From "Differential Equations, Dynamical Systems, and Linear Algebra" by Hirsch + Smale 1974

Chapter 5

Linear Systems and Exponentials of Operators

The object of this chapter is to solve the linear homogeneous system with constant coefficients

(1) $$x' = Ax,$$

where A is an operator on $\mathbf{R}^n$ (or an $n \times n$ matrix). This is accomplished with exponentials of operators.

This method of solution is of great importance, although in this chapter we can compute solutions only for special cases. When combined with the operator theory of Chapter 6, the exponential method yields explicit solutions for every system (1).

For every operator A, another operator e^A, called the *exponential of* A, is defined in Section 4. The function $A \to e^A$ has formal properties similar to those of ordinary exponentials of real numbers; indeed, the latter is a special case of the former. Likewise the function $t \to e^{tA}$ $(t \in \mathbf{R})$ resembles the familiar e^{ta}, where $a \in \mathbf{R}$. In particular, it is shown that the solutions of (1) are exactly the maps $x\colon \mathbf{R} \to \mathbf{R}^n$ given by

$$x(t) = e^{tA}K \qquad (K \in \mathbf{R}^n).$$

Thus we establish existence and uniqueness of solution of (1); "uniqueness" means that there is only one solution $x(t)$ satisfying a given initial condition of the form $x(t_0) = K_0$.

Exponentials of operators are defined in Section 3 by means of an infinite series in the operator space $L(\mathbf{R}^n)$; the series is formally the same as the usual series for e^a. Convergence is established by means of a special norm on $L(\mathbf{R}^n)$, the *uniform norm*. Norms in general are discussed in Section 2, while Section 1 briefly reviews some basic topology in $\mathbf{R}^n$.

Sections 5 and 6 are devoted to two less-central types of differential equations. One is a simple inhomogeneous system and the other a higher order equation of one variable. We do not, however, follow the heavy emphasis on higher order equations

of some texts. In geometry, physics, and other kinds of applied mathematics, one seldom encounters naturally any differential equation of order higher than two. Often even the second order equations are studied with more insight after reducing to a first order system (for example, in Hamilton's approach to mechanics).

§1. Review of Topology in $\mathbf{R}^n$

The inner product ("dot product") of vectors x and y in $\mathbf{R}^n$ is

$$\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n.$$

The Euclidean norm of x is $|x| = \langle x, x\rangle^{1/2} = (x_1^2 + \cdots + x_n^2)^{1/2}$. Basic properties of the inner product are

Symmetry: $\langle x, y\rangle = \langle y, x\rangle$;

Bilinearity: $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$,

$\langle \alpha x, y\rangle = \alpha\langle x, y\rangle, \quad a \in \mathbf{R}$;

Positive definiteness: $\langle x, x\rangle \geq 0$ and

$\langle x, x\rangle = 0$ if and only if $x = 0$.

An important inequality is

Cauchy's inequality: $\langle x, y\rangle \leq |x|\,|y|$.

To see this, first suppose $x = 0$ or $y = 0$; the inequality is obvious. Next, observe that for any λ

$$\langle x + \lambda y, x + \lambda y\rangle \geq 0$$

or

$$\langle x, x\rangle + \lambda^2\langle y, y\rangle + 2\lambda\langle x, y\rangle \geq 0.$$

Writing $-\langle x, y\rangle/\langle y, y\rangle$ for λ yields the inequality.

The basic properties of the norm are:

(1) $|x| \geq 0$ and $|x| = 0$ if and only if $x = 0$;
(2) $|x + y| \leq |x| + |y|$;
(3) $|\alpha x| = |\alpha|\,|x|$;

where $|\alpha|$ is the ordinary absolute value of the scalar α. To prove the triangle inequality (2), it suffices to prove

$$|x + y|^2 \leq |x|^2 + |y|^2 + 2|x|\,|y|.$$

Since

$$|x + y|^2 = \langle x + y, x + y\rangle = |x|^2 + |y|^2 + 2\langle x, y\rangle,$$

this follows from Cauchy's inequality.

Geometrically, $|x|$ is the length of the vector x and

$$\langle x, y\rangle = |x|\,|y| \cos\theta,$$

where θ is the angle between x and y.

The *distance* between two points $x, y \in \mathbf{R}^n$ is defined to be $|x - y| = d(x, y)$. It is easy to prove:

(4) $|x - y| \geq 0$ and $|x - y| = 0$ if and only if $x = y$;
(5) $|x - z| \leq |x - y| + |y - z|$.

The last inequality follows from the triangle inequality applied to

$$x - z = (x - y) + (y - z).$$

If $\epsilon > 0$ the *ϵ-neighborhood* of $x \in \mathbf{R}^n$ is

$$B_\epsilon(x) = \{y \in \mathbf{R}^n \mid |y - x| < \epsilon\}.$$

A *neighborhood* of x is any subset of $\mathbf{R}^n$ containing an ϵ-neighborhood of x.

A set $X \subset \mathbf{R}^n$ is *open* if it is a neighborhood of every $x \in X$. Explicitly, X is open if and only if for every $x \in X$ there exists $\epsilon > 0$, depending on x, such that

$$B_\epsilon(x) \subset X.$$

A sequence $\{x_k\} = x_1, x_2, \ldots$ in $\mathbf{R}^n$ *converges to the limit* $y \in \mathbf{R}^n$ if

$$\lim_{k\to\infty} |x_k - y| = 0.$$

Equivalently, every neighborhood of y contains all but a finite number of the points of the sequence. We denote this by $y = \lim_{k\to\infty} x_k$ or $x_k \to y$. If $x_k = (x_{k1}, \ldots, x_{kn})$ and $y = (y_1, \ldots, y_n)$, then $\{x_k\}$ converges to y if and only if $\lim_{k\to\infty} x_{kj} = y_j$, $j = 1, \ldots, n$. A sequence that has a limit is called *convergent*.

A sequence $\{x_k\}$ in $\mathbf{R}^n$ is a *Cauchy* sequence if for every $\epsilon > 0$ there exists an integer k_0 such that

$$|x_j - x_k| < \epsilon \qquad \text{if } k \geq k_0 \quad \text{and} \quad j \geq k_0.$$

The following basic property of $\mathbf{R}^n$ is called metric completeness:

A sequence converges to a limit if and only if it is a Cauchy sequence.

A subset $Y \subset \mathbf{R}^n$ is *closed* if every sequence of points in Y that is convergent has its limit in Y. It is easy to see that this is equivalent to: Y is closed if the complement $\mathbf{R}^n - Y$ is open.

Let $X \subset \mathbf{R}^n$ be any subset. A map $f: X \to \mathbf{R}^m$ is *continuous* if it takes convergent sequences to convergent sequences. This means: for every sequence $\{x_k\}$ in X with

$$\lim_{k\to\infty} x_k = y \in X,$$

it is true that

$$\lim_{k\to\infty} f(x_k) = f(y).$$

A subset $X \subset \mathbf{R}^n$ is *bounded* if there exists $\alpha > 0$ such that $X \subset B_\alpha(0)$.

A subset X is *compact* if every sequence in X has a subsequence converging to a point in X. The basic *theorem of Bolzano–Weierstrass* says:

A subset of $\mathbf{R}^n$ is compact if and only if it is both closed and bounded.

Let $K \subset \mathbf{R}^n$ be compact and $f: K \to \mathbf{R}^m$ be a continuous map. Then $f(K)$ is compact.

A nonempty compact subset of $\mathbf{R}$ has a maximal element and a minimal element. Combining this with the preceding statement proves the familiar result:

Every continuous map $f: K \to \mathbf{R}$, defined on a compact set K, takes on a maximum value and a minimum value.

One may extend the notions of distance, open set, convergent sequence, and other topological ideas to vector subspaces of $\mathbf{R}^n$. For example, if E is a subspace of $\mathbf{R}^n$, the distance function $d: \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ restricts to a function $d_E: E \times E \to \mathbf{R}$ that also satisfies (4) and (5). Then ϵ-neighborhoods in E may be defined via d_E and thus open sets of E become defined.

§2. New Norms for Old

It is often convenient to use functions on $\mathbf{R}^n$ that are similar to the Euclidean norm, but not identical to it. We define a *norm* on $\mathbf{R}^n$ to be any function $N: \mathbf{R}^n \to \mathbf{R}$ that satisfies the analogues of (1), (2), and (3) of Section 1:

(1) $N(x) \geq 0$ and $N(x) = 0$ if and only if $x = 0$;
(2) $N(x + y) \leq N(x) + N(y)$;
(3) $N(\alpha x) = |\alpha| N(x)$.

Here are some other norms on $\mathbf{R}^n$:

$$|x|_{\max} = \max\{|x_1|, \ldots, |x_n|\},$$

$$|x|_{\text{sum}} = |x_1| + \cdots + |x_n|.$$

Let $\mathcal{B} = \{f_1, \ldots, f_n\}$ be a basis for $\mathbf{R}^n$ and define the *Euclidean $\mathcal{B}$-norm*:

$$|x|_{\mathcal{B}} = (t_1^2 + \cdots + t_n^2)^{1/2} \qquad \text{if} \quad x = \sum_{j=1}^n t_j f_j.$$

In other words, $|x|_{\mathcal{B}}$ is the Euclidean norm of x in $\mathcal{B}$-coordinates $(t_1, \ldots, t_n)$.

The $\mathcal{B}$ *max-norm* of x is

$$| x |_{\mathcal{B},\max} = \max\{| t_1 |, \dots, | t_n |\}.$$

The basic fact about norms is the *equivalence of norms*:

Proposition 1 *Let $N: \mathbf{R}^n \to \mathbf{R}$ be any norm. There exist constants $A > 0$, $B > 0$ such that*

$$(4) \qquad A\,| x | \le N(x) \le B\,| x |$$

for all x, where $| x |$ is the Euclidean norm.

Proof. First, consider the max norm. Clearly,

$$(\max | x_j |)^2 \le \sum_j x_j^2 \le n(\max | x_j |)^2;$$

taking square roots we have

$$| x |_{\max} \le | x | \le \sqrt{n}\,| x |_{\max}.$$

Thus for the max norm we can take $A = 1/\sqrt{n}$, $B = 1$, or, equivalently,

$$\frac{1}{\sqrt{n}}| x | \le | x |_{\max} \le | x |.$$

Now let $N: \mathbf{R}^n \to \mathbf{R}$ be any norm. We show that N *is continuous*. We have

$$N(x) = N(\sum x_j e_j) \le \sum | x_j |\, N(e_j),$$

where $e_1, \dots, e_n$ is the standard basis. If

$$\max\{N(e_1), \dots, N(e_n)\} = M,$$

then

$$N(x) \le M \sum | x_j | \le Mn\,| x |_{\max}$$
$$\le Mn\,| x |.$$

By the triangle inequality,

$$| N(x) - N(y) | \le N(x - y)$$
$$\le Mn\,| x - y |.$$

This shows that N is continuous; for suppose $\lim x_k = y$ in $\mathbf{R}^n$:

$$| N(x_k) - N(y) | \le Mn\,| x_k - y |,$$

so $\lim N(x_k) = N(y)$ in $\mathbf{R}$.

Since N is continuous, it attains a maximum value B and a minimum value A on the closed bounded set

$$\{x \in \mathbf{R}^n \mid | x | = 1\}.$$

Now let $x \in \mathbf{R}^n$. If $x = 0$, (4) is obvious. If $| x | = \alpha \ne 0$, then

$$N(x) = \alpha N(\alpha^{-1}x).$$

Since $| \alpha^{-1}x | = 1$ we have

$$A \le N(\alpha^{-1}x) \le B.$$

Hence

$$A \le \alpha^{-1}N(x) \le B,$$

which yields (4), since $\alpha = | x |$.

Let $E \subset \mathbf{R}^n$ be a subspace. We define a *norm on* E to be any function

$$N: E \to \mathbf{R}$$

that satisfies (1), (2), and (3). In particular, every norm on $\mathbf{R}^n$ restricts to a norm on E. In fact, every norm on E is obtained from a norm on $\mathbf{R}^n$ by restriction. To see this, decompose $\mathbf{R}^n$ into a direct sum

$$\mathbf{R}^n = E \oplus F.$$

(For example, let $\{e_1, \dots, e_n\}$ be a basis for $\mathbf{R}^n$ such that $\{e_1, \dots, e_m\}$ is a basis for E; then F is the subspace whose basis is $\{e_{m+1}, \dots, e_n\}$.) Given a norm N on E, define a norm N' on $\mathbf{R}^n$ by

$$N'(x) = N(y) + | z |,$$

where

$$x = y + z,\ y \in E,\ z \in F,$$

and $| z |$ is the Euclidean norm of z. It is easy to verify that N' is a norm on $\mathbf{R}^n$ and $N' \restriction E = N$.

From this the equivalence of norms on E follows. For let N be a norm on E. Then we may assume N is restriction to E of a norm on $\mathbf{R}^n$, also denoted by N. There exist $A, B \in \mathbf{R}$ such that (4) holds for all x in $\mathbf{R}^n$, so it holds *a fortiori* for all x in E.

We now define a normed vector space (E, N) to be a vector space E (that is, a subspace of some $\mathbf{R}^n$) together with a particular norm N on E.

We shall frequently use the following corollary of the equivalence of norms:

Proposition 2 *Let (E, N) be any normed vector space. A sequence $\{x_k\}$ in E converges to y if and only if*

$$(5) \qquad \lim_{k\to\infty} N(x_k - y) = 0.$$

Proof. Let $A > 0$, $B > 0$ be as in (4). Suppose (5) holds. Then the inequality

$$0 \le | x_k - y | \le A^{-1}N(x_k - y)$$

shows that $\lim_{k\to\infty} | x_k - y | = 0$, hence $x_k \to y$. The converse is proved similarly.

Another useful application of the equivalence of norms is:

Proposition 3 *Let (E, N) be a normed vector space. Then the unit ball*

$$D = \{x \in E \mid N(x) \leq 1\}$$

is compact.

Proof. Let B be as in (4). Then D is a bounded subset of $\mathbf{R}^n$, for it is contained in

$$\{x \in \mathbf{R}^n \mid |x| \leq B^{-1}\}.$$

It follows from Proposition 2 that D is closed. Thus D is compact.

The Cauchy convergence criterion (of Section 1) can be rephrased in terms of arbitrary norms:

Proposition 4 *Let (E, N) be a normed vector space. Then a sequence $\{x_k\}$ in E converges to an element in E if and only if:*

(6) *for every $\epsilon > 0$, there exists an integer $n_0 > 0$ such that if $p > n \geq n_0$, then*

$$N(x_p - x_n) < \epsilon.$$

Proof. Suppose $E \subset \mathbf{R}^n$, and consider $\{x_k\}$ as a sequence in $\mathbf{R}^n$. The condition (6) is equivalent to the Cauchy condition by the equivalence of norms. Therefore (6) is equivalent to convergence of the sequence to some $y \in \mathbf{R}^n$. But $y \in E$ because subspaces are closed sets.

A sequence in $\mathbf{R}^n$ (or in a subspace of $\mathbf{R}^n$) is often denoted by an *infinite series* $\sum_{k=0}^{\infty} x_k$. This is merely a suggestive notation for the *sequence of partial sums* $\{s_k\}$, where

$$s_k = x_1 + \cdots + x_k.$$

If $\lim_{k\to\infty} s_k = y$, we write

$$\sum_{k=1}^{\infty} x_k = y$$

and say the *series* $\sum x_k$ *converges* to y. If all the x_k are in a subspace $E \subset \mathbf{R}^n$, then also $y \in E$ because E is a closed set.

A series $\sum x_k$ in a normed vector space (E, N) is *absolutely convergent* if the series of real numbers $\sum_{k=0}^{\infty} N(x_k)$ is convergent. This condition implies that $\sum x_k$ is convergent in E. Moreover, it is independent of the norm on E, as follows easily from equivalence of norms. Therefore it is meaningful to speak of absolute convergence of a series in a vector space E, without reference to a norm.

A useful criterion for absolute convergence is the *comparison test*: a series $\sum x_k$ in a normed vector space (E, N) converges absolutely provided there is a convergent series $\sum a_k$ of nonnegative real numbers a_k such that

$$N(x_k) \leq a_k; \qquad k = 1, 2, \ldots .$$

For

$$0 \leq \sum_{k=n+1}^{p} N(x_k) \leq \sum_{k=n+1}^{p} a_k;$$

hence $\sum_{k=0}^{\infty} N(x_k)$ converges by applying the Cauchy criterion to the partial sum sequences of $\sum N(x_k)$ and $\sum a_k$.

PROBLEMS

1. Prove that the norms described in the beginning of Section 2 actually are norms.
2. $|x|_p$ is a norm on $\mathbf{R}^n$, where
   $$|x|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p}; \qquad 1 \leq p < \infty.$$
   Sketch the unit balls in $\mathbf{R}^2$ and $\mathbf{R}^3$ under the norm $|x|_p$ for $p = 1, 2, 3$.
3. Find the largest $A > 0$ and smallest $B > 0$ such that
   $$A|x| \leq |x|_{\text{sum}} \leq B|x|$$
   for all $x \in \mathbf{R}^n$.
4. Compute the norm of the vector $(1, 1) \in \mathbf{R}^2$ under each of the following norms:
   (a) the Euclidean norm;
   (b) the Euclidean $\mathcal{B}$-norm, where $\mathcal{B}$ is the basis $\{(1, 2), (2, 2)\}$;
   (c) the max norm;
   (d) the $\mathcal{B}$-max norm;
   (e) the norm $|x|_p$ of Problem 2, for all p.
5. An *inner product* on a vector space E is any map $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$, denoted by $(x, y) \to \langle x, y\rangle$, that is symmetric, bilinear, and positive definite (see Section 1).
   (a) Given any inner product show that the function $\langle x, x\rangle^{1/2}$ is a norm.
   (b) Prove that a norm N on E comes from an inner product as in (a) if and only if it satisfies the "parallelogram law":
   $$N(x + y)^2 + N(x - y)^2 = 2(N(x)^2 + N(y)^2).$$
   (c) Let $a_1, \ldots, a_n$ be positive numbers. Find an inner product on $\mathbf{R}^n$ whose corresponding norm is
   $$N(x) = \left(\sum a_k x_k^2\right)^{1/2}.$$

(d) Let $\{e_1, \ldots, e_m\}$ be a basis for E. Show that there is a unique inner product on E such that

$$\langle e_i, e_j \rangle = \delta_{ij} \quad \text{for all} \quad i, j.$$

6. Which of the following formulas define norms on $\mathbf{R}^2$? (Let (x, y) be the coordinates in $\mathbf{R}^2$.)
(a) $(x^2 + xy + y^2)^{1/2}$; (b) $(x^2 - 3xy + y^2)^{1/2}$;
(c) $(|x| + |y|)^2$; (d) $\frac{1}{3}(|x| + |y|) + \frac{2}{3}(x^2 + y^2)^{1/2}$.

7. Let $U \subset \mathbf{R}^n$ be a bounded open set containing 0. Suppose U is *convex*: if $x \in U$ and $y \in U$, then the line segment $\{tx + (1-t)y \mid 0 \le t \le 1\}$ is in U. For each $x \in \mathbf{R}^n$ define

$$\sigma(x) = \text{least upper bound of } \{\lambda \ge 0 \mid \lambda x \in U\}.$$

Then the function

$$N(x) = \frac{1}{\sigma(x)}$$

is a norm on $\mathbf{R}^n$.

8. Let M_n be the vector space of $n \times n$ matrices. Denote the transpose of $A \in M_n$ by A^t. Show that an inner product (see Problem 5) on M_n is defined by the formula

$$\langle A, B \rangle = \operatorname{Tr}(A^t B).$$

Express this inner product in terms of the entries in the matrices A and B.

9. Find the orthogonal complement in M_n (see Problem 8) of the subspace of diagonal matrices.

10. Find a basis for the subspace of M_n of matrices of trace 0. What is the orthogonal complement of this subspace?

§3. Exponentials of Operators

The set $L(\mathbf{R}^n)$ of operators on $\mathbf{R}^n$ is identified with the set M_n of $n \times n$ matrices. This in turn is the same as $\mathbf{R}^{n^2}$ since a matrix is nothing but a list of n^2 numbers. (One chooses an ordering for these numbers.) Therefore $L(\mathbf{R}^n)$ is a vector space under the usual addition and scalar multiplication of operators (or matrices). We may thus speak of norms on $L(\mathbf{R}^n)$, convergence of series of operators, and so on.

A frequently used norm on $L(\mathbf{R}^n)$ is the *uniform norm*. This norm is defined in terms of a given norm on $\mathbf{R}^n = E$, which we shall write as $|x|$. If $T: E \to E$ is an operator, the uniform norm of T is defined to be

$$\|T\| = \max\{|Tx| \,\big|\, |x| \le 1\}.$$

In other words, $\|T\|$ is the maximum value of $|Tx|$ on the *unit ball*

$$D = \{x \in E \,\big|\, |x| \le 1\}.$$

The existence of this maximum value follows from the compactness of D (Section 1, Proposition 3) and the continuity of $T: \mathbf{R}^n \to \mathbf{R}^n$. (This continuity follows immediately from a matrix representation of T.)

The uniform norm on $L(\mathbf{R}^n)$ depends on the norm chosen for $\mathbf{R}^n$. If no norm on $\mathbf{R}^n$ is specified, the standard Euclidean norm is intended.

Lemma 1 *Let $\mathbf{R}^n$ be given a norm $|x|$. The corresponding uniform norm on $L(\mathbf{R}^n)$ has the following properties:*

(a) *If $\|T\| = k$, then $|Tx| \le k|x|$ for all x in $\mathbf{R}^n$.*
(b) $\|ST\| \le \|S\| \cdot \|T\|$.
(c) $\|T^m\| \le \|T\|^m$ *for all* $m = 0, 1, 2, \ldots$.

Proof. (a) If $x = 0$, then $|Tx| = 0 = k|x|$. If $x \ne 0$, then $|x| \ne 0$. Let $y = |x|^{-1}x$, then

$$|y| = \frac{1}{|x|}|x| = 1.$$

Hence

$$k = \|T\| \ge |Ty| = \frac{1}{|x|}|Tx|$$

from which (a) follows.

(b) Let $|x| \le 1$. Then from (a) we have

$$\begin{aligned} |S(Tx)| &\le \|S\| \cdot |Tx| \\ &\le \|S\| \cdot \|T\| \cdot |x| \\ &\le \|S\| \cdot \|T\|. \end{aligned}$$

Since $\|ST\|$ is the maximum value of $|STx|$, (b) follows.

Finally, (c) is an immediate consequence of (b).

We now define an important series generalizing the usual exponential series. For any operator $T: \mathbf{R}^n \to \mathbf{R}^n$ define

$$\exp(T) = e^T = \sum_{k=0}^{\infty} \frac{T^k}{k!}.$$

(Here $k!$ is *k factorial*, the product of the first k positive integers if $k > 0$, and $0! = 1$ by definition.) This is a series in the vector space $L(\mathbf{R}^n)$.

Theorem *The exponential series $\sum_{k=0}^{\infty} T^k/k!$ is absolutely convergent for every operator T.*

Proof. Let $\|T\| = \alpha \geq 0$ be the uniform norm (for some norm on $\mathbf{R}^n$). Then $\|T^k/k!\| \leq \alpha^k/k!$, by Lemma 1, proved earlier. Now the real series $\sum_{k=0}^{\infty} \alpha^k/k!$ converges to e^α (where e is the base of natural logarithms). Therefore the exponential series for T converges absolutely by the comparison test (Section 2).

We have also proved that

$$\|e^A\| \leq e^{\|A\|}.$$

We shall need the following result.

Lemma 2 *Let $\sum_{j=0}^{\infty} A_j = A$ and $\sum_{k=0}^{\infty} B_k = B$ be absolutely convergent series of operators on $\mathbf{R}^n$. Then $AB = C = \sum_{l=0}^{\infty} C_l$, where $C_l = \sum_{j+k=l} A_jB_k$.*

Proof. Let the nth partial sum of the series $\sum A_j$, $\sum B_k$, $\sum C_l$ be denoted respectively by α_n, β_n, γ_n. Then

$$AB = \lim_{n\to\infty} \alpha_n\beta_n,$$

while

$$C = \lim_{n\to\infty} \gamma_{2n}.$$

If $\gamma_{2n} - \alpha_n\beta_n$ is computed, it is found that it equals

$$\sum{}' A_jB_k + \sum{}'' A_jB_k,$$

where $\sum'$ denotes the sum over terms with indices satisfying

$$j + k \leq 2n, \qquad 0 \leq j \leq n, \qquad n + 1 \leq k \leq 2n,$$

while $\sum''$ is the sum corresponding to

$$j + k \leq 2n, \qquad n + 1 \leq j \leq 2n, \qquad 0 \leq k \leq n.$$

Therefore

$$\|\gamma_{2n} - \alpha_n\beta_n\| \leq \sum{}' \|A_j\|\cdot\|B_k\| + \sum{}'' \|A_j\|\cdot\|B_k\|.$$

Now

$$\sum{}' \|A_j\|\cdot\|B_k\| \leq \Big(\sum_{j=0}^{\infty} \|A_j\|\Big)\Big(\sum_{k=n+1}^{2n} \|B_k\|\Big).$$

This tends to 0 as $n \to \infty$ since $\sum_{j=0}^{\infty} \|A_j\| < \infty$. Similarly, $\sum'' \|A_j\|\cdot\|B_k\| \to 0$ as $n \to \infty$. Therefore $\lim_{n\to\infty}(\gamma_{2n} - \alpha_n\beta_n) = 0$, proving the lemma.

The next result is useful in computing with exponentials.

Proposition *Let P, S, T denote operators on $\mathbf{R}^n$. Then:*

(a) *if $Q = PTP^{-1}$, then $e^Q = Pe^TP^{-1}$;*

(b) *if $ST = TS$, then $e^{S+T} = e^Se^T$;*

(c) $e^{-S} = (e^S)^{-1}$;

(d) *if $n = 2$ and $T = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, then*

$$e^T = e^a \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}.$$

The proof of (a) follows from the identities $P(A + B)P^{-1} = PAP^{-1} + PBP^{-1}$ and $(PTP^{-1})^k = PT^kP^{-1}$. Therefore

$$P\left(\sum_{k=0}^{n} \frac{T^k}{k!}\right)P^{-1} = \sum_{k=0}^{n} \frac{(PTP^{-1})^k}{k!}$$

and (a) follows by taking limits. To prove (b), observe that because $ST = TS$ we have by the binomial theorem

$$(S + T)^n = n! \sum_{j+k=n} \frac{S^j}{j!}\frac{T^k}{k!}.$$

Therefore

$$e^{S+T} = \sum_{n=0}^{\infty}\left(\sum_{j+k=n} \frac{S^j}{j!}\frac{T^k}{k!}\right) = \left(\sum_{j=0}^{\infty} \frac{S^j}{j!}\right)\left(\sum_{k=0}^{\infty} \frac{T^k}{k!}\right)$$

by Lemma 2, which proves (b). Putting $T = -S$ in (b) gives (c).

The proof of (d) follows from the correspondence

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \leftrightarrow a + ib$$

of Chapter 3, which preserves sums, products, and real multiples. It is easy to see that it also preserves limits. Therefore

$$e^T \leftrightarrow e^a e^{ib},$$

where e^{ib} is the complex number $\sum_{k=0}^{\infty} (ib)^k/k!$. Using $i^2 = -1$, we find the real part of e^{ib} to be the sum of the Taylor series (at 0) for $\cos b$; similarly, the imaginary part is $\sin b$. This proves (d).

Observe that (c) implies that e^S is invertible for every operator S. This is analogous to the fact that $e^s \neq 0$ for every real number s.

As an example we compute the exponential of $T = \begin{bmatrix} a & 0 \\ b & a \end{bmatrix}$. We write

$$T = aI + B, \qquad B = \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix}.$$

Note that aI commutes with B. Hence

$$e^T = e^{aI}e^B = e^a e^B.$$

Now $B^2 = 0$; hence $B^k = 0$ for all $k > 1$, and

$$e^B = \sum_{k=0}^{1} \frac{1}{k!} B^k = I + B.$$

Thus

$$e^T = e^a(I + B) = e^a \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} = \begin{bmatrix} e^a & 0 \\ e^a b & e^a \end{bmatrix}.$$

We can now compute e^A for any 2×2 matrix A. We will see in Chapter 6 that can find an invertible matrix P such that the matrix

$$B = PAP^{-1}$$

has one of the following forms:

$$(1) \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}; \quad (2) \begin{bmatrix} a & -b \\ b & a \end{bmatrix}; \quad (3) \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix}.$$

We then compute e^B. For (1),

$$e^B = \begin{bmatrix} e^\lambda & 0 \\ 0 & e^\mu \end{bmatrix}.$$

For (2)

$$e^B = e^a \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{bmatrix}$$

as was shown in the proposition above. For (3)

$$e^B = e^\lambda \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

as we have just seen. Therefore e^A can be computed from the formula

$$e^A = e^{P^{-1}BP} = P^{-1}e^B P.$$

There is a very simple relationship between the eigenvectors of T and those of e^T:

If $x \in \mathbf{R}^n$ is an eigenvector of T belonging to the real eigenvalue α of T, then x is also an eigenvector of e^T belonging to e^α.

For, from $Tx = \alpha x$, we obtain

$$e^T x = \lim_{n\to\infty} \left(\sum_{k=0}^{n} \frac{T^k x}{k!} \right) = \lim_{n\to\infty} \left(\sum_{k=0}^{n} \frac{\alpha^k}{k!} x \right) = \left(\sum_{k=0}^{\infty} \frac{\alpha^k}{k!} \right) x = e^\alpha x.$$

We conclude this section with the observation that all that has been said for exponentials of operators on $\mathbf{R}^n$ also holds for operators on the *complex* vector space $\mathbf{C}^n$. This is because $\mathbf{C}^n$ can be considered as the real vector space $\mathbf{R}^{2n}$ by simply ignoring nonreal scalars; every complex operator is *a fortiori* a real operator. In addition, the preceding statement about eigenvectors is equally valid when complex eigenvalues of an operator on $\mathbf{C}^n$ are considered; the proof is the same.

PROBLEMS

1. Let N be any norm on $L(\mathbf{R}^n)$. Prove that there is a constant K such that
$$N(ST) \leq KN(S)N(T)$$
for all operators S, T. Why must $K \geq 1$?

2. Let $T: \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. Show that T is *uniformly continuous*: for all $\epsilon > 0$ there exists $\delta > 0$ such that if $|x - y| < \delta$ then
$$|Tx - Ty| < \epsilon.$$

3. Let $T: \mathbf{R}^n \to \mathbf{R}^n$ be an operator. Show that
$$\| T \| = \text{least upper bound} \left\{ \frac{|Tx|}{|x|} \,\middle|\, x \neq 0 \right\}.$$

4. Find the uniform norm of each of the following operators on $\mathbf{R}^2$:
$$\text{(a)} \begin{bmatrix} 3 & 0 \\ 0 & -4 \end{bmatrix} \quad \text{(b)} \begin{bmatrix} \frac{1}{2} & 0 \\ 10 & \frac{1}{2} \end{bmatrix} \quad \text{(c)} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

5. Let
$$T = \begin{bmatrix} \frac{1}{2} & 0 \\ 10 & \frac{1}{2} \end{bmatrix}.$$

(a) Show that

$$\lim_{n\to\infty} \| T^n \|^{1/n} = \tfrac{1}{2}.$$

(b) Show that for every $\epsilon > 0$ there is a basis $\mathcal{B}$ of $\mathbf{R}^2$ for which

$$\| T \|_{\mathcal{B}} < \tfrac{1}{2} + \epsilon,$$

where $\| T \|_{\mathcal{B}}$ is the uniform norm of T corresponding to the Euclidean $\mathcal{B}$-norm on $\mathbf{R}^2$.

(c) For any basis $\mathcal{B}$ of $\mathbf{R}^2$,

$$\| T \|_{\mathcal{B}} > \tfrac{1}{2}.$$

6. (a) Show that

$$\| T \| \cdot \| T^{-1} \| \geq 1$$

for every invertible operator T.

(b) If T has two distinct real eigenvalues, then

$$\| T \| \cdot \| T^{-1} \| > 1.$$

(*Hint*: First consider operators on $\mathbf{R}^2$.)

7. Prove that if T is an operator on $\mathbf{R}^n$ such that $\| T - I \| < 1$, then T is invertible and the series $\sum_{k=0}^{\infty} (I - T)^k$ converges absolutely to T^{-1}. Find an upper bound for $\| T^{-1} \|$.

8. Let $A \in L(\mathbf{R}^n)$ be invertible. Find $\epsilon > 0$ such that if $\| B - A \| < \epsilon$, then B is invertible. (*Hint*: First show $A^{-1}B$ is invertible by applying Problem 7 to $T = A^{-1}B$.)

9. Compute the exponentials of the following matrices ($i = \sqrt{-1}$):

(a) $\begin{bmatrix} 5 & -6 \\ 3 & -4 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

(e) $\begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$ (f) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 3 \end{bmatrix}$ (g) $\begin{bmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{bmatrix}$

(h) $\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$ (i) $\begin{bmatrix} 1+i & 0 \\ 2 & 1+i \end{bmatrix}$ (j) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

10. For each matrix T in Problem 9 find the eigenvalues of e^T.

11. Find an example of two operators A, B on $\mathbf{R}^2$ such that

$$e^{A+B} \neq e^A e^B.$$

12. If $AB = BA$, then $e^A e^B = e^B e^A$ and $e^A B = B e^A$.

13. Let an operator $A\colon \mathbf{R}^n \to \mathbf{R}^n$ leave invariant a subspace $E \subset \mathbf{R}^n$ (that is, $Ax \in E$ for all $x \in E$). Show that e^A also leaves E invariant.

14. Show that if $\| T - I \|$ is sufficiently small, then there is an operator S such that $e^S = T$. (*Hint*: Expand $\log(1 + x)$ in a Taylor series.) To what extent is S unique?

15. Show that there is no real 2×2 matrix S such that $e^S = \begin{bmatrix} -1 & 0 \\ 0 & -4 \end{bmatrix}$.

§4. Homogeneous Linear Systems

Let A be an operator on $\mathbf{R}^n$. In this section we shall express solutions to the equation:

$$(1) \qquad x' = Ax$$

in terms of exponentials of operators.

Consider the map $\mathbf{R} \to L(\mathbf{R}^n)$ which to $t \in \mathbf{R}$ assigns the operator e^{tA}. Since $L(\mathbf{R}^n)$ is identified with $\mathbf{R}^{n^2}$, it makes sense to speak of the derivative of this map.

Proposition

$$\frac{d}{dt} e^{tA} = A e^{tA} = e^{tA} A.$$

In other words, the derivative of the operator-valued function e^{tA} is another operator-valued function Ae^{tA}. This means the composition of e^{tA} with A; the order of composition does not matter. One can think of A and e^{tA} as matrices, in which case Ae^{tA} is their product.

Proof of the proposition.

$$\begin{aligned} \frac{d}{dt} e^{tA} &= \lim_{h\to 0} \frac{e^{(t+h)A} - e^{tA}}{h} \\ &= \lim_{h\to 0} \frac{e^{tA}e^{hA} - e^{tA}}{h} \\ &= e^{tA} \lim_{h\to 0} \left(\frac{e^{hA} - I}{h} \right) \\ &= e^{tA} A; \end{aligned}$$

that the last limit equals A follows from the series definition of e^{hA}. Note that A commutes with each term of the series for e^{tA}, hence with e^{tA}. This proves the proposition.