

# Mat1062: Introductory Numerical Methods for PDE

Mary Pugh

March 24, 2009

## 1 Ownership

These notes are the joint property of Rob Almgren and Mary Pugh.

## 2 An Example

Spectral methods are similar to finite element methods in that they involve representing a solution as a linear combination of basis functions and using an inner product to find the coefficients of the linear combination. The difference is that spectral methods use the eigenfunctions of the linear operator as the basis functions.

Consider the operation  $\mathcal{L} = \partial_{xx}$  on  $[0, L]$ . As we observed earlier, an operator is an operation plus boundary conditions. Different boundary conditions lead to different operators leading to different eigenfunctions. For example,

$$\begin{aligned}u(0) = u(L) = 0 &\implies \sin(j\pi x/L) \quad j \in \mathbb{N} \\u_x(0) = u_x(L) &\implies 1, \quad \cos(j\pi x/L) \quad j \in \mathbb{N} \\u(0) = u(L) &\implies \exp(ij\pi x/L) \quad j \in \mathbb{Z} \\u(0) = u_x(L) = 0 &\implies \sin((j + 1/2)\pi x/L) \quad j \in \mathbb{N}\end{aligned}$$

Here is an example of how we would use the basis to find a solution of an initial boundary value problem. Take  $u_t = u_{xx} + f$  on  $0 \leq x \leq L$ , with  $u = 0$  at  $x = 0, L$ . We write  $u(x, t)$  in terms of the eigenfunctions

$$\phi_j(x) = \sin\left(\frac{j\pi x}{L}\right).$$

Then any approximation of the form

$$u_N(x, t) = \sum_{j=1}^N c_j(t) \phi_j(x)$$

satisfies the boundary conditions.

We define the inner product to be

$$\langle u, v \rangle = \frac{2}{L} \int_0^L u(x) v(x) dx.$$

Then

$$\langle \phi_j, \phi_j \rangle = \frac{2}{L} \int_0^L \sin^2\left(\frac{j\pi x}{L}\right) dx = \frac{1}{L} \int_0^L \left(1 - \cos\left(\frac{j2\pi x}{L}\right)\right) dx = 1$$

and for  $j \neq k$ ,

$$\begin{aligned} \langle \phi_j, \phi_k \rangle &= \frac{2}{L} \int_0^L \sin\left(\frac{j\pi x}{L}\right) \sin\left(\frac{k\pi x}{L}\right) dx \\ &= \frac{1}{L} \int_0^L \left(\cos\left(\frac{(j-k)\pi x}{L}\right) - \cos\left(\frac{(j+k)\pi x}{L}\right)\right) dx = 0. \end{aligned}$$

Thus  $\langle \phi_j, \phi_k \rangle = \delta_{jk}$  as we hoped. Furthermore,

$$\mathcal{L}\phi_j = -\left(\frac{j\pi}{L}\right)^2 \phi_j$$

and so the ODE system for the  $c_j$  is

$$\frac{dc_k}{dt} = -\left(\frac{k\pi}{L}\right)^2 c_k + f_k(t)$$

with

$$f_k(t) = \frac{2}{L} \int_0^L f(x, t) \sin\left(\frac{k\pi x}{L}\right) dx.$$

In this way, solving the PDE has been reduced to solving infinitely many ODE.

The above is the exact same thing that you would have found via separation of variables. When you first learnt separation of variables, you were learning to work with the eigenfunctions of the Laplacian and were taking advantage of the fact that, for reasonable boundary conditions, the eigenfunctions are orthogonal to one another. It may not have been presented to you in this language.

### 3 Fourier Transforms

#### 3.1 Forward transform

Suppose  $u(x)$  is a real or complex function defined on a finite interval  $[0, L]$ ; it will be convenient to think of  $u$  as extended periodically outside this interval. The *Fourier coefficients* are the doubly infinite sequence of complex numbers

$$\hat{u}_k = \langle u, \phi_k \rangle = \frac{1}{L} \int_0^L u(x) e^{-ik2\pi x/L} dx, \quad \text{for } k = 0, \pm 1, \pm 2, \dots \quad (1)$$

(the “hat” is commonly used to denote the Fourier transform). Above, we use the inner product for complex-valued functions:

$$\langle f, g \rangle := \frac{1}{L} \int_0^L f(x) \overline{g(x)} dx$$

where the overline denotes complex conjugate. We use the integer  $k$  as the index of the mode. The physical wave number is  $\xi_k = 2\pi k/L$ , so the wavelength is  $\lambda_k = 2\pi/|\xi_k| = L/|k|$ . Thus the  $k$ th mode has  $k$  waves in the period interval.

The symmetries of the sequence  $\{\hat{u}_k\}$  mirror those of the function  $u$ . First, if  $u$  is real-valued, then  $\hat{u}_{-k} = \overline{\hat{u}_k}$ . Thus in that case, the values of  $\hat{u}_k$  with  $k < 0$  duplicate the ones with  $k > 0$ .

Suppose  $u$  is *even*:  $u(-x) = u(x)$ . Since we have extended it outside the interval  $[0, L]$  by assuming periodicity, this really means  $u(L-x) = u(x)$ . Then it is easy to see that  $\hat{u}_{-k} = \hat{u}_k$  (the sequence is even), and again the negative indices contain duplicate information. If  $u$  is real *and* even, then the coefficients are real and even. If  $u$  is smooth when extended periodically, then evenness requires zero derivative at  $x = 0$ ; this case is relevant for insulating Neumann boundary conditions on a finite interval.

If  $u$  is *odd*, so  $u(-x) = -u(x)$  or  $u(L-x) = -u(x)$ , then  $\hat{u}_{-k} = -\hat{u}_k$ . In particular,  $\hat{u}_0 = 0$ , since it is the average value of  $u(x)$ . If  $u$  is real as well, then all the Fourier coefficients are pure imaginary. This corresponds to zero Dirichlet conditions at the end of the interval.

Next, we compute some special “transform pairs,” that follow from the *orthogonality relations*

$$\frac{1}{L} \int_0^L e^{i\ell 2\pi x/L} e^{-ik 2\pi x/L} dx = \frac{1}{L} \int_0^L e^{i(\ell-k)2\pi x/L} dx = \delta_{k\ell} = \begin{cases} 1, & \text{if } k = \ell \\ 0, & \text{else.} \end{cases}$$

Thus if  $u(x)$  is a single complex exponential,

$$u(x) = e^{i\ell 2\pi x/L} \implies \hat{u}_k = \delta_{k\ell}.$$

The infinite sequence  $\{\hat{u}_k\}$  is all zeros, with a single value equal to one at index  $\ell$ . (We do not consider the case of non-integer values of  $\ell$ .) If  $u(x)$  is a sum of exponentials, then by linearity of the transform, we have

$$u(x) = a_1 e^{i\ell_1 2\pi x/L} + \dots + a_n e^{i\ell_n 2\pi x/L} \implies \hat{u}_k = a_1 \delta_{k\ell_1} + \dots + a_n \delta_{k\ell_n} \quad (2)$$

The sequence is all zeros, with values  $a_1, \dots, a_n$  at indices  $\ell_1, \dots, \ell_n$ .

In particular, using the de Moivre relation  $\cos x = (e^{ix} - e^{-ix})/2$ , we see

$$u(x) = \cos\left(\frac{2\pi}{L}(\ell x - \alpha L)\right) \implies \begin{cases} \hat{u}_\ell = \frac{1}{2} e^{-i\alpha 2\pi} \\ \hat{u}_{-\ell} = \frac{1}{2} e^{i\alpha 2\pi} \end{cases}$$

We have introduced  $\alpha$  as a *phase shift*. A pure cosine wave has  $\alpha = 0$ , giving  $\hat{u}_{\pm\ell} = 1/2$ . and a sine has  $\alpha = 1/4$ , giving  $\hat{u}_{\pm\ell} = \pm 1/2i$ ; you may check the symmetry relationships above. In general, *translation* in physical space corresponds to a *rotation* of the Fourier coefficients.

We now know what the Fourier transform is, when  $u(x)$  is any finite sum of complex exponentials. You may be willing to believe that if  $u(x)$  is an *infinite* sum of complex exponentials, then the same relationships hold:

$$u(x) = \sum_{\ell=-\infty}^{\infty} a_\ell e^{i\ell 2\pi x/L} \implies \hat{u}_k = a_k$$

assuming the numbers  $\{a_\ell\}_{\ell=-\infty}^{\infty}$  are such that the infinite sum exists.

### 3.2 Decay behavior

The decay of the coefficients  $\{\hat{u}_k\}$  as  $k \rightarrow \pm\infty$  is closely related to the smoothness of  $u(x)$ . The best possible behavior, and the reason that spectral methods work well for smooth problems, comes when  $u$  is very smooth. The smoothest that a function of a real variable can be is *real analytic*, meaning that it is the restriction to the real axis of a complex analytic function.

**Theorem:** *Suppose that  $u(x)$ , periodic in  $x$  with period  $L$ , can be extended to complex values of  $x$  so that it is analytic in a strip  $-\rho \leq$*

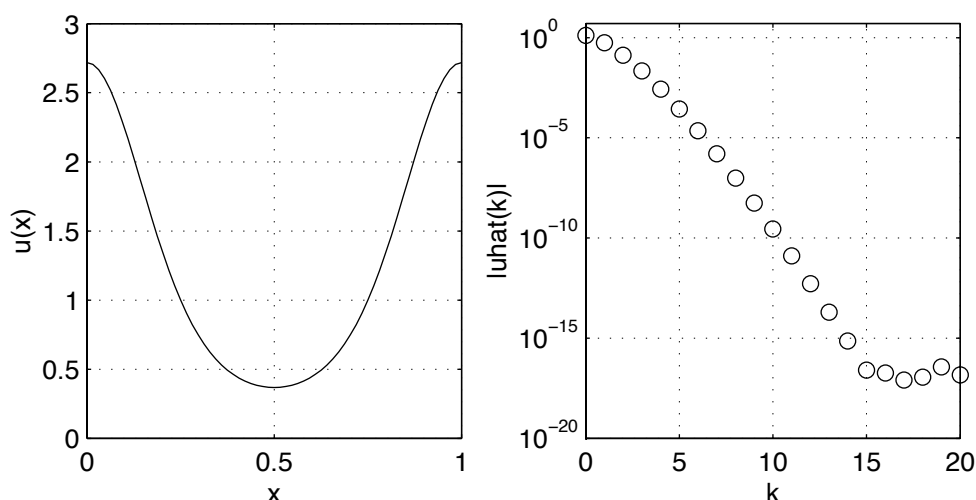


Figure 1: The very smooth function  $u(x) = \exp(\cos(2\pi x))$  (left picture) and its numerically computed Fourier coefficients (right). Since  $u$  is analytic in the entire complex plane, the coefficients decay faster than any exponential in  $k$ . Thus in this log-linear plot, the coefficients go to zero faster than any line, until they bottom out at machine precision around  $k = 15$ .

$\text{Im } x \leq \rho$ . Then the magnitudes of the Fourier coefficients  $\hat{u}_k$  decay at least exponentially with parameter controlled by  $\rho$ :

$$|\hat{u}_k| \sim \mathcal{O}(C e^{-c\rho|k|}), \quad |k| \rightarrow \infty.$$

where  $c = 2\pi/L$  and  $C$  depends on  $\rho$ ,  $L$  and  $u$  but not on  $k$ .

The trigonometric polynomials (2) are analytic in the entire complex plane, and the Fourier series are exactly zero beyond a finite index; this is certainly rapid decay. For more realistic examples, see Figures 1, 2.

**Proof:** If  $u(x)$  is analytic in  $0 \leq \text{Im } x \leq \rho$ , then the path of integration may be shifted upwards to  $\text{Im } x = \rho$  without changing the value of the integral, since the contributions along the vertical end segments cancel. Thus

$$\hat{u}_k = \frac{1}{L} \int_0^L u(x) e^{-ik2\pi x/L} dx = e^{2\pi k\rho/L} \cdot \frac{1}{L} \int_0^L u(x + i\rho) e^{-ik2\pi x/L} dx.$$

Now since  $u$  is analytic along the line  $\text{Im } x = \rho$ , it is continuous and therefore bounded. Thus the integral is bounded independently of  $k$ , and this

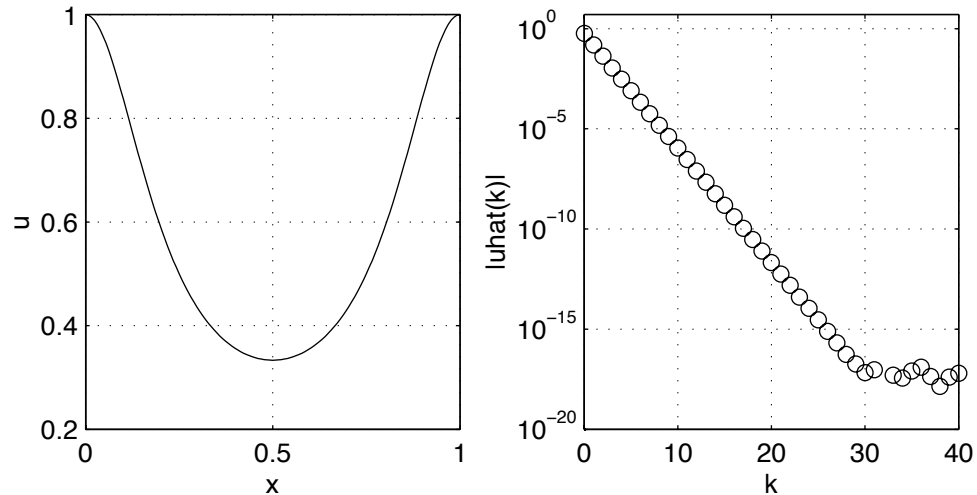


Figure 2: The real analytic function  $u(x) = 1/(2 - \cos(2\pi x))$ , together with its Fourier coefficients. Although this function is real analytic and looks about the same as the example in Figure 1, it has singularities in the complex plane at a finite distance from the real axis. Hence the magnitude of the Fourier coefficients decays only exponentially, rather than super-exponentially, and we need  $k = 30$  to get down to machine precision.

gives the result for  $k < 0$ , with  $c = 2\pi/L$ . The case  $k > 0$  follows by shifting the path down to  $\text{Im } \chi = -\rho$ . This finishes the proof.

An alternate proof would be to change variables to  $z = \exp(i2\pi x/L)$ . Then  $u(x)$  being analytic in a strip corresponds to  $v(z) = u((L/2\pi i) \log z)$  being analytic in an annulus, and the Fourier series for  $u$  is the Laurent series for  $v$ , whose coefficients decay rapidly.

In the theorem we included the edges of the strip, so there cannot be any singularities right at  $\text{Im } \chi = \pm\rho$ . If  $u$  has singularities in the complex plane then the decay rate of  $|\hat{u}_k|$  also has an algebraic term multiplying the exponential term that can be used to extract information about the nature of the closest singularities of  $u(x)$  to the real axis.<sup>1</sup>

<sup>1</sup>e.g., M. J. Shelley, "A study of singularity formation in vortex-sheet motion by a spectrally accurate vortex method," *J. Fluid Mech.* 244 (1992) 493–526.

```

function specplot( n )

L = 1;

x = ((1:n)-1)*L/n;
u = exp(cos(2*pi*x/L));

k = [ 0:(n/2)-1 0 -(n/2):-1 ];
uhat = fft(u)/n;

figure(1)
plot( [ x L ], [ u u(1) ] );
xlabel('x'); ylabel('u(x)'); grid on;

figure(2);
semilogy( 0:(n/2)-1, abs(uhat(1:(n/2))), 'o' );
xlabel('k'); ylabel('|uhat(k)|'); grid on
axis( [ 0 n/2 1e-20 5 ] );

```

Figure 3: The Matlab code that produced the plots in Figures 1 and 2. `fft` is used to compute the DFT, and we plot only the first half of the resulting sequence; since  $u$  is real (and even) the second half contains redundant information.

As an application of the theorem, consider

$$u_\epsilon(x) = \frac{1}{\epsilon^2 + \sin^2(\pi(x - 1/2))} \quad (3)$$

on  $[0, 1]$ . This is a periodic function, by construction, see the left plot in Figure 4. It has a maximum value of  $1/\epsilon^2$  achieved at  $x = 1/2$ . The smaller  $\epsilon$  is, the taller and narrower the peak at  $x = 1/2$  will be. If we were to integrate  $u_\epsilon$  we see that we get a function which has a sharp transition at  $x = 1/2$ , and so the following discussion is relevant not just for localized “blow-ups” but also for localized sharp fronts. We can use the Fourier coefficients to determine how many meshpoints we need to properly resolve  $u_\epsilon$  for a given value of  $\epsilon$ . Consider the related function  $1/(\epsilon^2 + x^2)$ . This function is real analytic and has two poles in the complex plane at  $\pm i\epsilon$ . Similarly,  $u_\epsilon$  has poles at  $.5 \pm i\epsilon$ . This means that by the above theorem, the Fourier coefficients will decay like

$$|\widehat{u}_{\epsilon k}| \approx e^{-|k|\epsilon}.$$

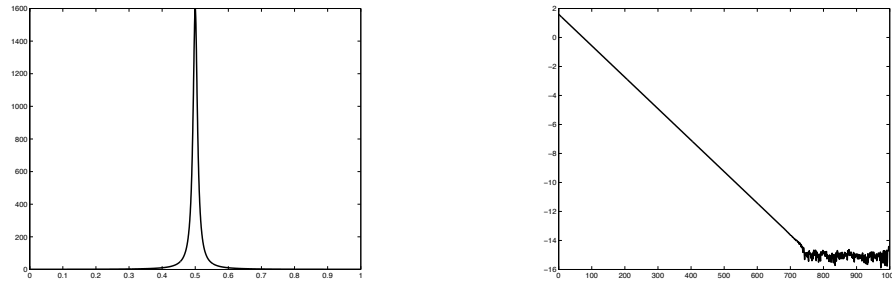


Figure 4: The function  $u_\epsilon$  as given by (3) with  $\epsilon = 1/40$ . The function is sampled on a uniform mesh with  $dx = .0005$ . Left plot: the graph of  $u_\epsilon$  as a function of  $x$ . Right plot: Because there are 2000 sample points, 1000 frequencies can be resolved. This is a plot of  $\log_{10}(|\widehat{u}_{\epsilon k}|)$  as a function of  $k$ . The Fourier coefficients reach the level of round-off around  $k = 800$ .

If we seek the wave number at which the coefficient is of order of round-off we find

$$e^{-|k_c|\epsilon} \approx 10^{-14} \implies k_c \approx 1/\epsilon$$

And so we need wave numbers up to  $\mathcal{O}(1/\epsilon)$  to fully resolve  $u_\epsilon$ . This points to one of the weaknesses of spectral methods. As we will see shortly, to resolve the  $N$ th wave number we need  $2N$  meshpoints, uniformly distributed. And so we have to use lots of points everywhere because of this fine structure localised in one area. This is unavoidable because spectral methods are based on eigenfunctions which are inherently non-localized objects. Using finite element or finite difference methods, we could concentrate meshpoints where they are most needed.<sup>2</sup> For the  $\epsilon$  used in Figure 4 we need enough meshpoints to be able to find the  $k = 800$  modes:  $\widehat{u}_{800}$  and  $\widehat{u}_{-800}$ . To see the effect of not having enough modes, consider Figure 5 Here, I took the Fourier Transform of  $u_\epsilon$  and then constructed a function  $v$  by using  $\widehat{u}_{\epsilon-100}$ ,  $\widehat{u}_{\epsilon-99}$ ,  $\dots$ ,  $\widehat{u}_{\epsilon 99}$ ,  $\widehat{u}_{\epsilon 100}$ . (That is,  $v$  is what one would get by applying a low-pass filter to  $u_\epsilon$ .) In the left plot of Figure 5 we see the effect of under-resolution on the graph near the local maximum. In the right plot of Figure

<sup>2</sup>This is, of course, more complicated than I'm making it sound. Finite elements and finite-difference methods have to work hard if the structure of interest is moving in time. And of course, if fine structures are appearing at multiple places in space, and moving in time (think of turbulent fluid flow) then you may as well resolve all points in space equally. Also, as we'll see, spectral methods are more accurate than finite element and finite difference methods. This means that you may end up having to use a lot more finite elements than you'd expected to, resulting in a scheme that's no faster than the spectral method.

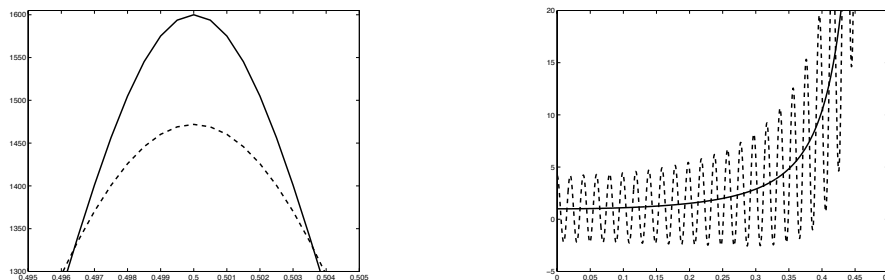


Figure 5: We took the Fourier transform of  $u_\epsilon$ , set  $\widehat{u}_{\epsilon k} = 0$  for all  $|k| > 100$ , and then applied an inverse Fourier transform resulting in a function  $v$ . Left plot: The solid line is  $u_\epsilon$  and the dashed line is  $v$ . The graph has been chosen to focus in near the peak of  $u_\epsilon$ . Right plot: The solid and dashed lines are as in the left plot. The graph has been chosen to focus in on the left flank of the peak.

5 we see the effect of under-resolution on the graph near the flanking flat areas. Note that there are significant oscillations.

If  $u(x)$  is not real analytic, then the Fourier coefficients have an algebraic rate of decay, with an exponent controlled by the differentiability of  $u$ . If  $u$  is infinitely differentiable but not real analytic, the coefficients decay faster than any power of  $k$ , but not exponentially (this is a rare case).

Note that we are referring to the smoothness of  $u$  after it has been extended to a periodic function on  $\mathbb{R}$ . If  $u(0) \neq u(L)$ , then the extension will be discontinuous and the Fourier coefficients will decay like  $1/|k|$ .

It is the *super-algebraic* decay of the Fourier coefficients for smooth functions that makes spectral methods so powerful. In the language of finite-difference methods with errors  $\sim h^p$ , these have order  $p = \infty$  (when everything works right). This is called *spectral accuracy*.

### 3.3 Inverse transform

Now let us consider the opposite of the above operation: Suppose we start with a function  $u(x)$ , compute the infinite sequence of coefficients  $\{\hat{u}_k\}_{k=-\infty}^{\infty}$ , and we form the infinite sum

$$u(x) = \sum_{k=-\infty}^{\infty} \hat{u}_k e^{ik2\pi x/L}.$$

As described above, if  $u(x)$  is smooth, then  $|\hat{u}_k|$  decays rapidly as  $|k| \rightarrow \pm\infty$ , and so this sum exists for each  $x$ . Furthermore, we have taken the liberty of naming the sum  $u(x)$  since indeed the infinite sum is equal to the original function if it is smooth.

Thus we have two representations of the *same* information. In the “physical space representation,” we have the function  $u(x)$ , consisting of its values (real or complex) at all the points  $x \in [0, L]$ . In the “Fourier representation,” we have an infinite sequence of complex numbers. You may think this is very strange, since the first form contains uncountably many different numbers, while the second has only countably many. In fact, the assumption of smoothness reduces the amount of information in the function  $u$  to countably many degrees of freedom.

A more formal way to say the above is to form the approximations to  $u$  consisting of only finitely many Fourier modes:

$$u_n(x) = \sum_{k=-n}^n \hat{u}_k e^{ik2\pi x/L}.$$

Certainly, for any finite  $n$ , this sum exists, and the question is whether  $u_n(x) \rightarrow u(x)$  in some reasonable sense as  $n \rightarrow \infty$ . If  $u$  is real analytic, then this convergence is true in *every* sense. In general, the exact nature of the convergence depends on precisely how smooth  $u$  is.<sup>3</sup>

- If  $u$  has continuous derivatives of order 2, then  $u_n \rightarrow u$  uniformly.
- If  $u$  is piecewise continuous with piecewise continuous first derivative, then  $u_n \rightarrow u$  at points where  $u$  is continuous. At points of discontinuity, the series converges to the average of the left and right values.
- If  $u$  is square-integrable, then  $\int |u_n - u|^2 \rightarrow 0$ . This is often the most suitable notion, since this guarantees the existence of the associated inner product. In addition, we have *Parseval's identity*

$$\frac{1}{L} \int_0^L |u(x)|^2 dx = \sum_{k=-\infty}^{\infty} |\hat{u}_k|^2.$$

---

<sup>3</sup>See Walter A. Strauss, *Partial Differential Equations: An Introduction*, John Wiley & Sons 1992.

These matters were very controversial throughout the 19th century; Joseph Fourier is credited with proposing that a series of *continuous* trigonometric polynomials could actually converge to a *discontinuous* function.

The situation is somewhat reminiscent of finite elements, since there also, we projected our function into a space spanned by a discrete set of basis functions, and we reconstructed the original as well as we could. Here, we can recover the original function *exactly*, assuming some smoothness, by taking more and more basis functions. That was also true for finite elements, but the geometric complexity increased as the space got larger.

Two final comments:

- We shall not consider the extension to functions defined on all  $\mathbb{R}$ , with suitable decay assumptions at  $\pm\infty$ ; then  $\hat{u}$  becomes a function of a continuous variable  $k$ . Periodicity in  $x$  translates to discreteness in  $k$ .
- Everything we have said can be extended to functions of several variables, defined on a rectangle  $[0, L_1] \times \cdots \times [0, L_d]$  instead of the interval  $[0, L]$ ; the wave index  $k$  becomes a multi-index  $(k_1, \dots, k_d)$ .