Elementary Differential Equations

M.V. Safonov
School of Mathematics, University of Minnesota, Minneapolis, MN, 55455

These lecture notes serve as supplementary material for the course
Math4512: Differential Equations with Applications.

1 Existence and Uniqueness for the Initial Value Problem.

Consider the differential equation of order n

$$y^{(n)} = f(t, y, y', \dots, y^{(n-1)}), \tag{1.1}$$

where f is a continuous function of all its variables, which is defined for $t \in I := (\alpha, \beta)$ and arbitrary values of $y, y', \cdots, y^{(n-1)}$. The **initial value problem**, or the **Cauchy problem** for this equation is the equations together with the initial conditions at some fixed point $t_0 \in I$;

$$y(t_0) = b_0, \quad y'(t_0) = b_1, \quad \dots, \quad y^{(n-1)}(t_0) = b_{n-1}, \tag{1.2}$$

where $b_0, b_1, \dots, b_{n-1}$ are prescribed constants. This means that if $y = y(t, C_1, C_2, \dots, C_n)$ is the general solution to (1.1), then one needs to choose n constants $C_1, C_1, \dots, C_n$ in order to satisfy n conditions in (1.2).

Lemma 1.1. The differential equation (1.1) is equivalent to the system

$$\frac{dy_1}{dt} = y_2, \ \frac{dy_2}{dt} = y_3, \ \cdots, \ \frac{dy_{n-1}}{dt} = y_n, \ \frac{dy_n}{dt} = f(t, y_1, \cdots, y_n), \tag{1.3}$$

in the following sense.
(i) If $y = y(t)$ satisfies (1.1), then $y_1 := y, y_2 := y', \dots, y_n := y^{(n)}$ satisfy (1.3).
(ii) If $(y_1, y_2, \dots, y_n)$ satisfy (1.3), then $y(t) := y_1(t)$ satisfies (1.1).

Proof is straightforward. □

Correspondingly, the Cauchy problem (1.1)–(1.2) is equivalent to

$$\frac{dY}{dt} = F(t, Y) \quad \text{for} \quad t \in I := (\alpha, \beta), \qquad Y(t_0) = Y_0, \tag{1.4}$$

where Y and F are vector functions: $Y := (y_1, y_2, \dots, y_n)$, $Y_0 := (b_0, b_1, \dots, b_{n-1})$, and
$F := (f_1, f_2, \dots, f_n)$ with $f_1 := y_2, f_2 := y_3, \dots, f_{n-1} := y_n, f_n = f(t, Y) = f(t, y_1, y_2, \dots, y_n)$.

In turn, it is easy to check that the Cauchy problem for the differential equation (1.4) with vector valued functions Y and F is equivalent to the integral equation.

Lemma 1.2. The vector functions $Y = Y(t)$ satisfies (1.4) if and only if it satisfies

$$Y(t) = Y_0 + \int_{t_0}^{t} F\big(s, Y(s)\big)ds \quad \text{for} \quad t \in I. \tag{1.5}$$

It is known that the problems (1.4) or (1.5) indeed have solutions for continuous $F(t, Y)$, but in order to guarantee uniqueness, one needs to impose some additional restrictions. Technically, the simplest condition of such kind is the **Lipschitz condition** with respect to Y:

$$|F(t, Y_1) - F(t, Y_2)| \le K \cdot |Y_1 - Y_2| \quad \text{for all} \quad t \in I \quad \text{and} \quad Y_1, Y_2 \in \mathbb{R}^n, \tag{1.6}$$

where $K = \text{const} \ge 0$. For example, this condition holds true if all the components f_j of F have bounded partial derivatives $\partial f_j / \partial x_k$.

Here we use the notation

$$|x| := \Big(\sum_j x_j^2\Big)^{1/2} \quad \text{for} \quad x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

We also need the **scalar**, or **inner product**

$$(x, y) := \sum_j x_j y_j \quad \text{for} \quad x = (x_1, x_2, \ldots, x_n),\ y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n.$$

Note that if $Y = (y_1, y_2, \ldots y_n)$ with differentiable components, then

$$\frac{d}{dt}\big(|Y|^2\big) = \frac{d}{dt}\Big(\sum_j y_j^2\Big) = 2\sum_j y_j \frac{dy_j}{dt} = 2\Big(Y, \frac{dY}{dt}\Big). \tag{1.7}$$

Lemma 1.3 (Cauchy-Schwartz inequality).

$$\big|(x, y)\big| \le |x| \cdot |y| \quad \text{for all} \quad x, y \in \mathbb{R}^n. \tag{1.8}$$

Proof. This is obvious if either $x = 0$ or $y = 0$. Otherwise, if both $x \ne 0$ and $y \ne 0$, then one can multiply x and y by appropriate nonzero constants to reduce the proof to the case $|x| = |y| = 1$. Then (1.7) is reduced to the equivalent relation

$$\big|(x, y)\big| = \Big|\sum_j x_j y_j\Big| \le \sum_j |x_j y_j| \le \frac{1}{2}\sum_j (x_j^2 + y_j^2) = \frac{1}{2}\big(|x|^2 + |y|^2\big) = 1 = |x| \cdot |y|.$$

□

Theorem 1.4 (Uniqueness). Under the Lipschitz condition (1.6), the problems (1.4) or (1.5) cannot have more than one solution.

Proof. Let functions $Y_1(t)$ and $Y_2(t)$ both satisfy (1.4). Combining together (1.6)–(1.8), we get the following estimates for the function $Y(t) := Y_1(t) - Y_2(t)$:

$$\Big|\frac{dY}{dt}\Big| = \Big|\frac{dY_1}{dt} - \frac{dY_2}{dt}\Big| = \big|F(t, Y_1) - F(t, Y_2)\big| \le K \cdot |Y_1 - Y_2| = K \cdot |Y|,$$
$$\frac{d}{dt}\big(|Y|^2\big) = 2\Big(Y, \frac{dY}{dt}\Big) \le 2\,|Y| \cdot \Big|\frac{dY}{dt}\Big| \le 2K \cdot |Y|^2.$$

Then the function

$$h(t) := e^{-2Kt}|Y(t)|^2 \ge 0, \quad \text{and} \quad \frac{dh}{dt} = e^{-2Kt}\Big(-2K \cdot |Y|^2 + \frac{d}{dt}\big(|Y|^2\big)\Big) \le 0.$$

Since also $h(t_0) = 0$, we must have $h(t) \equiv 0$ and $Y_1(t) \equiv Y_2(t)$ for $t \ge t_0$. Replacing t by $-t$, we also get $Y_1(t) \equiv Y_2(t)$ for $t \le t_0$. Hence Y_1 and Y_2 cannot be distinct. □

In order to prove the existence, we can assume without loss of generality that $t_0 = 0$, and consider the problem only for $t \ge 0$ (otherwise we can replace t by $-t$. We need some auxiliary fact from the advanced Calculus. Let $X = C\big([0, A], \mathbb{R}^n\big)$ denote the set of all continuous vector functions $Y(t)$ on the closed interval $[0, A]$ with values in $\mathbb{R}^n$, i.e. $Y = (y_1, y_2, \dots, y_n)$ with scalar components. Introduce the **distance** in X by the formula

$$d(Y, Z) := \sup_{[0,A]} |Y - Z| \quad \text{for} \quad Y, Z \in X. \tag{1.9}$$

It is easy to see that (X, d) is a **metric space**, which means:
(i) $d(Y, Z) \ge 0$, and $d(Y, Z) = 0$ if and only if $Y = Z$,
(ii) $d(Y, Z) = d(Z, Y)$,
(iii) (the triangle inequality) $d(Y, W) \le d(Y, Z) + d(Z, W)$ for all $Y, Z, W \in X$.

Theorem 1.5. The metric space (X, d) is **complete**, i.e. every **Cauchy sequence** $\{Y_k\}$ converges. This means that from $d(Y_j, Y_k) \to 0$ as $j, k \to \infty$ it follows $d(Y_k, Y) \to 0$ as $k \to \infty$ for some $Y \in X$.

Proof. We still rely on two more elementary facts: (i) $\mathbb{R}^n$ with the Euclidean distance $|x-y|$ is complete, and (ii) every continuous function $Y(t)$ on a compact $[0, A]$ is **uniformly continuous**: for every $\varepsilon > 0$, there exists $\delta > 0$, such that

$$|Y(t_1) - Y(t_2)| \le \varepsilon \quad \text{for all} \quad t_1, t_2 \in [0, A] \quad \text{with} \quad |t_1 - t_2| \le \delta. \tag{1.10}$$

For completeness, we provide the proofs of these facts in Section 4 below.

For every $t \in [0, A]$, we have $|Y_j(t) - Y_k(t)| \le d(Y_j, Y_k) \to 0$ as $j, k \to \infty$, so that $\{Y_j(t)\}$ is a Cauchy sequence in $\mathbb{R}^n$, therefore, there exists

$$Y(t) := \lim_{j\to\infty} Y_j(t) \quad \text{for every} \quad t \in [0, A]. \tag{1.11}$$

We have to show that $Y \in X$, i.e. $Y(t)$ is continuous and satisfies the above property (1.10).

Fix $\varepsilon > 0$ and then choose a large enough k, such that $d(Y_j, Y_k) \le \varepsilon/3$ for all $j \ge k$. Then

$$|Y(t) - Y_k(t)| = \lim_{j\to\infty} |Y_j(t) - Y_k(t)| \le \sup_{j\ge k} d(Y_j, Y_k) \le \frac{\varepsilon}{3} \quad \text{for every} \quad t \in [0, A].$$

Further, applying (1.10) to the functions Y_k with $\varepsilon/3$ in place of ε, we can get a constant $\delta > 0$, such that

$$|Y_k(t_1) - Y_k(t_2)| \le \frac{\varepsilon}{3} \quad \text{for all} \quad t_1, t_2 \in [0, A] \quad \text{with} \quad |t_1 - t_2| \le \delta.$$

For such t_1, t_2, we also have

$$|Y(t_1) - Y(t_2)| \le |Y(t_1) - Y_k(t_1)| + |Y_k(t_1) - Y_k(t_2)| + |Y_k(t_2) - Y(t_2)| \le \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

This means that Y satisfies (1.10) and belongs to X. □

Theorem 1.6 (Existence and Uniqueness). Under the Lipschitz condition (1.6), the problems (1.4) or (1.5) have a unique solution.

Proof. Since these two problems are equivalent, we can deal with integral equation (1.5), which means that Y is a fixed point for the transformation $\mathbb{T}$:

$$Y(t) = \mathbb{T}[Y](t) := Y_0 + \int_0^t F\big(s, Y(s)\big)ds \quad \text{for} \quad 0 \le t \le A. \tag{1.12}$$

It is convenient to introduce another distance on the metric space X as follows:

$$d_0(Y,Z) := \sup_{[0,A]} e^{-2Kt}|(Y-Z)(t)| \quad \text{for} \quad Y, Z \in X. \tag{1.13}$$

Comparing with the distance $d(Y,Z)$ in (1.9), one can see that $d_0 \le d \le e^{2KA} d_0$, so that previous Theorem 1.5 is equally applicable to the metric space (X, d_0). In terms of this distance, we have a pointwise estimate

$$|(Y-Z)(t)| \le d_0(Y,Z) \cdot e^{2Kt} \quad \text{for} \quad 0 \le t \le A. \tag{1.14}$$

We claim that $\mathbb{T}$ is a **contraction** in (X, d_0):

$$d_0(V_1, W_1) \le \frac{1}{2} d_0(V, W), \quad \text{where} \quad V_1 := \mathbb{T}[V], \quad W_1 := \mathbb{T}[W], \quad \text{for} \quad V, W \in X. \tag{1.15}$$

Indeed, by virtue of (1.12) and (1.14),

$$\big|(V_1 - W_1)(t)\big| = \big|\mathbb{T}[V](t) - \mathbb{T}[W](t)\big| = \left| \int_0^t \Big[F(s, V(s)) - F(s, W(s))\Big] ds \right|$$

$$\le K \int_0^t \big|(V-W)(s)\big| ds \le K d_0(V,W) \cdot \int_0^t e^{2Ks} ds \le \frac{1}{2} d_0(V,W) e^{2Kt},$$

and then (1.15) follows by definition of d_0 in (1.13).

Now we introduce the sequence

$$Y_0(t) \equiv Y_0, \quad \text{and then} \quad Y_{k+1} := \mathbb{T}[Y_k] \quad \text{for} \quad k = 0, 1, 2, \ldots.$$

Iterating the estimate (1.15), we obtain

$$d_0(Y_k, Y_{k+1}) \le 2^{-1} d_0(Y_{k-1}, Y_k) \le \cdots \le 2^{-k} C_0, \quad \text{where} \quad C_0 := d_0(Y_0, Y_1).$$

Further, using the triangle inequality, we also get

$$\begin{aligned} d_0(Y_k, Y_j) &\le d_0(Y_k, Y_{k+1}) + d_0(Y_{k+1}, Y_{k+2}) + \cdots + d_0(Y_{j-1}, Y_j) \\ &\le \big(2^{-k} + 2^{-k-1} + \cdots\big) C_0 \le 2^{1-k} C_0 \quad \text{for} \quad j > k. \end{aligned}$$

This means that $\{Y_k\}$ is a Cauchy sequence in (X, d_0), and by Theorem 1.5, there exists $Y \in X$ such that $d_0(Y_k, Y) \to 0$ as $k \to \infty$.

Finally, we claim that Y is the desired solution of (1.12), i.e. $Y = \mathbb{T}[Y]$. Indeed, denote $Z := \mathbb{T}[Y]$. They by (1.15), we have

$$d_0(Y_{k+1}, Z) \le \frac{1}{2} d_0(Y_k, Y) \to 0 \quad \text{as} \quad k \to \infty.$$

Then

$$d_0(Y, Z) \le d_0(Y, Y_k) + d_0(Y_k, Y_{k+1}) + d_0(Y_{k+1}, Z) \to 0 \quad \text{as} \quad k \to \infty.$$

But the left hand side does not depend on k, which means that $Y = Z = \mathbb{T}[Y]$.

Independently of Theorem 1.4, the uniqueness follows immediately from (1.15): if both V and W are fixed point for $\mathbb{T}$, then $V_1 = V$, $W_1 = W$, and the inequality $d_0(V, W) = d_0(V_1, W_1) \le 2^{-1} d_0(V, W)$ is only possible if $d_0(V, W) = 0$ and $V = W$. □

Since the problem (1.1)–(1.2) is equivalent to (1.4), one can re-formulate Theorem 1.5 as follows:

Theorem 1.7. Let the function f in (1.1) satisfies the Lipschitz condition with respect to the variables $y, y', \ldots, y^{(n-1)}$ on every closed subinterval $I_0 := [a_0, \beta_0] \subset I := (\alpha, \beta)$. Then for arbitrary $t_0 \in I$, and arbitrary $b_0, b_1, \ldots, b_{n-1}$ in $\mathbb{R}^1$, there exists a unique solution to the problem (1.1)–(1.2).

2 Linear Homogeneous Equations.

We first consider linear homogeneous equations with variable coefficients.

Theorem 2.1. Let $p_0, p_1, \ldots, p_n$ be continuous functions on an interval $I := (\alpha, \beta)$, $p_0 \ne 0$. Then for arbitrary $t_0 \in I$, and arbitrary $b_0, b_1, \ldots, b_n$ in $\mathbb{R}^1$, there exists a unique solution $y = y(t)$ of the equation

$$Ly = p_0 y^{(n)} + p_1 y^{(n-1)} + \cdots + p_{n-1} y' + p_n y = \sum_{j=0}^{n} p_{n-j} y^{(j)} = 0 \quad \text{on} \quad I, \tag{2.1}$$

where we set $y^{(0)} := y$, with the initial conditions

$$y(t_0) = b_0, \quad y'(t_0) = b_1, \quad \ldots, \quad y^{(n-1)}(t_0) = b_{n-1}. \tag{2.2}$$

Proof. Dividing by $P_0 \ne 0$, we reduce the proof to the case $p_0 \equiv 1$. Then one can rewrite (2.1) as follows:

$$y^{(n)} = f(t, y, y', \ldots, y^{(n-1)}) := -p_1(t) y^{(n-1)} - \cdots - p_{n-1}(t) y' - p_n(t) y. \tag{2.3}$$

On every closed subinterval $I_0 \subset I$, we have $|p_j| \le K$ for all j with a constant K depending on I_0. Then automatically the function f in (2.3) satisfies the Lipschitz condition with respect to the variables $y, y', \ldots, y^{(n-1)}$ on I_0, and the claim follows from Theorem 1.7. □

Definition 2.2. A **fundamental set of solutions** of the equation (2.1) is a set of solutions $y_1, y_2, \dots, y_n$, for which the **Wronskian**

$$W := W[y_1, y_2, \dots, y_n] := \begin{vmatrix} y_1 & y_2 & \dots & y_n \\ y_1' & y_2' & \dots & y_n' \\ \dots & \dots & \dots & \dots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_n^{(n-1)} \end{vmatrix} \tag{2.4}$$

is nonzero at some point $t_0 \in I$.

Theorem 2.3. The previous definition is equivalent to the statement that the solutions $y_1, y_2, \dots, y_n$ are linearly independent on I, i.e. the identity

$$y := c_1 y_1 + c_2 y_2 + \dots + c_n y_n \equiv 0 \quad \text{on} \quad I \tag{2.5}$$

with constants $c_1, c_2, \dots, c_n$ is possible only if $c_1 = c_2 = \dots = c_n = 0$.

Proof. Let $W(t_0) \neq 0$, and let the identity (2.5) hold true with some constants c_j. By differentiation, we get a system of n equations with n unknowns $c_1, c_2, \dots, c_n$:

$$\begin{cases} y & = c_1 y_1 + c_2 y_2 + \dots + c_n y_n \equiv 0, \\ y' & = c_1 y_1' + c_2 y_2' + \dots + c_n y_n' \equiv 0, \\ & \dots \\ y^{(n-1)} & = c_1 y_1^{(n-1)} + c_2 y_2^{(n-1)} + \dots + c_n y_n^{(n-1)} \equiv 0. \end{cases} \tag{2.6}$$

From Linear Algebra it is knows that this system has non-trivial solutions (not all c_j are zeroes) if and only if the determinant of the matrix of coefficients if zero. This determinant is exactly the Wronskian W in (2.4). Since $W(t_0) \neq 0$, we must have $c_1 = c_2 = \dots = c_n = 0$, which means that $y_1, y_2, \dots, y_n$ are linearly independent on I.

Now suppose that $W(t_0) = 0$ at some point $t_0 \in I$. At this point, the system (2.6) has nontrivial solution $c_1, c_2, \dots, c_n$. Then the function $y := c_1 y_1 + c_2 y_2 + \dots + c_n y_n$ satisfied the equation (2.1) with the initial conditions $y(t_0) = y'(t_0) = \dots = y^{(n-1)}(t_0) = 0$. By uniqueness of solution to the problem (2.1)–(2.2), we must have $y \equiv 0$ on I, hence $y_1, y_2, \dots, y_n$ are dependent on I. In this case, we also have $W \equiv 0$ on I. □

In the rest of this section, we consider the equations (2.1) **with constant coefficients** $p_0 \neq 0, p_1, \cdots, p_n$.

The **characteristic polynomial** of this equation is

$$p(r) = p_0 r^n + p_1 r^{n-1} + \dots + p_{n-1} r + p_n = \sum_{j=0}^{n} p_{n-j} r^j.$$

It is known from Algebra that any polynomial $p(r)$ can be represented in the form

$$p(r) = p_0 (r - r_1)^{m_1} (r - r_2)^{m_2} \cdots (r - r_s)^{m_s} = p_0 \prod_{j=1}^{s} (r - r_j)^{m_j}, \tag{2.7}$$

where r_j $(j = 1, 2, \cdots, s)$ are distinct (real or complex) zeros of $p(r)$ and m_j are their **multiplicities.** One can re-write (2.1) as follows:

$$Ly = p(D)y = \sum_{j=0}^{n} p_{n-j} D^j y = 0, \quad \text{where} \quad D := \frac{d}{dt}. \tag{2.8}$$

Lemma 2.4

$$p(D)e^{\lambda t} = p(\lambda)e^{\lambda t}.$$

This lemma is a particular case of the following one with $v = 1$.

Lemma 2.5

$$p(D)[e^{\lambda t}v(t)] = e^{\lambda t}p(D + \lambda)v(t). \tag{2.9}$$

Proof. We have

$$D[e^{\lambda t}v(t)] = \lambda e^{\lambda t}v(t) + e^{\lambda t}v'(t) = e^{\lambda t}(D + \lambda)v(t),$$

$$D^2[e^{\lambda t}v(t)] = D\left[D[e^{\lambda t}v(t)]\right] = D\left[e^{\lambda t}(D + \lambda)v(t)\right] = e^{\lambda t}(D + \lambda)^2 v(t),$$

etc.,

$$D^j[e^{\lambda t}v(t)] = e^{\lambda t}(D + \lambda)^j v(t)$$

for all j, and finally,

$$p(D)[e^{\lambda t}v(t)] = \sum_{j=0}^{n} p_{n-j} D^j[e^{\lambda t}v(t)] = e^{\lambda t}\sum_{j=0}^{n} p_{n-j}(D + \lambda)^j v(t) = e^{\lambda t}p(D + \lambda)v(t).$$

□

Corollary 2.6 If $p(r)$ has a root λ_0 of multiplicity m_0, then

$$p(D)[e^{\lambda_0 t}t^k] = 0 \quad \text{for} \quad k = 0, 1, \cdots, m_0 - 1.$$

Proof. We have $p(r) = q(r)(r - \lambda_0)^{m_0}$ for some polynomial $q(r)$. Hence

$$p(D)[e^{\lambda_0 t}t^k] = e^{\lambda_0 t}p(D + \lambda_0)t^k = e^{\lambda_0 t}q(D + \lambda_0)D^{m_0}t^k = 0$$

for all $k \le m_0 - 1$. □

Using this Corollary, we now prove the following

Theorem 2.7. Consider a linear homogeneous equation (2.1) with constant coefficients. Let its characteristic polynomial $p(r)$ be represented in the form (2.7) with distinct r_j. Then this equation has n linearly independent solutions

$$t^k e^{r_j t} \quad (j = 1, 2, \ldots s; \quad k = 0, 1, \ldots, m_j - 1), \tag{2.10}$$

so that its general solution is

$$y = \sum_{j=1}^{s} \sum_{k=0}^{m_j - 1} c_{j,k} t^k e^{r_j t}. \tag{2.11}$$

Proof. By Corollary 2.6, the functions $t^k e^{r_j t}$ in (2.10) satisfy the equation (2.1). It remains to show that these functions are linearly independent, i.e. the function y in (2.11) is identically 0 if and only if the coefficients $c_{j,k} = 0$ for all j, k. This means that for distinct $r_1, \cdots, r_s$ and arbitrary polynomials $P_1, \cdots, P_s$,

$$\sum_{j=1}^{s} e^{r_j t} P_j(t) \equiv 0 \quad \Longleftrightarrow \quad P_j(t) \equiv 0 \quad \text{for all} \quad j. \tag{2.12}$$

For $s = 1$, this statement is trivial. Suppose it is true for some $s \geq 1$ and arbitrary distinct $r_1, \cdots, r_s$, and show that it remains true for $s+1$ as well. Let

$$e^{r_1 t} P_1(t) + \cdots + e^{r_s t} P_s(t) + e^{r_{s+1} t} P_{s+1}(t) \equiv 0. \tag{2.13}$$

Note that $D^N P_{s+1} \equiv 0$ for large N. Dividing (2.13) by $e^{r_{s+1} t}$ and differentiating N times, we obtain

$$D^N \left[e^{(r_1 - r_{s+1})t} P_1(t) + \cdots + e^{(r_s - r_{s+1})t} P_s(t) \right] \equiv 0. \tag{2.14}$$

For $r \neq 0$, and any polynomial P,

$$D\left(e^{rt} P\right) = \left(e^{rt} P\right)' = e^{rt}\left(r + D\right) P = e^{rt}\left(rP + P'\right) = e^{rt} Q,$$

where the polynomial $Q = rP + P'$ has same degree as P, and $Q \equiv 0 \Longleftrightarrow P \equiv 0$. Repeating this operation N times, we rewrite (2.14) as follows:

$$e^{(r_1 - r_{s+1})t} Q_1(t) + \cdots + e^{(r_s - r_{s+1})t} Q_s(t) \equiv 0, \tag{2.15}$$

where $Q_j \equiv 0 \iff P_j \equiv 0$ for all $k \leq s$. By our assumption, (2.15) implies $Q_j \equiv 0$ for all $j \leq s$. Then $P_j \equiv 0$ for all $j \leq s$, and now (2.13) yields $P_{s+1} \equiv 0$, so that (2.12) holds true for $s+1$. By induction, it holds true for all s. □

Remark 2.8. If the set of distinct roots $\{r_j\}$ in (2.7) contains complex numbers, then they can be grouped by pairs $\alpha \pm i\beta$ of the same multiplicity. In this case, one can replace each pair of complex valued solutions in (2.10),

$$t^k e^{\alpha \pm i\beta} \quad \text{by} \quad t^k e^{\alpha t} \cos \beta t, \quad t^k e^{\alpha t} \sin \beta t.$$

Remark 2.9. In the case when every root r_j in (2.7) has multiplicity one, we have the fundamental set of solutions $\left\{ y_j(t) := e^{r_j},\ j = 1, 2, \ldots, n \right\}$. Then one can re-write the Wronskian in (2.4) as follows:

$$W := W[e^{r_1 t}, e^{r_2 t}, \ldots, e^{r_n t}] := \begin{vmatrix} e^{r_1 t} & e^{r_2 t} & \ldots & e^{r_n t} \\ r_1 e^{r_1 t} & r_2 e^{r_2 t} & \ldots & r_n e^{r_n t} \\ \ldots & \ldots & \ldots & \ldots \\ r_1^{n-1} e^{r_1 t} & r_2^{n-1} e^{r_2 t} & \ldots & r_n^{n-1} e^{r_n t} \end{vmatrix}$$

$$= e^{(r_1 + r_2 + \cdot + r_n)t} V_n, \quad \text{where} \quad V_n = V_n[r_1, r_2, \ldots, r_n] := \begin{vmatrix} 1 & 1 & \ldots & 1 \\ r_1 & r_2 & \ldots & r_n \\ \ldots & \ldots & \ldots & \ldots \\ r_1^{n-1} & r_2^{n-1} & \ldots & r_n^{n-1} \end{vmatrix}. \tag{2.16}$$

The determinant in (2.16) is called the **Vandermonde determinant**. The fact that $V_n \neq 0$, and therefore $W \neq 0$, follows immediately from the following explicit expression.

Theorem 2.10. The Vandermonde determinant in (2.16)

$$V_n[r_1, r_2, \ldots, r_n] = \prod_{1 \le j < k \le n} (r_k - r_j).$$

Proof. This is obviously true for $n = 2$: $V_2[r_1, r_2] = r_2 - r_1$. For $n \geq 3$, Consider V_n as a function of r_n with fixed $r_1, r_2, \ldots, r_{n-1}$. This is a polynomial of degree $n-1$ with $n-1$ roots $r_1, r_2, \ldots, r_{n-1}$, because if $r_n = r_j$, $1 \le j \le n-1$, then the j^{th} and n^{th} columns in (2.16) coincide, hence $V_n = 0$. Therefore,

$$V_n[r_1, r_2, \ldots, r_n] = (r_n - 1)(r_n - r_2) \cdots (r_n - r_{n-1}) \cdot V^* = \prod_{1 \le j < n} (r_n - r_j) \cdot V^*,$$

where V^* does not depend on r_n. One can see that V^* is the coefficient of r_n^{n-1}. Comparing with the decomposition of the determinant in (2.16) with respect to the last column, we conclude that $V^* = V[r_1, r_2, \ldots, r_{n-1}]$. The rest of the proof follows by induction. □

3 Linear Nonhomogeneous Equations.

Here we consider equations with constant coefficients.

Theorem 3.1. Consider a linear nonhomogeneous equation

$$Ly = p_0 y^{(n)} + p_1 y^{(n-1)} + \cdots + p_{n-1} y' + p_n y = e^{\lambda_0 t} Q_s \tag{3.1}$$

with constant coefficients $p_0 \neq 0, p_1, \ldots, p_n$, where Q_s is a polynomial of degree s. If λ_0 is a root of the characteristic polynomial $p(r) = p_0 r^n + p_1 r^{n-1} + \cdots + p_{n-1} r + p_n$ of multiplicity $m_0 \geq 0$ ($m_0 = 0$ if $p(\lambda_0) \neq 0$), then there exists a particular solution of this equation of the form

$$y_p(t) = e^{\lambda_0 t} t^{m_0} P_s(t), \tag{3.2}$$

where P_s is a polynomial of degree s.

Proof. Consider the linear mapping $\mathbb{T}P = e^{-\lambda_0 t} L \left[e^{\lambda_0 t} t^{m_0} P \right]$ on the set $\mathcal{P}_s$ of all polynomials of degree $\le s$. We can write $L = p(D) = q(D)(D - \lambda_0)^{m_0}$, where $q(\lambda_0) \neq 0$. By Lemma 2.5,

$$\mathbb{T}P = e^{-\lambda_0 t} q(D)(D - \lambda_0)^{m_0} \left[e^{\lambda_0 t} t^{m_0} P \right] = q(\lambda_0 + D) D^{m_0} \left[t^{m_0} P \right]. \tag{3.3}$$

It is easy to see that $\mathbb{T}\mathcal{P}_s \subseteq \mathcal{P}_s$, i.e. $\mathbb{T}$ is a linear mapping of the linear space $\mathcal{P}_s$ of finite dimension into itself. Moreover, since

$$q(\lambda_0 + D) = q(\lambda_0) + \frac{q'(\lambda_0)}{1!} D + \frac{q''(\lambda_0)}{2!} D^2 + \cdots + \frac{q^{(n)}}{n!} D^n, \qquad q(\lambda_0) \neq 0, \tag{3.4}$$

we have

$$\mathbb{T}P = q(\lambda_0 + D) D^{m_0} \left[t^{m_0} P \right] \equiv 0 \iff D^{m_0} \left[t^{m_0} P \right] \equiv 0 \iff P \equiv 0.$$

From Linear Algebra it is known that $\mathbb{T}$ is a one-to-one correspondence. Therefore, for given $Q_s \in \mathcal{P}_s$, there exists a unique $P_s \in \mathcal{P}_s$ satisfying $\mathbb{T}P_s = Q_s$. In other words, $L\left[e^{\lambda_0 t} t^{m_0} P_s\right] = e^{\lambda_0 t} Q_s$, which proves our statement. $\square$

Theorem 3.2. Consider a linear nonhomogeneous equation

$$Ly = p_0 y^{(n)} + p_1 y^{(n-1)} + \cdots + p_{n-1} y' + p_n y = e^{\alpha_0 t}\left[Q_s \cos\beta_0 t + Q_s^* \sin\beta_0 t\right] \tag{3.5}$$

with real constant coefficients $p_0 \neq 0, p_1, \cdots, p_n$, where Q_s and Q_s^* are polynomials of degree $\le s$, and α_0, β_0 are real numbers. If $\lambda_0 = \alpha_0 + i\beta_0$ is a root of the characteristic polynomial $p(r)$ of multiplicity $m_0 \ge 0$, then there exists a particular solution of this equation in the form

$$y_p(t) = e^{\alpha_0 t} t^{m_0}\left[P_s \cos\beta_0 t + P_s^* \sin\beta_0 t\right], \tag{3.6}$$

where P_s and P_s^* are polynomials of degree $\le s$.

This theorem follows from the previous one by re-writing $e^{\alpha_0 t}\cos\beta_0$ and $e^{\alpha_0 t}\sin\beta_0$ through $e^{(\alpha_0 \pm i\beta_0)t}$ as in Remark 2.8.

Linear nonhomogeneous equations $Ly = f$ with constant coefficients and zero initial conditions at time $t = 0$ can be considered as a particular case of systems with "input" $f(t) \dashrightarrow$ "output" $y(t)$, satisfying two properties:
(i) Linearity: if $f_j(t) \dashrightarrow y_j(t)$, then $\sum c_j f_j(t) \dashrightarrow \sum c_j y_j(t)$, where c_j are constants. By approximation, this property is also extended to integrals as limits of Riemann sums.
(ii) Delay property: if $f_j(t) \dashrightarrow y_j(t)$, then $f_j(t-\tau) \dashrightarrow y_j(t-\tau)$ for every $\tau > 0$.

Let $h(t)$ be the "output" for unit "input":

$$u_0(t) := \begin{cases} 0, & t < 0, \\ 1, & t \ge 0. \end{cases} \qquad \dashrightarrow \qquad h(t). \tag{3.7}$$

If $f(t)$ is a smooth function for $t \ge 0$, then

$$f(t) = f(0)\cdot u_0(t) + \int_0^t f'(\tau)d\tau = f(0)\cdot u_0(t) + \int_0^\infty f'(\tau)u_0(t-\tau)\,d\tau, \qquad t \ge 0.$$

By the properties (i), (ii), we have $f(t) \dashrightarrow y(t)$, where

$$y(t) = f(0)\cdot h(t) + \int_0^\infty f'(\tau)h(t-\tau)\,d\tau, \qquad t \ge 0.$$

Since $h(t) \equiv 0$ for $t < 0$, after integrating by parts, we obtain the **Duhamel's integral**

$$y(t) = (f * h')(t) := \int_0^t f(\tau)h'(t-\tau)\,d\tau, \qquad t \ge 0. \tag{3.8}$$

In application to our equations, this result can be formulated as follows.

Theorem 3.3 (Duhamel). Let $h(t)$ be the solution of the equation with constant coefficients

$$Lh = p_0 h^{(n)} + p_1 h^{(n-1)} + \cdots + p_{n-1} h' + p_n h = 0 \quad \text{for} \quad t \ge 0,$$

with the initial conditions $h(0) = h'(0 = \cdots h^{(n-1)}(0) = 0$. Then the function $y(t)$ in (3.8) is the solution of $Ly = f$ for $t \ge 0$, with the initial conditions $y(0) = y'(0 = \cdots y^{(n-1)}(0) = 0$.

The following example is just for demonstration, the suggested method here is not the shortest possible.

Example 3.4. We will apply this theorem to the problem

$$y'' + y = \sin t, \quad y(0) = y'(0) = 0. \tag{3.9}$$

We have $h(t) = 1 - \cos t$, $h'(t) = \sin t$, hence

$$y(t) = (\sin t) * (\sin t) = \int_0^t \sin(t-s) \sin s \, ds.$$

The equality

$$\sin\alpha \cdot \sin\beta = \frac{1}{2}[\cos(\alpha - \beta) - \cos(\alpha + \beta)].$$

implies

$$y(t) = \frac{1}{2}\int_0^t [\cos(t-2s) - \cos t]\, ds = \frac{1}{2}\left[\frac{\sin(t-2s)}{-2} - s\cos t\right]_{s=0}^{s=t} = \frac{1}{2}(\sin t - t\cos t).$$

Example 3.4. Solve the problem

$$Ly = y'' + 4y' + 4y = t^{-2}e^{-2t} \quad \text{for} \quad t > 0, \qquad y(1) = y'(1) = 0.$$

Note that the right hand side is not integrable near the point $t = 0$, but we still can use formula (3.8) with 1 in place of 0 as the lower limit of the integral. We have

$$h(t) = \frac{1}{4} - \Big(\frac{1}{4} + \frac{t}{2}\Big)e^{-2t}, \qquad h'(t) = te^{-2t};$$

$$y(t) = \int_1^t f(\tau)h'(t-\tau)\, d\tau = \int_1^t \tau^{-2}e^{-2\tau}(t-\tau)e^{-2(t-\tau)} d\tau = e^{-2t}\int_1^t (\tau^{-2}t - \tau^{-1}) d\tau$$

$$= e^{-2t}\Big(-\frac{t}{\tau} - \ln\tau\Big)\Big|_{\tau=1}^{\tau=t} = e^{-2t}(-1 - \ln t + t).$$

In this example, a shorter way is to write $y = (c_1 + c_2 t)e^{-2t} + y_p$, find a particular solution in the form $y_p = e^{-2t}v$:

$$Ly_p = (D+2)^2(e^{-2t}v) = e^{-2t}D^2 v = t^{-2}e^{-2t}, \quad v'' = t^{-2}, \quad v' = -t^{-1}, \quad v = -\ln t,$$

and then find the constants c_1 and c_2 from the initial conditions.

4 Proof of Auxiliary Results.

Here we give the proofs of statements (i) and (ii) in the beginning of the proof of Theorem 1.5, in a more general setting. These facts are well known, and they can be formulated in a few equivalent ways. We choose the formulations which are most convenient for our applications. In this sections, we deviate from our notations $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$: in what follows below, $x_1, x_2, \ldots$ are points in $\mathbb{R}^n$.

Theorem 4.1 (Bolzano–Weierstrass). Every bounded sequence $\{x_j\} = \{x_1, x_2, \ldots\}$ in $\mathbb{R}^n$ has a convergent subsequence $\{x_{j_k}\},\ k = 1, 2, \ldots$, i.e.

$$\text{there exists} \quad x_0 = \lim_{k\to\infty} x_{j_k}. \tag{4.1}$$

Proof. We first consider the case $n = 1$. Since $\{x_j\}$ is bounded, there exists an interval $I_0 := [a_0.b_0]$, which contains all the points x_j. Divide I_0 into two equal parts at the center point $c_0 := (a_0 + b_0)/2 \in I_0$:

$$I_0 = I_0' \cup I_0'', \quad \text{where} \quad I_0' := [a_0, c_0], \quad I_0'' := [c_0, b_0].$$

Then at least one of subintervals I_0' or I_0'' contains infinitely many x_j (more precisely, x_j for infinitely many indices j). Denote $I_1 := [a_1, b_1]$ one of such intervals. Continuing in a similar manner, we get a sequence of nested intervals

$$I_0 \supset I_1 \supset I_2 \supset \cdots \supset I_k := [a_k, b_k] \supset \cdots,$$

where each interval I_k contains infinitely many x_j. By construction, $\{a_k\}$ is non-decreasing, $\{b_k\}$ is non-increasing, $a_k < b_k$, and $b_k - a_k = 2^{-k}(b_0 - a_0) \to 0$ as $k \to \infty$. These are bounded monotone sequences, hence

$$\text{there exists} \quad x_0 = \lim_{k\to\infty} a_k = \lim_{k\to\infty} b_k. \tag{4.2}$$

Now choose a sequence $1 \le j_1 < j_2 < \cdots < j_k < \cdots$, such that $x_{j_k} \in I_k$ for every k. It is easy to see that the point x_0 in (4.2) also belongs to I_k for every k. Therefore,

$$|x_{j_k} - x_0| \le b_k - a_k = 2^{-k}(b_0 - a_0) \to 0 \quad \text{as} \quad k \to \infty,$$

which means that we have the convergence in (4.1).

In the general case of arbitrary dimension n, if $\{x_j\}$ is a bounded sequence in $\mathbb{R}^n$, then we use the previous argument to select a subsequence of $\{x_j\}$, for which their first coordinates converge in $\mathbb{R}^1$. Out of resulting subsequence, choose another subsequence, for which their second coordinates converge in $\mathbb{R}^1$. Proceeding in a similar way, we end up with a sequence for which all n coordinates converge in $\mathbb{R}^1$, which is equivalent to the convergence in $\mathbb{R}^n$. □

Closed and compact subsets in $\mathbb{R}^n$ can be defined by means of sequences as follows.

Definition 4.2 (a). A subset $K \subseteq \mathbb{R}^n$ is **closed** if from $\{x_j\} \subseteq K$ and $x_j \to x_0$ as $j \to \infty$ it follows $x_0 \in K$.

(b). A subset $K \subseteq \mathbb{R}^n$ is **compact** if every sequence $\{x_j\}$ contains a convergent subsequence x_{j_k}, and its limit x_0 belongs to K.

Theorem 4.3. A subset $K \subseteq \mathbb{R}^n$ is **compact** if and only if it is bounded and closed.

In one direction, this theorem follows immediately from Theorem 4.1 and the above definitions. We need it only in this direction, namely, we need the fact that the closed interval $[0, A]$ is compact. It is also easy to prove it in the opposite directions.

Definition 4.4. A sequence $\{x_j\}$ in $\mathbb{R}^n$ is a **Cauchy sequence** if $|x_i - x_j| \to 0$ as $i, j \to \infty$. More formally, for every $\varepsilon > 0$, there exists a natural number $m = m(\varepsilon)$ such that

$$|x_i - x_j| \le \varepsilon \quad \text{for all} \quad i, j \ge m. \tag{4.3}$$

One can also re-write (4.3) as follows:

$$\varepsilon_m := \sup_{i,j \ge m} |x_i - x_j| \to 0 \quad \text{as} \quad m \to \infty. \tag{4.4}$$

Theorem 4.5. $\mathbb{R}^n$ with the Euclidean distance $|x-y|$ is **complete**, i.e. every Cauchy sequence $\{x_j\}$ in $\mathbb{R}^n$ converges.

Proof. Let m_1 be the constant m in (4.3) corresponding to $\varepsilon = 1$. Then from (4.3) it follows:

$$|x_j| \le |x_{m_1}| + |x_{m_1} - x_j| \le |x_{m_1}| + 1 \quad \text{for all} \quad j \ge m_1.$$

Therefore, the sequence $\{x_j\}$ is bounded:

$$|x_j| \le M := \max \Big\{ |x_1|, |x_2|, \ldots, |x_{m_1 - 1}|, |x_{m_1}| + 1 \Big\} \quad \text{for all} \quad j \ge 1.$$

By Theorem 4.1, there exists a subsequence $\{x_{j_k}\}$ convergent to $x_0 \in \mathbb{R}^n$. Finally, we can use (4.4) with $i = m$ and $j = j_k$:

$$|x_m - x_0| = \lim_{k \to \infty} |x_m - x_{j_k}| \le \varepsilon_m \to 0 \quad \text{as} \quad m \to \infty.$$

This means that the whole sequence $\{x_j\}$ also converges to x_0. □

One can also define continuity in by means of sequences.

Definition 4.6 (a). Let $f(x)$ be a function defined on a set $K \subseteq \mathbb{R}^n$. We say that $f(x)$ is **continuous** at the point $x_0 \in K$, if from $\{x_j\} \subseteq K$ and $x_j \to x_0$ it follows that $f(x_j) \to f(x_0)$ as $j \to \infty$. The function $f(x)$ is continuous on K if it is continuous at every point $x_0 \in K$.

(b). The function $f(x)$ is **uniformly continuous** on K if from $\{x_j\}, \{y_j\} \subseteq K$ and $|x_j - y_j| \to 0$ it follows that $|f(x_j) - f(y_j)| \to 0$ as $j \to \infty$.

Theorem 4.7. If a function $f(x)$ is continuous on a compact $K \subset \mathbb{R}^n$, then it is uniformly continuous on K.

Proof. Suppose that this statement fails. Then there is a constant $\varepsilon > 0$ and two sequences $\{x_j\}, \{y_j\} \subseteq K$, such that

$$|x_j - y_j| \to 0 \quad \text{as} \quad j \to \infty, \quad \text{but} \quad |f(x_j) - f(y_j)| \ge \varepsilon > 0 \quad \text{for all} \quad j. \tag{4.5}$$

By Theorem 4.1, we can choose a subsequence $x_{j_k} \to x_0 \in K$ as $k \to \infty$. Then also $y_{j_k} \to x_0$, and since $f(x)$ is continuous at the point x_0, we must have $f(x_{j_k}) \to f(x_0)$ and $f(y_{j_k}) \to f(x_0)$ as $j \to \infty$. Therefore,

$$|f(x_{j_k}) - f(y_{j_k})| \le |f(x_{j_k}) - f(x_0)| + |f(y_{j_k}) - f(x_0)| \to 0 \quad \text{as} \quad k \to \infty,$$

in contradiction to (4.5) with $j = j_k$. This contradiction proves the statement. □