \phi & = & \dfrac{1}{\sqrt{2 \ \ kr}} e^{i(kr-\frac{3\pi}{4})} \left\{ \displaystyle\int_I v_\alpha \, e^{-ikx\cos\theta} \, dx \right\} + \quad (r^{-\frac{3}{2}}). \end{array}$$

The curly braces separate out the angular dependent part of the field which we now define as the far-field diffraction coefficients F and Γ corresponding to ϕ_f and ϕ . In terms of the $L_2(I)$ inner product,

$$\Gamma(\alpha,\theta) = \langle \, v_\alpha \mid f_{\pi-\theta} \, \rangle \qquad \text{and} \qquad F(\alpha,\theta) = ik\sin\theta \langle \, \psi_\alpha \mid f_{\pi-\theta} \, \rangle.$$

The symmetry property established in the previous section is also pertinent here. Consider U again: this operator is clearly unitary so following on from relations 3.5 we see that

$$\Gamma(\alpha,\theta) = \Gamma(\quad - \alpha, \quad - \theta) \quad \text{and} \quad F(\alpha,\theta) = F(\quad - \alpha, \quad - \theta).$$

3.8 The ual Problems Linked

Here we give an algebraic equation relating ψ_α to v_α, v_0 and v_π. We start by establishing a property of K which we can state more generally for a linear integral operator generated by a kernel of the form $k(x,t) = s(x-t)$. We consider the effect of differentiating $K\phi$ where $\phi \in C^1(I)$ is arbitrary. Suppose, without loss of generality, $I = [-1,1]$ then we can manipulate as follows by a change of integration variable,

$$\begin{aligned} \frac{d}{dx}\int_{-1}^{1} s(x-t)\phi(t)\, dt &= \frac{d}{dx}\int_{x-1}^{x+1} s(q)\phi(x-q)\, dq \\ &= \big[s(q)\phi(x-q)\big]_{x-1}^{x+1} + \int_{x-1}^{x+1} s(q)\phi'(x-q)\, dq \\ &= s(x+1)\phi(-1) - s(x-1)\phi(1) + \int_{-1}^{1} s(x-t)\phi'(t)\, dt. \end{aligned}$$

If ϕ is zero at the endpoints of the domain this implies that

$$K\phi = K \quad \phi.$$

By factoring the differential operator in the ψ_α equation (equation 3.4) we can commute one factor through K by using the above property which is valid because ψ_α is zero at the endpoints of I, and w

A similar result holds for the far-field diffraction coefficient $F(\alpha,\theta)$. Taking the inner product of equation 3.6 with $f_{\pi-\theta}$ and using relation 3.8 we find

$$\sin\alpha(\cos\theta - 1)F(\alpha,\theta) = \sin\theta(1+\cos\alpha)(\Gamma(\alpha,\theta) - c_\alpha\Gamma(\ ,\theta)), \tag{3.9}$$

an algebraic equation for $F(\alpha,\theta)$ in terms of $\Gamma(\alpha,\theta)$, $\Gamma(\ ,\theta)$ and c_α. Clearly we might have started from equation 3.7 and in so doing found a parallel relation to this involving d_α. Between these we can eliminate $\Gamma(\alpha,\theta)$ and on substitution for c_α and d_α it turns out that

$$2(\cos\alpha + \cos\theta)\Gamma(\ ,\ \ )F(\alpha,\theta) = \sin\alpha\sin\theta(\Gamma(\alpha,\ \ )\Gamma(\ ,\theta) - \Gamma(\alpha,\ \ )\Gamma(\ ,\theta)).$$

This relation has a deeper embedding property in that we get $F(\alpha,\theta)$ in terms of Γ for observation or incidence angles of and alone. In fact once $F(\alpha,\theta)$ has been established, 3.9 can then be used to calculate $\Gamma(\alpha,\theta)$ which leads to the embedding result for Γ. This can be decoupled from the equation for ψ_α as we show in the next section.

3.9 Embeddin

We identify here the structural components which lead to embedding. The principal difference between this problem and example 1 of the theory section is that the interval, I, is taken here to be centred on the origin; in fact without losing generality we can choose $I = (-1,1)$, I has been non-dimensionalised by the wave-number k.

A problem, which turns out to be benign, is that $v_\alpha \notin L_2([-1,1])$ due to its singular nature. From the analysis of section 3.3 we know that qualification here is that vectors in S have the projection property only for the subset $E_0 \quad E$ with which their inner product is defined. That everything in S is continuous means that, given our definition of E, $E_0 = E$.

The equation for v_α can then be written in the notation of chapter 2 as

$$(\kappa - \quad (l) - \quad ^*(m))\phi = v \in S \tag{3.1 }$$

where, for all $x \in [-1, 1]$,

$$\begin{aligned}
\kappa &= \quad , \\
l(x) &= \tfrac{-i}{2} H_0^{(1)}(k(x+1)), \\
m(x) &= \tfrac{i}{2} \overline{H}_0^{(1)}(k(x+1)), \\
\phi(x) &= e^{-ik\cos\alpha} v_\alpha(x) \quad \text{and} \\
v(x) &= e^{-ik\cos\alpha} f_\alpha(x).
\end{aligned}$$

The eigenvalue of v under P is $e^{i2k\cos\alpha}$ and, if $w_0 \in S$ corresponds to an angle β, the constant α_0 is straightforward to evaluate, and we find

$$\alpha_0 = ik(\cos\beta - \cos\alpha).$$

By substituting these quantities into equation 2.6 we arrive at the embedding result for v_α. We can benefit from the symmetry relation 3.5 by choosing $w_1 = Uw_0$ which correspond to incident angles β and $\quad - \beta$ for some $\beta \neq \quad /2$. We find, after noting that $\langle\, \phi \mid w_0 \,\rangle = e^{-ik(\cos\alpha - \cos \quad)}\Gamma(\alpha, \quad - \beta)$ and so on, that

$$\begin{aligned}
2\cos\beta\,\Gamma(\beta,\beta)\ v_\alpha = {} & \Gamma(\beta,\alpha)(\cos\alpha + \cos\beta - ik(\cos^2\alpha - \cos^2\beta) \quad (v))\ v \\
& -\Gamma(\beta, \quad - \alpha)(\cos\alpha - \cos\beta - ik(\cos^2\alpha - \cos^2\beta) \quad (v))CC
\end{aligned}$$

solution of a one-dimensional equation on a finite interval. We find the link between solutions of the two problems so it is possible to compute one from the other, which begins the embedding theme. It is continued because the one-dimensional equations involve the same integral operator which is of generalised convolution

Chapter 4

The Numerical Solution

In this chapter we solve the diffraction problem numerically. This means that for a given k and α, we compute approximations to the far-field diffraction pattern on $[\ ,\ ]$ and the diffracted field on some subset of $\mathbb{R}^2$. The embedding results for α can be discretised to give a computational saving independent of the numerical method used. However, it is desirable that the discrete system preserves embedding. The following commutative diagram, in the notation of the theory section, clarifies what we mean by this,

$$\begin{tikzcd} \psi_0, \psi_1 \arrow[r] \arrow[d] & \phi \arrow[d] \\ \widetilde{\psi}_0, \widetilde{\psi}_1 \arrow[r] & \widetilde{\phi} \end{tikzcd}$$

where arrows to the right represent embedding and arrows down represent discretisation.

This can be achieved by careful implementation of a basic collocation method, which has the additional benefit of being extremely simple. We write a collocation discretisation of equation 3.3 and it turns out we can identify the structural components which lead to embedding. This enables us to write the discrete equivalent of equation 3.1 and then direct substitution into 2.6 brings us to the discrete embedding formula.

4.1 iscretisation

We choose to approximate functions on the interval $[-1, 1]$ at the set of points

$$x_j = \frac{2j - n}{n+1},\ j \in \{\ , 1 \ldots n\},$$

the half-width of a discrete interval being $\delta = \frac{1}{n+1}$. Equation 3.3 can then be approximated by the matrix system

$$\widetilde{K}\widetilde{v}_\alpha = \widetilde{f}_\alpha, \tag{4.1}$$

where

$$\begin{aligned}
\widetilde{f}_{\alpha j} &= f_\alpha(x_j) \quad \text{and} \\
\widetilde{K}_{jl} &= \frac{i}{2}\int_{x_l-\delta}^{x_l+\delta} H_0^{(1)}(k|x_j - t|)\ dt + \quad (\delta
\end{aligned}$$

where

$$
\begin{aligned}
\kappa &= \quad , \\
l_j &= \frac{-i}{2}\int_{-\delta}^{\delta} H_0^{(1)}(k(2\delta j - t))\ dt + \quad (\delta^2), \quad j \in \{\ , 1 \ldots n\}, \\
m_j &= \frac{i}{2}\int_{-\delta}^{\delta} \overline{H}_0^{(1)}(k(2\delta j - t))\ dt + \quad (\delta^2), \quad j \in \{1, 2 \ldots n\}, \quad m_0 = \quad , \\
v &= e^{-ik\cos\alpha\delta n}\ \widetilde{f}_\alpha \quad \text{and} \\
\phi &= e^{-ik\cos\alpha\delta n}\ \widetilde{v}_\alpha .
\end{aligned}
$$

The eigenvalue of v under P is $e^{i2k\cos\alpha\delta n}$. In the notation of the theory section, we suppose $w_0 \in S$ corresponds to incident angle β, the constant α_0 can then be evaluated, we find

$$\alpha_0 = 1 - e^{-i2\delta k(\cos \quad - \cos\alpha)}.$$

As in the continuous case, we take advantage of the symmetry and choose $w_1 = \widetilde{U}w_0$ corresponding to angle $\quad - \beta$. Substitution of these quantities into equation 2.6 reveals, after a little manipulation,

$$
\begin{aligned}
2\sin(2\delta k\cos\beta)\widetilde{\Gamma}(\beta,\beta)\widetilde{v}_\alpha &= \widetilde{\Gamma}(\beta,\alpha)[\sin(2\delta k\cos\alpha) + \sin(2\delta k\cos\beta) + \\
&\quad\ i(\cos(2\delta k\cos\alpha) - \cos(2\delta k\cos\beta))(2 \quad (v) - 1)]\widetilde{v} \\
&- \widetilde{\Gamma}(\beta, \quad - \alpha)[\sin(2\delta k\cos\alpha) - \sin(2\delta k\cos\beta) + \\
&\quad\ i(\cos(2\delta k\cos\alpha) - \cos(2\delta k\cos\beta))(2 \quad (v) - 1)]\widetilde{U}\widetilde{v}\ ,
\end{aligned}
$$

where

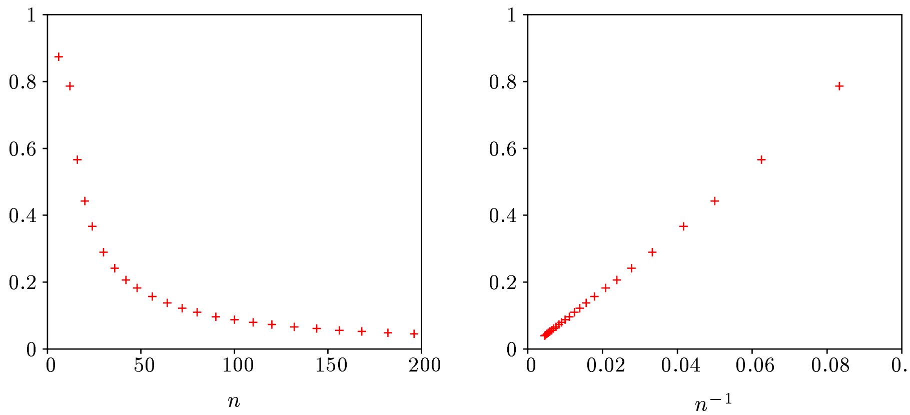
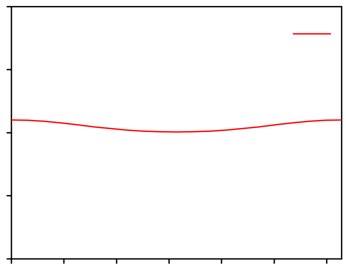
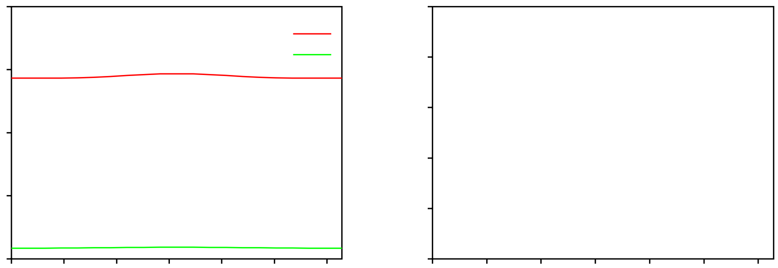


Figure 4.1: n dependence of the near-field RMS error, E





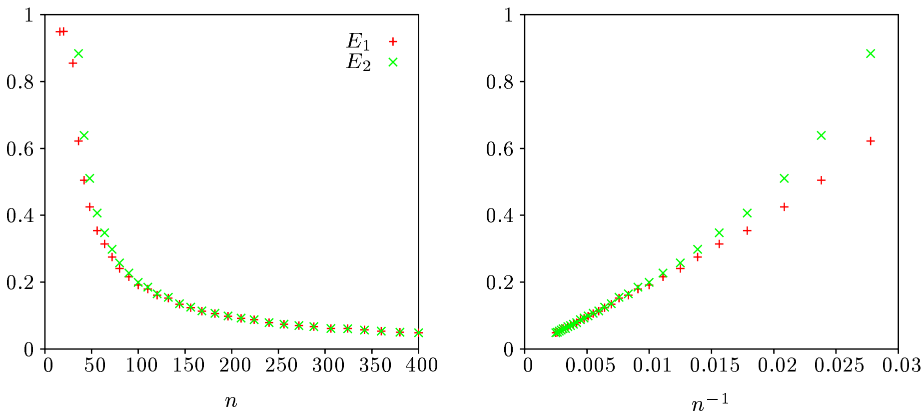
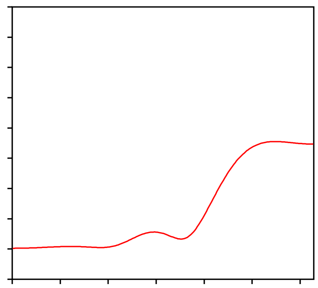


Figure 4.7: n dependence of the near-field RMS error, $k = 32$, $\alpha = \ .3$.

illustrates.

4.5 emarks

To summarise, we have discretised the integral equation for v_α in order to approximate it numerically. This has been done using a basic collocation method, but in two ways; one which draws on results of the theory section and preserves embedding, and one which is more satisfactory because it naturally removes the



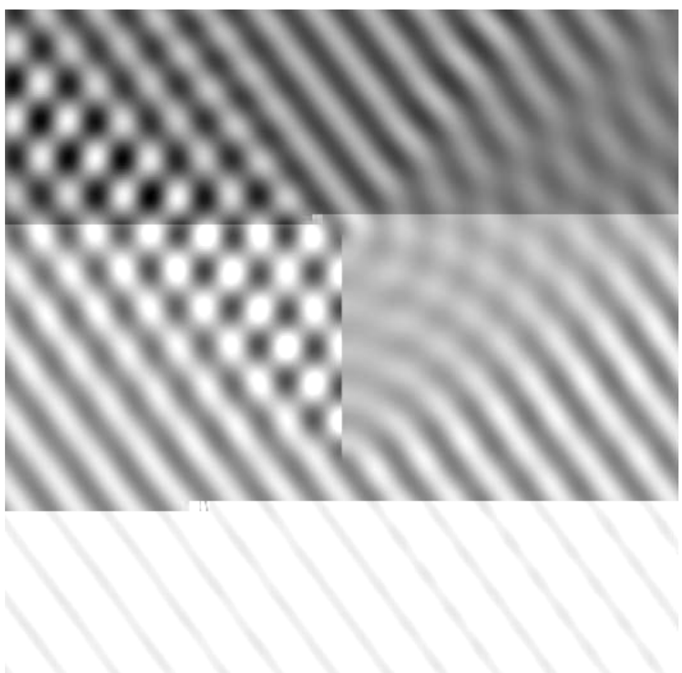


Figure 4.13: Re(ϕ) on $[-3,3] \times [-3,3]$, $k = 16$, $\alpha =\ .7$.

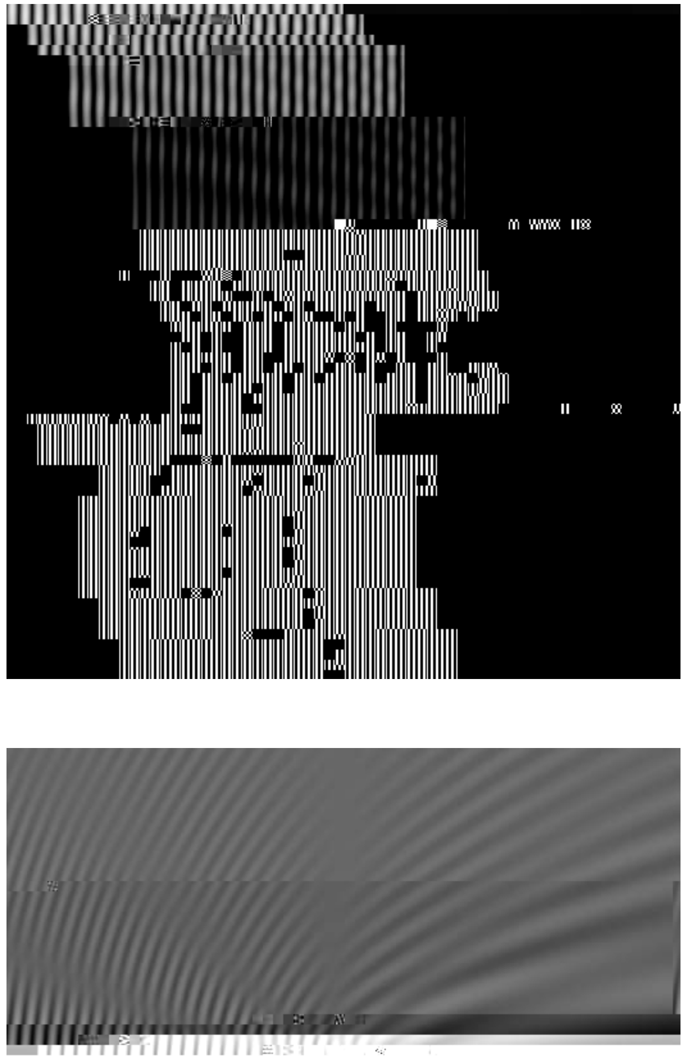


Figure 4.14: $|\phi|$ on $[-1\ ,1\ ]\times[-1\ ,1\ ]$, $k=8$, α

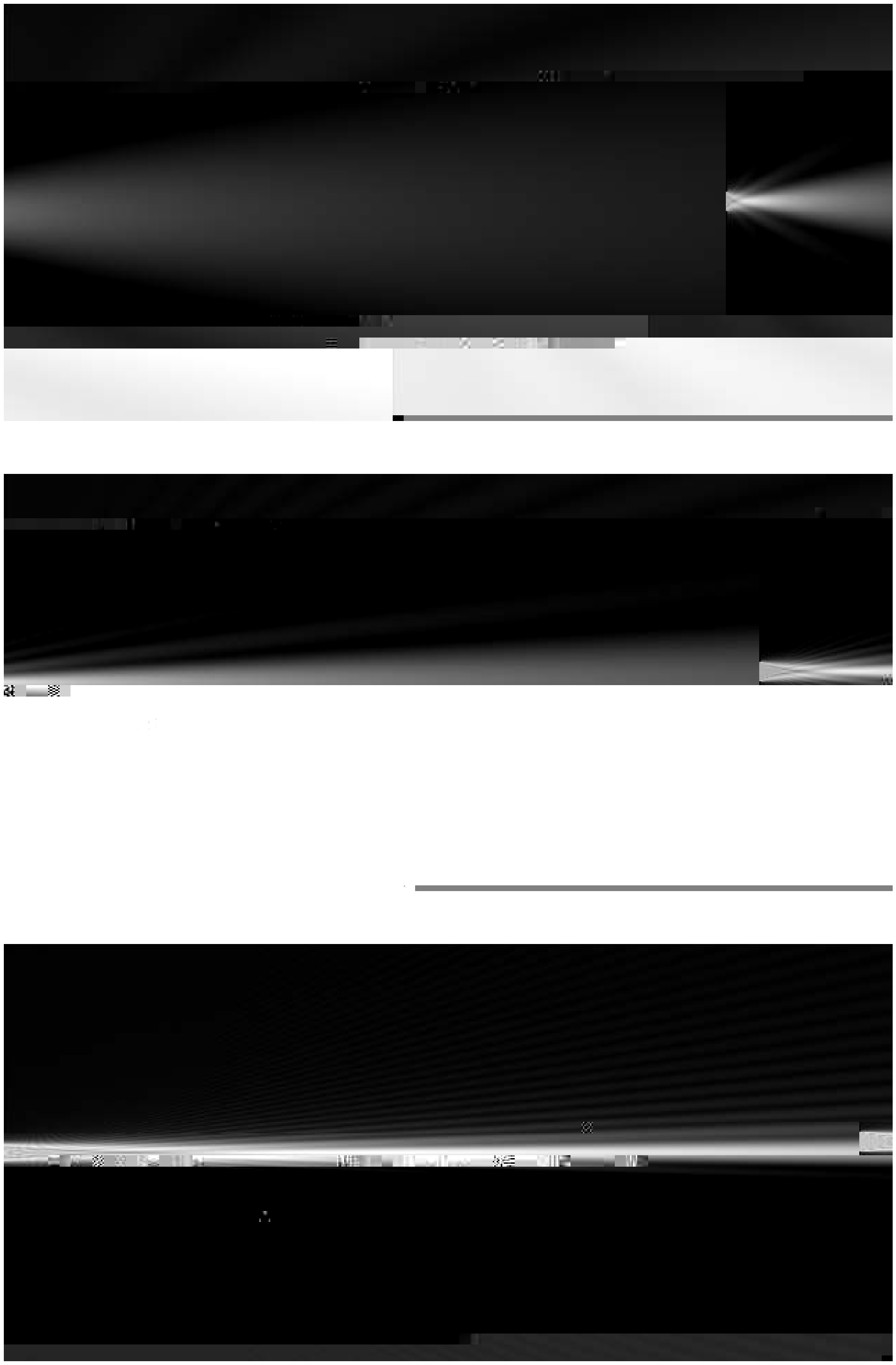


Figure 4.15: $|\phi \ |$ on $[-2$

Chapter 5

Conclusions

5.1 Overview

In this report we have derived an embedding result in an abstract setting which coincides with a result of Porter [8] in the continuous case and provides an analogous result for Toeplitz matrices in the discrete case. The application of this to the numerical solution of a simple diffraction problem illustrates how structure is preserved by a basic collocation discretisation of the governing integral equation. Our investigation into this numerical method has shown that, for this application, another discretisation isligati3B44A3SiR4SnRJw369m.93Jw3application,MT con

5.3 Further Work

This study has led to the identification of a few areas where further investigation may be useful.

- The core of the project was the theoretical work which stands alone now the intended application has been shown to be un-beneficial. As a general framework this is currently unsatisfactory because there exists a result for the principle examples (1 and 2 of the theory section) which it does not yet describe. Specifically, if $A = \kappa - \ (l) - \ ^*(m)$ is invertible and $\kappa \neq$ then the inverse is *finitely generate* by, i.e. expressible in terms of, ψ_0 and ψ_1 which satisfy

  $$A\psi_0 = l \quad \text{and} \quad A^*\psi_1 = m.$$

  This is analogous to example 3 of section 2.2 which applies to the simpler convolution type operator. In the continuous case this result is attributable to Gokhberg and Feldman [2] (mentioned already in section 1.2), in fact they discover it by analogy with its discrete counterpart. Porter [8] shows, again in the continuous case, that this result can be derived from the main embedding formula we give in this report, which suggests the results are intimately connected.

  nifying the discrete and continuous versions of this result thus seems a distinct possibility, and can only lead to a deeper insight into the nature of these embedding results and hence their extensions and applications.

- It may be that there exist other methods of discretisation which, when applied to integral operators generated by a difference kernel, lead to a Toeplitz matrix equation. If such a method also exhibited desirable error characteristics the discrete embedding formula could find application. If one considers implementing Galerkin's method using a finite set, say $\{w_i\}$, to approximate the solution of $A\phi = f$ where $A = \kappa - \ (l) - \ ^*(m)$ in the notation of the theory section, one quickly arrives at the discrete equation

  $$\sum_{i=0}^{n} \left[\kappa\langle\, w_i \,|\, w_j \,\rangle + \langle\, l \,|\ \ ^*(w_i)w_j \,\rangle + \langle\ \ ^*(w_j)w_i \,|\, m \,\rangle\right] \phi_i = \langle\, f \,|\, w_j \,\rangle,$$

  $$\text{where} \quad j \in \{\ ,1 \ldots n\} \quad \text{and} \quad \widetilde{\phi} = \sum_{i=0}^{n} \phi_i w_i.$$

  Which places quite specific demands on $\{w_i\}$ so that a search can be conducted in a systematic way.

- The numerical error analysis showed some interesting results for the approximate embedding formula, in some circumstances improving the accuracy over the non-embedded calculation. Such results may be worth investigating analytically because once understood they could lead to numerical methods with improved error performance for such equations.

Appendi A

Conjugate Linear Maps

A.1 Introduction

We list here some elementary properties of a conjugate linear map defined on a unitary space. By conjugate linear we mean that for all $x, y \in$ and $\lambda, \mu \in \mathbb{C}$

$$A(\lambda x + \mu y) = \overline{\lambda} A x + \overline{\mu} A y.$$

A.2 The Adjoint

We define the adjoint of A, which we write as A^*, so that

$$\langle\, Ax \,|\, y \,\rangle = \langle\, A^* y \,|\, x \,\rangle$$

for all $x, y \in$. It is easy to show that A^* is also conjugate linear. If is a linear transform on then A and A are conjugate linear and $(A \quad)^* = \quad ^*A^*$ because

$$\begin{aligned} \langle\, A \quad x \,|\, y \,\rangle &= \langle\, A^* y \,|\, \quad x \,\rangle \\ &= \langle\, \quad ^*A^* y \,|\, x \,\rangle. \end{aligned}$$

If is conjugate linear, A and A are linear and $(A \quad)^* = \quad ^*A^*$ because

$$\begin{aligned} \langle\, A \quad x \,|\, y \,\rangle &= \langle\, A^* y \,|\, \quad x \,\rangle \\ &= \langle\, x \,|\, \quad ^*A^* y \,\rangle. \end{aligned}$$

A.3 Unitary Conju ate Linear Maps

If P is conjugate linear and $P^*P = I$ we refer to P as unitary. If P has an eigenvector v with eigenvalue λ then

$$\langle\, v \,|\, v \,\rangle = \langle\, P^*Pv \,|\, v \,\rangle = \langle\, Pv \,|\, Pv \,\rangle = \lambda\overline{\lambda}\langle\, v \,|\, v \,\rangle.$$

i.e. $|\lambda| \neq 1 \Rightarrow v =$.

If, then, $v \neq$, $Pv = \lambda v \Rightarrow P^*Pv = \overline{\lambda}P^*v \Rightarrow v = \overline{\lambda}P^*v \Rightarrow \lambda v = P^*v$. So that eigenvectors of P are eigenvectors of P^* with the same eigenvalue. A consequence of this is that if v, w are eigenvectors of P with eigenvalues λ, μ, then

$$\lambda\langle\, v \mid w\, \rangle = \langle\, Pv \mid w\, \rangle = \langle\, P^*w \mid v\, \rangle = \mu\langle\, w \mid v\, \rangle.$$

Bibliography

[1] **J.H.Carr and .E.Stelzriede**, *Diffraction of ater ave by brea ater* S Nat. Bur. Stds. 521, 1 9-125, 1952.

[2] **I.Gohberg and I.Feldman** (translation from Russian by F.Goldware), *Convolution equation an projection metho*

[12] **N.R.T.Biggs and D.Porter** , *Wave Diffraction through a Perforate Barrier of Non-Zero Thic ne* , Q.Jl Mech. appl. Math. 54(4), 523-547, 2 1.

[13] **N.R.T.Biggs and D.Porter** , *Wave Scattering by a Perforate Duct*, Q.Jl Mech. appl. Math. 55(2), 249-272, 2 2.