b. Compute $|R_A|$ and $|L_A|$ in terms of $|A|$.

2.6.6 a. As in Exercise 2.6.5, find the matrices for the linear transformations $L_A : \text{Mat}\,(3,3) \to \text{Mat}\,(3,3)$ and $R_A : \text{Mat}\,(3,3) \to \text{Mat}\,(3,3)$ when A is a 3×3 matrix.

b. Repeat when A is an $n \times n$ matrix.

c. In the $n \times n$ case, compute $|R_A|$ and $|L_A|$ in terms of $|A|$.

2.6.7 Let V be the vector space of C^1 functions on $(0,1)$. Which of the following are subspaces of V?

a. $\{\, f \in V \mid f(x) = f'(x) + 1 \,\}$ b. $\{\, f \in V \mid f(x) = xf'(x) \,\}$

c. $\left\{\, f \in V \;\middle|\; f(x) = (f'(x))^2 \,\right\}$

Exercise 2.6.8: By "identified to $\mathbb{R}^3$ via the coefficients" we mean that

$$p(x) = a + bx + cx^2 \in P_2$$

is identified to

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Exercise 2.6.8, part c: The pattern should become clear after the first three.

2.6.8 Let P_2 be the space of polynomials of degree at most two, identified to $\mathbb{R}^3$ via the coefficients. Consider the mapping $T : P_2 \to P_2$ given by

$$T(p)(x) = (x^2 + 1)p''(x) - xp'(x) + 2p(x).$$

a. Verify that T is linear, i.e., that $T(ap_1 + bp_2) = aT(p_1) + bT(p_2)$.

b. Choose the basis of P_2 consisting of the polynomials $p_1(x) = 1$, $p_2(x) = x$, $p_3(x) = x^2$. Denote by $\Phi_{\{p\}} : \mathbb{R}^3 \to P_2$ the corresponding concrete-to-abstract linear transformation. Show that the matrix of $\Phi_{\{p\}}^{-1} \circ T \circ \Phi_{\{p\}}$ is $\begin{bmatrix} 2 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

c. Using the basis $1, x, x^2, \dots, x^n$, compute the matrices of the same differential operator T, viewed first as an operator from P_3 to P_3, then from P_4 to P_4, $\dots$, P_n to P_n (polynomials of degree at most $3, 4, \dots, n$).

Exercise 2.6.9 says that any linearly independent set can be extended to form a basis. In French treatments of linear algebra, this is called the *theorem of the incomplete basis*; it plus induction can be used to prove all the theorems of linear algebra in Chapter 2.

2.6.9 a. Let V be a finite-dimensional vector space, and let $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ be linearly independent vectors. Show that there exist $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n \in V$ such that $\mathbf{v}_1, \dots, \mathbf{v}_n$ is a basis of V.

b. Let V be a finite-dimensional vector space, and let $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ be a set of vectors that span V. Show that there exists a subset $i_1, i_2, \dots, i_m$ of $\{1, 2, \dots, k\}$ such that $\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_m}$ is a basis of V.

2.6.10 Let $A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix}$. Using the "standard" identification of $\text{Mat}\,(2,2)$ with $\mathbb{R}^4$, what is the dimension of

$$\text{Span}\,(A, B, AB, BA) \quad \text{in terms of } a \text{ and } b?$$

2.7 Eigenvectors and eigenvalues

When Werner Heisenberg discovered 'matrix' mechanics in 1925, he didn't know what a matrix was (Max Born had to tell him), and neither Heisenberg nor Born knew what to make of the appearance of matrices in the context of the atom. (David Hilbert is reported to have told them to go look for a differential equation with the same *eigenvalues, if that would make them happier. They did not follow Hilbert's well-meant advice and thereby may have missed discovering the Schrödinger wave equation.)*—M. R. Schroeder, *Mathematical Intelligencer*, Vol. 7, No. 4

In Section 2.6 we discussed the change of basis matrix. We change bases when a problem is easier in a different basis. Most often this comes down to some problem being easier in an *eigenbasis*: a basis of *eigenvectors*. Before defining the terms, let's give an example.

Equation 2.7.1: Note that for large n, the second term in equation 2.7.1 is negligible, since

$$\frac{1-\sqrt{5}}{2} \approx -0.618.$$

For instance, $(\frac{1-\sqrt{5}}{2})^{1000}$ starts with at least 200 zeros after the decimal point. But the first term grows exponentially, since

$$\frac{1+\sqrt{5}}{2} \approx 1.618.$$

Assume $n = 1000$. Using logarithms base 10 to evaluate the first term, we see that

$$\begin{aligned}\log_{10} a_{1000} &\approx \log_{10} \frac{5+\sqrt{5}}{10} + \left(1000 \times \log_{10} \frac{1+\sqrt{5}}{2}\right)\\ &\approx -.1405 + 1000 \times .20899\\ &\approx 208.85,\end{aligned}$$

so a_{1000} has 209 digits.

Example 2.7.1 (Fibonacci numbers). *Fibonacci numbers* are the numbers $1, 1, 2, 3, 5, 8, 13, \ldots$ defined by $a_0 = a_1 = 1$ and $a_{n+1} = a_n + a_{n-1}$ for $n \geq 1$. We propose to prove the formula

$$a_n = \frac{5+\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n + \frac{5-\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n. \qquad 2.7.1$$

Equation 2.7.1 is quite amazing: it isn't even obvious that the right side is an integer! The key to understanding it is the matrix equation

$$\begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_n \end{bmatrix}. \qquad 2.7.2$$

The first equation says $a_n = a_n$, and the second says $a_{n+1} = a_n + a_{n-1}$. What have we gained? We see that

$$\begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^2 \begin{bmatrix} a_{n-2} \\ a_{n-1} \end{bmatrix} = \cdots = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \qquad 2.7.3$$

This looks useful, until you start computing the powers of the matrix, and discover that you are just computing Fibonacci numbers the old way. Is there a more effective way to compute the powers of a matrix?

Certainly there is an easy way to compute the powers of a *diagonal* matrix; you just raise all the diagonal entries to that power:

$$\begin{bmatrix} c_1 & 0 & \ldots & 0 \\ 0 & c_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & c_m \end{bmatrix}^n = \begin{bmatrix} c_1^n & 0 & \ldots & 0 \\ 0 & c_2^n & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & c_m^n \end{bmatrix}. \qquad 2.7.4$$

We will see that we can turn this to our advantage. Let

$$P = \begin{bmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{bmatrix}, \text{ so } P^{-1} = \frac{1}{4\sqrt{5}}\begin{bmatrix} \sqrt{5}-1 & 2 \\ \sqrt{5}+1 & -2 \end{bmatrix} \qquad 2.7.5$$

and "observe" that if we set $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, then

$$P^{-1}AP = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \text{ is diagonal.} \qquad 2.7.6$$

Equation 2.7.6: The matrices A and $P^{-1}AP$ represent the same linear transformation, in different bases. Such matrices are said to be *conjugate*. In this case, it is much easier to compute with $P^{-1}AP$.

This has the following remarkable consequence:

$$(P^{-1}AP)^n = (P^{-1}A\underbrace{P)(P^{-1}}_{I}AP)\ldots(P^{-1}A\underbrace{P)(P^{-1}}_{I}AP) = P^{-1}A^nP, \qquad 2.7.7$$

which we can rewrite as

$$A^n = P(P^{-1}AP)^n P^{-1}. \qquad 2.7.8$$

So applying equations 2.7.4, 2.7.6, and 2.7.8 leads to

$$\overbrace{\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n}^{A^n} = \overbrace{\begin{bmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{bmatrix}}^{P} \overbrace{\begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{bmatrix}}^{(P^{-1}AP)^n} \overbrace{\begin{bmatrix} \frac{\sqrt{5}-1}{4\sqrt{5}} & \frac{2}{4\sqrt{5}} \\ \frac{\sqrt{5}+1}{4\sqrt{5}} & \frac{-2}{4\sqrt{5}} \end{bmatrix}}^{P^{-1}}$$

$$= \frac{1}{2\sqrt{5}} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n (\sqrt{5}-1) + \left(\frac{1-\sqrt{5}}{2}\right)^n (1+\sqrt{5}) & 2\left(\frac{1+\sqrt{5}}{2}\right)^n - 2\left(\frac{1-\sqrt{5}}{2}\right)^n \\ \left(\frac{1+\sqrt{5}}{2}\right)^{n+1} (\sqrt{5}-1) + \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} (\sqrt{5}+1) & 2\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - 2\left(\frac{1-\sqrt{5}}{2}\right)^{n+1} \end{bmatrix}$$

This confirms that our mysterious and miraculous equation 2.7.1 for a_n is correct: if we multiply this value of $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n$ by $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, the matrix equation 2.7.3 says that the top line of the product is a_n and indeed we get

$$a_n = \frac{5+\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n + \frac{5-\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n. \qquad 2.7.9$$

So instead of computing $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n$ to determine a_n, we only need to compute the numbers $\left(\frac{1+\sqrt{5}}{2}\right)^n$ and $\left(\frac{1-\sqrt{5}}{2}\right)^n$; the problem has been *decoupled* (see Remark 2.7.4). $\triangle$

How did we hit on the matrix P (equation 2.7.5) for the Fibonacci numbers? Clearly that choice of matrix was not random. Indeed, the columns of P are *eigenvectors* for $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, and the diagonal entries of the handy matrix $P^{-1}AP$ are their corresponding *eigenvalues*.

Definition 2.7.2 (Eigenvector, eigenvalue, multiplicity). Let V be a complex vector space and $T : V \to V$ a linear transformation. A nonzero vector $\mathbf{v}$ such that

$$T\mathbf{v} = \lambda\mathbf{v} \qquad 2.7.10$$

for some number λ is called an *eigenvector* of T. The number λ is the corresponding *eigenvalue*. The *multiplicity* of an eigenvalue λ is the dimension of the *eigenspace* $\{\, \mathbf{v} \mid T\mathbf{v} = \lambda\mathbf{v} \,\}$.

It is obvious but of great importance that if $T\mathbf{v} = \lambda\mathbf{v}$, then $T^k\mathbf{v} = \lambda^k\mathbf{v}$ (for instance, $TT\mathbf{v} = T\lambda\mathbf{v} = \lambda TV = \lambda^2\mathbf{v}$).

Definition 2.7.3 (Eigenbasis). A basis for a complex vector space V is an *eigenbasis* of V for a linear transformation T if each element of the basis is an eigenvector of T.

Definition 2.7.2: In order to have an eigenvector, a matrix must map a vector space *to itself*: $T\mathbf{v}$ is in the codomain, and $\lambda\mathbf{v}$ is in the domain. So *only square matrices can have eigenvectors*.

Often an eigenvalue is defined as a root of the characteristic polynomial (see Definition 4.8.18). Our approach to eigenvectors and eigenvalues was partially inspired by the paper "Down with determinants" by Sheldon Axler. We are not in favor of jettisoning determinants, which we use heavily in Chapters 4, 5, and 6 as a way to measure volume. But determinants are more or less uncomputable for large matrices. The procedure given later in this section for finding eigenvectors and eigenvalues is more amenable to computation. However, it is still not a practical algorithm even for medium-sized matrices; the algorithms that actually work are the QR algorithm or Jacobi's method for symmetric matrices; see Exercise 3.29.

Recall that the Greek letter λ is pronounced "lambda".

An eigenvector must be nonzero, but 0 can be an eigenvalue: this happens if and only if $\mathbf{v} \neq \mathbf{0}$ is in the kernel of T.