Introduction. We have considered a general homogeneous linear equation and discussed its *fundamental matrices*, which we labeled generically as $\Psi(t)$, and picked one with special properties, labeled $\Phi(t)$. As we will see, in case the coefficient matrix of the homogenous equation is constant, this latter matrix has an alternative representation as an exponential infinite series, hence it is called the *matrix exponential*. But first, let’s see how to use a fundamental matrix to solve a non-homogeneous system of linear equations. The method we will use for that is called *variation of parameters*. Apparently this simple and ingenious idea was first introduced by Lagrange.

Fundamental Matrices

Recall that a fundamental matrix for a homogeneous system of linear ODEs

$$ x' = Ax, \quad A = \begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{pmatrix} $$

is an $n \times n$ matrix $\Psi(t)$ whose columns are a set of $n$ linearly independent solutions of the system of ODEs, say:

$$ \Psi(t) = \begin{bmatrix} x_1(t) & \cdots & x_n(t) \end{bmatrix} $$

In case Eq. (9.1) can be solved on an interval containing the origin, we have seen that there is a particular fundamental matrix defined by

$$ \Phi(t) = \Psi(t)\left[\Psi(0)\right]^{-1} $$

$$ = \begin{bmatrix} x_1(t) & \cdots & x_n(t) \end{bmatrix} \begin{bmatrix} x_1(0) & \cdots & x_n(0) \end{bmatrix}^{-1} $$

that is convenient for solving IVPs. It is easy to show that:

$$ \Phi' = A\Phi, \quad \Phi(0) = I_n $$

Thus, the $j$-th column of $\Phi(t)$ is the unique solution of the IVP

---

1 Based on a course taught jointly with Michael Moody.
where \( e_j \) is the \( j^{th} \) standard basis vector (i.e., all entries are 0 except of 1 in the \( j^{th} \) row).

This proves that \( \Phi(t) \), which is uniquely determined by the set of IVPs defined by Eq. (9.4), can be derived from any fundamental set \( \{x_1, \ldots, x_n\} \) using Eq. (9.3). And we have seen that the general IVP \( x(t_0) = x^0 \) can now be expressed simply as

\[
x(t) = \Phi(t) \Phi^{-1}(t_0) x^0
\]

with some possible limitations on the interval of definition, depending on the common region of continuity for the entries of the coefficient matrix \( A \). We extend this idea to non-homogeneous equations in the next section.

**Variation of Parameters**

Here’s a general approach to solving a non-homogeneous system like

\[
x' = A(t)x + g(t), \quad A = A_{\text{non}}, \quad g = \begin{pmatrix} g_1(t) \\ \vdots \\ g_2(t) \end{pmatrix}
\]

where \( A(t) \) and \( g(t) \) are continuous for \( \alpha < t < \beta \). Your text discusses two others that can often be applied equally well, but we discuss this method because of its general applicability. We begin with \( \Psi(t) \), a fundamental matrix for the homogeneous system

\[
x' = A(t)x
\]

We have seen that the general solution of the system Eq. (9.8) is of the form

\[
x(t) = \Psi(t)c, \quad c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}
\]

for constant vectors \( c \). This suggests that we try to solve Eq. (9.7) with solutions of the form

\[
x(t) = \Psi(t)u(t), \quad u = \begin{pmatrix} u_1(t) \\ \vdots \\ u_2(t) \end{pmatrix}
\]
This approach is called \textit{variation of parameters} because we are allowing the parameters (components) of $u$ in Eq. (9.10) to vary as functions of time. Inserting Eq. (9.10) into Eq. (9.7), we get

\[
\frac{d}{dt}(Ψu) = Ψ' u + Ψ u' = AΨu + g(t)
\]

\[
⇒ AΨu + Ψ u' = AΨu + g(t)
\]

\[
⇒ Ψ u' = g(t)
\]

(9.11)

The second line of Eq. (9.11) follows from the fact that $Ψ' = AΨ$, since the columns of $Ψ(t)$ are solutions of Eq. (9.8).

As we know from the theory of the Wronskian, the fundamental matrix $Ψ$ is invertible for all values $t ∈ (α, β)$, so we may write

\[
u' = Ψ^{-1}(t)g(t)
\]

\[
⇒ u(t) = u(t_0) + \int_{t_0}^{t} Ψ^{-1}(s)g(s)ds
\]

(9.12)

where $u(t_0)$ is an arbitrary constant vector. Inserting this solution back into Eq. (9.10) yields

\[
x(t) = Ψ(t)u(t_0) + Ψ(t)\int_{t_0}^{t} Ψ^{-1}(s)g(s)ds
\]

(9.13)

which is seen to be the sum of, respectively, the general solution for Eq. (9.8) and a particular solution for Eq. (9.7).

\textbf{Solving a Non-homogeneous IVP}

Eq. (9.13) leads immediately to the solution of the IVP $x(t_0) = x^0$. Setting

\[
Ψ(t_0)u(t_0) = x^0
\]

(9.14)

we obtain

\[
x(t) = Ψ(t)Ψ^{-1}(t_0)x^0 + Ψ(t)\int_{t_0}^{t} Ψ^{-1}(s)g(s)ds
\]

This formulation yields the same result for any fundamental matrix $Ψ(t)$, but we will see in the next section that it may be computationally less onerous in the case of a constant coefficient matrix $A$ to use the form:

\[
x(t) = Φ(t)Φ^{-1}(t_0)x^0 + Φ(t)\int_{t_0}^{t} Φ^{-1}(s)g(s)ds
\]

(9.15)
We show below that, when the coefficient matrix $A$ is constant, an important property of $\Phi$ is that $\Phi(s)\Phi(t) = \Phi(s+t)$, therefore, $\Phi^{-1}(t) = \Phi(-t)$. In that case, this permits us to re-write the general solution more conveniently as

$$x(t) = \Phi(t)\Phi(-t_0)x^0 + \Phi(t)\int_{t_0}^{t} \Phi(-s)g(s)ds$$

$$(9.16)$$

This formulation allows us to eliminate one matrix multiplication and avoid inverting a variable matrix $\Psi(t)$. Instead, the smaller price we pay is that we must invert and multiply $\Psi(t)$ by $\Psi^{-1}(0)$ to obtain $\Phi(t)$. Before moving on, notice again that the first expression on the second line of Eq. (9.16) is the general solution of the homogenous portion of Eq. (9.7) and the final expression (the integral) is a particular solution of the inhomogeneous equation. The same is true of Eq. (9.16).

**Example:**

$$x' = \begin{pmatrix} 2 & -5 \\ 1 & -2 \end{pmatrix} x + \begin{pmatrix} -\cos t \\ \sin t \end{pmatrix}$$

$$(9.17)$$

Here $g = (-\cos t, \sin t)^T$. The coefficient matrix has complex eigenvalues: $\pm i$, and corresponding eigenvectors

$$\begin{pmatrix} 5 \\ 2 \end{pmatrix} \pm \begin{pmatrix} 0 \\ -1 \end{pmatrix} i$$

$$(9.18)$$

Recall that if Eq. (9.1) has constant coefficients and a pair of eigenvalues are $\lambda \pm i\mu$ and corresponding eigenvectors are $a \pm ib$, then two real and linearly-independent solutions are given by

$$u_1 = e^{\lambda t} \left[ a \cos \mu t - b \sin \mu t \right]$$

$$u_2 = e^{\lambda t} \left[ a \sin \mu t + b \cos \mu t \right]$$

$$(9.19)$$

Thus we may write as a fundamental matrix for Eq. (9.17)

$$\Psi(t) = \begin{pmatrix} 5\cos t & 5\sin t \\ 2\cos t + \sin t & 2\sin t - \cos t \end{pmatrix}, \quad \det \Psi(t) = -5$$

$$(9.20)$$

We deduce that
\[
\Phi(t) = \begin{pmatrix}
5\cos t & 5\sin t \\
2\cos t + \sin t & 2\sin t - \cos t
\end{pmatrix}
\begin{pmatrix}
5 & 0 \\
2 & -1
\end{pmatrix}^{-1}
= \begin{pmatrix}
\cos t + 2\sin t & -5\sin t \\
\sin t & -2\sin t + \cos t
\end{pmatrix}
\]

(9.21)

Inverting the matrix for \(\Phi\), we obtain,

\[
\Phi^{-1}(t) = \begin{pmatrix}
\cos t - 2\sin t & 5\sin t \\
-\sin t & 2\sin t + \cos t
\end{pmatrix}
\]

(9.22)

The preceding equation illustrates a general fact that \(\Phi^{-1}(t) = \Phi(-t)\) when the coefficient matrix \(A\) is constant, as we shall see in the discussion of the matrix exponential below. It is this property that makes \(\Phi\) the fundamental matrix of choice in Eqs. (9.13) and (9.16).

Looking ahead to plugging \(\Phi^{-1}\) into Eq. (9.13), we compute:

\[
\Phi^{-1}(t)g(t) = \begin{pmatrix}
\cos t - 2\sin t & 5\sin t \\
-\sin t & 2\sin t + \cos t
\end{pmatrix}
\begin{pmatrix}
-\cos t \\
\sin t
\end{pmatrix}
= \begin{pmatrix}
-\cos^2 t + 2\cos t \sin t + 5\sin^2 t \\
2\cos t \sin t + 2\sin^2 t
\end{pmatrix}
= \begin{pmatrix}
2 - 3\cos 2t + \sin 2t \\
1 - \cos 2t + \sin 2t
\end{pmatrix}
\]

(9.23)

then integrate

\[
\int \Phi^{-1}(t)g(t)\,dt = \begin{pmatrix}
\int (2 - 3\cos 2t + \sin 2t)\,dt \\
\int (1 - \cos 2t + \sin 2t)\,dt
\end{pmatrix}
= \begin{pmatrix}
2t - \frac{3}{2}\sin 2t - \frac{1}{2}\cos 2t \\
t - \frac{1}{2}\sin 2t - \frac{1}{2}\cos 2t
\end{pmatrix}
\]

(9.24)

Finally, Eqs. (9.20) and (9.21) can be folded into Eq. (9.16) to express the general solution of the inhomogeneous IVP \(x(t_0) = x^0\). The method is complicated but straightforward.

**The Matrix Exponential – The Extraordinary Properties of \(\Phi(t)\)**

It is not immediately obvious, but for the case of a constant coefficient matrix \(A\), the fundamental matrix \(\Phi(t)\) defined by Eq. (9.4) behaves like the exponential function \(e^{\lambda t}\), interpreted in a formal sense explained below. By analogy with the power series for the
exponential function \(e^{\alpha t}\), we define a function by means of a similar-looking infinite series of matrices:

\[
\Phi(t) = I_n + tA + \frac{t^2}{2!} A^2 + \frac{t^3}{3!} A^3 + \cdots \tag{9.25}
\]

The questions asked are whether the series on the right-hand side of Eq. (9.25) converges, and if so, does the resulting sum behave anything like the scalar exponential function? The answer to both questions is yes, and, as explained afterward, the function \(\Phi(t)\) has these properties:

1. \(\Phi(0) = e^{0,A} = I_n\) \hspace{1cm} (9.26)
2. \(e^{sA} e^{tA} = \Phi(s)\Phi(t) = \Phi(s+t) = e^{(s+t)A}\) \hspace{1cm} (9.27)
3. \(\Phi(-t) = [\Phi(t)]^{-1}\) \hspace{1cm} (9.28)
4. \(\Phi'(t) = A\Phi(t) = \Phi(t)A\) \hspace{1cm} (9.29)

An explanatory sketch follows. Probably the simplest conceptual approach is to begin by defining an infinite sequence of matrix polynomials \(F_0, F_1, \ldots, F_k, \ldots\)

\[
F_k(t) = \sum_{j=0}^{k} \frac{t^j}{j!} A^j \tag{9.30}
\]

where by convention \(A^0 = I_n\) is the \(n\)-by-\(n\) identity matrix, and \(0! = 0^0 = 1\). Because \(A\) is a square matrix, the functions \(F_k(t)\) are well-defined matrix polynomials.

Observe that for each \(k, F_k\) is an \(n\)-by-\(n\) matrix. So, for each \(k\), you can think of Eq. (9.30) as representing \(n^2\) individual sequences – one for each entry of \(F_k\). Then what we ask is whether each of those \(n^2\) entries approaches a limit as \(k \to \infty\)? In other words, does each of the \(n^2\) infinite series represented by the matrix series on the right-hand side of Eq. (9.25) converge to a finite limit?

The answer to the preceding question is yes, the series of Eq. (9.25) converges absolutely to a well-defined differentiable function of \(t\) for all values of \(t\). The proof depends on some elementary facts about the operator norm and the convergence of infinite series, which we state but do not prove. First, the operator norm \(\|A\|_{\infty,n}\) of an \(n \times n\) matrix \(A\) measures the maximum proportionate amount that \(A\) can stretch any vector of \(\mathbb{R}^n\). Writing simply \(\|\cdot\| = \|\cdot\|_{\infty,n}\), the critical property for this application is
that if $A$ and $B$ are any two $n \times n$ constant matrices, $\|AB\| \leq \|A\|\|B\|$, therefore $\|A^t\| \leq \|A\|^t$.

It follows that

$$\|\Phi(t)\| \leq 1 + \frac{t}{1!} \|A\| + \frac{t^2}{2!} \|A\|^2 + \frac{t^3}{3!} \|A\|^3 + \cdots = e^{tA} \quad (9.31)$$

Second, the latter series is the well-known power series for the real-valued exponential function, which can be differentiated term-by-term. Third, we take it on faith that the convergence of Eq. (9.31) implies the convergence and term-by-term differentiability of the matrix series defined by Eq. (9.25). After you have studied some analysis, the preceding argument will seem more natural and down-to-earth.

So let’s proceed now on the plausible assumption that the series defining the function $\Phi(t)$ of Eq. (9.25) converges and is differentiable term-by-term for all values $t \in \mathbb{R}$. **Item 1** follows simply by setting $t = 0$ in Eq. (9.25). **Item 2** can be obtained by collecting increasing powers and using the binomial theorem in the algebraic product of the two series:

$$= I_n + \left( \frac{s + t}{1!} \right) A + \left( \frac{s^2 + st + t^2}{2!} \right) A^2 + \left( \frac{s^3 + s^2 t + s t^2 + t^3}{3!} \right) A^3 + \cdots \quad (9.32)$$

These operations are justified by the absolute convergence of Eq. (9.25). **Item 3** follows by setting $s = -t$ and applying Items 1 and 2, which, importantly, also shows that $\Phi(t)$ is invertible for all $t \in \mathbb{R}$. **Item 4** follows upon differentiating Eq. (9.25):

$$\frac{d}{dt} \Phi = A + \frac{t}{1!} A^2 + \frac{t^2}{2!} A^3 + \cdots = A \left( I_n + tA + \frac{t^2}{2!} A^2 + \cdots \right) = A\Phi \quad (9.33)$$

Eq. (9.33) shows that the columns of $\Phi$ satisfy Eq. (9.1). And Item 1 shows that Eq. (9.33) satisfies the initial condition $\Phi(0) = I_n$. Hence, by the fundamental existence and Uniqueness Theorem for linear systems, the matrix $\Phi(t)$ defined by Eq. (9.25) is identical to the solution of the IVP (9.4). Accordingly we may write

$$\Phi(t) = e^{At} = \exp(At) \quad (9.34)$$
and refer to $\Phi(t)$ as the matrix exponential function for $A$.\(^2\)

**Example.** Find the general solution for

$$x' = Ax, \quad A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad (9.35)$$

We find by successive multiplication that

$$A^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} \quad (9.36)$$

Therefore,

$$\exp(tA) = \exp \left( t \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \right) = I + \sum_{n=1}^{\infty} \frac{A^n t^n}{n!}$$

$$= I + \sum_{n=1}^{\infty} \begin{pmatrix} \frac{\lambda^n t^n}{n!} & \frac{n\lambda^{n-1} t^n}{n!} \\ 0 & \frac{\lambda^n t^n}{n!} \end{pmatrix} = \begin{pmatrix} 1 + \sum_{n=1}^{\infty} \frac{\lambda^n t^n}{n!} & \sum_{n=1}^{\infty} \frac{\lambda^{n-1} t^n}{n!} \\ 0 & 1 + \sum_{n=1}^{\infty} \frac{\lambda^n t^n}{n!} \end{pmatrix}$$

$$= \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}$$

Therefore, the general solutions for Eq. (9.35) is,

$$x(t) = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad (9.38)$$

The coefficient matrix $A$ has a deficient eigenspace, so this problem can also be solved by the technique described in an earlier lecture beginning with the trial solution

$$x = (a + b) e^{\lambda t}. \text{ Try it!}$$

---

\(^2\) Your text introduces the matrix exponential in the same manner as here, but more briefly.