

III. LINEAR DIFFERENTIAL EQUATIONS

In this chapter we specialize to the class of linear differential equations; we will find it possible to develop both general theories of the properties of solutions and explicit formulae for solutions in special cases. Linear problems are always much simpler than nonlinear ones, but before becoming too enthusiastic at our success in attacking them we should recall an old saw: dividing problems into linear and nonlinear ones is like dividing organisms into chickens and non-chickens. There are a lot of non-linear problems out there.

3.1 General properties

We will in general suppose that $A(t)$ is a given $n \times n$ matrix with entries $a_{ij}(t)$, $i, j = 1, \dots, n$, defined and continuous on an open interval $I \subset \mathbb{R}$, and that $u : I \rightarrow \mathbb{R}^n$ is continuous, and study solutions of the initial value problem

$$x(t_0) = x^0, \quad (3.1a)$$

$$x'(t) = A(t)x(t) + u(t), \quad (3.1b)$$

where $t_0 \in I$ and $x^0 \in \mathbb{R}^n$. We begin with the existence and uniqueness theorem already introduced in Section 2.4.

Theorem 3.1: *The initial value problem (3.1) has a unique solution defined for all $t \in I$ and depending continuously on t , t_0 , and x^0 .*

Proof: The problem (3.1) fits into the standard scheme of Chapter II if we take $D = I \times \mathbb{R}^n$ and define $f : D \rightarrow \mathbb{R}^n$ by $f(t, x) = A(t)x + u(t)$. The function f is then locally Lipschitzian, since if $K \subset D$ is compact, then $K \subset J \times \mathbb{R}^n$ for some closed interval $J \subset I$, and for $(t, x^1), (t, x^2) \in K$ and $k = \sup_{t \in J} \|A(t)\|$,

$$|f(t, x^1) - f(t, x^2)| = |A(t)(x^1 - x^2)| \leq k|x^1 - x^2|.$$

Hence existence, uniqueness, and continuity of the solution follow from Theorems 2.1 and 2.11.

It remains to show that the maximal interval of definition satisfies $I_{(t_0, x^0)} = I$, whatever the initial conditions, i.e., that the domain E of the solution $\hat{x}(t; t_0, x^0)$ is $I \times I \times \mathbb{R}^n$. Suppose otherwise; in particular, suppose that $I = (a, b)$ and $I_{(t_0, x^0)} = (a', b')$ with $b' < b$. From the integral equation, we see that the solution $x(t)$ satisfies

$$|x(t)| \leq |x^0| + \left| \int_{t_0}^t [A(s)x(s) + u(s)] ds \right| \leq |x^0| + \int_{t_0}^{b'} |u(s)| ds + \int_{t_0}^t \|A(s)\| |x(s)| ds$$

and hence from Gronwall's inequality,

$$|x(t)| \leq \left[|x^0| + \int_{t_0}^{b'} |u(s)| ds \right] \exp \left[\int_{t_0}^{b'} \|A(t)\| dt \right] \equiv M$$

for $t \in [t_0, b')$. Thus $(t, x(t))$ remains in the compact set $[t_0, b'] \times \{x \mid |x| \leq M\} \subset D$ on $[t_0, b')$, contradicting Theorem 2.19. ■

If the function $u(t)$ is identically zero on I , then the linear system (3.1b) becomes

$$x'(t) = A(t)x(t) \quad (3.2)$$

and is called *homogeneous*. The solutions of a homogeneous linear equation have a particularly simple structure:

Theorem 3.2: *The set of all solutions of (3.2) is an n -dimensional vector space over the real numbers.*

Proof: The set of solutions of (3.2) is clearly a vector space, that is, if $x(t)$ and $y(t)$ are solutions and α and β are real numbers, then $\alpha x(t) + \beta y(t)$ is also a solution. Choose $t_0 \in I$ and for $i = 1, \dots, n$ let x^i be the solution of (3.2) satisfying the initial condition $x^i(t_0) = e^i$, where $[e^1, \dots, e^n]$ is the standard basis in \mathbb{R}^n defined by $e_j^i = \delta_{ij}$. The functions $x^i(t)$ are linearly independent over \mathbb{R} —that is, no linear combination of these functions can vanish identically on I unless all the coefficients are zero—since if

$$x(t) \equiv \sum_{i=1}^n \alpha_i x^i(t) \quad (3.3)$$

satisfies $x(t) = 0$ for all $t \in I$, then $x(t_0) = (\alpha_1, \dots, \alpha_n) = 0$. Moreover, these functions span the solution set. For let $y(t)$ be any solution of (3.2) and write $y(t_0) = \alpha \equiv (\alpha_1, \dots, \alpha_n)$; then $x(t)$ as given by (3.3) solves the same initial value problem as does $y(t)$ ((3.2) with $x(t_0) = \alpha$), and hence $y(t) = x(t)$ on I by the uniqueness in Theorem 3.1. Thus the solutions $x^1(t), \dots, x^n(t)$ form a basis for the solution space, which must therefore have dimension n . ■

A set of n solutions of (3.2) which is linearly independent and thus forms a basis for the solution space is called a *fundamental set* of solutions. We shall frequently find it convenient to consider $n \times n$ matrices, each of whose columns is a solution, that is, matrices $X(t)$ which satisfy

$$X'(t) = A(t)X(t). \quad (3.4)$$

Such a matrix $X(t)$ in which the columns form a fundamental set of solutions is called a *fundamental matrix* for the homogeneous system (3.2).

We summarize some properties of these fundamental matrices in

Proposition 3.3: *Let $X(t)$ satisfy (3.4), and denote the n columns of the matrix $X(t)$ by $x^1(t), \dots, x^n(t)$. Then:*

(a) *The following four conditions are equivalent: (i) x^1, \dots, x^n are linearly dependent functions (so that X is not a fundamental matrix); (ii) $\det X(t) = 0$ for all $t \in I$; (iii) for some $t_0 \in I$, $\det X(t_0) = 0$; (iv) for some $t_0 \in I$, $x^1(t_0), \dots, x^n(t_0)$ are linearly dependent vectors.*

(b) *If $X(t)$ is a fundamental matrix then every solution of (3.2) has the form $x(t) = X(t)c$ for some constant vector c .*

(c) If $X(t)$ is a fundamental matrix and C is an invertible (constant) matrix, then $Y(t) = X(t)C$ is fundamental; conversely, if $X(t)$ and $Y(t)$ are fundamental, then $X(t) = Y(t)C$ for some constant matrix C .

(d) Let $W(t) = \det X(t)$. Then $W'(t) = [\text{tr } A(t)]W(t)$, and if $t_0 \in I$,

$$W(t) = W(t_0) \exp \left[\int_{t_0}^t \text{tr } A(s) ds \right], \quad (3.5)$$

where $\text{tr } A$ denotes the trace of the matrix A .

Proof: (a) If (i) holds, then $\sum \alpha_i x^i(t) \equiv 0$ for some α , so for any t the columns of $X(t)$ are linearly dependent, proving (ii). Clearly (ii) implies (iii) and (iii) is equivalent to (iv). Finally, if (iv) holds, then $\sum \alpha_i x^i(t_0) = 0$ for some α ; since the solution of (3.2) with initial condition $x(t_0) = 0$ is unique, necessarily $\sum \alpha_i x^i(t) = 0$ for all t , i.e., (i) holds.

(b) This simply restates the fact that the columns of $X(t)$ form a basis for the space of solutions.

(c) Since $Y(t)$ satisfies (3.4) and $\det Y(t) = \det X(t) \det C \neq 0$, Y is fundamental. To verify the second statement, we note that since $X(t)$ is invertible we may define $C(t) = X(t)^{-1}Y(t)$. Then $Y(t) = X(t)C(t)$ and hence $Y'(t) = X'(t)C(t) + X(t)C'(t)$. From (3.4) we may substitute $Y'(t) = A(t)Y(t) = A(t)X(t)C(t)$ and $X'(t)C(t) = A(t)X(t)C(t)$ in this last equation, and conclude that $X(t)C'(t) = 0$. Since $X(t)$ is invertible, $C'(t) = 0$.

(d) Let S_n denote the permutation group on n elements, σ_π the sign associated with $\pi \in S_n$. Then

$$W(t) = \sum_{\pi \in S_n} \sigma_\pi \prod_{j=1}^n x_j^{\pi(j)}(t),$$

so that

$$\begin{aligned} W' &= \sum_{i=1}^n \sum_{\pi \in S_n} \sigma_\pi (x_i^{\pi(i)})' \prod_{j \neq i} x_j^{\pi(j)} \\ &= \sum_{i,k=1}^n A_{ik} \sum_{\pi \in S_n} \sigma_\pi x_k^{\pi(i)} \prod_{j \neq i} x_j^{\pi(j)} \\ &= \sum_{i=1}^n A_{ii} \sum_{\pi \in S_n} \sigma_\pi \prod_j x_j^{\pi(j)} + \sum_{\substack{i,k=1 \\ i \neq k}}^n A_{ik} \sum_{\pi \in S_n} \sigma_\pi x_k^{\pi(i)} x_k^{\pi(k)} \prod_{j \neq i,k} x_j^{\pi(j)}. \end{aligned}$$

In the last line we have first displayed terms in the previous sum for which $i = k$; these yield precisely $[\text{tr } A]W$, as desired. It remains to show that the second sum in this line, over terms with $k \neq i$, vanishes. To do so we fix k and i and, in the sum over permutations, pair the term for the permutation π with the term for the permutation in which $\pi(i)$ and $\pi(k)$ are exchanged; these terms differ only in sign and hence cancel. To verify the second statement we notice that the right hand side of (3.5) satisfies the same o.d.e. as W and agrees with W at t_0 , and apply uniqueness of solutions.

Note that the formula (3.5) gives an independent proof of the equivalence of (ii) and (iii) in (a). ■

We now return briefly to the problem of solving the *inhomogeneous* linear equation (3.1b). It is easy to verify

Lemma 3.4: *Let $y(t)$ be a solution of (3.1b). If $x(t)$ is a solution of the homogeneous equation (3.2), then*

$$\tilde{y}(t) = y(t) + x(t) \quad (3.6)$$

is also a solution of (3.1b). Moreover, every solution $\tilde{y}(t)$ of (3.1b) has the form (3.6) for some choice of $x(t)$.

Suppose now that we can solve two problems: that we can (1) find *all* solutions (or a fundamental matrix) for the homogeneous equation (3.2), and (2) find *one* solution $y(t)$ of the inhomogeneous equation (3.1b), often called a “particular” solution. Then the freedom to choose $x(t)$ in (3.6) allows us to solve the inhomogeneous problem with *any* initial condition: a solution $\tilde{y}(t)$ of (3.1b) satisfying $\tilde{y}(t_0) = y^0$ is obtained from (3.6) by choosing $x(t)$ to satisfy $x(t_0) = y^0 - y(t_0)$.

Moreover, once the first problem above has been solved, the second may be “reduced to quadratures” (i.e., reduced to the evaluation of specific integrals) by the method of *variation of parameters*, which we now describe. Let $X(t)$ be a fundamental matrix for the homogeneous problem. We look for a solution $y(t)$ of the inhomogeneous problem in the form

$$y(t) = X(t)c(t), \quad (3.7)$$

where $c : I \rightarrow \mathbb{R}^n$ is differentiable. Then y is a solution if $y'(t) = A(t)y(t) + u(t)$, and a little algebra shows that this is equivalent to $X(t)c'(t) = u(t)$. For example, we may take

$$c(t) = \int_{t_0}^t X(s)^{-1}u(s) ds. \quad (3.8)$$

Equations (3.7) and (3.8) may be regarded as providing a “particular” solution to (3.1b). Alternatively, we may add a constant vector to $c(t)$ in (3.8) and thus obtain explicitly a solution satisfying the initial condition $y(t_0) = y^0$, in the form

$$y(t) = X(t) \left[\int_{t_0}^t X(s)^{-1}u(s) ds + X(t_0)^{-1}y^0 \right].$$

We now turn to the problem of explicit computation of solutions (or of fundamental matrices) for homogeneous equations and of determining qualitative properties of solutions. These questions may be given more or less explicit answers in two cases: when A is constant and when A is periodic.

3.2 Linear Equations with Constant Coefficients

In this section we discuss the the homogeneous equation (3.2) in the case where $A(t)$ is a constant matrix A . Although we remain primarily interested in differential equations defined over the reals, it is convenient to consider also complex vectors and complex matrices. We define vector and matrix norms over the complex numbers just as we did over the reals: for $x \in \mathbb{C}^n$, $|x| = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$, and for C an $n \times n$ complex matrix,

$$\|C\| = \sup_{x \in \mathbb{C}^n, x \neq 0} \frac{|Cx|}{|x|}.$$

Both of these norms are positive definite and homogeneous, and satisfy the triangle inequality; the matrix norm also satisfies $|C_{ij}| \leq \|C\| \leq |C|_2$ (see Remark 2.7(a,b)).

We first consider the problem of exponentiating a matrix.

Definition 3.1: If B is an $n \times n$ (possibly complex) matrix we define the *exponential* of B , denoted e^B or $\exp B$, by

$$\exp B = \sum_{k=0}^{\infty} \frac{B^k}{k!}. \quad (3.9)$$

We must of course show that the exponential is well defined by (3.9). Let C_N denote a partial sum of (3.9):

$$C_N = \sum_{k=0}^N \frac{B^k}{k!}$$

Then if $N \geq M$,

$$\|C_N - C_M\| = \left\| \sum_{k=M+1}^N \frac{B^k}{k!} \right\| \leq \sum_{k=M+1}^N \frac{\|B\|^k}{k!}. \quad (3.10)$$

Because $\sum \|B\|^k/k!$ is convergent (it is the power series for $\exp \|B\|$), the sequence of partial sums $\{C_N\}$ is Cauchy in norm, i.e., $\|C_N - C_M\| \rightarrow 0$ as $N, M \rightarrow \infty$. By the last property of the norm mentioned above, the sequence of matrix elements $(C_N)_{ij}$ is also Cauchy, and hence converges; the limits of these sequences define the matrix $\exp B$.

We now summarize some properties of this matrix exponential.

Proposition 3.5: (a) If $BC = CB$, then $\exp(B + C) = \exp B \exp C$.

(b) $\frac{d}{dt} \exp tB = B \exp tB$.

Proof: (a) Just as for scalars:

$$\begin{aligned} \exp(B + C) &= \sum_{k=0}^{\infty} \frac{(B + C)^k}{k!} \\ &= \sum_{k=0}^{\infty} \sum_{i=0}^k \frac{1}{k!} \binom{k}{i} B^i C^{k-i} \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{B^i}{i!} \frac{C^j}{j!} \\ &= \exp B \exp C. \end{aligned}$$

(b) We first verify the formula for $t = 0$, calculating the derivative as the limit of difference quotients. Since $\exp 0 = I$ we must show that $\lim_{h \rightarrow 0} [h^{-1}(\exp hB - I) - B] = 0$. But from (3.9) we have

$$\|h^{-1}(\exp hB - I) - B\| = \left\| \sum_{k=1}^{\infty} \frac{h^{k-1} B^k}{k!} - B \right\| \leq h \|B\|^2 \sum_{k=0}^{\infty} \frac{|h|^k \|B\|^k}{(k+2)!} \leq h \|B\|^2 \exp |h| \|B\|,$$

from which the result follows immediately. To evaluate the derivative at a general point t_0 we write $\exp tB = \exp(t - t_0)B \exp t_0B$ by (a) and then use $d/dt = d/d(t - t_0)$. ■

From this proposition we immediately conclude that $\exp tA$ is a fundamental matrix for the equation $x' = Ax$: $\exp tA$ is invertible, with inverse $\exp -tA$, by (a) of the proposition. It is easy to verify that the solution satisfying the initial condition $x(t_0) = x^0$ is $[\exp(t - t_0)A]x^0$. That is, *the matrix exponential gives an immediate solution to the homogeneous linear o.d.e. with constant coefficients*. It remains to produce a computable form for this matrix, and to determine the qualitative properties of the solutions.

The key idea is to reduce A to a simpler form by the use of a similarity transformation. Recall that two $n \times n$ matrices A and B are *similar* if $B = P^{-1}AP$ for some invertible matrix P . Similar matrices have the same eigenvalues, the same determinant, and the same trace. It is useful to keep in mind an alternative view of the similarity relationship, in which A is identified with a linear transformation from \mathbb{R}^n to \mathbb{R}^n in the natural way: for $x \in \mathbb{R}^n$, $Ax \in \mathbb{R}^n$ is defined by matrix multiplication. In this view the matrix elements of A may be recovered from the action of A on the standard basis $[e^1, \dots, e^n]$ via the formula $Ae^i = \sum_j A_{ji}e^j$. If then $B = P^{-1}AP$ and the matrix P has (linearly independent) columns $[u^1, \dots, u^n]$, then $Au^i = \sum_j B_{ji}u^j$, that is, B is the matrix of the linear transformation in the new basis $[u^1, \dots, u^n]$.

In the simplest case A is *diagonalizable*, that is, similar to a diagonal matrix: $D = P^{-1}AP$ with

$$D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}.$$

In the alternate language above, a column u^i of P satisfies $Au^i = \lambda_i u^i$ and hence is an *eigenvector* of A , with λ_i the corresponding *eigenvalue*; thus A is diagonalizable if and only if there exists a basis of \mathbb{R}^n made up of eigenvectors of A . Since $A = PDP^{-1}$ and hence

$$A^n = (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1}) = PD^n P^{-1},$$

it follows from (3.9) that

$$\exp tA = P \exp[tD]P^{-1} = P \begin{bmatrix} e^{t\lambda_1} & & & 0 \\ & e^{t\lambda_2} & & \\ & & \ddots & \\ 0 & & & e^{t\lambda_n} \end{bmatrix} P^{-1}.$$

Thus $\exp tA$ is computable once we know P and the λ_i , which are obtained by standard calculations: the eigenvalues are the roots of the n^{th} degree polynomial $\det[A - \lambda I]$, and u^i is obtained by solving the system of linear equations $[A - \lambda_i I]u^i = 0$.

Various conditions are known to guarantee diagonalizability of A . Because there is at least one eigenvector for each eigenvalue and eigenvectors for distinct eigenvalues are linearly independent, A is diagonalizable if A has n distinct eigenvalues. It may also be shown that a real matrix A is diagonalizable if it is symmetric ($A^T = A$) or, more generally for complex A , if A is normal ($AA^* = A^*A$, where A^* is the Hermitian adjoint of A). Unfortunately, not all matrices are diagonalizable; for example, $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not. We may reduce such matrices to *Jordan canonical form*.

Definition 3.2: A *Jordan block* is a (possibly complex) square matrix in which all diagonal entries are equal, all entries immediately above the diagonal are one, and all other entries are zero:

$$J = \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \lambda & \ddots \\ 0 & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix}. \quad (3.11)$$

A (square) matrix is in *Jordan form* if it has a block decomposition in which all diagonal blocks are Jordan blocks and all other blocks are zero:

$$J = \begin{bmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_m \end{bmatrix}, \quad (3.12)$$

where each J_i represents a Jordan block. Notationally, we will assume that there are m blocks and that J_i is an $n_i \times n_i$ block with diagonal entries equal to λ_i , so that $\sum_{i=1}^m n_i = n$, the dimension of J . Since $\det[J - \lambda I] = \prod_{i=1}^m (\lambda_i - \lambda)^{n_i}$, the λ_i are the eigenvalues of J . The same eigenvalue may appear in several blocks.

Theorem 3.6: *Every $n \times n$ matrix A is similar to a matrix J in Jordan form. Equivalently, there is a basis $\{u^{ij} \mid i = 1, \dots, m, j = 1, \dots, n_i\}$ for \mathbb{C}^n satisfying*

$$\begin{aligned} Au^{i1} &= \lambda_i u^{i1}, \\ Au^{ij} &= \lambda_i u^{ij} + u^{i,j-1}, \quad \text{if } j > 1; \end{aligned} \tag{3.13}$$

elements of this basis are called *generalized eigenvectors*. Moreover:

(a) J is uniquely determined by A except for possible reordering of the blocks, that is, m and the set of pairs $\{(\lambda_i, n_i)\}$ is determined.

(b) If A is real but some eigenvalues are complex then the blocks corresponding to complex eigenvalues may be grouped in pairs $(J_i, J_{i'})$, and the generalized eigenvectors may be chosen, so that $\lambda_i = \bar{\lambda}_{i'}$, $n_i = n_{i'}$, and $u^{ij} = \bar{u}^{i'j}$ for $j = 1, \dots, n_i$.

We will not prove the Jordan form theorem but give a procedure for constructing generalized eigenvectors in an appendix to this section.

Now consider a matrix A and suppose $A = PJP^{-1}$ with J in Jordan form. Again, $\exp tA = P \exp[tJ]P^{-1}$, and we must compute $\exp tJ$. It is easy to see that, if J is given by (3.12), then

$$e^{tJ} = \begin{bmatrix} e^{tJ_1} & & & 0 \\ & e^{tJ_2} & & \\ & & \ddots & \\ 0 & & & e^{tJ_m} \end{bmatrix}, \tag{3.14}$$

which reduces the problem to that of exponentiating a Jordan block J_i . Now $J_i = \lambda_i I + D_i$, where D_i is an $n_i \times n_i$ matrix in which all entries D_{jk} are zero except for a diagonal of ones immediately above the diagonal. It is easy to compute powers of D_i ,

$$\begin{aligned} D_i &= \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ & 0 & 1 & 0 & \\ & & 0 & 1 & \ddots & \vdots \\ & & & 0 & \ddots & 0 \\ 0 & & & & \ddots & 1 \\ & & & & & 0 \end{bmatrix}, & D_i^2 &= \begin{bmatrix} 0 & 0 & 1 & \dots & 0 \\ & 0 & 0 & 1 & \\ & & 0 & 0 & \ddots & \vdots \\ & & & 0 & \ddots & 1 \\ 0 & & & & \ddots & 0 \\ & & & & & 0 \end{bmatrix}, \\ & \vdots \\ D_i^{n_i-1} &= \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \\ & 0 & 0 & 0 & \\ & & 0 & 0 & \ddots & \vdots \\ & & & 0 & \ddots & 0 \\ 0 & & & & \ddots & 0 \\ & & & & & 0 \end{bmatrix}, & D_i^r &= 0, \quad r \geq n_i, \end{aligned}$$

and hence its exponential:

$$e^{tD_i} = \sum_{k=0}^{n_i-1} \frac{t^k D_i^k}{k!} = \begin{bmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{n_i-1}}{(n_i-1)!} \\ & 1 & t & \frac{t^2}{2} & \\ & & 1 & t & \ddots \\ & & & 1 & \ddots \\ 0 & & & & \ddots \\ & & & & t \\ & & & & 1 \end{bmatrix}. \quad (3.15)$$

Finally, since the $n_i \times n_i$ identity matrix I commutes with D_i and $e^{t\lambda_i I} = e^{t\lambda_i} I$, we have

$$e^{tJ_i} = e^{t\lambda_i} e^{tD_i} = \begin{bmatrix} e^{t\lambda_i} & te^{t\lambda_i} & \frac{t^2}{2}e^{t\lambda_i} & \cdots & \frac{t^{n_i-1}}{(n_i-1)!}e^{t\lambda_i} \\ & e^{t\lambda_i} & te^{t\lambda_i} & \frac{t^2}{2}e^{t\lambda_i} & \\ & & e^{t\lambda_i} & te^{t\lambda_i} & \ddots \\ & & & e^{t\lambda_i} & \ddots \\ 0 & & & & \ddots \\ & & & & te^{t\lambda_i} \\ & & & & e^{t\lambda_i} \end{bmatrix}. \quad (3.16)$$

Thus, from (3.14), $\exp tJ$ is a block diagonal matrix with blocks of the form (3.16).

Now the fundamental matrix e^{tA} is given by from the formula $e^{tA} = Pe^{tJ}P^{-1}$. In fact, however, by Proposition 3.3, $e^{tA}P = Pe^{tJ}$ is an equally good fundamental matrix. We have proved

Theorem 3.7: *Suppose that A is an $n \times n$ matrix and that $J = P^{-1}AP$ is in Jordan form. Then a fundamental matrix for the homogeneous system $x' = Ax$ is given by*

$$X(t) = Pe^{tJ},$$

where e^{tJ} may be calculated from (3.14) and (3.16).

Now it follows from (3.14) and (3.16) that the matrix $X(t) = Pe^{tJ}$ may be described as follows: the column of $X(t)$ which corresponds to the j^{th} column in the i^{th} Jordan block has the form

$$x^{ij}(t) = e^{t\lambda_i} \sum_{k=1}^j \frac{t^{j-k}}{(j-k)!} u^{ik}. \quad (3.17)$$

Since every solution of the system is a linear combination of these columns, we have

Corollary 3.8: *Every component of every solution of the system $x' = Ax$ is a linear combination of functions $t^k e^{t\lambda}$, where (i) λ is an eigenvalue of A , and (ii) k is at most one less than the maximum size of the Jordan blocks corresponding to λ , and this maximum value of k actually occurs in some component of some solution.*

Up to this time we have ignored the fact that some eigenvalues λ_i of A , as well as the corresponding generalized eigenvectors u^{ij} , may in fact be complex, even though the matrix A is real. A real matrix A , with real initial conditions, will of course lead to real solutions. The above discussion, including the conclusion of Corollary 3.8 which now refers to a complex linear combination of complex functions, is correct, but it is convenient to rewrite this conclusion so that only real functions are involved.

The simplest approach is to work directly from (3.17). Suppose that the eigenvalue λ_i for the block J_i is complex, $\lambda_i = \alpha_i + i\beta_i$, with $\beta_i \neq 0$, and that the corresponding generalized eigenvectors are $u^{ij} = v^{ij} + iw^{ij}$. We know from Theorem 3.6 that we may choose P so that J contains a block $J_{i'}$ with eigenvalue $\lambda_{i'} = \alpha_i - i\beta_i$ and generalized eigenvectors $u^{i'j} = v^{ij} - iw^{ij}$. We may then derive a new fundamental matrix $Y(t)$ by replacing the pair of columns x^{ij} and $x^{i'j}$ of $X(t)$ by the columns

$$y^{ij} \equiv \frac{x^{ij} + x^{i'j}}{2} = e^{t\alpha_i} \sum_{k=1}^j \frac{t^{j-k}}{(j-k)!} [\cos t\beta_i v^{ik} - \sin t\beta_i w^{ik}] \quad (3.18a)$$

and

$$z^{ij} \equiv \frac{x^{ij} - x^{i'j}}{2i} = e^{t\alpha_i} \sum_{k=1}^j \frac{t^{j-k}}{(j-k)!} [\cos t\beta_i w^{ik} + \sin t\beta_i v^{ik}]. \quad (3.18b)$$

Then Corollary 3.8 becomes

Corollary 3.9: *If A is real, then every component of every solution of the system $x' = Ax$ is a linear combination of functions $t^k e^{t\lambda}$, where λ is a real eigenvalue of A , and of functions $t^k e^{t\alpha} \cos t\beta$ and $t^k e^{t\alpha} \sin t\beta$, where $\alpha + i\beta$ is a complex eigenvalue of A ; in each case, k is at most one less than the maximum size of the Jordan blocks corresponding to λ or $\alpha + i\beta$, respectively, and this maximum value of k actually occurs in some component of some solution.*

The fundamental matrix $Y(t)$ and the conclusions of Corollary 3.9 may also be derived by reducing the real matrix A to *real canonical form* rather than to Jordan form. We sketch briefly the details of this reduction. Suppose that the Jordan form of A has m' blocks J_i for which the eigenvalue λ_i is real and m'' pairs of blocks $J_i, J_{i'}$ related as in Theorem 3.6(b), so that $m = m' + 2m''$, and suppose that each complex eigenvalue λ_i and the corresponding generalized eigenvectors u^{ij} are decomposed into real and imaginary parts as above. Now (3.13) becomes

$$\begin{aligned} Av^{ij} &= \alpha_i v^{ij} - \beta_i w^{ij} + v^{i,j-1}, \\ Aw^{ij} &= \alpha_i w^{ij} + \beta_i v^{ij} + w^{i,j-1}, \end{aligned}$$

where the last term in each equation is missing if $j = 1$. We choose a basis for \mathbb{R}^n as follows: (i) for each block J_i with real eigenvalue, the basis includes the generalized eigenvectors $[u^{i1}, \dots, u^{in_i}]$; (ii) for each pair $J_i, J_{i'}$, the basis includes the $2n_i$ vectors $[v^{i1}, w^{i1}, \dots, v^{in_i}, w^{in_i}]$, in that order. If Q is the matrix whose columns are these basis

vectors, then $A = QKQ^{-1}$, where K is in *real canonical form*. This means that

$$K = \begin{bmatrix} K_1 & & & 0 \\ & K_2 & & \\ & & \ddots & \\ 0 & & & K_{m'+m''} \end{bmatrix}. \quad (3.19)$$

and each block K_i is either a Jordan block with real eigenvalue or a $2n_i \times 2n_i$ block which itself has a decomposition into 2×2 blocks:

$$K_i = \begin{bmatrix} \Lambda_i & I & & 0 \\ & \Lambda_i & I & \\ & & \Lambda_i & \ddots \\ 0 & & & \ddots & I \\ & & & & \Lambda_i \end{bmatrix}; \quad (3.20)$$

here I is the 2×2 identity matrix and $\Lambda_i = \begin{bmatrix} \alpha_i & \beta_i \\ -\beta_i & \alpha_i \end{bmatrix}$.

A fundamental matrix for the system is now $Y(t) = Qe^{tK}$, and as in (3.14) it suffices to compute e^{tK_i} . When K_i is a Jordan block we use (3.16); it remains only to compute the exponential of a block of the form (3.20). Write $K_i = E_i + D_i$, where (in 2×2 blocks)

$$E_i = \begin{bmatrix} \Lambda_i & & & 0 \\ & \Lambda_i & & \\ 0 & & \ddots & \\ & & & \Lambda_i \end{bmatrix},$$

$$D_i = \begin{bmatrix} 0 & I & & 0 \\ & 0 & I & \\ & & 0 & \ddots \\ 0 & & & \ddots & I \\ & & & & 0 \end{bmatrix}.$$

Since E_i and D_i commute, $e^{tK_i} = e^{tE_i}e^{tD_i}$, and e^{tD_i} is computed just as for the ordinary Jordan block; the result is as in (3.15) but with each entry replaced by that entry times a 2×2 identity matrix. On the other hand,

$$e^{tE_i} = \begin{bmatrix} e^{\Lambda_i} & & & 0 \\ & e^{\Lambda_i} & & \\ 0 & & \ddots & \\ & & & e^{\Lambda_i} \end{bmatrix}; \quad (3.21)$$

write $\Lambda_i = \alpha_i I + \beta_i \Gamma$, with $\Gamma = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and use the fact that $\Gamma^2 = -I$ to obtain

$$\begin{aligned} e^{t\Lambda_i} &= e^{t\alpha_i} e^{t\beta_i \Gamma} = e^{t\alpha_i} \left[I + t\beta_i \Gamma - \frac{(t\beta_i)^2}{2} I - \frac{(t\beta_i)^3}{3!} \Gamma + \frac{(t\beta_i)^4}{4!} I + \dots \right] \\ &= e^{t\alpha_i} \begin{bmatrix} \cos t\beta_i & \sin t\beta_i \\ -\sin t\beta_i & \cos t\beta_i \end{bmatrix}. \end{aligned} \quad (3.22)$$

Putting these results together we find that e^{tK_i} has the 2×2 block decomposition

$$e^{tK_i} = \begin{bmatrix} e^{t\Lambda_i} & te^{t\Lambda_i} & \frac{t^2}{2}e^{t\Lambda_i} & \cdots & \frac{t^{n_i-1}}{(n_i-1)!}e^{t\Lambda_i} \\ & e^{t\Lambda_i} & te^{t\Lambda_i} & \frac{t^2}{2}e^{t\Lambda_i} & \\ & & e^{t\Lambda_i} & te^{t\Lambda_i} & \ddots \\ & & & e^{t\Lambda_i} & \ddots \\ 0 & & & & \ddots \\ & & & & te^{t\Lambda_i} \\ & & & & e^{t\Lambda_i} \end{bmatrix}, \quad (3.23)$$

with $e^{t\Lambda_i}$ given by (3.22). The columns of the fundamental matrix $Y(t) = Qe^{tK}$ may now be computed as in (3.17) and shown to be given by x^{ij} , as in (3.17), for λ_i real, or by y^{ij} and z^{ij} , as in (3.18), for λ_i complex. Corollary 3.9 follows from this result.

We finally want to see what qualitative properties of solutions can be derived from Corollary 3.9. We are particularly interested in *asymptotic* behavior, that is, the properties of the solution as $t \rightarrow \infty$ or $t \rightarrow -\infty$, and in the possibility of *periodic* solutions.

Theorem 3.10: *Every solution $x(t)$ of the system $x' = Ax$ satisfies $\lim_{t \rightarrow \infty} x(t) = 0$ if and only if $\operatorname{Re} \lambda < 0$ for all eigenvalues λ of A . Every solution is bounded on all intervals of the form $[a, \infty)$ if and only if (i) $\operatorname{Re} \lambda \leq 0$ for all eigenvalues λ , and (ii) every Jordan block in the Jordan form of A for which the eigenvalue λ satisfies $\operatorname{Re} \lambda = 0$ has dimension 1. Similar results hold, with the inequalities reversed, for $\lim_{t \rightarrow -\infty} x(t) = 0$ or for intervals $(-\infty, b]$.*

Proof: Consider the first statement of the theorem. If the condition $\operatorname{Re} \lambda < 0$ is satisfied by all eigenvalues then the asymptotic decrease of all solutions is an immediate consequence of Corollary 3.9. Conversely, if $\operatorname{Re} \lambda_i \geq 0$ for some i then the solution x^{i1} does not decrease at infinity; specifically, if $u_k^{i1} \neq 0$, then $x_k^{i1} = e^{t\lambda_i} u_k^{i1}$ does not vanish as t goes to ∞ . The proof of the second statement is similar. ■

Periodic solutions of our system are solutions which satisfy, for some fixed $\tau \in \mathbb{R}$ with $\tau \neq 0$,

$$x(t + \tau) = x(t), \quad \text{for all } t \in \mathbb{R}. \quad (3.24)$$

A constant solution is trivially periodic; (3.24) holds for all $\tau \in \mathbb{R}$. If x is not constant, then the minimum positive τ for which (3.24) holds is called the *period* of the solution x ; all τ satisfying (3.24) are integer multiples of this (minimal) period. Suppose that A has a pure imaginary eigenvalue $\lambda = i\beta$; since $-i\beta$ is also an eigenvalue, we may suppose that $\beta > 0$. The solutions y^{i1} and z^{i1} of (3.18) for blocks with eigenvalue λ are periodic, with period $\frac{2\pi}{\beta}$. All periodic solutions are linear combinations of these solutions and constant solutions, but a variety of possibilities for such linear combinations occur when several of the β_i have rational ratios. To summarize these we introduce some notation. Let $S \subset \{1, \dots, m\}$ index the set of Jordan blocks J_i of A with purely imaginary eigenvalues $\lambda_i = i\beta_i$, $\beta_i > 0$. S is divided into equivalence classes S_1, \dots, S_M by the relation: $i \sim j$ if and only if β_i/β_j is rational. It is easy to see that for each S_l there is a positive real number ω_l such that $\beta_i = q_i\omega_l$ for $i \in S_l$, with the q_i integers with greatest common divisor

equal to 1. We let $S_0 \subset \{1, \dots, m\}$ index the set of Jordan blocks with eigenvalue 0; S_0 may be empty.

Theorem 3.11: (a) If 0 is not an eigenvalue of A , then every solution of the system $x' = Ax$ of the form

$$x(t) = \sum_{i \in S_l} [c_i y^{i1} + d_i z^{i1}] \quad (3.25)$$

satisfies (3.24) with $\tau = 2\pi/\omega_l$. If $x(t)$ is not the zero solution then it has (minimal) period $2\pi/q'\omega_l$, where q' is the greatest common divisor of the integers q_i , $i \in S_l$, for which c_i or d_i is non-zero. Moreover, every periodic solution is of the form (3.25) for some S_l .

(b) If 0 is an eigenvalue of A then the above statements hold with (3.25) replaced by

$$x(t) = \sum_{i \in S_l} [c_i y^{i1} + d_i z^{i1}] + \sum_{i \in S_0} e_i x^{i1} \quad (3.26)$$

and with the condition “ $x(t)$ is not the zero solution” replaced by “ $x(t)$ is not constant”.

Remark 3.1: The theorem classifies all periodic solutions as follows: to each l and each q' which is a greatest common divisor of some subset of the q_i , $i \in S_l$, there is associated a subspace $V_{lq'}$ of the solution space which essentially consists of solutions of (minimal) period $\tau_{lq'} = 2\pi/q'\omega_l$. However, within $V_{lq'}$ there are proper subspaces of exceptional solutions for which $\tau_{lq'}$ is not the minimal period: the zero subspace, possibly subspaces of non-zero constant solutions (if A has zero as an eigenvalue), and possibly subspaces of solutions whose period divides $\tau_{lq'}$. Note that the intersection of two subspaces $V_{lq'}$ and $V_{l_1q'_1}$ is the subspace of constant solutions.

Proof: It is straightforward to verify that (3.25) or (3.26) satisfies (3.24) with the indicated τ and has the indicated minimal period. To see that no other periodic solutions exist note that if $x(t) = Y(t)c$ is periodic then so is $y(t) = Q^{-1}x(t) = e^{tK}c$. A component of $y(t)$ has one of the forms

$$G(t)e^{t\lambda} \quad \text{or} \quad e^{t\alpha}[G_1(t)\cos t\beta + G_2(t)\sin t\beta],$$

where λ is real and G , G_1 , and G_2 are polynomials; this component is constant if and only if it is of the first form, λ is zero, and G constant, and is non-constant and periodic if and only if it is of the second form, α is zero, and G_1 and G_2 constant. Finally, a vector with each component periodic or constant will be periodic if and only if the periods of the non-constant components have rational ratios. ■

3.3: Computation of Jordan Canonical Form

We outline one procedure for the computation of the Jordan canonical form of an $n \times n$ matrix A . Because we must find simultaneously all generalized eigenvectors associated with a single eigenvalue, it will be convenient to change our notational convention for the Jordan blocks and generalized eigenvectors from that adopted in Section 3.2.

STEP 1: Find the eigenvalues of A by finding roots of the equation

$$\det[A - \lambda I] = 0.$$

We suppose that there are p *distinct* eigenvalues, which we label μ_1, \dots, μ_p .

We now fix an eigenvalue μ_i and find a corresponding family of generalized eigenvectors. These will be labeled as v^{ijk} ; here $j = 1, \dots, q(i)$ indexes the different Jordan blocks corresponding to the eigenvalue μ_i , and $k = 1, \dots, r(i, j)$ indexes the generalized eigenvectors for the block J_{ij} . The vectors should satisfy

$$\begin{aligned} Bv^{ijk} &= v^{ij(k-1)} & k \geq 2, \\ Bv^{ij1} &= 0. \end{aligned}$$

where we have written $B = A - \mu_i I$.

If we number the Jordan blocks in inverse order of size, so that $r(i, 1) \geq r(i, 2) \geq \dots \geq r(i, q(i))$, then we may usefully arrange the corresponding generalized eigenvectors in a tableau:

$$\begin{array}{ccccccc} v^{i11} & v^{i21} & \dots & v^{ij1} & \dots & v^{iq(i)1} & \\ v^{i12} & v^{i22} & \dots & v^{ij2} & & & \\ \vdots & \vdots & & \vdots & & & \\ v^{i1r(i,j)} & v^{i2r(i,j)} & \dots & v^{ijr(i,j)} & & & \\ \vdots & \vdots & & & & & \\ v^{i1r(i,2)} & v^{i2r(i,2)} & & & & & \\ \vdots & & & & & & \\ v^{i1r(i,1)} & & & & & & \end{array}$$

Here the vectors in the j^{th} column of the tableau are the generalized eigenvectors for the j^{th} Jordan block J_{ij} . Multiplication by B carries a vector in the tableau into the vector immediately above it. The rows of the tableau also have a natural interpretation. The vectors in the first row are eigenvectors; they form a basis for the eigenspace of A corresponding to μ_i ; equivalently, they form a basis for the null space of B . The vectors in the top two rows all satisfy $B^2v = 0$, and form a basis for the null space of B^2 . Similarly, the vectors in the top j rows ($j = 1, \dots, r(i, 1)$) form a basis for the null space of B^j . The null space of a power B^r for $r > r(i, 1)$ is the same as the null space of $B^{r(i,1)}$. Let $s(j)$ denote the number of elements in row j of the tableau and $S(j) = \sum_{i=1}^j s(i)$ the nullity of B^j . (The $s(j)$ will be determined during the construction.)

STEP 2: Determine $s(1) = S(1)$, and find the a basis $[w^1, \dots, w^{S(1)}]$ for the null space of B (eigenspace of A), by solving the system of equations $Bw = 0$. Let W^1 denote the $n \times S(1)$ matrix whose columns are $[w^1, \dots, w^{S(1)}]$.

It is helpful to specify our solution procedure. We first reduce B to its row-echelon form R^0 by row operations (equivalently, we determine an invertible E^0 with $R^0 = E^0 B$.) In R^0 there are $S(1)$ *free* columns and $n - S(1)$ *pivot* columns; the corresponding variables w_i are called free or pivot variables. To obtain a solution we choose the free variables arbitrarily and solving for the pivot variables; to obtain the basis vector w^i we choose the i^{th} free variable to be 1 and all other free variables to be zero.

The remainder of the construction is essentially inductive, but for clarity we will describe the first inductive step separately. Recall that we have found some basis for the null space of B , that is, for the span of the vectors in the first row of the tableau, but we must still determine the specific basis $[\dots v^{ij^1} \dots]$. Now the vectors in the second row of the tableau have the property that their images under B lie in the null space of B , which we have determined above to be spanned by $[w^1, \dots, w^{S(1)}]$, that is, a vector w in this row satisfies

$$Bw = \sum_{k=1}^{S(1)} c_k w^k. \quad (3.27)$$

Moreover, if we knew the $s(2)$ vectors in the second row we would also (by applying B to these) know the first $s(2)$ vectors in the first row; we could then choose the remaining vectors in the first row to be any vectors in the null space of B independent of these first vectors. These remarks motivate

STEP 3 (INDUCTIVE STEP—FIRST APPLICATION): Solve the homogeneous equations $Bw = W^1 c$ for the unknowns $w_1, \dots, w_n, c_1, \dots, c_{S(1)}$. Let $S(2)$ denote the dimension of the solution space. $s(1)$ independent solutions are given by $(w^k, 0)$; let $s(2) = S(2) - S(1)$ and let $[(w^{S(1)+1}, c^{S(1)+1}), \dots, (w^{S(2)}, c^{S(2)})]$ be $s(2)$ additional linearly independent solutions. Then $[w^1, \dots, w^{S(2)}]$ is a basis for the null space of B^2 ; let W^2 denote the $n \times S(2)$ matrix whose columns are $[w^1, \dots, w^{S(2)}]$. Finally, define the vectors v^{ij^1} for $s(2) < j \leq s(1)$ to be any $s(1) - s(2)$ of the vectors $w^1, \dots, w^{S(1)}$ which are linearly independent of the vectors $\sum_{k=1}^{S(1)} c_k^{S(1)+j} w^k$, $j = 1, \dots, s(2)$.

The algebra necessary for this calculation is simplified if at step 2 we have followed the solution procedure described there. The equations to be solved can be written in standard form as

$$[B \ W^1] \begin{bmatrix} w \\ -c \end{bmatrix} = 0.$$

We have already reduced B to row-echelon form, so that if we perform the same operations on W^1 we will have partially reduced $[B \ W^1]$. Thus we need only perform the operations necessary to complete the reduction of the last $s(1)$ columns, arriving at a new row-echelon matrix $R^1 = E^1 [B \ W^1]$. Among the last $s(1)$ columns of R^1 there will be $s(2)$ free columns, used to generate the solutions $(w^{S(1)+j}, c^{S(1)+j})$ of the system. The columns of W^1 which

are pivot columns in R^1 are then linearly independent of the vectors

$$Bw^{S(1)+j} = \sum_{k=1}^{S(1)} c_k^{S(1)+j} w^k$$

and we can choose these columns as the vectors v^{ij^1} for $s(2) < j \leq s(1)$.

We now generalize this procedure. The $k = 1$ instance of the inductive step below is the step we have just described.

STEP 3 (INDUCTIVE STEP): Suppose that we have determined

- (i) The integers $s(1), \dots, s(k)$;
- (ii) An $n \times S(k)$ matrix W^k such that, for any $j \leq k$, columns $1, \dots, S(j)$ of W^k form a basis for the null space of B^j .
- (iii) A row reduction procedure for the matrix $[B \ W^{k-1}]$; that is, an invertible matrix E^{k-1} with $R^{k-1} = E^{k-1}[B \ W^{k-1}]$ in row-echelon form.
- (iv) All generalized eigenvectors u^{ij^k} for $j > s(k)$.

Now solve the system

$$[B \ W^k] \begin{bmatrix} w \\ -c \end{bmatrix} = 0,$$

for the unknowns $(w_1, \dots, w_n, c_1, \dots, c_{S(k)})$. To do so, first apply the previously determined row reduction to the last $s(k)$ columns of $[B \ W^k]$ to reduce this matrix partially, then complete the row reduction to obtain $R^k = E^k [B \ W^k]$. Let $s(k+1)$ be the number of free columns among the last $s(k)$. If $s(k+1) = 0$ then the induction will terminate at the current step; otherwise, let $[(w^{S(k)+1}, c^{S(k)+1}), \dots, (w^{S(k+1)}, c^{S(k+1)})]$ be the solutions corresponding to these free columns and adjoin columns $[w^{S(k)+1}, \dots, w^{S(k+1)}]$ to W^k to form W^{k+1} . In either case, define the vectors $v^{ij^{(k+1)}}$ for $s(k+1) < j \leq s(k)$ to be the columns of W^k corresponding to the pivot columns among the last $s(k)$ columns of R^k . This completes the induction step.

We make two closing remarks. First, the total number of generalized eigenvectors associated with the eigenvalue μ_i is equal to the algebraic multiplicity of μ_i , that is, to the number of factors $\mu_i - \lambda$ of the polynomial $\det[A - \lambda I]$. This can serve as a check on the algorithm. Second, if A is a real matrix and μ_i is a complex eigenvalue then $\bar{\mu}_i$ is also an eigenvalue; after determining the generalized eigenvectors v^{ij^k} for μ_i one simply takes \bar{v}^{ij^k} as the generalized eigenvectors for $\bar{\mu}_i$. The real canonical form for A is then immediately determined by splitting these generalized eigenvectors into their real and imaginary parts.

3.4 Linear Equations with Periodic Coefficients: Floquet Theory

We now consider the homogeneous system

$$x' = A(t)x \quad (3.28a)$$

in the case in which the matrix $A(t)$ is continuous and periodic:

$$A(t + \tau) = A(t), \quad t \in \mathbb{R}. \quad (3.28b)$$

Two important second order equations which attain this form when reduced to a 2×2 system are Mathieu's equation

$$x'' + (a + 16q \cos 2t)x = 0$$

(a and q are constants), which arises in the discussion of the vibrations of an elliptical membrane, and the more general Hill's equation

$$x'' + p(t)x = 0,$$

where $p(t)$ is periodic (frequently taken to have period π). The latter equation was studied by G. W. Hill (*Acta. Math.* **VIII** (1886)) in connection with the motion of the lunar perigee. The Floquet theory which we will present does not yield solutions to these systems, but does provide some information about the structure of such solutions.

Theorem 3.12: *Let $X(t)$ be a fundamental matrix for the system (3.28). Then*

$$X(t) = P(t)e^{tR} \quad (3.29)$$

for some continuous matrix P of period τ and some constant matrix R .

Proof: Let $Y(t) = X(t + \tau)$. Then Y satisfies

$$Y'(t) = X'(t + \tau) = A(t + \tau)X(t + \tau) = A(t)Y(t),$$

and hence is itself a fundamental matrix for the system. By Proposition 3.3 there is an invertible constant matrix C such that $X(t + \tau) = X(t)C$; we will show shortly that there exists a matrix R such that $C = \exp \tau R$. Now define $P(t) = X(t)e^{-tR}$; (3.29) follows immediately from this definition, as does the periodicity of P :

$$P(t + \tau) = X(t + \tau)e^{-\tau R}e^{-tR} = P(t). \quad \blacksquare$$

To complete the proof we verify

Lemma 3.13: *Every invertible $n \times n$ matrix C may be written as the exponential of some $n \times n$ matrix B : $C = \exp B$.*

Proof: We refer to any such matrix B as a logarithm of C and write $B = \log C$, although we do not attempt to define precisely any single valued logarithm function. Let $J = PCP^{-1}$ be the Jordan form of C ; it suffices to find $K = \log J$, for then $C = P^{-1}JP = \exp(P^{-1}KP)$ and we take $B = P^{-1}KP$. Similarly, it suffices to find $K_i = \log J_i$ for each block J_i of J . Hence we may assume from the beginning that C is a Jordan block: $C = \lambda I + D$, with $\lambda \neq 0$ (since C is invertible) and $D_{jk} = \delta_{j,k-1}$.

Now formally,

$$B = \log C = \log[(\lambda I)(I + \lambda^{-1}D)] = \log(\lambda I) + \log(I + \lambda^{-1}D). \quad (3.30)$$

We choose $\mu \in \mathbb{C}$ satisfying $e^\mu = \lambda$ and define $\log(\lambda I)$ in (3.30) to be μI , and define $\log(I + \lambda^{-1}D)$ by the standard power series $\log(1+z) = g(z) = -\sum_{k=1}^{\infty} (-z)^k/k$. That is, we finally define

$$B = \mu I + \sum_{k=1}^{\infty} \frac{(-1)^{k-1} \lambda^{-k} D^k}{k}. \quad (3.31)$$

Because $D^k = 0$ for $k \geq n$, the infinite sum here is in fact finite, B is certainly well defined, and $\exp B$ may be calculated by substituting (3.31) into the exponential power series and rearranging. From the formal power series identity $e^{g(z)} = 1+z$ we conclude that $\exp B = C$. ■

Remark 3.2: (a) The logarithm B of the matrix C is certainly not unique, since we may always add a multiple of $2\pi i$ to the μ chosen in the proof of Lemma 3.13 (and may add different multiples in different Jordan blocks, rearrange the basis vectors for Jordan blocks before doing so, etc.) Thus the matrices R and $P(t)$ which appear in the decomposition (3.29) of the fundamental matrix $X(t)$ are not unique. However, the eigenvalues of $C = \exp \tau R$, called the *characteristic multipliers* of $A(t)$, and the Jordan block structure of C , are in fact uniquely determined. For C is obtained from $X(t)$ by the formula $C = X(t)^{-1}X(t+\tau)$, and if $X_1(t)$ is another fundamental matrix and $C_1 = X_1(t)^{-1}X_1(t+\tau)$, then since $X_1(t) = X(t)M$ for some invertible M we have

$$X_1(t+\tau) = X(t+\tau)M = X(t)CM = X_1(t)M^{-1}CM,$$

that is, $C_1 = M^{-1}CM$ is similar to C and hence has the same Jordan form. The eigenvalues of R itself, called *characteristic exponents* of $A(t)$, are not uniquely determined. (In the reference by Cronin it is shown that the Jordan form of R is uniquely determined, up to a shift in the eigenvalues by a multiple of $2\pi i$.)

(b) Even when A is real the matrix P in (3.29) may be complex. Coddington and Levinson point out, however, that one may always obtain a decomposition of the form (3.29) with a *real* matrix $P(t)$ of period 2τ .

(c) It is clear from (3.29) that the asymptotic behavior of solutions is determined by the Jordan form of R or of C . In particular: (i) All solutions $x(t)$ approach 0 as $t \rightarrow \infty$ if and only if $\operatorname{Re} \rho < 0$ for all eigenvalues ρ of R , or equivalently if and only if $|\lambda| < 1$ for all eigenvalues λ of C , i.e., all characteristic multipliers of A ; (ii) All solutions remain bounded as $t \rightarrow \infty$ if and only if $\operatorname{Re} \rho \leq 0$ for all eigenvalues ρ of R , and eigenvalues ρ with

$\operatorname{Re} \rho = 0$ occur only in Jordan blocks of dimension 1, or equivalently if and only if $|\lambda| \leq 1$ for all eigenvalues λ of C , and eigenvalues λ with $|\lambda| = 1$ occur only in Jordan blocks of dimension 1. Similar results hold for the asymptotics as $t \rightarrow -\infty$.

(d) We may also say a little about periodic solutions. Suppose that τ is the minimal period for A . We note first that P must also have minimal period τ , since from (3.29) we may write $X' = P'e^{tR} + PRe^{tR} = AX = APe^{tR}$, so that $A = P'P^{-1} + PRP^{-1}$. We are frequently interested in finding solutions of this same period. Observe that if $R = QJQ^{-1}$, with J in Jordan form, then another fundamental matrix is $P(t)Qe^{tJ}$, and that since $P(t)Q$ has the same periodicity as $P(t)$ we may simply assume that R in (3.29) is in Jordan form. Now if e^{tR} contains columns with (not necessarily minimal) period $m\tau$ (i.e., with frequency a multiple of $2\pi/m\tau$), then the corresponding columns of $X(t)$ will also have (again, not necessarily minimal) period $m\tau$. This is the only way to produce solutions of period τ . There are various ways of producing solutions of period rationally related to τ , or even (since some elements of $P(t)$ may be constant) of period not rationally related to τ .