Linear operators
(a.k.a. linear transformations, linear maps, …)

Basics *(Secs. 15–16, parts of 22–23)*

Definition: If $\mathcal{V}$ and $\mathcal{U}$ are vector spaces (same $\mathcal{F}$), a function $\underline{A} : \mathcal{V} \to \mathcal{U}$ is *linear* if

$$\forall \vec{v}_1,\ \vec{v}_2 \in \mathcal{V},\ \forall \lambda \in \mathcal{F} :\ \underline{A}(\lambda \vec{v}_1 + \vec{v}_2) = \lambda \underline{A}(\vec{v}_1) + \underline{A}(\vec{v}_2).$$

DEFINITION:

$\mathcal{V}$ = *domain* of $\underline{A}$ = dom $\underline{A}$.

$\mathcal{U}$ = *codomain* of $\underline{A}$.

REMARKS:

(1) Exactly as for "subspace", there are various equivalent definitions of "linear"; e.g.,

$$\underline{A}\left(\sum_{j=1}^{N} \lambda^j \vec{v}_j\right) = \sum_{j=1}^{N} \lambda^j \underline{A}(\vec{v}_j).$$

(2) It's traditional to omit parentheses: $\underline{A}\vec{v} \equiv \underline{A}(\vec{v})$.

(3) Special case: $\mathcal{U} = \mathcal{F}$. Then $\underline{A}$ is called a *linear functional*. [Postponed to Chap. 7.]

INFINITE-DIMENSIONAL EXAMPLES
[Cf. Milne pp. 37–40]

1. $\mathcal{V}$ = polynomials = $\mathcal{U}$. $\underline{A} = \dfrac{d}{dt}$.

2. $\mathcal{V}$ = (space of integrable functions on $[0,1]$) $\equiv \mathcal{L}^1(0,1)$.
$\underline{A}f \equiv \int_0^1 f(x)\,dx$ defines a linear functional on $\mathcal{V}$. (So does $\int_0^1 f(x)\,w(x)\,dx$ for suitable fixed w.)

3. $\mathcal{V}$ = (space of continuous functions on $\mathbf{R}$) $\equiv \mathcal{C}^0(-\infty,\infty)$. $\underline{A}f \equiv f(0)$.

4. Old example: $\underline{A} = +\dfrac{d^2}{dt^2} + \omega^2$.
A suitable domain is $\mathcal{V} = \mathcal{C}^2(0,1)$; then $\mathcal{U} = \mathcal{C}^0(0,1)$ (or larger).

$$\mathcal{C}^2(0,1) \equiv \{\, f : f,\ f',\ f'' \text{ all continuous on } (0,1)\}.$$

We may wish to restrict the domain by imposing boundary conditions, such as

$$f(0) = 0 = f(1). \tag{C$_1$}$$

(For now we consider only *homogeneous* BC; hence we get a subspace $\mathcal{V}_1$ of $\mathcal{C}^2(0,1)$.)
Motivations:

(A) BC may make solution of $\underline{A}f = g$ *unique* (for $g \in \mathcal{U}$). See first exercise on next assignment. This part of the ODE problem is pure linear algebra (unlike the second exercise).

(B) If f_1 and $f_2 \in \mathcal{V}_1$, then

$$f_1 \cdot (\underline{A}f_2) = (\underline{A}f_1) \cdot f_2 \quad \text{— i.e.,}$$

$$\int_0^1 f_1(t)\, \overline{[f_2''(t) + \omega^2 f_2(t)]}\, dt = \int_0^1 [f_1''(t) + \omega^2 f_1(t)]\, \overline{f_2(t)}\, dt.$$

[Proof: Integrate by parts. All endpoint terms vanish because of (C$_1$).] This property is closely related to the facts which you will verify by brute force in the *second* exercise of the assignment. Its significance will become clearer later.

REMARK: Different BC — e.g.,

$$f'(0) = 0 = f'(1) \tag{C$_2$}$$

define a different domain, hence a *different operator.* (Its eigenfunctions are cosines instead of sines.)

MATRIX FORMULATION

Let $\mathcal{V}$ and $\mathcal{U}$ be finite-dimensional. Choose bases:

$$\{\vec{v}_k\}_{k=1}^n \subset \mathcal{V}, \qquad \{\vec{u}_j\}_{j=1}^m \subset \mathcal{U}.$$

Now $\underline{A}\vec{v}_k \in \operatorname{span}\{\vec{u}_j\}$, so $\underline{A}\vec{v}_k = \sum_{j=1}^m A^j{}_k \vec{u}_j$ for some numbers $A^j{}_k$. For $\vec{v} = \sum_{k=1}^n \lambda^k \vec{v}_k \in \mathcal{V}$, calculate $\underline{A}\vec{v}$ by linearity:

$$\underline{A}\Bigl(\sum_k \lambda^k \vec{v}_k\Bigr) = \sum_k \lambda^k \underline{A}\vec{v}_k = \sum_k \lambda^k \sum_j A^j{}_k \vec{u}_j$$
$$= \sum_{j=1}^m \Bigl(\sum_{k=1}^n A^j{}_k \lambda^k\Bigr) \vec{u}_j\,.$$

Thus

$$\lambda^k \mapsto \sum_{k=1}^{n} A^j{}_k \lambda^k \equiv A^j{}_k \lambda^k \qquad (\dagger)$$

is the coordinate expression of the mapping $\underline{A} : \mathcal{V} \to \mathcal{U}$. ($\dagger$) is a linear map of $\mathbf{R}^n \to \mathbf{R}^m$. In matrix notation, it is $\lambda \mapsto A\lambda$. Here I abuse notation by equating $\underline{A}$ with its matrix; one must keep in mind that *the matrix depends on the bases chosen.* (I shall try to use an underline for operators and no underline for matrices.)

The correspondence between operators and matrices is one-to-one and onto (see an exercise of the current assignment).

COROLLARY 1. *The linear maps from $\mathbf{R}^n$ to $\mathbf{R}^m$ (or $\mathbf{C}^n$ to $\mathbf{C}^m$) are precisely the functions given by homogeneous first-degree formulas:*

$$y^j = \sum_{k=1}^{n} a^j{}_k x^k \quad (j = 1, \dots, m).$$

PROOF: Choose the natural bases [e.g., $(\vec{e}_j)^i = \delta^i_j$]. Then each vector $\vec{v}$ *is* its coordinate representation, the column vector λ. We have a special case of the situation just discussed at length.

COROLLARY 2. *The matrix of a composite linear operator*

$$\underline{A} \circ \underline{B} \qquad (\underline{B} : \mathcal{V}_1 \to \mathcal{V}_2, \quad \underline{A} : \mathcal{V}_2 \to \mathcal{V}_3)$$

is the product, AB, of the corresponding matrices.

PROOF: Substitute $(\dagger)_B$ into $(\dagger)_A$ and interchange sums.

COROLLARY 3. *The kth column of the matrix A is the coordinate representation of $\underline{A}\vec{v}_k$ (where $\vec{v}_k$ is the kth element of the basis used for $\mathcal{V}$).*

COMMENTS ON NOTATION: There are two conventions, each with its advantages. (Here I introduce, in both cases, the famous or notorious *summation convention*, that the $\sum$ may be omitted when an index is *repeated*, since this almost always is associated with a sum over the natural range of the index.)

1. All indices "down". The fundamental convention is **row before column**. (An $m \times n$ matrix has m rows and n columns, and the row index comes first on the symbol A_{jk} .) In normal matrix multiplication, *adjacent* indices are summed:

$$(AB)_{jl} = A_{jk} B_{kl}\,; \qquad (A\lambda)_j = A_{jk}\lambda_k\,.$$

2. Basis index down, coordinate index up: $\vec{v} = \lambda^k \vec{v}_k$. Therefore the matrix is $A^j{}_k$. Sum over any *repeated up-down pair*:

$$(AB)^j{}_l = A^j{}_k B^k{}_l\,; \qquad (A\lambda)^j = A^j{}_k \lambda^k.$$

   Unfortunately, the order of indices tends to get lost in typesetting.

   Simmonds provides the mnemonic

$$\text{Row} = \text{Roof}, \qquad \text{Column} = \text{Cellar}.$$

Inner product case: If $\mathcal{U}$ is an inner product space and $\{\hat{u}_j\}$ is an *ON* basis, then (in the real case)

$$A^j{}_k = \hat{u}_j \cdot (\underline{A}\vec{v}_k) \equiv \langle j|\underline{A}|k\rangle \text{ or } \langle \hat{u}_j, \underline{A}\vec{v}_k\rangle.$$

In the complex case, in the "mathematicians' notation" $A^j{}_k$ equals $(\underline{A}\vec{v}_k) \cdot \hat{u}_j$.

Remark: All of this algebraic formalism also applies to infinite-dimensional Hilbert space. It should be noted, however, that in ∞-dim. vector spaces, matrices tend to lose their usefulness, for two reasons:

1. Explicit calculations become difficult; abstract methods gain a comparative advantage.

2. Linear operators must satisfy special conditions to make formal manipulations valid for infinite sums:

   (a) $\underline{A}(\lambda^j \vec{v}_j) = \lambda^j (\underline{A}\vec{v}_j)$? Only if $\underline{A}$ is continuous (alias *bounded*).

   (b) Even if $\underline{A}$ and $\underline{B}$ are bounded, the interchange needed to prove Corollary 2 might be invalid; $A^j{}_k B^k{}_l$ might not converge. Even if true, this may be hard to prove.