Functions of Difference Matrices Are Toeplitz Plus Hankel

SIAM REVIEW
Vol. 56, No. 3, pp. 525–546

Functions of Difference Matrices Are Toeplitz Plus Hankel*

Gilbert Strang[†]
Shev MacNamara[†]

Abstract. When the heat equation and wave equation are approximated by $\boldsymbol{u}_t = -\boldsymbol{K}\boldsymbol{u}$ and $\boldsymbol{u}_{tt} = -\boldsymbol{K}\boldsymbol{u}$ (discrete in space), the solution operators involve $e^{-\boldsymbol{K}t}$, $\sqrt{\boldsymbol{K}}$, $\cos(\sqrt{\boldsymbol{K}}t)$, and $\operatorname{sinc}(\sqrt{\boldsymbol{K}}t)$. We compute these four matrices and find accurate approximations with a variety of boundary conditions. The second difference matrix $\boldsymbol{K}$ is Toeplitz (shift-invariant) for Dirichlet boundary conditions, but we show why $e^{-\boldsymbol{K}t}$ also has a Hankel (anti-shift-invariant) part. Any symmetric choice of the four corner entries of $\boldsymbol{K}$ leads to Toeplitz plus Hankel in all functions $f(\boldsymbol{K})$. Overall, this article is based on diagonalizing symmetric matrices, replacing sums by integrals, and computing Fourier coefficients.

Key words. Toeplitz, Hankel, Laplacian, exponential, Bessel function

AMS subject classifications. 97N99, 65-01, 65N06

DOI. 10.1137/120897572

1. Introduction.

In teaching numerical methods for partial differential equations, we begin with the heat equation and the wave equation. Our model problems are posed on an interval $0 \le x \le 1$ with zero boundary conditions. The second derivative u_{xx} is replaced by second differences at the mesh points $x = h, 2h, \ldots, Nh$. The second difference matrix with $1, -2, 1$ on its diagonals is denoted by $-\boldsymbol{K}$:

$$\begin{array}{lllll} (1.1) & \text{Heat equation} & \dfrac{\partial}{\partial t}\boldsymbol{u} = \dfrac{\partial^2}{\partial x^2}\boldsymbol{u} & \text{becomes} & \dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{u} = -\dfrac{\boldsymbol{K}}{h^2}\boldsymbol{u}; \\ & \text{Wave equation} & \dfrac{\partial^2}{\partial t^2}\boldsymbol{u} = \dfrac{\partial^2}{\partial x^2}\boldsymbol{u} & \text{becomes} & \dfrac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{u} = -\dfrac{\boldsymbol{K}}{h^2}\boldsymbol{u}. \end{array}$$

Time remains continuous. We choose signs so that $\boldsymbol{K}$ is positive definite, corresponding to $-\frac{\mathrm{d}^2}{\mathrm{d}x^2}$. Constant diagonals in $\boldsymbol{K}$ reflect constant coefficients in the differential equations, so $\boldsymbol{K}$ is an $N \times N$ tridiagonal Toeplitz matrix [27, 3, 15]:

$$(1.2) \qquad \boldsymbol{K} = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}, \qquad h = \frac{1}{N+1}.$$

[*] Received by the editors November 5, 2012; accepted for publication (in revised form) July 15, 2013; published electronically August 7, 2014.
http://www.siam.org/journals/sirev/56-3/89757.html

[†] Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139 (gs@math.mit.edu, shev.mac@gmail.com). The work of the first author was supported by the National Science Foundation (grant 1023152 on engineered networks) and by a MathWorks professorship. The work of the second author was supported by a Fulbright Scholarship and an MIT Energy Initiative Fellowship.

Finite differences and linear finite elements both produce this well-loved matrix. Its Cholesky decomposition $\boldsymbol{K} = \boldsymbol{LDL}^T$ is known, with pivots in $\boldsymbol{D}$ and multipliers in $\boldsymbol{L}$. Its eigenvalues and eigenvectors are a central example used in teaching computational science [20] and produce the spectral decomposition $\boldsymbol{K} = \boldsymbol{V\Lambda V}^T$ in (3.2) to (3.4)—the starting point of this paper.

The model problems (1.1) reflect the reality of modern numerical analysis: we often discretize in space and solve ordinary differential equations in time. In solving the time-dependent problem, accuracy can be monitored and increased at will. So the semidiscrete approximations in (1.1) are the crucial steps, and their solutions involve exponentials of matrices:

$$\text{Heat equation:}\quad \boldsymbol{u}(t) = e^{-\boldsymbol{K}t/h^2}\boldsymbol{u}(0);\qquad \text{Wave equation:}\quad \boldsymbol{u}(t) = \cos\left(\sqrt{\boldsymbol{K}}t/h\right)\boldsymbol{u}(0).$$

These equations are easy to write, but we did not know the actual structure of the matrix exponentials. This article concerns that structure.

A small note: Our original study began with $e^{\boldsymbol{A}}$, not $e^{-\boldsymbol{K}}$. We had a simple graph (a line of nodes) with adjacency matrix $\boldsymbol{A} = 2\boldsymbol{I} - \boldsymbol{K}$. Then $e^{\boldsymbol{A}}$ counts the walks between nodes in the graph, weighted by their lengths. ($\boldsymbol{A}^n$ always counts the walks of length n. Longer walks are weighted by $1/n!$ in $e^{\boldsymbol{A}}$.) This measure of *communicability* on a graph was introduced by Estrada, Hatano, Benzi, and Higham [7, 8, 9]. Since $\boldsymbol{A}$ differs from $-\boldsymbol{K}$ only by $2\boldsymbol{I}$, their exponentials differ only by a factor e^2, and we can study both at once.

2. Toeplitz and Hankel Matrices.

A striking feature is that all functions of $\boldsymbol{K}$ are *Toeplitz* plus *Hankel*: the entries are constant along each diagonal plus constant along each antidiagonal. The Toeplitz part dominates, reflecting the shift invariance of the differential equation itself. (With no boundaries, a shift in the initial function $\boldsymbol{u}(0)$ produces the same shift in $\boldsymbol{u}(t)$ at all times.) The Hankel part represents a shift in the opposite direction! That part must be caused by the boundaries, and we will try to explain it.

When the input $(0, 1, 0, 0)$ is shifted forward, you see how constant diagonals in $\boldsymbol{T}$ produce a forward shift in the output vectors. Hankel shifts backwards:

$$\textbf{Toeplitz}\qquad \begin{bmatrix} \mathrm{b} & \mathrm{a} & & \\ \mathrm{c} & \mathrm{b} & \mathrm{a} & \\ & \mathrm{c} & \mathrm{b} & \mathrm{a} \\ & & \mathrm{c} & \mathrm{b} \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathrm{a} & \\ \mathrm{b} & \mathrm{a} \\ \mathrm{c} & \mathrm{b} \\ & \mathrm{c} \end{bmatrix}\qquad \text{forward shift;}$$

$$\textbf{Hankel}\qquad \begin{bmatrix} & & \mathrm{a} & \mathrm{b} \\ & \mathrm{a} & \mathrm{b} & \mathrm{c} \\ \mathrm{a} & \mathrm{b} & \mathrm{c} & \\ \mathrm{b} & \mathrm{c} & & \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} & \mathrm{a} \\ \mathrm{a} & \mathrm{b} \\ \mathrm{b} & \mathrm{c} \\ \mathrm{c} & \end{bmatrix}\qquad \text{backward shift.}$$

Projections onto the eigenvectors of $\boldsymbol{K}$ (discrete sines) produce Toeplitz plus Hankel matrices for all matrix functions $f(\boldsymbol{K})$ (see sections 8, 10, and 11).

We begin (in section 3) with the square root of the discrete Laplacian. The doubly infinite Toeplitz matrix $\sqrt{\boldsymbol{K}_{\infty,\infty}}$ has neat coefficients; its entries must be familiar to some, but they were new to us. For problems on a half line (singly infinite) or an interval, $\sqrt{\boldsymbol{K}_{\infty}}$ and $\sqrt{\boldsymbol{K}}$ have a Hankel part.

The matrix exponential $e^{-\boldsymbol{K}}$ comes next (in section 4). The function e^{-z} is everywhere analytic, and our Bessel approximation is superexponentially accurate.

The Toeplitz part comes from the fundamental solution to the heat equation on an infinite interval. The Hankel part comes from the old idea of placing "image" sources across the endpoints $x = 0$ and $x = 1$ to cancel the actual source at those boundaries (see section 5). Thus the boundary conditions are responsible for the Hankel part.

The wave equation $u_{tt} = u_{xx}$ has solutions $u(x,t) = f(x-t) + g(x+t)$. These waves are Toeplitz; boundary reflections bring in Hankel waves. Their form in the discrete case (see section 6) comes from $e^{i\sqrt{\boldsymbol{K}}t/h}$ or, better, from the real matrix function $\cos(\sqrt{\boldsymbol{K}}t/h)$. The other matrix that appears in this second-order problem (multiplying initial velocity) is $\operatorname{sinc}(\sqrt{\boldsymbol{K}}t/h)$, which might provide a starting point from which to consider *waves on graphs*, a topic that deserves more attention (see [11] for an *edge* Laplacian).

In all these examples, we want to go beyond a line of nodes. The first step is the usual five-point approximation to the two-dimensional Laplacian $-\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}$ on a square grid, which gives the matrix $\mathbb{K}$ of size N^2. We combine $\boldsymbol{K}$ in the x-direction with $\boldsymbol{K}$ in the y-direction, and it is no surprise that $e^{-\mathbb{K}}$ and $\cos(\sqrt{\mathbb{K}})$ are accessible. The more difficult goal, not attempted here, is to treat more general graphs, which have become the fundamental framework for models in discrete applied mathematics.

The *graph Laplacian* matrix (sometimes known as a Kirchhoff matrix) is defined as $\boldsymbol{L} \equiv \boldsymbol{D} - \boldsymbol{A}$. The diagonal matrix $\boldsymbol{D}$ records the *degree* of each node, and $\boldsymbol{A}$ is the *adjacency matrix*. The (i,j) entry of $\boldsymbol{A}$ is 1 if nodes i and j share an edge, and zero otherwise [4]. When the graph is a regular two-dimensional mesh, $\boldsymbol{L}$ is a finite difference approximation to the usual continuum Laplacian [4, 20]. In referring to $\boldsymbol{K}$ or $\mathbb{K}$ as a "graph Laplacian" (see section 9), we are not quite accurate; the correct choice is a singular matrix $\boldsymbol{B}$ or $\mathbb{B}$ with zero row sums:

$$\boldsymbol{B} = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix} = (\text{degree matrix}) - (\text{adjacency matrix}).$$

This corresponds to Neumann boundary conditions $\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}x}(0) = 0$ and $\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}x}(1) = 0$. The eigenvectors $\boldsymbol{v}$ of $\boldsymbol{B}$ are discrete cosines instead of discrete sines. Those vectors retain the crucial property that $\boldsymbol{v}\boldsymbol{v}^T$ is Toeplitz plus Hankel. Therefore, all functions of $\boldsymbol{B}$ have this $\boldsymbol{T} + \boldsymbol{H}$ property. $\boldsymbol{B}$ itself is $\boldsymbol{K} + (\boldsymbol{B} - \boldsymbol{K})$.

The *complete graph* with edges between all pairs of nodes is particularly simple. The Laplacian matrix $\boldsymbol{L}$ has $N-1$ as diagonal entries and -1 everywhere else. Then $\boldsymbol{L}^2$ equals N times $\boldsymbol{L}$, so $\sqrt{\boldsymbol{L}}$ is simply $\boldsymbol{L}/\sqrt{N}$.

Finally, we consider the one-way wave equation $u_t = u_x$ (see section 7) and the following centered first difference matrix $\boldsymbol{F}$:

$$\boldsymbol{F} = \begin{bmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ & & & -1 & 0 \end{bmatrix}.$$

First come the eigenvalues (pure imaginary). Then the eigenvectors lead to "alternating Hankel" matrices that were new to us. The exact solution is $u_0(x+t)$. The Taylor series for that shift is exactly $\exp(t\frac{\mathrm{d}}{\mathrm{d}x})$, but with those alternating signs, the matrix exponential of F shows only weak convergence for the semidiscrete $\boldsymbol{u}_t = \boldsymbol{F}\boldsymbol{u}/2h$. The convergence to a shift operator looks terrible until you test it for a smooth $\boldsymbol{u}_0$.

We have certainly not exhausted the subject of this paper. In the future we hope to combine $\boldsymbol{K}$ with $\boldsymbol{F}$, *diffusion with convection.* Those matrices are tridiagonal and Toeplitz, but this does not ensure that *functions* of $\boldsymbol{K} + c\boldsymbol{F}$ are Toeplitz plus Hankel. (That much stronger requirement depends on all the eigenvectors.)

A big encouragement for the authors—and we hope for the readers—is the elementary nature of the mathematics used here. We are working with known eigenvalues and eigenvectors. Every function $f(\boldsymbol{K})$ has eigenvalues $f(\lambda)$, with the same eigenvectors as $\boldsymbol{K}$. As $N \to \infty$, sums approach integrals, and these matrices are computed exactly. The steps for $\sqrt{\boldsymbol{K}}$ are (3.2) to (3.7), and there are two different limits:

1. *Singly infinite* $f(\boldsymbol{K}_\infty)$ is Toeplitz plus Hankel. Rows and columns are numbered from 1 to ∞. In the continuous problem, the right endpoint $x = 1$ moves to infinity, leaving one boundary point.
2. *Doubly infinite* $f(\boldsymbol{K}_{\infty,\infty})$ is purely Toeplitz. Rows and columns are numbered from $-\infty$ to ∞. The left endpoint $x = 0$ moves to minus infinity, leaving no boundary.

In both cases bandedness is lost and the square root (for example) is full. We can report the numbers that appear on every row of the doubly infinite matrix $\sqrt{\boldsymbol{K}_{\infty,\infty}}$ and go unchanged down all of its diagonals. Each row contains

$$\dots \frac{-1}{5\times 7} \quad \frac{-1}{3\times 5} \quad \frac{-1}{1\times 3} \quad 1 \quad \frac{-1}{1\times 3} \quad \frac{-1}{3\times 5} \quad \frac{-1}{5\times 7} \quad \dots, \quad \text{all multiplied by } \frac{4}{\pi}. \tag{2.1}$$

Surely this matrix has been seen somewhere else. For the finite and singly infinite matrices with boundaries, the same numbers are seen down antidiagonals in the Hankel part. A useful general rule is that circulants grow into doubly infinite Toeplitz matrices, while the properties of $\boldsymbol{T}_N + \boldsymbol{H}_N$ extend to the singly infinite $\boldsymbol{T}_\infty + \boldsymbol{H}_\infty$. This touches on the classic study of connections between infinite matrix limits and differential operators [5, 3].

The analysis of a Toeplitz matrix (with entries t_j on diagonal j) is always connected to its *symbol* $\sum t_j e^{ij\theta}$ [27, 3]. To find $\boldsymbol{T}^{-1}$, $e^{\boldsymbol{T}}$, or $\sqrt{\boldsymbol{T}}$, we work with the symbol. For the $-1, 2, -1$ matrix $\boldsymbol{K}$, the symbol is $2 - 2\cos\theta$. This locates the eigenvalues of $\boldsymbol{K}$ in (3.3) on the interval from 0 to 4. The numbers in (2.1) are the Fourier cosine coefficients of $\sqrt{2 - 2\cos\theta}$.

Functions of banded matrices are often approximately banded in the sense of fast decay away from the diagonal, as in the nice examples of Iserles [16] and of Benzi and Razouk [1]. More generally, Higham explained beautifully the subject of matrix functions [14], and contour integrals are successful in computing functions (including the square root) of a matrix times a vector [26, 13]. We focus on the very special space of Toeplitz plus Hankel matrices. Previous work on these matrices includes connections to Fredholm integral equations, spectral properties, and displacement rank [24, 10, 2].

3. The Square Root of *K*.

If $\boldsymbol{K}$ did not have special eigenvalues and eigenvectors, our computations could not go very far. The matrix corresponds to $-\frac{\mathrm{d}^2}{\mathrm{d}x^2}$ with zero (Dirichlet) boundary conditions at $x = 0$ and $x = 1$. For this differential operator we know that the eigenfunctions are sines:

$$-\frac{\mathrm{d}^2}{\mathrm{d}x^2}\ \sin(k\pi x) = k^2\pi^2\ \sin(k\pi x). \tag{3.1}$$

The unit eigenvectors $\boldsymbol{v}_k$ in the discrete case sample just the first N sine functions at the mesh points $x = h, 2h, \dots, Nh$:

$$(3.2) \qquad \textbf{Eigenvectors} \qquad \boldsymbol{v}_k = \sqrt{\frac{2}{N+1}}(\sin(k\pi h), \sin(2k\pi h), \dots, \sin(Nk\pi h))^T.$$

Since $\boldsymbol{K}$ is symmetric, the $\boldsymbol{v}_k$ are orthogonal. Then $\boldsymbol{K}\boldsymbol{v}_k = \lambda_k \boldsymbol{v}_k$:

$$(3.3) \qquad \textbf{Eigenvalues of } \boldsymbol{K} \qquad \lambda_k = 2 - 2\cos(k\pi h), \qquad k = 1, \dots, N.$$

The matrix $\boldsymbol{K}$ is constructed from its eigenvalues in $\boldsymbol{\Lambda}$ and its eigenvectors in the columns of $\boldsymbol{V}$:

$$(3.4) \qquad \textbf{Spectral theorem} \qquad \boldsymbol{K} \;=\; \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T \;=\; \sum_1^N \lambda_k \boldsymbol{v}_k \boldsymbol{v}_k^T.$$

All linear algebra textbooks present this fundamental theorem. Diagonalization has separated $\boldsymbol{K}$ into a sum of rank-one symmetric matrices $\lambda_k \boldsymbol{v}_k \boldsymbol{v}_k^T$. (Multiplying by $\boldsymbol{v}_j$, orthogonality gives $\boldsymbol{v}_k^T \boldsymbol{v}_j = 0$ except for the jth term. Then $\boldsymbol{K}\boldsymbol{v}_j = \lambda_j \boldsymbol{v}_j$ as required. We also see the singular value decomposition [21, 20, 15], since $\boldsymbol{K}$ is symmetric positive definite.)

Notice that $\boldsymbol{K}^2 = (\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T)(\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T) = \boldsymbol{V}\boldsymbol{\Lambda}^2\boldsymbol{V}^T$. The entries of *any function* $f(\boldsymbol{K})$ come from $\boldsymbol{V} f(\boldsymbol{\Lambda}) \boldsymbol{V}^T$ [14]. As in (3.4) this is a combination of $\boldsymbol{v}_k \boldsymbol{v}_k^T$:

$$\textbf{Matrix function} \qquad f(\boldsymbol{K})_{m,n} = \frac{2}{N+1} \sum_{k=1}^N f(\lambda_k) \sin(mk\pi h) \sin(nk\pi h).$$

The crucial point is that $\sqrt{\boldsymbol{K}}$ has the same eigenvectors $\boldsymbol{v}_k$ with eigenvalues $\sqrt{\lambda_k}$. A half-angle identity gives that square root:

$$(3.5) \qquad \sqrt{\lambda_k} = \sqrt{2 - 2\cos(\theta_k)} = 2\sin\left(\frac{\theta_k}{2}\right) \qquad \text{with} \qquad \theta_k = k\pi h = \frac{k\pi}{N+1}.$$

Now the product of sines yields the splitting we hoped for, into Toeplitz plus Hankel. The whole paper depends on this elementary identity for $\sin A \sin B$:

$$\begin{aligned} \left(\sqrt{\boldsymbol{K}}\right)_{m,n} &= \frac{2}{N+1} \sum_{k=1}^N 2\sin\left(\frac{\theta_k}{2}\right) \sin(m\theta_k) \sin(n\theta_k) \\ (3.6) \qquad &= \frac{2}{N+1} \sum_{k=1}^N \sin\left(\frac{\theta_k}{2}\right) \Big(\cos((\boldsymbol{m}-\boldsymbol{n})\theta_k) - \cos((\boldsymbol{m}+\boldsymbol{n})\theta_k) \Big). \end{aligned}$$

The dependence on $m - n$, which is constant down each diagonal of $\sqrt{\boldsymbol{K}}$, signals Toeplitz. The dependence on $m + n$, which is constant down every antidiagonal, signals Hankel. The sum over N terms is closely approximated, and much improved, when it is replaced by an integral (the limit as $N \to \infty$):

$$(3.7) \qquad \left(\sqrt{\boldsymbol{K}}\right)_{m,n} \approx \frac{2}{\pi} \int_0^\pi \sin\left(\frac{\theta}{2}\right) \Big(\cos((\boldsymbol{m}-\boldsymbol{n})\theta) - \cos((\boldsymbol{m}+\boldsymbol{n})\theta) \Big) \mathrm{d}\theta.$$

Eigenvalues of $\boldsymbol{K}$	$\lambda_k = 2 - 2\cos(k\pi h), \qquad k = 1, \dots, N$
Eigenvectors of $\boldsymbol{K}$	$\boldsymbol{v}_k = \sqrt{\frac{2}{N+1}}(\sin(k\pi h), \sin(2k\pi h), \dots, \sin(Nk\pi h))^T$
Function of $\boldsymbol{K}$	$f(\boldsymbol{K})_{m,n} = \frac{2}{N+1}\sum_{k=1}^{N} f(\lambda_k)\sin(mk\pi h)\sin(nk\pi h)$
Doubly infinite $\sqrt{\boldsymbol{K}_{\infty,\infty}}$	$a_p = 4/\pi(1-p^2)$ in (3.9)
Singly infinite	$\left(\sqrt{\boldsymbol{K}_\infty}\right)_{m,n} = a_{m-n} - a_{m+n}$ in (3.7)
Finite square root	$\left(\sqrt{\boldsymbol{K}}\right)_{m,n} \approx a_{m-n} - a_{m+n}$ with aliasing (3.12)
Doubly infinite $e^{-t\boldsymbol{K}_{\infty,\infty}}$	$b_p = e^{-2t} I_p(2t) =$ modified Bessel in (4.5)
Singly infinite	$(\exp(-t\boldsymbol{K}_\infty))_{m,n} = b_{m-n} - b_{m+n}$
Finite heat equation	$(\exp(-t\boldsymbol{K}))_{m,n} \approx b_{m-n} - b_{m+n}$ with aliasing
Doubly infinite $\cos(\sqrt{\boldsymbol{K}_{\infty,\infty}})$	$c_p = J_{2p}(2) =$ Bessel in (6.6)
Doubly infinite $\operatorname{sinc}(\sqrt{\boldsymbol{K}_{\infty,\infty}})$	$s_p/\pi = s_0/\pi - \sum_{k=1}^{p} J_{2k-1}(2)$ in (6.9)
Finite wave equation	$\cos(\sqrt{\boldsymbol{K}})\,\boldsymbol{u}_0 + \operatorname{sinc}(\sqrt{\boldsymbol{K}})\boldsymbol{v}_0, \quad t = h = 1$ in (6.2)

Now we see that *the crucial numbers are the Fourier cosine coefficients* of $\sin(\frac{\theta}{2})$. So we must consider the periodic even function $f(\theta)$ on $[-\pi, \pi]$:

$$f(\theta) = \left|\sin\left(\frac{\theta}{2}\right)\right|. \tag{3.8}$$

This changes slope from $-\frac{1}{2}$ to $\frac{1}{2}$ at $\theta = 0$. That discontinuity in slope means $1/p^2$ decay in the pth Fourier coefficient [25]. For this half-angle function $f(\theta)$, the integrals in (3.7) are easily computed:

$$\begin{aligned} a_p &= \frac{2}{\pi}\int_0^\pi \sin\left(\frac{\theta}{2}\right)\cos(p\theta)\mathrm{d}\theta \\ &= \frac{1}{\pi}\int_0^\pi \left(\sin\left(\frac{1+2p}{2}\theta\right) + \sin\left(\frac{1-2p}{2}\theta\right)\right)\mathrm{d}\theta \\ &= \frac{2}{\pi}\left(\frac{1}{1+2p} + \frac{1}{1-2p}\right) \;=\; \frac{4}{\pi}\frac{1}{(1-2p)(1+2p)}. \end{aligned} \tag{3.9}$$

The denominators 1×3, 3×5, and 5×7 enter the infinite square root matrix that was anticipated in (2.1). The sum of entries is zero along a row $(\dots, a_1, a_0, a_1, \dots)$ of $\sqrt{\boldsymbol{K}_{\infty,\infty}}$:

$$a_0 + 2a_1 + 2a_2 + \cdots = \frac{4}{\pi}\left(1 + \left(\frac{1}{3} - \frac{1}{1}\right) + \left(\frac{1}{5} - \frac{1}{3}\right) + \cdots\right) = 0. \tag{3.10}$$

The doubly infinite matrix is singular and purely Toeplitz, with the all-ones vector in its nullspace. (The function $f(\theta) = |\sin(\frac{\theta}{2})|$ touches zero at $\theta = 0$, where (3.10) adds up its cosine series.) This corresponds to the fact that, with no boundaries, constant functions are in the nullspace of the second derivative and its positive square root. We are seeing, in one dimension, *the square root* (2.1) *of the discrete Laplacian*.

We cannot expect such perfection in the finite case, for the $N \times N$ matrix $\sqrt{\boldsymbol{K}_N}$. It becomes important to include the Hankel part $\boldsymbol{H}_N$, together with the Toeplitz part $\boldsymbol{T}_N$ from (3.6). The approximate square root (still using the integral in (3.7) rather than the sum) has entries $a_{m-n} - a_{m+n}$. MATLAB confirms that the first rows of $\sqrt{\boldsymbol{K}_\infty} = \boldsymbol{T}_\infty + \boldsymbol{H}_\infty$ are

$$\frac{4}{\pi} \text{ times } \begin{pmatrix} 1 & -\frac{1}{3} & -\frac{1}{15} & -\frac{1}{35} & \cdots \\ -\frac{1}{3} & 1 & -\frac{1}{3} & -\frac{1}{15} & \cdots \end{pmatrix} + \begin{pmatrix} \frac{1}{15} & \frac{1}{35} & \frac{1}{63} & \frac{1}{99} & \cdots \\ \frac{1}{35} & \frac{1}{63} & \frac{1}{99} & \frac{1}{143} & \cdots \end{pmatrix}.$$

One important point: $\sqrt{\boldsymbol{K}_N}$ is symmetric across its main antidiagonal as well as its main diagonal (thus, it is *centrosymmetric*). This is because high-frequency cosines agree with low-frequency cosines at the N sample points θ_k (*aliasing*):

$$\cos\left(\frac{p\pi}{N+1}\right) = \cos\left(\frac{(2N+2-p)\pi}{N+1}\right). \tag{3.11}$$

Then the entries a_{m+n} of the Hankel part of the exact $\sqrt{\boldsymbol{K}_N}$ are reflected across the antidiagonal, where $m + n = N + 1$. The lower frequency gives an integral closer to the sum from 1 to N. Therefore, we choose the Hankel part of the approximation to be

$$\boldsymbol{H}_{m,n} = \begin{cases} -a_{m+n} & \text{if } m + n \le N + 1, \\ -a_{2N+2-m-n} & \text{otherwise.} \end{cases} \tag{3.12}$$

Figure 3.1 shows the largest error among the entries of $\sqrt{\boldsymbol{K}_N}$.

Note 1. The square root of $|\frac{\mathrm{d}^2}{\mathrm{d}x^2}|$ corresponds to multiplication (in transform space) by the absolute value $|\theta|$. Its trigonometric analogue in our discrete case is multiplication by $2|\sin\frac{\theta}{2}|$. Both quantities are compared in Figure 3.2.

Note 2. By remarkable chance, the integral a_0 of $|\sin\frac{\theta}{2}|$ is the first example chosen by Weideman [25] to illustrate the distance from the Riemann sum (the trapezoidal rule) with N terms. The speed of convergence is astonishing when Riemann sums approach the integrals of periodic analytic functions.

For this example, Weideman pointed out that *Mathematica* will give an explicit expression for the sums in (3.6):

$$(\sqrt{\boldsymbol{K}})_{m,n} = \frac{1}{N+1}\left(A_{m-n} - A_{m+n}\right),$$

where

$$A_p = (-1)^{p-1} - \frac{1}{2}\cot\left(\frac{2p-1}{4+4N}\pi\right) + \frac{1}{2}\cot\left(\frac{2p+1}{4+4N}\pi\right).$$

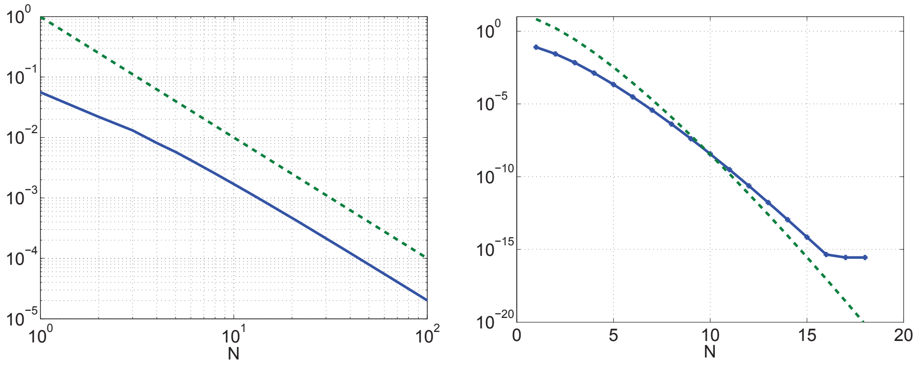

Fig. 3.1 *Left: The approximation of* $\sqrt{\boldsymbol{K}_N}$ *by* $a_{m-n} - a_{m+n}$*, with* a_p *from* (3.9)*, and aliasing* (3.12)*. We observe second-order accuracy (the dotted line has reference slope* -2*). Right: The approximation of* $\exp(-\boldsymbol{K}_N)$ *via* (4.4) *and aliasing. The dotted line is Weideman's estimate of exponential accuracy (!) for a periodic integral* [25]*. We used MATLAB's sqrtm and expm as reference solutions. Both graphs show the largest error among the entries of the matrix.*

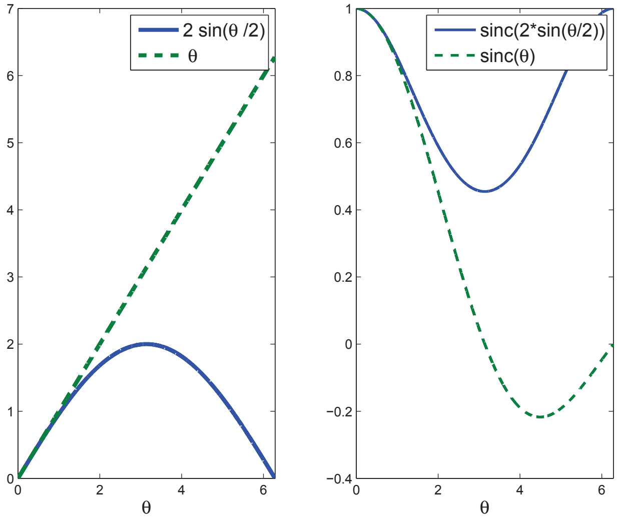

Fig. 3.2 *For small* θ*,* $2\sin(\theta/2)$ *and* $\operatorname{sinc}(2\sin(\theta/2))$ *are close to* θ *and* $\operatorname{sinc}(\theta)$*.*

Note 3. The valuable paper [23] (in this issue!) by Trefethen and Weideman gives a clear picture of the analysis. They credit Poisson as the first to study very fast convergence of numerical integration for analytic functions. (The exponential $e^{-\boldsymbol{K}}$ will be our analytic example.) The simplest approximation says that the difference between the sum and the integral comes from all the aliasing terms $a_N, a_{2N}, \ldots$ that are captured exactly by the sum and are absent in the integral for a_0:

$$\text{sum} - \text{integral} = a_N + a_{2N} + a_{3N} + \cdots .$$

For $e^{-\boldsymbol{K}}$ those aliasing terms drop off like $(cN)^{-N}$ (this is better than exponential in

N) and the convergence is very fast. For $\sqrt{\boldsymbol{K}}$ we see only the ordinary trapezoidal convergence rate N^{-2}, because $f(\theta)$ is not analytic.

Note 4. On a square grid in two dimensions, the Kronecker sum

$$\mathbb{K} = \boldsymbol{K} \oplus \boldsymbol{K} = (\boldsymbol{K} \otimes \boldsymbol{I}) + (\boldsymbol{I} \otimes \boldsymbol{K})$$

becomes the usual five-point approximation to the Laplacian $-\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}$. Here $\otimes$ is the Kronecker product, producing matrices of size N^2 from $\boldsymbol{K}$ and $\boldsymbol{I}$. For $\mathbb{K}$, the components of the N^2 unit eigenvectors $\boldsymbol{v}_{k,l}$ are products of sines:

$$(\boldsymbol{v}_{k,l})_{m,n} = \frac{2}{N+1} \sin(km\pi h) \sin(ln\pi h) \quad \text{with} \quad k, l, m, n = 1, \dots, N. \tag{3.13}$$

The eigenvalues now involve two angles θ and ϕ:

$$\lambda_{k,l} = (2 - 2\cos\theta_k) + (2 - 2\cos\phi_l) = 4\sin^2\left(\frac{\theta_k}{2}\right) + 4\sin^2\left(\frac{\phi_l}{2}\right). \tag{3.14}$$

The same steps, (3.5) to (3.8), lead us to a function of two variables for $\sqrt{\mathbb{K}}$,

$$f(\theta, \phi) = \left(\sin^2\left(\frac{\theta}{2}\right) + \sin^2\left(\frac{\phi}{2}\right)\right)^{1/2}, \tag{3.15}$$

with singularity at $\theta = \phi = 0$. Its Fourier cosine coefficients a_{pq}, which will multiply $(\cos p\theta)(\cos q\phi)$, have no form as simple as $c/(1 - 4p^2)$.

Note 5. A nice property of Kronecker forms is that the exponential of a Kronecker sum $\boldsymbol{A} \oplus \boldsymbol{B}$ is the Kronecker product $e^{\boldsymbol{A}} \otimes e^{\boldsymbol{B}}$. Also, Kronecker products are bilinear. In our example with $\mathbb{K} = \boldsymbol{K} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{K}$,

$$e^{-\mathbb{K}} = e^{-\boldsymbol{K}} \otimes e^{-\boldsymbol{K}} = (\boldsymbol{T} + \boldsymbol{H}) \otimes (\boldsymbol{T} + \boldsymbol{H}) = \boldsymbol{T} \otimes \boldsymbol{T} + \boldsymbol{T} \otimes \boldsymbol{H} + \boldsymbol{H} \otimes \boldsymbol{T} + \boldsymbol{H} \otimes \boldsymbol{H}.$$

We see the Toeplitz ($\boldsymbol{T}$) and Hankel ($\boldsymbol{H}$) exponentials again. This connection gives us an easy, accurate approximation to $e^{-\mathbb{K}}$ from our approximation to $e^{-\boldsymbol{K}}$.

4. The Heat Equation and $\mathbf{e^{-tK}}$.

Once again take the eigenvalues $\lambda_k = 2 - 2\cos\theta_k$ and the eigenvectors $\boldsymbol{v}_k = \sqrt{\frac{2}{N+1}}(\sin\theta_k, \sin 2\theta_k, \dots, \sin N\theta_k)^T$; then $e^{-t\boldsymbol{K}}$ has the same eigenvectors with eigenvalues $e^{-t\lambda_k}$. The entries of the matrix exponential (the heat kernel) are

$$(e^{-t\boldsymbol{K}})_{m,n} = \frac{2}{N+1} \sum_{k=1}^{N} e^{2t\cos\theta_k - 2t} \sin(m\theta_k)\sin(n\theta_k). \tag{4.1}$$

Again we use the identity $\sin A \sin B = \frac{1}{2}(\cos(A - B) - \cos(A + B))$:

$$(e^{-t\boldsymbol{K}})_{m,n} = \frac{e^{-2t}}{N+1} \sum_{k=1}^{N} e^{2t\cos\theta_k} \Bigg(\cos((m-n)\theta_k) - \cos((m+n)\theta_k) \Bigg). \tag{4.2}$$

Those sums are very closely approximated by integrals [25, 23]. For $p = m - n$ and $p = m + n$ the limits as $N \to \infty$ are

$$b_p = \frac{e^{-2t}}{\pi} \int_0^{\pi} e^{2t\cos\theta} \cos(p\theta) \, \mathrm{d}\theta. \tag{4.3}$$

The infinite matrix $\exp(-t\boldsymbol{K}_\infty)$ has entries $b_{m-n} - b_{m+n}$ and is symmetric positive definite. With $t = 1$, the entries along the first rows of $\boldsymbol{T} + \boldsymbol{H}$ are

$$\begin{pmatrix} 0.3085 & 0.2153 & 0.0932 & \dots \\ 0.2153 & 0.3085 & 0.2153 & \dots \end{pmatrix} - \begin{pmatrix} 0.0932 & 0.0288 & 0.0069 & \dots \\ 0.0288 & 0.0069 & 0.0013 & \dots \end{pmatrix}.$$

For finite N, cosines obey the aliasing equation (3.11). As before, the sums in (4.2) are the same for $p = m + n$ as for $P = (2N + 2) - (m + n)$. The true $e^{-t\boldsymbol{K}}$ is centrosymmetric. We choose the smaller of p and P, so the Hankel part of the approximation is symmetric across the main antidiagonal $m + n = N + 1$:

$$e^{-t\boldsymbol{K}} \approx \boldsymbol{T} + \boldsymbol{H} \text{ with entries } b_{m-n} - b_{m+n} \qquad \text{for } m + n \le N + 1. \tag{4.4}$$

This example has one more beautiful feature. The integral in (4.3) is the celebrated representation of a *modified Bessel function* of the first kind I_p [16, 25, 6], with integer p. So those integrals are exactly

$$b_p = e^{-2t} I_p(2t). \tag{4.5}$$

For $e^{-\boldsymbol{K}t/h^2}$, t in (4.5) changes to t/h^2.

A remarkable point about the approximation is that the role of x (on the spatial mesh) is played by the *order p of the Bessel function!* We don't ordinarily consider finite differences with respect to p. But we do use recursion formulas.

The same steps apply to the model Schrödinger equation $u_t = -iu_{xx}$ and its semidiscrete approximation $u_t = i\boldsymbol{K}u/h^2$. That factor i changes the entries in $\boldsymbol{T} + \boldsymbol{H}$ from modified Bessel I_p to the ordinary Bessel coefficients J_p in (6.6). The second-order accuracy from our tridiagonal $\boldsymbol{K}$ stands in contrast to the spectral (infinite-order) accuracy achievable [22, 19] with full matrices.

5. Images and Hankel Matrices.

We want to understand the unexpected appearance of Hankel matrices, which produce the very opposite of shift invariance. If the entries of a vector $\boldsymbol{x}$ shift down, then the entries of $\boldsymbol{Hx}$ shift up:

$$\text{Hankel} \quad \begin{bmatrix} & & \mathrm{a} & \mathrm{b} \\ & \mathrm{a} & \mathrm{b} & \mathrm{c} \\ \mathrm{a} & \mathrm{b} & \mathrm{c} & \\ \mathrm{b} & \mathrm{c} & & \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} & \mathrm{a} \\ \mathrm{a} & \mathrm{b} \\ \mathrm{b} & \mathrm{c} \\ \mathrm{c} & \end{bmatrix} \quad \text{backward shift.}$$

The first clue is that this effect must come from the boundary conditions. An initial value problem on the whole line will be shift-invariant and purely Toeplitz if the coefficients in the differential or difference equation are constant.

Recall the trick of image sources for the heat equation [12], where they are placed to cancel the effect of the original sources at the boundary. Suppose the original is a point source $u_0(x) = \delta(x - a)$ at $x = a$. If there is just one boundary, at $x = 0$, place an image source $-\delta(x + a)$ at $x = -a$. When we solve the heat equation on the whole line starting from $u(0) = \delta(x - a) - \delta(x + a)$, the solution remains zero at $x = 0$ (by symmetry):

$$u(x,t) = \frac{1}{\sqrt{4\pi t}}\left(e^{-(x-a)^2/4t} - e^{-(x+a)^2/4t}\right) \qquad \text{and } u(0,t) = 0. \tag{5.1}$$

The second exponential from the image source is *anti-shift-invariant.* When the source point $x = a$ moves to the right, its image point $x = -a$ moves to the left. This accounts for Hankel.

For problems on a finite interval, we need infinitely many image points. Across $x = 0$, an image at $x = -a$ balances the original source at $x = a$. But then both the source and that image have to be balanced across $x = 1$ (by images at $2 - a$ and at $2 + a$). In the end, we extend $u_0 = \delta(x - a)$ to a sequence of + and − images with period 2 on the whole line:

$$U_0(x) = \sum_{-\infty}^{\infty} \delta(x + 2n - a) - \sum_{-\infty}^{\infty} \delta(x - 2n + a) \qquad \text{for} \qquad \infty < x < \infty. \tag{5.2}$$

Now we solve the heat equation $U_t = U_{xx}$ starting from this $U_0(x)$. At time zero, $U = u = \delta(x - a)$ on the interval $0 \le x \le 1$. At all times the boundary conditions $U(0, t) = 0$ and $U(1, t) = 0$ are satisfied, from the symmetry of sources and image sources. Across $x = 0$, each point $x = 2n + a$ has an image point $x = -2n - a$; across $x = 1$ we see the image point $x = 2 - (2n + a)$.

If we shift the original sources by changing a, the negative image sources at $2n - a$ move in the *opposite direction*. This gives the Hankel part of the (continuous) solution operator.

A corresponding discrete theory for a point source $u_0 = \delta$ explains the Hankel part of the finite matrix $\exp(-\boldsymbol{K})$. When δ has $N - 1$ zero components and 1 in position j, its extension $\boldsymbol{u}$ needs an image -1 in position $-j$. By analogy with $U_0(x)$ in (5.2), extend $\boldsymbol{u}$ to have period $2(N + 1)$ on the whole discrete line:

$$\boldsymbol{u}_k = \begin{cases} 1 & \text{for } k = n(2N + 2) + j, \\ -1 & \text{for } k = n(2N + 2) - j. \end{cases} \tag{5.3}$$

Certainly, $\boldsymbol{K}_{\infty,\infty}\boldsymbol{u}$ (with infinite Toeplitz matrix) agrees with the N components of $\boldsymbol{K}\delta$. Moreover, $\exp(-\boldsymbol{K}_{\infty,\infty})\boldsymbol{u}$ agrees with $\exp(-\boldsymbol{K})\delta$ and satisfies the boundary conditions of zero at positions 0 and $N + 1$.

The Hankel part of the finite matrix corresponds to the negative images in $\boldsymbol{u}$. When j changes, the initial vector δ moves one way and the images move the other way—the opposite of shift invariance.

6. The Wave Equation.

For the wave equation $u_{tt} = u_{xx}$, eigenvalues are imaginary. Energy is conserved, not lost. The eigenfunctions with boundary conditions $u(0, t) = u(1, t) = 0$ are still $\sin k\pi x$. But the eigenvalues change from $-k^2\pi^2$ for the heat equation to $\pm ik\pi$ for the wave equation.

Starting from $u_0(x)$ with velocity $u_t = v_0(x)$, d'Alembert found left and right waves coming from u_0 and a spreading wave coming from v_0:

$$u(x, t) = \frac{1}{2}(u_0(x + t) + u_0(x - t)) + \frac{1}{2}\int_{x-t}^{x+t} v_0(s)\, \mathrm{d}s. \tag{6.1}$$

Compare with the solution to the semidiscrete equation $u_{tt} = -\boldsymbol{K}u/h^2$:

$$u(t) = \cos(\sqrt{\boldsymbol{K}}t/h)\, u_0 + h\boldsymbol{K}^{-1/2} \sin(\sqrt{\boldsymbol{K}}t/h)\, v_0. \tag{6.2}$$

The matrix cosine and sine come from the matrix exponentials $\exp(\pm i\sqrt{\boldsymbol{K}}t/h)$. The form (6.2) separates the waves evolving from u_0 and v_0. There are no fractional powers of $\boldsymbol{K}$ because of the factor $\boldsymbol{K}^{-1/2}$, which corresponds to the integration in (6.1) just as $\sqrt{\boldsymbol{K}}$ corresponds to $|\frac{\mathrm{d}}{\mathrm{d}x}|$. In (6.3) below, this $\boldsymbol{K}^{-1/2}$ factor moves us from $\mathbf{sin}(\sqrt{\boldsymbol{K}})$ to $\mathbf{sinc}(\sqrt{\boldsymbol{K}})$.

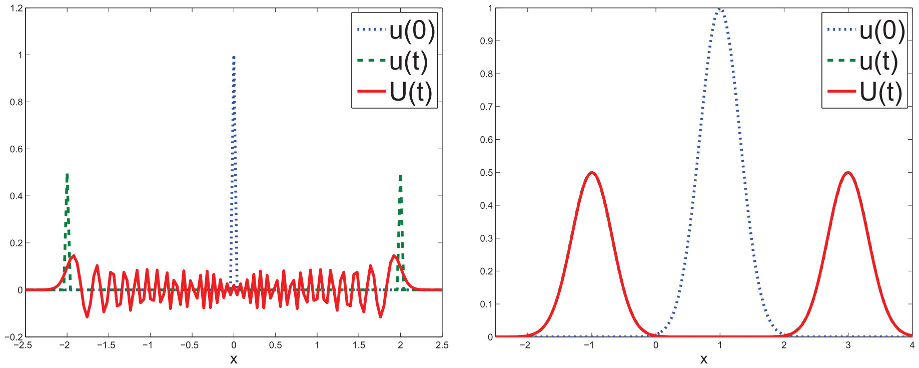

Fig. 6.1 *Waves spreading left and right from $u(0)$ have large oscillations in U when $u(0) = \delta$. The error $|U - u|$ is too small to see when $u(0)$ is smooth.*

To find good approximations to $\cos(\sqrt{\boldsymbol{K}}t/h)$ and $\boldsymbol{K}^{-1/2}\sin(\sqrt{\boldsymbol{K}}t/h)$, we start with $t/h = 1$. The cosine and sinc matrices still have the eigenvectors $\boldsymbol{v}_k$:

$$\cos(\sqrt{\boldsymbol{K}}) = \sum_{k=1}^{N} \cos(\sqrt{\lambda_k})\, \boldsymbol{v}_k \boldsymbol{v}_k^T \qquad \text{and}$$

$$\text{(6.3)} \qquad \operatorname{sinc}(\sqrt{\boldsymbol{K}}) = \boldsymbol{K}^{-1/2}\sin(\sqrt{\boldsymbol{K}}) = \sum_{k=1}^{N} \frac{\sin(\sqrt{\lambda_k})}{\sqrt{\lambda_k}}\, \boldsymbol{v}_k \boldsymbol{v}_k^T .$$

Recall from (3.5) that $\sqrt{\lambda_k} = 2\sin(\frac{\theta_k}{2})$ with $\theta_k = \frac{k\pi}{N+1}$. Then the entries in the matrix cosine become

(6.4)

$$\cos(\sqrt{\boldsymbol{K}})_{m,n} = \frac{1}{N+1}\sum_{k=1}^{N} \cos\left(2\sin\left(\frac{\theta_k}{2}\right)\right)\Big(\cos((m-n)\theta_k) - \cos((m+n)\theta_k)\Big).$$

This sum is exponentially close to an integral because $\cos(2\sin(\frac{\theta}{2}))$ is real analytic. We recognize the Toeplitz and Hankel parts as Fourier cosine coefficients:

$$\text{(6.5)} \quad \cos(\sqrt{\boldsymbol{K}})_{m,n} \approx \frac{1}{\pi}\int_0^{\pi} \cos\left(2\sin\left(\frac{\theta}{2}\right)\right)\Big(\cos((m-n)\theta) - \cos((m+n)\theta)\Big)\,\mathrm{d}\theta.$$

The integrals produce Bessel function values $J_{2p}(2)$ [6] as entries in $\cos(\sqrt{\boldsymbol{K}_{\infty,\infty}})$:

$$\text{(6.6)} \qquad \textbf{Cosine coefficients} \qquad c_p = \frac{1}{\pi}\int_0^{\pi} \cos\left(2\sin\left(\frac{\theta}{2}\right)\right)\cos(p\theta)\,\mathrm{d}\theta = J_{2p}(2).$$

For the matrix $\cos(\sqrt{\boldsymbol{K}}t/h)$ in (6.2), the entries of $\boldsymbol{T}$ and $\boldsymbol{H}$ become $J_{2p}(2t/h)$.

Unlike the (parabolic) heat equation, the (hyperbolic) wave equation is not smoothing. Figure 6.1 shows an initial spike and a Gaussian, both splitting into left and right waves $(u_0(x+t)+u_0(x-t))/2$. For the difference equation, U stays close to the Gaussians, but it oscillates badly for the spikes. Figure 6.2 shows the corresponding results from initial velocities $v_0(x)$.

Now we turn to the sinc matrix that multiplies the initial velocity v_0 in (6.2). To see why the sinc function $(\sin\theta)/\theta$ appears, look at the final term in the exact

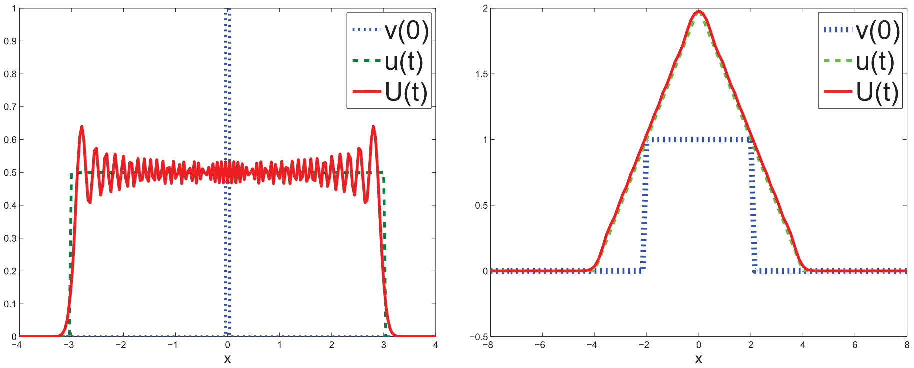

Fig. 6.2 *Oscillations when $v(0) = \delta(x)$ and the exact $u(t)$ is a box function. Extra smoothness and accuracy when $v(0) =$ box and $u(t) =$ box $*$ box $=$ hat function.*

solution (6.1). That term is a convolution of v_0 with a box function, equal to 1 on $[-t, +t]$. In Fourier space this is multiplication by the transform of the box function, and that Fourier transform is exactly $(\sin\theta t)/\theta = t\,\mathrm{sinc}(\theta t)$, a sinc function.

The discrete case involves $\mathrm{sinc}(\sqrt{\lambda}) = \mathrm{sinc}(2\sin(\theta/2))$. We want to find its Fourier cosine coefficients, the integrals s_p in (6.8). Those integrals look alarming at first, but they arise naturally so there must be some hope. You will see next that Hung Cheng has found a way.

To evaluate $\mathrm{sinc}(\boldsymbol{K}^{1/2}) = \boldsymbol{K}^{-1/2}\sin(\boldsymbol{K}^{1/2}) = \boldsymbol{I} - \boldsymbol{K}/3! + \boldsymbol{K}^2/5! - \cdots$, the eigenvectors $\boldsymbol{v}_k$ are more useful than this infinite series. From the spectral theorem (6.3),

$$\mathrm{sinc}(\sqrt{\boldsymbol{K}_N}) = \sum_{k=1}^{N} \mathrm{sinc}(\sqrt{\lambda_k})\; \boldsymbol{v}_k \boldsymbol{v}_k^T. \tag{6.7}$$

As in (3.10) the m, n entry of this matrix is approximately $(s_{m-n} - s_{m+n})/\pi$ with

$$s_p = \int_0^\pi \mathrm{sinc}\left(2\sin\left(\frac{\theta}{2}\right)\right)\cos(p\theta)\,\mathrm{d}\theta. \tag{6.8}$$

Cheng's Lemma. *Each integral s_p comes from s_0 and p values of Bessel functions at $x = 2$:*

$$s_p = s_0 - \pi\sum_{k=1}^{p} J_{2k-1}(2) \;=\; \pi\sum_{k=p+1}^{\infty} J_{2k-1}(2). \tag{6.9}$$

The first four integrals are $s_0 \approx 2.2396$, $s_1 \approx 0.4278$, $s_2 \approx 0.0227$, and $s_3 \approx 0.0006$.

Simplified proof. It was a key insight of Hung Cheng (in private correspondence) to compute *differences* of the integrals s_p. He worked with $s_p - s_0$, and we noticed that $s_p - s_{p-1}$ is even simpler:

$$s_p - s_{p-1} = \int_0^\pi \mathrm{sinc}\left(2\sin\frac{\theta}{2}\right)(\cos p\theta - \cos(p-1)\theta)\,\mathrm{d}\theta. \tag{6.10}$$

The integrand is an even function of period 2π. So (6.10) equals half of the integral from $-\pi$ to π, and one quarter of the integral from -2π to 2π. Now write $\theta = 2\phi$ and

$\mathrm{d}\theta = 2\mathrm{d}\phi$ to shrink back to $-\pi \le \phi \le \pi$:

$$s_p - s_{p-1} = \frac{1}{2}\int_{-\pi}^{\pi} \frac{\sin(2\sin\phi)}{2\sin\phi}\,(\cos 2p\phi - \cos((2p-2)\phi))\,\mathrm{d}\phi. \tag{6.11}$$

This difference of cosines is $-2\sin\phi\,\sin((2p-1)\phi)$. After canceling $2\sin\phi$ (this is the nice step), we are left with

$$s_p - s_{p-1} = -\frac{1}{2}\int_{-\pi}^{\pi} \sin(2\sin\phi)\sin((2p-1)\phi)\mathrm{d}\phi \quad = \quad -\pi J_{2p-1}(2). \tag{6.12}$$

Then $(s_p - s_{p-1}) + (s_{p-1} - s_{p-2}) + \cdots + (s_1 - s_0)$ produces $s_p - s_0$ in (6.9). We learned from *Mathematica* that these numbers are related to Struve functions. □

Signal Speed. An important feature of the wave equation $u_{tt} = u_{xx}$ is that the signal speed is finite. No information about the initial values at X reaches the point x before the time $t = |x - X|$. This is apparent from the d'Alembert solution (6.1). Equivalently, the solution at x, t depends only on initial values in the interval from $x - t$ to $x + t$.

How is this property reflected in the semidiscrete wave equation $u_{tt} = -\boldsymbol{K}u/h^2$? Not exactly. The infinite Toeplitz matrices $\boldsymbol{C}_\infty = \cos(\sqrt{\boldsymbol{K}_\infty})$ and $\boldsymbol{S}_\infty = \mathrm{sinc}(\sqrt{\boldsymbol{K}_\infty})$ are *not banded*. This would be expecting too much. What we do expect is rapid decay beyond the appropriate diagonals. Since the matrices appear at time h (when $t/h = 1$ in the exponentials of $\pm i\sqrt{\boldsymbol{K}}t/h$), the full matrices should be "morally tridiagonal."

Notice that a fully discrete approximation (if it is explicit) does have finite signal speed. That speed depends on the Courant–Friedrichs–Lewy (CFL) number $r = \Delta t/\Delta x$. The simplest finite difference approximation to $u_{tt} = u_{xx}$ would use $\boldsymbol{K}$ and also $\boldsymbol{K}^*$, the second difference matrix in time:

$$-\boldsymbol{K}^*U(x,t) = U(x, t + \Delta t) - 2U(x,t) + U(x, t - \Delta t).$$

Working backwards in $\boldsymbol{K}^*U/(\Delta t)^2 = \boldsymbol{K}U/(\Delta x)^2$, each value of $U(x,t)$ comes from earlier values on and inside a triangle with vertex at (x,t). The sides of the triangle have slopes $\pm 1/r$. Therefore, $U(x,t)$ uses only initial values $u_0(x)$ and $v_0(x)$ between $x - t/r$ and $x + t/r$. The fully discrete problem deals strictly with banded matrices.

The consistency of the difference equation does not guarantee convergence of U to u. If $\Delta t > \Delta x$ so that $r > 1$, then U is using initial values on a smaller interval than $[x-t, x+t]$ for the true solution. Without using the needed information from u_0 and v_0, the approximation U cannot converge to u. The CFL requirement is $r \le 1$.

7. First Differences and One-Way Waves.

At this point we abandon $\boldsymbol{K}$ and look at first differences. Our equation becomes $u_t = u_x$, whose solution is $u_0(x+t)$, a pure translation of the initial function $u_0(x)$. If there is a left endpoint $x = 0$, we don't want a boundary condition there: the wave is arriving and the solution $u_0(t)$ at that point is already known. If there is a right endpoint $x = 1$, then an inflow condition is appropriate. When the condition is $u(1,t) = 0$ so that nothing enters, the solution $u(x,t)$ is zero for $x + t \ge 1$.

In the semidiscrete problem, one-sided differences are a natural choice:

$$\boldsymbol{\Delta} = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \\ & & & & -1 \end{bmatrix} \qquad \text{and} \qquad u_t = \frac{\boldsymbol{\Delta}}{h}u.$$

The forward difference matrix $\boldsymbol{\Delta}$ is upper triangular, looking on the "upwind" side of each mesh point for information. (Looking downwind with backward differences $-\boldsymbol{\Delta}^T$ would be a disaster, as all waves are coming from the right side.) The missing 1 in the last row of $\boldsymbol{\Delta}$ correctly reflects the boundary condition $u = 0$ at the right endpoint $x = 1$.

The matrix $\boldsymbol{\Delta}$ is Toeplitz and so are all its powers. We think of $\boldsymbol{\Delta}$ as $\boldsymbol{S} - \boldsymbol{I}$, where the upper triangular shift has $\boldsymbol{S}^N = 0$. The entries on the jth diagonal of $\boldsymbol{\Delta}^n = (\boldsymbol{S} - \boldsymbol{I})^n$ are binomial coefficients $\binom{n}{j}$ times $(-1)^{n+j}$. Then the entries on the jth diagonal of the matrix exponential $e^{\boldsymbol{\Delta}}$ are

$$(e^{\boldsymbol{\Delta}})_{i,i+j} = \sum_{n=0}^{\infty} \frac{(-1)^{n+j}}{n!}\binom{n}{j} = \frac{1}{e}\frac{1}{j!} \qquad \text{for } j = 0, \dots, N-1. \tag{7.1}$$

The infinite matrix $\exp(\boldsymbol{\Delta}_\infty)$ would have these Poisson probabilities on all diagonals $j \geq 0$.

Notice that we do *not* use the eigenvectors of $\boldsymbol{\Delta}$ to study $e^{\boldsymbol{\Delta}}$. The space of eigenvectors is only one-dimensional. The matrix is not diagonalizable. The matrix $\boldsymbol{\Delta}$ happens to be already in Jordan form, with a single Jordan block, and its only eigenvector is $(1, 0, \dots, 0)$ with $\lambda = -1$. We need to approximate $u_t = u_x$ by another difference matrix to make life interesting again.

A more accurate approximation of $\partial u/\partial x$, and a more exciting choice of the difference matrix, comes from *centered first differences*:

$$\boldsymbol{F} = \begin{bmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ & & & -1 & 0 \end{bmatrix} \qquad \text{and} \qquad u_t = \frac{\boldsymbol{F}}{2h}\, u.$$

This matrix is still Toeplitz, but its powers will not have constant diagonals. The effect of the boundary rows in $\boldsymbol{F}$ will move into the interior of $\boldsymbol{F}^n$, one step at a time. Consistent with the rest of this paper, we want an approximation for $e^{\boldsymbol{F}}$. Since $\boldsymbol{F}$ is antisymmetric, its eigenvalues will be imaginary and its eigenvectors will be orthogonal. Introduce the diagonal matrices $\boldsymbol{D} = \operatorname{diag}(i, i^2, \dots, i^N)$ and $\boldsymbol{D}^{-1} = \bar{\boldsymbol{D}}$. Then $\boldsymbol{D}^{-1}\boldsymbol{F}\boldsymbol{D}$ multiplies row m by $(-i)^m$ and column n by i^n:

$$\boldsymbol{G} = \boldsymbol{D}^{-1}\boldsymbol{F}\boldsymbol{D} = i \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ & & & 1 & 0 \end{bmatrix}.$$

The eigenvalues of $\boldsymbol{G}$ are already known to be i times $2\cos\theta_k$, with $\theta_k = k\pi/(N+1) = k\pi h$ as before. These are also the eigenvalues μ_k of $\boldsymbol{F}$ since the two matrices are similar.

The eigenvectors of $\boldsymbol{G}$ are the same discrete sine vectors $\boldsymbol{v}_k$ that have appeared throughout this paper. The eigenvectors of $\boldsymbol{F}$ are then $\boldsymbol{w}_k = \boldsymbol{D}\boldsymbol{v}_k$:

$$\boldsymbol{D}^{-1}\boldsymbol{F}\boldsymbol{D}\boldsymbol{v}_k = \mu_k \boldsymbol{v}_k \;\text{ means that }\; \boldsymbol{F}\boldsymbol{D}\boldsymbol{v}_k = \mu_k \boldsymbol{D}\boldsymbol{v}_k \;\text{ and }\; \boldsymbol{F}\boldsymbol{w}_k = \mu_k \boldsymbol{w}_k.$$

Explicitly, the eigenvectors of $\boldsymbol{F}$ are $\boldsymbol{w}_k = (i\sin\theta_k, i^2\sin 2\theta_k, \ldots, i^N \sin N\theta_k)^T$, to be normalized so that $||\boldsymbol{w}_k||^2 = \overline{\boldsymbol{w}}_k^T \boldsymbol{w}_k = 1$. Then the rank-one projection matrices $\boldsymbol{w}_k\overline{\boldsymbol{w}}_k^T$ onto these unit eigenvectors are almost Toeplitz plus Hankel, but the "Hankel" part has become "alternating Hankel," which was completely new to us:

$$\begin{aligned}(\boldsymbol{w}_k\overline{\boldsymbol{w}}_k^T)_{m,n} &= \frac{2}{N+1} i^{m-n}\sin(m\theta_k)\sin(n\theta_k)\\ &= \frac{1}{N+1} i^{m-n}\Big(\cos((m-n)\theta_k) - \cos((m+n)\theta_k)\Big).\end{aligned} \tag{7.2}$$

For each $\boldsymbol{w}_k\overline{\boldsymbol{w}}_k^T$, the Toeplitz part has constant entries $f_p = (i^p \cos p\theta_k)/(N+1)$ on diagonal $p = m-n$. The alternating Hankel part has constant absolute value, but the factor $i^{m-n} = i^{m+n}i^{-2n}$ produces alternating signs down each antidiagonal $p = m+n$:

$$\frac{-1}{N+1} i^{m+n} i^{-2n} \cos((m+n)\theta_k) \quad = \quad (-1)^{n+1} f_p.$$

These projections $\boldsymbol{w}_k\overline{\boldsymbol{w}}_k^T$ are "Toeplitz plus alternating Hankel," and so is $e^{\boldsymbol{F}}$:

$$\textbf{Exponential of F} \qquad e^{\boldsymbol{F}} = \sum_1^N e^{2i\cos(\theta_k)}\, \boldsymbol{w}_k\overline{\boldsymbol{w}}_k^T. \tag{7.3}$$

The eigenvalues have $|e^{\mu_k}| = 1$, so that $e^{\boldsymbol{F}}$ is a unitary matrix. (Exponentials of antisymmetric matrices are unitary, just as $|e^{i\theta}| = 1$.) This sum is close to an integral, also of the form $\boldsymbol{T} + \boldsymbol{AH}$:

(7.4)
$$(e^{\boldsymbol{F}})_{m,n} \approx \frac{1}{\pi}\int_0^\pi e^{2i\cos(\theta)}\Big(i^{m-n}\cos((m-n)\theta) \;-\; (-1)^n i^{m+n}\cos((m+n)\theta)\Big)\,\mathrm{d}\theta.$$

The alternating signs down each antidiagonal of $\boldsymbol{AH}$ must reflect the fact that the original $\boldsymbol{F}$ uncouples even indices from odd indices. The entries along the diagonal $m-n = p$ of $\boldsymbol{T}$ and the antidiagonal $m+n = p$ of $\boldsymbol{AH}$ are t_p and $(-1)^{n+1}t_p$:

$$t_p = \frac{i^p}{\pi}\int_0^\pi e^{2i\cos(\theta)}\cos(p\theta)\,\mathrm{d}\theta = (-1)^p J_p(2). \tag{7.5}$$

For $e^{\boldsymbol{F}t/2h}$ the Bessel functions J_p are evaluated at $2t/2h$. *The space variable always turns up in the order p.*

Recall that the exact solution to $u_t = u_x$ is $u_0(x+t)$. With boundary condition $u = 0$ at $x = 1$, the solution is zero for $x + t \geq 1$. Then the solution operator at time $t = 2h$ is simply a shift by $2h$. The exponential $e^{\boldsymbol{F}}$ approximates that shift, since it is the solution operator at $t = 2h$ in the semidiscrete problem $\boldsymbol{u}_t = \boldsymbol{F}\boldsymbol{u}/2h$.

How close is $e^{\boldsymbol{F}}$ to a perfect shift? Not close, if the initial function $\boldsymbol{u}_0$ is a point source. In that case we are looking at a single column of $e^{\boldsymbol{F}}$, which shows oscillating entries. But if the initial function $\boldsymbol{u}_0$ is a Gaussian (therefore much smoother), then this is translated in a coherent way. Figures 6.1 and 6.2 show similar phenomena for the two-way wave equation.

We will leave untouched the classical question of the convergence of the discretized operator to the true solution operator. Before closing, we include some quick tests to determine whether a matrix has either of special forms $\boldsymbol{T} + \boldsymbol{H}$ or $\boldsymbol{T} + \boldsymbol{AH}$.

8. Toeplitz Plus Hankel Matrices.

Matrices of this form fill a subspace TH of the N^2-dimensional space of $N \times N$ matrices. What is the dimension of that subspace, and what tests on a matrix $\boldsymbol{M}$ will confirm that it can be written as $\boldsymbol{M} = \boldsymbol{T} + \boldsymbol{H}$? We quickly recall the answers presented in [2]:

$$\begin{aligned} &\boldsymbol{M} = \boldsymbol{T} + \boldsymbol{H} \quad \text{if and only if} \\ &\boldsymbol{M}_{i-1,j} + \boldsymbol{M}_{i+1,j} = \boldsymbol{M}_{i,j-1} + \boldsymbol{M}_{i,j+1} \quad \text{for } 1 < i < N,\ 1 < j < N. \end{aligned} \tag{8.1}$$

A Toeplitz matrix satisfies those $(N-2)^2$ conditions because $\boldsymbol{T}_{i-1,j} = \boldsymbol{T}_{i,j+1}$ and $\boldsymbol{T}_{i,j-1} = \boldsymbol{T}_{i+1,j}$. A Hankel matrix satisfies the same conditions because $\boldsymbol{H}_{i-1,j} = \boldsymbol{H}_{i,j-1}$ and $\boldsymbol{H}_{i,j+1} = \boldsymbol{H}_{i+1,j}$. Therefore, the sum satisfies the conditions (8.1) and they produce a subspace TH of dimension $N^2 - (N-2)^2 = 4N - 4$ for $N > 1$.

To find basis matrices for this subspace, we could try to choose separately the $2N-1$ diagonals of $\boldsymbol{T}$ and the $2N-1$ antidiagonals of $\boldsymbol{H}$. But that gives $4N-2$ parameters. There must be a two-dimensional intersection of T and H, and it turns out that there is. The *all-ones matrix* is both Toeplitz and Hankel, as is the *checkerboard matrix* with entries $(-1)^{i+j}$.

Note: The splitting in this paper comes from $m - n$ and $m + n$ in the cosines of (3.6). $\boldsymbol{K}$ itself is pure Toeplitz, but our method puts the checkerboard part into $\boldsymbol{H}$ and not $\boldsymbol{T}$. The website math.mit.edu/highdegree develops other splittings, with source codes.

If the matrix $\boldsymbol{M}$ is required to be symmetric, that removes the entries of $\boldsymbol{T}$ on $N-1$ lower diagonals as independent parameters. The dimension of this subspace STH drops from $4N-4$ to $3N-3$. Correspondingly, the tests (8.1) only apply in the upper triangular part $1 < i < j < N$. Those $(N-2)(N-3)/2$ conditions act on the $N \times N$ symmetric matrices (dimension $N(N+1)/2$) to leave the correct subspace STH.

If, in addition, the Hankel part $\boldsymbol{H}$ is required to be centrosymmetric, that removes $N-1$ more parameters. The lower antidiagonals are reflections of the upper antidiagonals, as in the matrices of this paper. And our examples had the further condition that both parts $\boldsymbol{T}$ and $\boldsymbol{H}$ came from Fourier coefficients of the same function.

For $\boldsymbol{M} = \boldsymbol{T} + \boldsymbol{AH}$, Toeplitz plus alternating Hankel, there is a new and equally quick set of tests:

$$\begin{aligned} &\boldsymbol{M} = \boldsymbol{T} + \boldsymbol{AH} \quad \text{if and only if} \\ &\boldsymbol{M}_{i-1,j} + \boldsymbol{M}_{i,j-1} = \boldsymbol{M}_{i+1,j} + \boldsymbol{M}_{i,j+1} \quad \text{for } 1 < i < N,\ 1 < j < N. \end{aligned} \tag{8.2}$$

Toeplitz matrices $\boldsymbol{T}$ pass this test, as before. Alternating Hankel matrices $\boldsymbol{AH}$ pass because the test gives $0 = 0$. Therefore, $\boldsymbol{M} = \boldsymbol{T} + \boldsymbol{AH}$ will pass. The separate subspaces T and AH again have a two-dimensional intersection, spanned by the Toeplitz matrices with first row $1, 0, -1, 0, \ldots$ and first row $0, 1, 0, -1, \ldots$. Then the subspace TAH has the same dimension $4N - 4 = \dim(\mathrm{T}) + \dim(\mathrm{AH}) - \dim(\mathrm{T} \cap \mathrm{AH})$ as TH.

For fast computations with special subspaces of matrices, Morf and Kailath introduced the fruitful idea of *displacement rank* [17, 18]. For the shift matrix $\boldsymbol{Z}$ with 1's along the first subdiagonal, $\boldsymbol{ZRZ}^T$ is a "displacement" of $\boldsymbol{R}$ by one row and column. When $\boldsymbol{R}$ is Toeplitz, the difference $\boldsymbol{R} - \boldsymbol{ZRZ}^T$ has rank 2. For non-Toeplitz matrices, that displacement rank measures the difficulty of solving $\boldsymbol{Rx} = \boldsymbol{b}$.

Other subspaces (H, TH, TAH) are associated with other matrices $\boldsymbol{Z}$ coming from the tests (8.1) and (8.2). We are not sure (but can hope) that this subspace TAH will appear somewhere else again.

9. Graph Laplacian *B* with Neumann Boundary Conditions. The Laplacian matrix for a line of nodes is not the invertible matrix $\boldsymbol{K}$, but the singular matrix $\boldsymbol{B}$, with zero row sums:

$$\boldsymbol{B}=\begin{bmatrix}1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1\end{bmatrix} = \begin{matrix}\text{degree matrix}\\ \text{minus}\\ \text{adjacency matrix.}\end{matrix} \tag{9.1}$$

Nodes 0 and $N+1$ are no longer "grounded." Rows 1 and N correspond to a Neumann condition $\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}x}=0$ at $x=0$ and $x=1$. In fact, the eigenvalues of $\boldsymbol{B}_N$ are exactly the eigenvalues $2-2\cos\theta_k$ of $\boldsymbol{K}_{N-1}$ with $\theta_k=k\pi/N$, together with the new eigenvalue $\lambda_0=0$.

With zero slope at the endpoints, the differential equation $-\boldsymbol{u}_{xx}=\lambda\boldsymbol{u}$ yields the eigenfunctions $u=\cos k\pi x$. Then the eigenvectors of $\boldsymbol{B}$ are *discrete cosines* instead of discrete sines. Again we sample the exact eigenfunctions, but now at *half-integer* multiples of the step $h=1/N$. The eigenvector for $\lambda_0=0$ is $\boldsymbol{q}_0=(1,1,\dots,1)^T/\sqrt{N}$. The other $N-1$ eigenvectors are

$$\boldsymbol{q}_k=\sqrt{\frac{2}{N}}\left(\cos\left(\frac{1}{2}k\pi h\right),\cos\left(\frac{3}{2}k\pi h\right),\dots,\cos\left(\left(N-\frac{1}{2}\right)k\pi h\right)\right)^T. \tag{9.2}$$

All matrix functions $f(\boldsymbol{B})$ come from the spectral theorem $\boldsymbol{B}=\boldsymbol{Q\Lambda Q}^T$:

$$f(\boldsymbol{B})=\sum_{k=0}^{N-1} f(\lambda_k)\,\boldsymbol{q}_k\boldsymbol{q}_k^T. \tag{9.3}$$

Now introduce the entries (cosines) of these eigenvectors. Replace each $\cos A\cos B$ by $[\cos(A-B)+\cos(A+B)]/2$ to see that $f(\boldsymbol{B})$ is Toeplitz plus Hankel. Then approximate the sum by an integral:

$$f(\boldsymbol{B})_{m,n}\approx\frac{1}{\pi}\int_0^\pi f(2-2\cos\theta)\Big(\cos((m-n)\theta)\;+\;\cos((m+n)\theta)\Big)\,\mathrm{d}\theta. \tag{9.4}$$

The Hankel part of the approximation to $f(\boldsymbol{B})$ has *opposite sign* to the Hankel part of $f(\boldsymbol{K})$! For the matrix $e^{-\boldsymbol{B}t}$ that solves the heat equation, the entries will again use values $b_p=e^{-2t}I_p(2t)$ of the modified Bessel function of the first kind in (4.5):

$$e^{-\boldsymbol{B}}\approx\boldsymbol{T}+\boldsymbol{H}=b_{m-n}+b_{m+n}\qquad\text{for}\quad m+n\le N+1. \tag{9.5}$$

For the heat equation on $[0,1]$ with Neumann conditions at the endpoints, the same placement of images will succeed, but now the image sources are positive (to make the solution an even function across $x=0$ and also across $x=1$). The images are still responsible for the Hankel part, shifting in the opposite direction from $u_0(x)$. The Hankel part has a plus sign, not a minus sign, exactly as in (9.5) for the matrix case.

What we have done for $\boldsymbol{K}$ and $\boldsymbol{B}$ extends to three more matrices, all involving changes in the corners of $\boldsymbol{K}$:

$\boldsymbol{C}$	with	$\boldsymbol{C}_{1,N}=\boldsymbol{C}_{N,1}=-1$	(periodic circulant),
$\boldsymbol{M}$	with	$\boldsymbol{M}_{1,1}=1$ and $\boldsymbol{M}_{N,N}=2$	(mixed Neumann–Dirichlet matrix),
$\boldsymbol{S}$	with	$\boldsymbol{S}_{1,1}=2$ and $\boldsymbol{S}_{N,N}=1$	(mixed Dirichlet–Neumann matrix).

Periodic Case. The boundary conditions $\boldsymbol{u}(0) = \boldsymbol{u}(1)$ and $\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}x}(0) = \frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}x}(1)$ produce a circulant matrix $\boldsymbol{C}$ with $-1, 2, -1$ on its cyclic diagonals. All functions of circulant matrices are still circulants (therefore Toeplitz). The eigenvectors are the columns of the Fourier matrix [20]. If we prefer to work with real eigenvectors (as we do), those come as above from sampling the continuous eigenfunctions $\sin 2\pi kx$ and $\cos 2\pi kx$. Separately, these lead to Hankel parts as they did for $\boldsymbol{K}$ and $\boldsymbol{B}$. Combined, the two Hankel parts cancel because of opposite signs.

We now have two eigenvectors (sine and cosine) for a typical frequency k, and half as many frequencies. The usual step from sum to integral yields a good approximation for the middle diagonals of $\boldsymbol{C}_N$ and an exact value for $\boldsymbol{C}_{\infty,\infty}$:

$$(f(\boldsymbol{C}_N))_{m,n} \;\approx\; c_{m-n} \;=\; (f(\boldsymbol{C}_{\infty,\infty}))_{m,n}. \tag{9.6}$$

The entries c_p are simply the Fourier coefficients of the function $f(2 - 2\cos\theta)$. For the exponential $e^{-\boldsymbol{C}}$ we see modified Bessel again, but there is only the Toeplitz part:

$$(e^{-\boldsymbol{C}t/h})_{m,n} \approx e^{-2} I_{m-n}(2t/h). \tag{9.7}$$

For the finite matrix, we use these values for the middle diagonals $|m-n| \le (N+1)/2$ and complete the matrix as a circulant.

Neumann–Dirichlet. The boundary conditions $\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}x}(0) = 0$ and $\boldsymbol{u}(1) = 0$ change $\boldsymbol{K}_{1,1} = 2$ to $\boldsymbol{M}_{1,1} = 1$ and they leave $\boldsymbol{M}_{N,N} = 2$. There is an important change in the eigenvectors: For the continuous problem they become $\cos((k+\frac{1}{2})\pi x)$, and so they behave like a cosine at $x = 0$ and like a sine at $x = 1$. The eigenvalues are $(k+\frac{1}{2})^2\pi^2$.

The eigenvalues and eigenvectors of the finite matrix $\boldsymbol{M}$ have $M = N + \frac{1}{2}$ where $\boldsymbol{K}$ and $\boldsymbol{B}$ had $N+1$ and N:

$$\lambda_k = 2 - 2\cos\theta_k \qquad \text{with} \quad \theta_k = \frac{(k-\frac{1}{2})\pi}{M}, \tag{9.8}$$

$$\boldsymbol{y}_k = \sqrt{\frac{2}{M}} \left(\cos\left(\frac{1}{2}\theta_k\right), \cos\left(\frac{3}{2}\theta_k\right), \ldots, \cos\left(\left(N - \frac{1}{2}\right)\theta_k\right) \right)^T. \tag{9.9}$$

The m, n component of $f(\boldsymbol{M}_N) = \sum f(\lambda_k)\boldsymbol{y}_k\boldsymbol{y}_k^T$ is

$$\frac{2}{M}\sum_{1}^{N} f(\lambda_k)\cos\left(\left(m-\frac{1}{2}\right)\left(k-\frac{1}{2}\right)\frac{\pi}{M}\right)\cos\left(\left(n-\frac{1}{2}\right)\left(k-\frac{1}{2}\right)\frac{\pi}{M}\right). \tag{9.10}$$

Then the integral approximation to $f(\boldsymbol{M})_{m,n}$ is

$$\frac{2}{\pi}\int_0^{\pi} f(2-2\cos\theta)\Big(\cos((m-n)\theta) + \cos((m+n-1)\theta)\Big)\mathrm{d}\theta. \tag{9.11}$$

Notice the change to $m + n - 1$ in the Hankel part.

This approximation is good above the main antidiagonal, where $\boldsymbol{M}$ imitates $\boldsymbol{B}$ (Neumann). Below that antidiagonal, $\boldsymbol{M}$ imitates $\boldsymbol{K}$ (Dirichlet). There is aliasing in the terms $\cos((m+n-1)\theta_k) = \cos((\bar{m}+\bar{n}-1)\theta_k)$, where $\bar{m} = N+1-m$ and $\bar{n} = N+1-n$. So we have the option of changing to $\bar{m}, \bar{n}$ in the case $m+n > N+1$. The sum is the same, but the integrand is less oscillatory. For the exponential of $-\boldsymbol{M}$ the approximation is still Toeplitz plus Hankel, but no longer centrosymmetric. We see $e^{-\boldsymbol{B}}$ above the antidiagonal and $e^{-\boldsymbol{K}}$ below:

$$(e^{-\boldsymbol{M}})_{m,n} \approx e^{-2}I_{m-n}(2) + e^{-2}I_{m+n-1}(2) \qquad \text{for } m+n \le N+1, \tag{9.12}$$

$$(e^{-\boldsymbol{M}})_{m,n} \approx e^{-2}I_{m-n}(2) + e^{-2}I_{2N+2-m-n}(2) \quad \text{for } m+n > N+1. \tag{9.13}$$

The Dirichlet–Neumann matrix is the reflection of Neumann–Dirichlet.

10. Resolvents.

The resolvent $\boldsymbol{R}(z) \equiv (z\boldsymbol{I} - \boldsymbol{A})^{-1}$ of a matrix $\boldsymbol{A}$ offers another way to understand the Toeplitz plus Hankel property of all functions $f(\boldsymbol{A})$. When the resolvent is Toeplitz plus Hankel for all z, the Cauchy integral formula

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

shows that $f(\boldsymbol{A})$ is likewise Toeplitz plus Hankel. If the contour Γ_k encloses one simple eigenvalue λ_k, the projection

$$\boldsymbol{P}_k = \boldsymbol{v}_k\boldsymbol{v}_k^T = \frac{1}{2\pi i}\int_{\Gamma_k} (z\boldsymbol{I} - \boldsymbol{A})^{-1}\,\mathrm{d}z$$

is also $\boldsymbol{T}+\boldsymbol{H}$. (For a repeated eigenvalue, $\boldsymbol{P}_k$ is the projection onto the eigenspace.) Thus, we have three equivalent conditions ($1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 1$) for a symmetric matrix to have the "TH property":

1. All analytic functions $f(\boldsymbol{A})$ are Toeplitz plus Hankel.
2. The resolvent $\boldsymbol{R}(z) \equiv (z\boldsymbol{I} - \boldsymbol{A})^{-1}$ is Toeplitz plus Hankel for (almost) all z.
3. The projections onto all eigenspaces of $\boldsymbol{A}$ are Toeplitz plus Hankel.

The exceptions in condition 2 are the eigenvalues of $\boldsymbol{A}$, where $z\boldsymbol{I} - \boldsymbol{A}$ is singular. To repeat for emphasis: This "TH property" is much stronger than merely the requirement that A itself is $\boldsymbol{T}+\boldsymbol{H}$.

The resolvent (which is involved in studying the Helmholtz equation $-u_{xx} - k^2u = f(x)$) is the Laplace transform of the exponential:

$$(z\boldsymbol{I} - \boldsymbol{A})^{-1} = \int_0^\infty e^{t\boldsymbol{A}}e^{-zt}\mathrm{d}t\,.$$

Our exact expression for $e^{-\boldsymbol{K}t}$ as a finite sum allows us to find the resolvent of $-\boldsymbol{K}$ exactly, as a finite sum of Laplace transforms. Alternatively, we have $e^{-\boldsymbol{K}t}$ approximately, in terms of modified Bessel functions of the first kind, I_p. The key Laplace transform L is that of $e^{-2t}I_p(2t)$:

$$L(p,z) = \frac{2^p}{\sqrt{z^2+4z}\;\;(z+2+\sqrt{z^2+4z})^p}.$$

Hence, the resolvent of $-\boldsymbol{K}$ is approximately

$$\Big((z\boldsymbol{I} + \boldsymbol{K})^{-1}\Big)_{m,n} \approx \;\; L(|m-n|, z) - L(m+n, z)\,. \tag{10.1}$$

As usual, we make use of symmetry and centrosymmetry for $m+n > N+1$. The approximation (10.1) is very accurate; for example, with $z = 5+i$, the approximation is close to machine precision at $N = 15$.

11. The Four Corners Theorem.

We had expected that only special values in the $(1,1)$, $(1,N)$, $(N,1)$, and (N,N) corners of our second difference matrix would ensure that all matrix functions are $\boldsymbol{T}+\boldsymbol{H}$. However, MATLAB experiments told us we were wrong. For random values in the four corners (maintaining symmetry by $\boldsymbol{A}_{1,N} = \boldsymbol{A}_{N,1}$) we computed the eigenvectors $\boldsymbol{v}_k$. In every case, the projections $\boldsymbol{v}_k\boldsymbol{v}_k^T$ passed the test (8.1). Thus, $\boldsymbol{A} = \sum \lambda_k\boldsymbol{v}_k\boldsymbol{v}_k^T$ surely has the TH property.

At this point we finally understood why every $\boldsymbol{M} = \boldsymbol{v}_k\boldsymbol{v}_k^T$ passes the T+H test. Subtract $2\boldsymbol{M}_{i,j}$ from both sides of (8.1), which does not change the test:

$$\boldsymbol{M}_{i-1,j} - 2\boldsymbol{M}_{i,j} + \boldsymbol{M}_{i+1,j} = \boldsymbol{M}_{i,j-1} - 2\boldsymbol{M}_{i,j} + \boldsymbol{M}_{i,j+1}. \tag{11.1}$$

At each interior position (i, j), the second difference $\boldsymbol{\Delta}^2$ down column j of $\boldsymbol{M}$ must equal the second difference along row i. Now apply this test to $\boldsymbol{M} = \boldsymbol{v}\boldsymbol{v}^T$. It requires

$$\boldsymbol{v}(j)\Delta^2\boldsymbol{v}(i) = \boldsymbol{v}(i)\Delta^2\boldsymbol{v}(j) \quad \text{for } 1 < i < N,\ 1 < j < N. \tag{11.2}$$

Each interior row of $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ is the statement that $\Delta^2\boldsymbol{v}(j) = -\lambda\boldsymbol{v}(j)$. Arbitrary entries in the corners of $\boldsymbol{A} = \boldsymbol{A}^T$ have no effect. (If we try to change other entries in row 1, $\boldsymbol{A}$ loses symmetry, and if $\boldsymbol{A}$ has more than the three diagonals, all our $\boldsymbol{T} + \boldsymbol{H}$ reasoning fails.) The test (11.2) is passed, because it becomes

$$\boldsymbol{v}(j)\lambda\boldsymbol{v}(i) = \boldsymbol{v}(i)\lambda\boldsymbol{v}(j) \quad \text{for } 1 < i < N,\ 1 < j < N. \tag{11.3}$$

We conclude that every $\boldsymbol{v}\boldsymbol{v}^T$ is $\boldsymbol{T} + \boldsymbol{H}$, and therefore all functions of $\boldsymbol{A}$ are $\boldsymbol{T} + \boldsymbol{H}$.

The changes in $\boldsymbol{A}_{1,1}$ and $\boldsymbol{A}_{N,N}$ correspond to *Robin boundary conditions* like $\frac{\mathrm{d}u}{\mathrm{d}x}(0) = au(0)$. Perhaps a four corner change corresponds to Robin conditions linking $x = 0$ and $x = 1$ (which we have never seen).

In relation to boundary value problems, we note that explicit formulas for eigenvalues and inverses of tridiagonal Toeplitz matrices with four perturbed corners have been studied elsewhere [28]. The eigenvectors are not simple sines or cosines.

For an integral operator $\boldsymbol{A}u(x) = \int P(x, y)u(y)\mathrm{d}y$ with kernel $P(x, y)$, the $\boldsymbol{T}+\boldsymbol{H}$ test (11.1) in the continuous case would become $P_{xx} = P_{yy}$. Solutions to this wave equation have exactly the form $P = f(x - y) + g(x + y)$. Then the kernel is Toeplitz plus Hankel.

12. Conclusions.

The eigenvalues and eigenvectors of a symmetric matrix $\boldsymbol{K}$ give a formula for all functions $f(\boldsymbol{K})$. When those eigenvectors are discrete sines or cosines, every $f(\boldsymbol{K})$ is Toeplitz plus Hankel. Moreover, entries of $f(\boldsymbol{K})$ can be approximated by an integral. For second difference matrices $\boldsymbol{K}$, this paper identifies those integrals as Bessel coefficients, which give exact solutions to basic finite difference equations on a half-line $x \geq 0$ or on the whole line.

The formulas for $e^{-\boldsymbol{K}t/h^2}$, $\cos(\sqrt{\boldsymbol{K}}t/h)$, and $\operatorname{sinc}(\sqrt{\boldsymbol{K}}t/h)$ allow a much more precise estimate for discretization errors than merely $\mathcal{O}(h^2)$.

Acknowledgment. We are particularly grateful to Joel Moses and Hung Cheng.

REFERENCES

[1] M. Benzi and N. Razouk, *Decay bounds and O(N) algorithms for approximating functions of sparse matrices*, Electron. Trans. Numer. Anal., 28 (2007), pp. 16–39.

[2] R. Bevilacqua, N. Bonanni, and E. Bozzo, *On algebras of Toeplitz plus Hankel matrices*, Linear Algebra Appl., 223/224 (1995), pp. 99–118.

[3] A. Böttcher and S. M. Grudsky, *Spectral Properties of Banded Toeplitz Matrices*, SIAM, Philadelphia, 2005.

[4] D. Cvetkovic, P. Rowlinson, and S. Simic, *An Introduction to the Theory of Graph Spectra*, London Mathematical Society Student Texts, Cambridge University Press, Cambridge, UK, 2010.

[5] E. B. Davies, *Spectral Theory and Differential Operators*, Cambridge University Press, Cambridge, UK, 1995.

[6] *Digital Library of Mathematical Functions*, http://dlmf.nist.gov/, National Institute of Standards and Technology (4 April 2014).

[7] E. Estrada and N. Hatano, *Communicability in complex networks*, Phys. Rev. E, 77 (2008), 036111.

[8] E. Estrada, N. Hatano, and M. Benzi, *The physics of communicability in complex networks*, Phys. Rep., 514 (2012), pp. 89–119.

[9] E. Estrada and D. J. Higham, *Network properties revealed through matrix functions*, SIAM Rev., 52 (2010), pp. 696–714.

[10] D. Fasino, *Spectral properties of Toeplitz-plus-Hankel matrices*, Calcolo, 33 (1996), pp. 87–98.

[11] J. Friedman and J.-P. Tillich, *Wave equations for graphs and the edge-based Laplacian*, Pacific J. Math., 2 (2004), pp. 229–266.

[12] R. Haberman, *Applied Partial Differential Equations*, Prentice–Hall, Englewood Cliffs, NJ, 2013.

[13] N. Hale, N. J. Higham, and L. N. Trefethen, *Computing A^α, $\log(A)$, and related matrix functions by contour integrals*, SIAM J. Numer. Anal., 46 (2008), pp. 2505–2523.

[14] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008.

[15] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 2013.

[16] A. Iserles, *How large is the exponential of a banded matrix?*, J. New Zealand Math. Soc., 29 (2000), pp. 177–192.

[17] T. Kailath, S.-Y. Kung, and M. Morf, *Displacement rank of a matrix*, Bull. Amer. Math. Soc., 1 (1979), pp. 769–773.

[18] T. Kailath and A. H. Sayed, *Displacement structure: Theory and applications*, SIAM Rev., 37 (1995), pp. 297–386.

[19] P. L. Nash and J. A. C. Weideman, *High accuracy representation of the free propagator*, Appl. Numer. Math., 59 (2009), pp. 2937–2949.

[20] G. Strang, *Computational Science and Engineering*, Wellesley-Cambridge Press, Wellesley, MA, 2007.

[21] G. Strang, *Introduction to Linear Algebra*, Wellesley-Cambridge Press, Wellesley, MA, 2009.

[22] L. N. Trefethen, *Spectral Methods in MATLAB*, SIAM, Philadelphia, 2000.

[23] L. N. Trefethen and J. A. C. Weideman, *The exponentially convergent trapezoidal rule*, SIAM Rev., 56 (2014), pp. 385–458.

[24] J. N. Tsitsiklis and B. C. Levy, *Integral Equations and Resolvents of Toeplitz Plus Hankel Kernels*, unpublished MIT Report LIDS-P-1170, 1981.

[25] J. A. C. Weideman, *Numerical integration of periodic functions: A few examples*, Amer. Math. Monthly, 109 (2002), pp. 21–36.

[26] J. A. C. Weideman and L. N. Trefethen, *Parabolic and hyperbolic contours for computing the Bromwich integral*, Math. Comp., 76 (2007), pp. 1341–1356.

[27] H. Widom, *Toeplitz matrices*, in Studies in Real and Complex Analysis, I. I. Hirschman, ed., Prentice–Hall, Englewood Cliffs, NJ, 1965.

[28] W.-C. Yueh and S. Cheng, *Explicit eigenvalues and inverses of tridiagonal Toeplitz matrices with four perturbed corners*, ANZIAM J., 49 (2008), pp. 361–387.