Summary:

- One of the main goals of this course is to understand how to write functions as series in a collection of mutually orthogonal functions which are such that the original partial differential equation (or whatever other problem we are dealing with) becomes simple.
- This bears some analogies to the process of diagonalising matrices and writing arbitrary vectors as linear combinations of the eigenvectors of a matrix.
- This process – whether for matrices or for the differential operators which shall be our main concern here – works best when we have an *inner product*, which gives us a way of generalising the notion of *projection* and hence allows us to compute the coefficients in the series expansions mentioned above.

EXAMPLE. Consider the matrix from last time, $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.[1] We recall that this has eigenvalues 3 and 1 and corresponding eigenvectors $\mathbf{e}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\mathbf{e}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, which are orthonormal. Now suppose that we wish to solve the equation $A\mathbf{x} = \mathbf{y}$, for some given vector $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Now since $\mathbf{e}_1$ and $\mathbf{e}_2$ span $\mathbf{R}^2$, there are numbers a_1 and a_2 such that $\mathbf{y} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2$; in fact, we may write

$$\mathbf{y} \cdot \mathbf{e}_1 = (a_1\mathbf{e}_1 + a_2\mathbf{e}_2) \cdot \mathbf{e}_1 = a_1\mathbf{e}_1 \cdot \mathbf{e}_1 + a_2\mathbf{e}_2 \cdot \mathbf{e}_1 = a_2,$$

and similarly $\mathbf{y} \cdot \mathbf{e}_2 = a_2$; thus

$$\mathbf{y} = (\mathbf{y} \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{y} \cdot \mathbf{e}_2)\mathbf{e}_2. \tag{1}$$

Similarly, we may write $\mathbf{x} = b_1\mathbf{e}_1 + b_2\mathbf{e}_2$; then the equation $A\mathbf{x} = \mathbf{y}$ becomes

$$A(b_1\mathbf{e}_1 + b_2\mathbf{e}_2) = b_1A\mathbf{e}_1 + b_2A\mathbf{e}_2 = 3b_1\mathbf{e}_1 + b_2\mathbf{e}_2 = a_1\mathbf{e}_1 + a_2\mathbf{e}_2.$$

Since $\{\mathbf{e}_1, \mathbf{e}_2\}$ forms a basis for $\mathbf{R}^2$, we see that we must have $3b_1 = a_1$, $b_2 = a_2$, i.e.,

$$\mathbf{x} = \frac{1}{3}(\mathbf{y} \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{y} \cdot \mathbf{e}_2)\mathbf{e}_2.$$

The main point here, though, is not this last formula, but rather that if we take what was originally a difficult problem (solving a system of equations) and rewrite it using the eigenvectors $\mathbf{e}_1$ and $\mathbf{e}_2$, it becomes a very simple problem. For the problems in partial differential equations which we wish to tackle, rewriting them in terms of (what we shall term) *eigenfunctions* is often about the only real way to approach the problem (at least if what we want is a formula for the solution, which is usually the case for us in this class).

COMMENTARY. To extend the above example to the problems in partial differential equations which we wish to treat, we see that we need to extend the notion of dot product to functions, in such a way that our expansion formulas will look like equation (1) above. The following is a particular example of such an extension. (We shall have occasion to use others, but they will be closely related to this one.)

DEFINITION. Suppose that $f, g : [a, b] \to \mathbf{C}$ are integrable[2]. We define their *inner product* to be the complex number[3]

$$(f, g) = \int_a^b f(x)\overline{g(x)}\,dx.$$

[1] In class I used a slightly different matrix, $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Either matrix can be used to make the points here.

[2] 'Integrable' for us means that their Riemann integral exists, which in particular means that they are bounded. It is enough to think of continuous or piecewise continuous functions for the moment.

[3] The integral here was motivated in the notes from last Thursday's lecture. The complex conjugate here can be motivated by observing that if $z = \alpha + i\beta$ is a complex number, then $z\overline{z} = \alpha^2 + \beta^2$, which is the square of the distance from $(0, 0)$ to (α, β) in the plane; in other words, $\sqrt{z\overline{z}}$ represents the length of z when considered as a vector in the plane.

We also define the L^2 *norm*[4] of f to be

$$\|f\| = \left(\int_a^b |f(x)|^2 \, dx \right)^{\frac{1}{2}},$$

where $|f(x)| = \sqrt{f(x)\overline{f(x)}}$ is the modulus of the complex number $f(x)$.

PROPERTIES. The inner product defined above satisfies the following properties (see the review sheet on linear algebra):

1. $(\alpha f + \beta g, h) = \alpha(f,h) + \beta(g,h)$ for all $\alpha, \beta \in \mathbf{C}$ and all integrable f, g, h.
2. $(f,g) = \overline{(g,f)}$ for all integrable f, g.
3. $(f,f) \geq 0$ for all integrable f, and $(f,f) = 0$ if and only if f is zero except on a set of content zero[5].

These can be proved directly from the definition of the inner product; for example, the first property may be proved as follows:

$$\begin{aligned}(\alpha f + \beta g, h) &= \int_a^b (\alpha f(x) + \beta g(x))\overline{h(x)} \, dx \\ &= \int_a^b \alpha f(x)\overline{h(x)} + \beta g(x)\overline{h(x)} \, dx \\ &= \alpha \int_a^b f(x)\overline{h(x)} \, dx + \beta \int_a^b g(x)\overline{h(x)} \, dx = \alpha(f,h) + \beta(g,h).\end{aligned}$$

As noted in the review sheet on linear algebra, the first and second properties show that $(f, \alpha g + \beta h) = \overline{\alpha}(f,h) + \overline{\beta}(g,h)$ (this can also be shewn more simply directly). We see that the inner product is *linear* in the first argument and *conjugate linear* in the second.

BESSEL'S INEQUALITY. The inner product also satisfies the following: suppose that $\{e_1, e_2, \ldots\}$ is a collection of pairwise orthonormal integrable functions; in other words, that

$$(e_i, e_j) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}.$$

Let f be any integrable function. Then we have

$$\sum_i |(f, e_i)|^2 \leq \|f\|^2. \tag{2}$$

Intuitively, this may be understood as follows. If our above inner product works as desired, the quantities (f, e_i) will be the coefficients in the expansion of f in terms of the e_i, or what amounts to the same thing, the (scalar) projections of f along the e_i; the above relation says that the sum of the squares of these projections can be no greater than the square of the length of the function f itself. This can be understood by considering the example in $\mathbf{R}^2$ of $\mathbf{e}_1 = \mathbf{i}$, $\mathbf{e}_2 = \mathbf{j}$, $\mathbf{x} = \mathbf{i} + 2\mathbf{j} - \mathbf{k}$; in this case the left-hand expression will be $1^2 + 2^2 = 5$ while the right-hand will be $1^2 + 2^2 + (-1)^2 = 6$.

This example suggests something else: note that the inequality is strict because we did not include enough vectors in our set – had we also defined $\mathbf{e}_3 = \mathbf{k}$, then the left-hand side would have become $1^2 + 2^2 + (-1)^2 = 6$, the same as the right-hand side. Thus perhaps equality in (2) holds exactly when our collection $\{e_1, e_2, \ldots\}$ has 'enough functions' in some sense – enough to write f as a series in the e_i. This turns out to be quite close to the truth, as we shall discuss on Thursday.

[4]It is possible to define L^p norms for all $p \geq 1$. These are important in advanced analysis but we shall not need them here.

Looking forward. Our goal, as stated multiple times, is to understand how to expand arbitrary functions in series of suitable orthonormal sets of functions. We have seen that the trigonometric functions on appropriate intervals give orthogonal sets, and it turns out that basically all functions we shall be interested in dealing with can be expanded in series of trigonometric functions; these are called *Fourier series* and are the topic of chapter 1 in the textbook. However, later on in the course we shall be interested in series in more general orthogonal sets, such as those arising from Bessel functions, Legendre polynomials, and spherical harmonics. It turns out that all of these arise as solutions to ordinary differential equations obtained by separating variables for one of our standard partial differential equations (Laplace's equation, the heat equation, and the wave equation) in different coordinate systems. In particular, the trigonometric functions arise from separating variables for Laplace's equation in rectangular coordinates, the Bessel functions arise when doing so in cylindrical coordinates, and the Legendre polynomials and spherical harmonics arise when doing so in spherical polar coordinates. Thus we shall first take some time to write down Laplace's equation in these three coordinate systems and discuss the kinds of ordinary differential equations which arise when looking for their separated solutions. This will lead us to the topic of *Sturm-Liouville problems*, and general expansions in terms of eigenfunctions of so-called *self-adjoint differential operators*. This will then allow us to discuss expansions in the above-mentioned functions.