

Linear Algebra Notes — Lecture 9/22

Abstract Vector Spaces, Bases, Linear Maps & Fibonacci

Lecture pages 104–120

Contents

Overview	1
1 Examples: things we can add and scale (pp. 104–105)	2
2 Definition of a vector space (p. 105)	2
3 Proofs of (9)–(13) (pp. 106–107)	3
4 More examples; other number systems (pp. 107–108)	4
5 Span, basis, coordinates, dimension (pp. 108–110)	5
6 Subspaces (pp. 111–112)	6
7 Dimension (p. 113)	8
8 Linear maps (pp. 114–115)	8
9 Coordinates and isomorphisms (pp. 115–117)	9
10 Fibonacci numbers via linear algebra (pp. 117–120)	11
Corrections to the handwritten notes	14

Overview

This lecture takes everything we know about $\mathbb{R}^n$ (span, independence, basis, dimension, linear maps) and shows that it works for *any* collection of objects that can be added and scaled: matrices, polynomials, sequences, functions. The payoff at the end is a formula for the Fibonacci numbers, found by treating sequences as vectors.

Every blank and question in the handwritten notes is answered in a **blue box**; orange boxes are cautions. Typos in the original are listed at the end.

1 Examples: things we can add and scale (pp. 104–105)

A *vector* is anything we can add and scale, with the usual rules. Four families to keep in mind:

1. $m \times n$ matrices. If A, B are $m \times n$ then $2A + B$ is again $m \times n$ (add entry by entry, scale every entry).

2. Polynomials. Take $f(x) = a_0 + a_1x + \cdots + a_nx^n$ and $g(x) = b_0 + b_1x + \cdots + b_mx^m$ with, say, $m < n$. Add like coefficients; where g has no term, treat its coefficient as 0:

$$(f + g)(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_m + b_m)x^m + a_{m+1}x^{m+1} + \cdots + a_nx^n, \\ 2f = 2a_0 + 2a_1x + \cdots + 2a_nx^n.$$

Example: $(1 + 2x + x^3) + (4 - x) = 5 + x + x^3$.

3. Sequences of numbers. (x_n) means the list x_n for every integer n (on p. 105 the index runs over *all* integers; later, on p. 117, only $k \geq 0$). For instance $(2^n) = (\dots, 2^{-1}, 2^0, 2^1, 2^2, \dots)$. Operations are term by term:

$$2(x_n) + (y_n) = (2x_n + y_n).$$

4. Functions $f : X \rightarrow \mathbb{R}$ on a fixed set X (the *domain*). Operations are point by point:

$$(f + g)(x) = f(x) + g(x), \quad (2f)(x) = 2f(x).$$

Sequences are a special case: a sequence is a function from the integers to $\mathbb{R}$.

2 Definition of a vector space (p. 105)

Definition

A **vector space** is a set V (whose elements are called *vectors*) with two operations, i.e. functions:

(a) **addition** $V \times V \rightarrow V$, written $(u, v) \mapsto u + v$;

(b) **scaling** $\mathbb{R} \times V \rightarrow V$, written $(c, v) \mapsto cv$,

such that for all $u, v, w \in V$ and all $c, d \in \mathbb{R}$:

Axiom	In words
1 $u + v = v + u$	order of addition doesn't matter
2 $(u + v) + w = u + (v + w)$	grouping doesn't matter
3 there is a zero vector 0 with $v + 0 = v$ for all v	adding 0 does nothing
4 for each v there is $-v$ with $v + (-v) = 0$	every vector can be undone
5 $c(u + v) = cu + cv$	scaling distributes over vector addition
6 $(c + d)u = cu + du$	scaling distributes over number addition
7 $(cd)u = c(du)$	scaling twice = scaling by the product
8 $1v = v$	scaling by 1 does nothing

Notation for $V \times V$. $V \times V$ is the set of ordered pairs (u, v) with $u, v \in V$; $\mathbb{R} \times V$ is the set of pairs (number, vector). Saying addition is a function $V \times V \rightarrow V$ means: the sum of two vectors in V is always again in V ("closure").

Two different zeros appear below: the *number* $0 \in \mathbb{R}$ and the *zero vector* $0 \in V$. When it matters we write the zero vector as $\mathbf{0}$.

Consequences of axioms 1–8

These are not extra assumptions; they follow from 1–8:

- (9) the zero vector is unique (12) $c\mathbf{0} = \mathbf{0}$
 (10) v uniquely determines $-v$ (13) $-v = (-1)v$
 (11) $0u = \mathbf{0}$

3 Proofs of (9)–(13) (pp. 106–107)

(9) The zero vector is unique

Suppose $\mathbf{0}'$ is another vector with $v + \mathbf{0}' = v$ for all v . Then

$$\mathbf{0} \stackrel{(3) \text{ for } \mathbf{0}'}{=} \mathbf{0} + \mathbf{0}' \stackrel{(1)}{=} \mathbf{0}' + \mathbf{0} \stackrel{(3) \text{ for } \mathbf{0}}{=} \mathbf{0}'.$$

(10) v uniquely determines $-v$

Suppose v' also satisfies $v + v' = \mathbf{0}$. We show $v' = -v$:

$$-v \stackrel{(3)}{=} -v + \mathbf{0} = -v + (v + v') \stackrel{(2)}{=} (-v + v) + v' = \mathbf{0} + v' \stackrel{(1), (3)}{=} v'.$$

(We used $-v + v = v + (-v) = \mathbf{0}$ by (1) and (4).)

(11) $0u = 0$ (completing the proof in the notes)

First, using (8) then (6):

$$0u + u \stackrel{(8)}{=} 0u + 1u \stackrel{(6)}{=} (0 + 1)u = 1u \stackrel{(8)}{=} u.$$

So $0u + u = u$. Now add $-u$ to both sides:

$$(0u + u) + (-u) = u + (-u) \implies 0u + (u + (-u)) = 0 \implies 0u + 0 = 0 \implies 0u = 0.$$

(12) $c0 = 0$ (completing the proof in the notes)

Same trick, now with distributivity (5):

$$c0 + cv \stackrel{(5)}{=} c(0 + v) \stackrel{(3)}{=} cv.$$

So $c0 + cv = cv$. Add $-(cv)$ to both sides: $c0 + 0 = 0$, i.e. $c0 = 0$.

(13) $-u = (-1)u$ (completing the proof in the notes)

$$(-1)u + u \stackrel{(8)}{=} (-1)u + 1u \stackrel{(6)}{=} (-1 + 1)u = 0u \stackrel{(11)}{=} 0.$$

So $(-1)u$ is an additive inverse of u . By (10) there is only one, so $(-1)u = -u$.

The pattern. To show something equals 0 , show $X + u = u$ and then cancel u by adding $-u$.

4 More examples; other number systems (pp. 107–108)

Any subspace of $\mathbb{R}^n$ is a vector space. The 8 axioms hold in $\mathbb{R}^n$, so they hold for vectors inside a subspace too, and a subspace is closed under $+$ and scaling.

Other scalars. In the definition, $\mathbb{R}$ can be replaced by any *field*: a number system where you can add, subtract, multiply, and divide by anything nonzero, with the usual rules. Examples:

- $\mathbb{C}$, the complex numbers (next lecture);
- $\mathbb{Q}$, the rational numbers (fractions);
- $\mathbb{Z}/p\mathbb{Z}$ for p prime: the integers $\{0, 1, \dots, p-1\}$ with arithmetic “mod p ” (e.g. in $\mathbb{Z}/5\mathbb{Z}$, $3 + 4 = 2$ and $3 \cdot 4 = 2$). When p is prime every nonzero element has an inverse (in $\mathbb{Z}/5\mathbb{Z}$, $3 \cdot 2 = 1$, so $1/3 = 2$).

Modules. If the “numbers” only form a ring (you can’t always divide), the same definition gives a *module*. Examples of such rings:

- $\mathbb{Z}$, since $1/2 \notin \mathbb{Z}$;
- $\mathbb{Z}/n\mathbb{Z}$ with n not prime: in $\mathbb{Z}/6\mathbb{Z}$, $2 \cdot 3 = 0$, so 2 can’t have an inverse;
- polynomials, since $1/x$ is not a polynomial.

Modules behave less nicely (e.g. not every module has a basis), which is why this course sticks to fields.

5 Span, basis, coordinates, dimension (pp. 108–110)

These notions carry over from $\mathbb{R}^n$ with *essentially* the same definitions:

1. linear combination, subspace;
2. linear independence, **span**;
3. linear map T , image of T , kernel of T ;
4. **basis**, **coordinates**, **dimension**.

The ones in bold (green in the notes) need small changes, because a vector space may be infinite-dimensional.

Definition (span)

Let S be a set of vectors in V . Then $\text{span } S$ = all linear combinations of vectors in S . Here S need not be finite, but a linear combination is always a *finite* sum $c_1v_1 + \cdots + c_kv_k$ with $v_1, \dots, v_k \in S$.

Why finite sums only? An infinite sum like $1 + x + x^2 + x^3 + \cdots$ is not a polynomial and doesn't even have a meaning in a general vector space, where only finitely many additions make sense.

Example: even monomials (p. 109)

$S = \{1, x^2, x^4, x^6, \dots\}$, the monomials of even degree, inside the space of all polynomials. Then

$$\text{span } S = \{\text{polynomials with no odd-degree terms}\}.$$

For example $3 - x^2 + 5x^6 \in \text{span } S$, but $1 + x \notin \text{span } S$.

Definition (basis, possibly infinite)

A **basis** of V is a (possibly infinite) set of vectors $\beta \subseteq V$ such that

1. $V = \text{span } \beta$, and
2. β is independent: for any finitely many distinct $v_1, \dots, v_k \in \beta$, $c_1v_1 + \cdots + c_kv_k = \mathbf{0} \implies c_1 = \cdots = c_k = 0$.

Example. A basis of the space of all polynomials in x with real coefficients (things like $a_0 + a_1x + \cdots + a_mx^m$) is

$$\beta = \{1, x, x^2, x^3, \dots, x^n, \dots\}.$$

Notation: this space is $\mathbb{R}[x]$.

Theorem (p. 110)

$\{1, x, x^2, \dots, x^n, \dots\}$ is independent.

Proof (with the details filled in)

Suppose $c_1x^{i_1} + \dots + c_kx^{i_k} = 0$ as a polynomial, i.e. for every x , with distinct exponents $i_1 < i_2 < \dots < i_k$. We show all $c_j = 0$ by induction on k (the number of terms).

- **Divide by x^{i_1}** (allowed for $x \neq 0$):

$$c_1 + c_2x^{i_2-i_1} + \dots + c_kx^{i_k-i_1} = 0 \quad \text{for all } x \neq 0.$$

All exponents $i_j - i_1$ are ≥ 1 .

- **Set $x = 0$.** The left side is a polynomial, hence continuous, so the equation also holds at $x = 0$ (take the limit $x \rightarrow 0$). Every term except c_1 vanishes at 0, so $c_1 = 0$.
- **Induction.** Now $c_2x^{i_2} + \dots + c_kx^{i_k} = 0$ has only $k - 1$ terms, so by the induction hypothesis $c_2 = \dots = c_k = 0$. (Base case $k = 1$: $c_1x^{i_1} = 0$ for all x ; take $x = 1$ to get $c_1 = 0$.)

Corollary

$\{1, x, x^2, \dots\}$ is a basis of $\mathbb{R}[x]$, the vector space of polynomials with real coefficients.

Why: every polynomial is by definition a finite combination of these monomials, so they span; they are independent by the theorem.

6 Subspaces (pp. 111–112)

Definition

A **subspace** of a vector space V is a nonempty subset S such that

- $x, y \in S \Rightarrow x + y \in S$ (closed under addition);
- $c \in \mathbb{R}, x \in S \Rightarrow cx \in S$ (closed under scaling).

Note. Nonempty guarantees $\mathbf{0} \in S$: pick any $x \in S$; then $0x = \mathbf{0} \in S$ by rule (11). (Lay's definition says "contains $\mathbf{0}$ " directly.) A subspace of a vector space is itself a vector space: the axioms hold because they already hold in V .

Example: $\mathbb{R}_{\leq n}[x] = \text{polynomials of degree } \leq n$

Why is this a subspace of $\mathbb{R}[x]$?

- It contains $\mathbf{0}$ (the zero polynomial has no terms of degree $> n$).
- Adding two polynomials of degree $\leq n$ only adds coefficients of $1, x, \dots, x^n$, so no power

above x^n appears: the sum has degree $\leq n$.

- Scaling multiplies each coefficient by c , which can't create new powers: degree stays $\leq n$.

Q: Do polynomials of degree exactly n form a subspace? No

Adding two of them can cancel the top term:

$$(a_0 + a_1x + \cdots + a_nx^n) + (-a_nx^n) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1},$$

which has degree $< n$. Concretely, $(x^2 + x) + (-x^2) = x$. So the set is not closed under addition. (It also fails to contain the zero polynomial.)

Q: What is a basis of $\mathbb{R}_{\leq n}[x]$?

$\beta = \{1, x, x^2, \dots, x^n\}$. It spans, because every polynomial of degree $\leq n$ is $a_0 \cdot 1 + a_1x + \cdots + a_nx^n$. It is independent, as part of the independent set from p. 110.

Example: upper triangular matrices

Do upper triangular $n \times n$ matrices form a subspace of all $n \times n$ matrices? Yes

Upper triangular means every entry below the diagonal is 0. Adding two such matrices gives $0 + 0 = 0$ below the diagonal, and scaling gives $c \cdot 0 = 0$. The zero matrix is upper triangular.

Example:

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} + 5 \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 7 \\ 0 & 13 \end{bmatrix} \quad (\text{still upper triangular}).$$

Example: a basis of $M_{m \times n}(\mathbb{R})$

$M_{m \times n}(\mathbb{R})$ is the vector space of $m \times n$ real matrices.

Basis of $M_{m \times n}(\mathbb{R})$

β = the matrices with exactly one nonzero entry, equal to 1. Call them E_{ij} : the 1 sits in row i , column j . There are mn of them. Every matrix is the combination of the E_{ij} with its own entries as coefficients:

$$\begin{bmatrix} 2 & 1 & 3 \\ 5 & 1 & 7 \end{bmatrix} = 2E_{11} + 1E_{12} + 3E_{13} + 5E_{21} + 1E_{22} + 7E_{23},$$

$$\text{where } E_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \dots, E_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

They are independent: $\sum c_{ij}E_{ij}$ is the matrix with entries c_{ij} , which is 0 only if every $c_{ij} = 0$. So $\dim M_{m \times n} = mn$.

7 Dimension (p. 113)

Definition

$\dim V$ = the number of vectors in a basis of V .

This makes sense because of a theorem (as in $\mathbb{R}^n$): any two bases of V have the same number of elements.

Space	Basis	Dimension
$\mathbb{R}_{\leq n}[x]$	$\{1, x, \dots, x^n\}$	$n + 1$ (count the powers $x^0, \dots, x^n$)
$\mathbb{R}[x]$	$\{1, x, x^2, \dots\}$	infinite
$M_{m \times n}(\mathbb{R})$	$\{E_{ij}\}$	mn
upper triangular $n \times n$	$\{E_{ij} : i \leq j\}$	$\frac{n^2 + n}{2}$

Why upper triangular has dimension $\frac{n^2+n}{2}$

The free entries are those on or above the diagonal:

$$\begin{bmatrix} a_{11} & \cdots & * \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{bmatrix}.$$

There are n diagonal entries. Of the remaining $n^2 - n$ off-diagonal entries, by symmetry half are above the diagonal: $\frac{n^2 - n}{2}$. Total:

$$n + \frac{n^2 - n}{2} = \frac{n^2 + n}{2} = \frac{n(n + 1)}{2}.$$

Check: $n = 2$ gives 3 (entries a_{11}, a_{12}, a_{22}); $n = 3$ gives 6.

8 Linear maps (pp. 114–115)

Definition

A map $T : V \rightarrow W$ between vector spaces is **linear** if

$$T(c_1 v_1 + c_2 v_2) = c_1 T(v_1) + c_2 T(v_2) \quad \text{for all } v_1, v_2 \in V, c_1, c_2 \in \mathbb{R}.$$

Kernel and image are defined as before: $\text{Ker } T = \{v \in V : T(v) = \mathbf{0}\}$ and $\text{Im } T = \{T(v) : v \in V\}$.

Transpose: $T : M_{n \times n}(\mathbb{R}) \rightarrow M_{n \times n}(\mathbb{R})$, $T(A) = A^T$, is linear

The question mark in the notes asks whether $(c_1A_1 + c_2A_2)^T = c_1A_1^T + c_2A_2^T$. Yes: transposing only moves entries (the (i, j) entry of A^T is the (j, i) entry of A). So the (i, j) entry of both sides is $c_1(A_1)_{ji} + c_2(A_2)_{ji}$.

Derivative: $T : \mathbb{R}[x] \rightarrow \mathbb{R}[x]$, $T(f) = f'$, is linear

$(c_1f_1 + c_2f_2)' = c_1f_1' + c_2f_2'$ by the sum and constant-multiple rules of calculus.

Kernel: $\text{Ker } T = \{f \in \mathbb{R}[x] : f' = 0\}$ = the constant polynomials (the only polynomials with zero derivative).

Image: all of $\mathbb{R}[x]$, since every polynomial has an antiderivative that is a polynomial.

So T is onto but not one-to-one. That is impossible for a map from a finite-dimensional space to itself (rank–nullity), and it shows $\mathbb{R}[x]$ is infinite-dimensional.

Integral: $T : \mathbb{R}[x] \rightarrow \mathbb{R}[x]$, $T(f) = \int_c^x f(t) dt$ (c a fixed number), is linear

By linearity of the integral:

$$T(c_1f_1 + c_2f_2) = \int_c^x (c_1f_1(t) + c_2f_2(t))dt = c_1 \int_c^x f_1(t) dt + c_2 \int_c^x f_2(t) dt = c_1T(f_1) + c_2T(f_2).$$

(The notes write $\int_c^x f(x) dx$; the integration variable should get its own name, t , since x is already the upper limit.)

Extra facts: this T is one-to-one, because $(T(f))' = f$ by the fundamental theorem of calculus, so $T(f) = 0$ forces $f = 0$. Its image is the polynomials vanishing at $x = c$. Here T is one-to-one but not onto: again only possible in infinite dimensions.

9 Coordinates and isomorphisms (pp. 115–117)

Definition

A vector space with a finite basis is **finite dimensional**.

Theorem (coordinate map)

Let $\beta = \{v_1, \dots, v_n\}$ be a basis of V , and let $T : V \rightarrow \mathbb{R}^n$ be the *coordinate map* $T(x) = [x]_\beta$, that is,

$$T(c_1v_1 + \dots + c_nv_n) = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Then T is linear and invertible.

Proof (p. 117, completed)

Well defined. Every $x \in V$ is a combination of the v_i (they span), and in only one way: if $\sum c_i v_i = \sum d_i v_i$ then $\sum (c_i - d_i) v_i = \mathbf{0}$, so $c_i = d_i$ by independence.

Addition.

$$(c_1 v_1 + \cdots + c_n v_n) + (d_1 v_1 + \cdots + d_n v_n) = (c_1 + d_1) v_1 + \cdots + (c_n + d_n) v_n,$$

which T sends to

$$\begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = T(\text{first}) + T(\text{second}).$$

Scaling. $a(c_1 v_1 + \cdots + c_n v_n) = (ac_1) v_1 + \cdots + (ac_n) v_n \mapsto a(c_1, \dots, c_n)^T$.

Inverse. $T^{-1}(c_1, \dots, c_n) = c_1 v_1 + \cdots + c_n v_n$; applying T then T^{-1} , or the reverse, gives back what you started with.

Definitions

A linear invertible map is a **linear isomorphism**. If $T : V \rightarrow W$ is a linear isomorphism, V and W are **isomorphic**.

Example (p. 116). $\mathbb{R}_{\leq n}[x]$ is isomorphic to $\mathbb{R}^{n+1}$, via the coordinate map for $\beta = \{1, x, \dots, x^n\}$:

$$a_0 + a_1 x + \cdots + a_n x^n \longmapsto (a_0, a_1, \dots, a_n)^T.$$

Key point. In isomorphic vector spaces linear algebra works the same way: an isomorphism matches up sums with sums and scalar multiples with scalar multiples. So spans, independence, bases and dimensions correspond exactly. Questions about polynomials can be answered by computing in $\mathbb{R}^{n+1}$.

Theorem (p. 116)

If $T : V \rightarrow W$ is linear and invertible, then $T^{-1} : W \rightarrow V$ is linear.

Proof

Let $w_1, w_2 \in W$, and put $v_1 = T^{-1}(w_1)$, $v_2 = T^{-1}(w_2)$, so $T(v_1) = w_1$ and $T(v_2) = w_2$. By linearity of T ,

$$T(c_1 v_1 + c_2 v_2) = c_1 w_1 + c_2 w_2.$$

Apply T^{-1} to both sides:

$$T^{-1}(c_1 w_1 + c_2 w_2) = c_1 v_1 + c_2 v_2 = c_1 T^{-1}(w_1) + c_2 T^{-1}(w_2).$$

(The notes argue the same thing by contradiction: if the two sides were different, applying the one-to-one map T would give different results, but both give $c_1 w_1 + c_2 w_2$.)

10 Fibonacci numbers via linear algebra (pp. 117–120)

The sequence. $(f_k)_{k \geq 0}$ with $f_0 = 0$, $f_1 = 1$ and $f_{k+2} = f_k + f_{k+1}$:

$$f_2 = 0 + 1 = 1, \quad f_3 = 1 + 1 = 2, \quad f_4 = 1 + 2 = 3, \quad f_5 = 5, \quad f_6 = 8, \dots$$

Goal: a formula for f_k as a function of k , without computing all earlier terms.

Setup. $\mathcal{S}$ = the vector space of all real sequences $(x_k)_{k \geq 0}$, with term-by-term operations (Section 1).

Definition

$\text{Fib} = \{(x_k) \in \mathcal{S} : x_{k+2} = x_k + x_{k+1} \text{ for all } k \geq 0\}$.

Fib is a subspace of $\mathcal{S}$

Scaling (as in the notes): if $(x_k) \in \text{Fib}$ then $2(x_k) = (2x_k)$, and $2x_{k+2} = 2x_k + 2x_{k+1}$ (multiply the recursion by 2). So yes; the same works for any c .

Addition: if $(x_k), (y_k) \in \text{Fib}$, add the two recursions: $x_{k+2} + y_{k+2} = (x_k + y_k) + (x_{k+1} + y_{k+1})$, so $(x_k + y_k) \in \text{Fib}$.

Zero: the zero sequence satisfies $0 = 0 + 0$.

The map to $\mathbb{R}^2$.

$$T : \text{Fib} \rightarrow \mathbb{R}^2, \quad (x_k) \mapsto \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}.$$

Why T is linear and invertible

Linear: it just reads off two entries, and sequence operations are term by term.

Invertible: the first two terms determine everything, since $x_2 = x_0 + x_1$, $x_3 = x_1 + x_2$, and so on. So $T^{-1}(a, b)$ is the unique sequence in Fib starting a, b . Hence $\dim \text{Fib} = \dim \mathbb{R}^2 = 2$.

A basis of Fib: $\{T^{-1}(e_1), T^{-1}(e_2)\}$

An isomorphism sends a basis to a basis, and $\{e_1, e_2\}$ is a basis of $\mathbb{R}^2$. Compute each by running the recursion:

$$T^{-1}(e_1) : x_0 = 1, x_1 = 0 \implies (1, 0, 1, 1, 2, 3, 5, 8, \dots),$$

$$T^{-1}(e_2) : x_0 = 0, x_1 = 1 \implies (0, 1, 1, 2, 3, 5, 8, \dots) = (f_k), \text{ the Fibonacci sequence itself.}$$

(The 9/22 version of the notes writes $T^{-1}(e_1) = (1, 0, 1, 0, 1, 0, \dots)$; that's wrong since $x_3 = x_1 + x_2 = 0 + 1 = 1$. The 9/24 version corrects it.)

A better basis: geometric sequences (p. 119)

Look for solutions of the form $(r^k)_{k \geq 0}$ with $r \neq 0$. Plug into the recursion:

$$r^{k+2} = r^k + r^{k+1} \iff r^2 - r - 1 = 0 \iff r_{\pm} = \frac{1 \pm \sqrt{1^2 - 4(-1)}}{2} = \frac{1 \pm \sqrt{5}}{2}.$$

(Dividing by r^k is allowed because $r \neq 0$.) So (r_+^k) and (r_-^k) are in Fib, with $r_+ \approx 1.618$ (the golden ratio) and $r_- \approx -0.618$.

Are they independent? Yes

Two vectors are independent exactly when neither is a multiple of the other. If $(r_+^k) = c(r_-^k)$ for all k , then $k = 0$ gives $1 = c$, and $k = 1$ gives $r_+ = c r_- = r_-$, which is false. So $(r_+^k) \neq c(r_-^k)$ for every c , and the same argument works the other way round. Two independent vectors in the 2-dimensional space Fib form a basis.

Intuition: why not stop at the basis $T^{-1}(e_1), T^{-1}(e_2)$?

That basis is perfectly valid, but it doesn't give us anything new. The goal (p. 117) was a *formula* for f_k : plug in $k = 100$ and get the answer without computing the 99 terms before it. Look at what the first basis is:

$$T^{-1}(e_1) = (1, 0, 1, 1, 2, 3, 5, \dots), \quad T^{-1}(e_2) = (0, 1, 1, 2, 3, 5, 8, \dots) = (f_k).$$

In this basis $(f_k) = 0 \cdot T^{-1}(e_1) + 1 \cdot T^{-1}(e_2)$. That's true but circular: it says "Fibonacci = Fibonacci." Neither basis sequence has a formula for its k -th term; the only way to get their terms is to run the recursion, which is what we wanted to avoid.

Geometric sequences (r^k) are different: their k -th term is just a number raised to the k -th power. So once (f_k) is written as a combination of two of them, a formula for f_k comes for free. Every basis can represent every sequence in Fib. We switch bases to get one whose vectors we can compute directly.

Where $r^2 - r - 1 = 0$ comes from. It isn't a new representation of the sequences. It is the *test* for which geometric sequences lie in Fib: plugging r^k into the recursion and dividing by r^k leaves $r^2 = r + 1$, and only two values of r pass.

The first step wasn't wasted. The isomorphism T is what tells us $\dim \text{Fib} = 2$. Without it, finding two geometric solutions wouldn't show they account for *every* Fibonacci-like sequence. Knowing the dimension is 2, two independent solutions must form a basis. So:

1. the isomorphism tells you how big the space is;
2. the geometric sequences give a basis you can compute with;
3. solving for $c_{\pm}$ (below) is a change of coordinates: (f_k) has coordinates $(0, 1)$ in the first basis and $(\frac{1}{\sqrt{5}}, -\frac{1}{\sqrt{5}})$ in the second.

This is the same idea as eigenvectors later: on an eigenvector, A^k is just multiplication by λ^k , just as each geometric sequence advances by multiplying by r .

Find the coordinates of (f_k) . Write $(f_k) = c_+(r_+^k) + c_-(r_-^k)$ and use the two known terms:

$$0 = f_0 = c_+ r_+^0 + c_- r_-^0 = c_+ + c_-, \quad 1 = f_1 = c_+ r_+ + c_- r_-.$$

Solving for c_+, c_- (the step the notes leave open)

The first equation gives $c_- = -c_+$. Substituting into the second: $c_+(r_+ - r_-) = 1$. Since $r_+ - r_- = \frac{1+\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2} = \sqrt{5}$,

$$c_+ = \frac{1}{\sqrt{5}}, \quad c_- = -\frac{1}{\sqrt{5}}.$$

Result (Binet's formula)

$$f_k = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^k - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2} \right)^k.$$

Check: $k = 2$: $\frac{r_+^2 - r_-^2}{\sqrt{5}} = \frac{(r_+ - r_-)(r_+ + r_-)}{\sqrt{5}} = \frac{\sqrt{5} \cdot 1}{\sqrt{5}} = 1 = f_2$. ✓

Growth rate (p. 120)

Why $f_k/r_+^k \rightarrow c_+$

Divide the formula by r_+^k :

$$\frac{f_k}{r_+^k} = c_+ + c_- \left(\frac{r_-}{r_+} \right)^k.$$

Since $|r_-/r_+| \approx 0.618/1.618 \approx 0.38 < 1$, the second term goes to 0. So $f_k/r_+^k \rightarrow c_+ = 1/\sqrt{5}$.

So f_k grows like $c_+ r_+^k = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^k$. In fact, because $|c_- r_-^k| < \frac{1}{2}$, f_k is exactly that number rounded to the nearest integer. For example $\frac{1}{\sqrt{5}}(1.618)^{10} \approx 55.004$ and $f_{10} = 55$.

Intuition: why this limit means f_k grows exponentially

The limit is a statement about a *ratio*. Saying $f_k/r_+^k \rightarrow c_+$ means that, for large k ,

$$f_k \approx c_+ r_+^k,$$

with the approximation getting better and better in relative terms (equivalently, $f_k/(c_+ r_+^k) \rightarrow 1$). So f_k and $c_+ r_+^k$ are essentially the same size. The right-hand side is exponential: each step multiplies it by $r_+ \approx 1.618$. Hence f_k grows exponentially too, at that rate.

k	$c_+ r_+^k$	f_k
10	≈ 55.004	55
20	≈ 6765.00003	6765

What other limits would have meant.

- If f_k grew *slower* than r_+^k (like k^2 , or like 1.5^k), the ratio would go to 0.

- If f_k grew *faster* (like 2^k), the ratio would go to ∞ .
- The ratio settling on a positive, finite number is exactly what says f_k grows at the *same rate* as r_+^k , no faster and no slower.

Side consequence: $f_{k+1}/f_k \rightarrow r_+$. Consecutive Fibonacci numbers have ratios approaching the golden ratio, e.g. $6765/4181 \approx 1.618$.

Corrections to the handwritten notes

Page	As written	Should be	Why
108	span S = combinations of vectors in V	... of vectors in S	span is built from S
110	$i_1 \leq i_2 \leq \dots \leq i_k$	$i_1 < i_2 < \dots < i_k$	exponents must be distinct
111	subspace: closed under + and scaling	...and nonempty (or contains $\mathbf{0}$)	the empty set would otherwise qualify
114	$T(A) = A^J$	$T(A) = A^T$	transpose
115	$\int_c^x f(x) dx$	$\int_c^x f(t) dt$	x is already the upper limit
118	$T^{-1}(e_1)$ $(1, 0, 1, 0, 1, 0, \dots)$	$= (1, 0, 1, 1, 2, 3, 5, \dots)$	$x_3 = x_1 + x_2 = 1$ (fixed in 9/24 notes)