ODE to Joy: Notes of Dr. Bar-Natan 267 Lectures

Samer Seraj
Trinity College, University of Toronto

Contents

Definition. An **ordinary** differential equation (ODE) is an equation where the variable is a function of one variable.

Definition. A **partial** differential equation (PDE) is an equation where the variable is an equation of more than one variable, and the terms may involve partial derivatives.

Definition. The **order** of an ODE is the highest n^{th} derivative that appears in it.

For future reference, in this document we define $\mathbb{N} = \mathbb{Z}_{\geq 0} = (0, 1, 2, \ldots)$.

1 Bag of Tricks

Type 0. The primitive:

$$y' = f(x) \implies y = \int f(x) + c \text{ for any } c \in \mathbb{R}.$$

Type 1. First order linear homogeneous:

$$a(x)y' + b(x)y = 0 \implies \frac{y'}{y} = p(x), \text{ where } p(x) = -\frac{b(x)}{a(x)}$$

$$\implies \log|y| = \int \frac{y'}{y} = \int p(x) + d \text{ for any } d \in \mathbb{R}$$

$$\implies y = ce^{\int p(x)} \text{ for any non-zero } c \in \mathbb{R}.$$

Type 2. First order linear non-homogeneous:

$$a(x)y' + b(x)y = c(x) \implies y' + p(x)y = q(x), \text{ where } p(x) = \frac{b(x)}{a(x)}, q(x) = \frac{c(x)}{a(x)}.$$

Multiply both sides of the equation by an **integrating factor** μ so that the LHS is a derivative: $\mu y' + \mu p(x)y = \mu q(x)$. We want, by Leibniz's Rule, any μ such that $\mu' = \mu p(x)$, which we solved in Type 1. Then

$$(\mu y)' = \mu y' + \mu' y = \mu y' + \mu p(x)y = \mu q(x) \implies \mu y = \int \mu q(x) + c \implies y = \frac{1}{\mu}\int \mu q(x) + \frac{c}{\mu} \text{ for any } c \in \mathbb{R}.$$

Note. Remembering the formula is not recommended. Use the process each time.

Type 3. Separable equations:

$$y' = f(x)g(y) \iff m(y)y' = n(x).$$

Find $\begin{cases} M : M'(x) = m(x) \\ N : N'(x) = n(x) \end{cases}$. Then $M(y(x)) = \int m(y)y'dx = \int n(x)dx + c = N(x) + c$, for any $c \in \mathbb{R}$. This is an algebraic equation and therefore is left to the reader to solve in each specific case.

Type 4. Autonomous equations (case of separable): $y' = f(y)$.

Note. In the case that the regular method is not convenient, it may be helpful to solve for y^{-1} instead.

Type 5. Homogeneous equations: $y' = F\left(\frac{y}{x}\right)$.

Let $v(x) = \frac{y(x)}{x}$. Then $y(x) = xv(x) \implies v + xv' = y' = F(v) \implies xv' = f(v)$, where $f(v) = F(v) - v$. This is separable.

2 Exact Equations

Definition. Let A be an open rectangle of $\mathbb{R}^2$. Let $M, N : A \to \mathbb{R}$ be continuous functions. Then the differential equation

$$M(x,y) + N(x,y) \cdot y' = 0$$

is called **exact** if there exists a **potential** function $F : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\frac{\partial \psi(x,y)}{\partial x} = M(x,y) \text{ and } \frac{\partial \psi(x,y)}{\partial y} = N(x,y).$$

Criterion. For a differential equation as in the definition above, a potential function ψ exists if and only if

$$M_y = \frac{\partial M(x,y)}{\partial y} = \frac{\partial N(x,y)}{\partial x} = N_x.$$

Proof. Suppose that a potential function ψ exists. Then

$$\begin{cases} \dfrac{\partial \psi(x,y)}{\partial x} = M(x,y) \implies \dfrac{\partial^2 \psi(x,y)}{\partial y \partial x} = \dfrac{\partial M(x,y)}{\partial y} \\[2ex] \dfrac{\partial \psi(x,y)}{\partial y} = N(x,y) \implies \dfrac{\partial^2 \psi(x,y)}{\partial x \partial y} = \dfrac{\partial N(x,y)}{\partial x} \end{cases}.$$

By Schwarz' theorem,

$$\frac{\partial^2 \psi(x,y)}{\partial y \partial x} = \frac{\partial^2 \psi(x,y)}{\partial x \partial y} \implies \frac{\partial M(x,y)}{\partial y} = \frac{\partial N(x,y)}{\partial x}.$$

Conversely, suppose that it holds that $\frac{\partial M(x,y)}{\partial y} = \frac{\partial N(x,y)}{\partial x}$. Then we construct a potential ψ. Let

$$\psi = \int M dx + \int \left(N - \frac{\partial}{\partial y}\int M dx\right) dy.$$

Now we have

$$\frac{\partial\psi}{\partial y} = \frac{\partial}{\partial y}\int M dx + N - \frac{\partial}{\partial y}\int M dx = N \qquad \text{and} \qquad \frac{\partial\psi}{\partial x} = M + \frac{\partial}{\partial x}\int \left(N - \frac{\partial}{\partial y}\int M dx\right) dy,$$

so we need that the second term in the second equation is 0. It suffices to show that

$$\frac{\partial}{\partial x}\left(N - \frac{\partial}{\partial y}\int M dx\right) = 0,$$

since independence of the integrand from x implies independence of the integral from x.
By Schwarz's theorem and the hypothesis, we have, as desired,

$$\frac{\partial}{\partial x}\left(N - \frac{\partial}{\partial y}\int M dx\right) = \frac{\partial N}{\partial x} - \frac{\partial^2}{\partial x \partial y}\int M dx = \frac{\partial N}{\partial x} - \frac{\partial^2}{\partial y \partial x}\int M dx = \frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} = 0.$$

■

Algorithm. It is impractical to memorize the formula used to construct ψ. The following process is natural:

1. Verify that the equation is exact by checking the sufficient direction of the above criterion.
2. Then $\psi_x = M \implies \psi = \int M dx + g(y)$, for some function g entirely dependent on y.
3. Then $N = \psi_y = \frac{\partial}{\partial y}\int M dx + g'(y) \implies g(y) = \frac{\partial}{\partial x}\left(N - \frac{\partial}{\partial y}\int M dx\right)$.
4. Thus $\psi = \int M dx + \int \left(N - \frac{\partial}{\partial y}\int M dx\right) dy$.
5. Finally by the chain rule, $0 = M + Ny' = \psi_x + \psi_y y' = \frac{d}{dx}\psi(x,y) \implies \psi(x,\phi(x)) = c$, for any $c \in \mathbb{R}$, which implicitly defines all solutions $y = \phi$.

Integration Factors. If an equation is not exact, it may be possible to turn it into an exact equation by multiplying by a function μ:

$$M + Ny' = 0 \implies \mu M + \mu N y' = 0.$$

This id exact if and only if $(\mu M)_y = (\mu N)_x$, which is difficult to manipulate without further information. We consider a special case of two cases: assume that μ is dependent on only x or only y.

Case 1: Assume that μ is dependent on only x:

$$(\mu M)_y = (\mu N)_x \implies \mu M_y = \mu_x N + \mu M_x \implies \frac{\mu'}{\mu} = \frac{M_y - N_x}{N}.$$

In the case that the RHS is dependent only on x, it is a Type 1 equation and we are done.

Case 2: Assume that μ is dependent on only y:

$$(\mu M)_y = (\mu N)_x \implies \mu_y M + \mu M_y = \mu N_x \implies \frac{\mu'}{\mu} = \frac{N_x - M_y}{M}.$$

In the case that the RHS is dependent only on y, it is a Type 1 equation and we are done.

3 Existence and Uniqueness Theorem

Definition. Let A be a closed interval in $\mathbb{R}$. A function $f : A \to \mathbb{R}$ is said to be **Lipschitz** if

$$\exists K \in \mathbb{R}, \forall x, y \in A, |f(x) - f(y)| \leq K \cdot |x - y|.$$

Theorem. Let $x_0, y_0 \in \mathbb{R}$, and let $I = [-a + x_0, a + x_0], J = [-b + y_0, b + y_0]$ for given $a, b \in \mathbb{R}^+$. Let $f : I \times J \to \mathbb{R}$ be a given function that is **uniformly** Lipschitz in the second variable; i.e.

$$\exists K \in \mathbb{R}, \forall x \in I, \forall y_1, y_2 \in J, |f(x, y_1) - f(x, y_2)| \leq K \cdot |y_1 - y_2|,$$

and overall continuous. Then $\exists \delta \leq a, \exists! \phi : [-\delta + x_0, \delta + x_0] \to \mathbb{R}$ such that

$$\phi(x_0) = \phi_0 \text{ and } \forall |x| \leq \delta, \phi'(x) = f(x, \phi(x)).$$

Note. This is called uniform because the constant K is independent of the first variable t.

Proof. Suppose the hypotheses. Inspired by the fact that if ϕ is a solution, then

$$\phi'(x) = f(x, \phi(x)) \implies \phi(x) = y_0 + \int_{x_0}^{x} f(t, \phi(t))dt,$$

we define a sequence of functions $\{\phi_n : [-\delta + x_0, \delta + x_0] \to \mathbb{R}\}_{n \in \mathbb{N}}$ by

$$\phi_0 = y_0 \qquad \text{and} \qquad \forall n \in \mathbb{N}, \phi_{n+1} = y_0 + \int_{x_0}^{x} f(t, \phi_n(t))dt,$$

where $\delta = \min\left(a, \dfrac{b}{M}\right)$ and $M = \sup\{|f(x, y)| : x \in I, y \in J\}$. We will prove the theorem in four steps:

1. The sequence is well-defined, meaning every term exists.
2. The sequence converges to a function.
3. The function at the point of convergence satisfies the differential equation and initial condition.
4. The function is the unique solution.

The proofs are as follows (partially adapted from the wiki pages):

1. *Proof.* For each $n \in \mathbb{N}$, let $g : I \to \mathbb{R}^2$ be defined by $\forall x \in I, g(x) = (x, \phi_n(x))$. The non-trivial part of proving that each expression exists is proving that $\forall n \in \mathbb{N}, f(x, \phi_n(x)) = f(g(x))$ is integrable on I. This is true if the domain is rectifiable, and $f \circ g$ is continuous and bounded (Munkres' *AOM*, p.112). I is a closed interval in $\mathbb{R}$, so it is rectifiable. f is defined to be continuous and g is continuous because both of its component functions are continuous (Munkres' *AOM*, p.28) (each ϕ_n can be shown to be continuous by simple induction since it is a constant plus the integral of a continuous function (Spivak's *Calculus*, p.270)); thus, $f \circ g$ is continuous by the chain rule. We will now prove boundedness; i.e.

   $$\forall n \in \mathbb{N} \forall x \in [-\delta + x_0, \delta + x_0], |\phi_n(x) - y_0| \leq .$$

   We will show this by induction. ϕ_0 is constant which completes the induction basis. Suppose that the proposition is true for some ϕ_n. Then

   $$|\phi_{n+1} - y_0| = \left|\int_{x_0}^{x} f(t, \phi_n(t))dt\right| \leq \left|\int_{x_0}^{x} |f(t, \phi_n(t))|\, dt\right| \leq \left|\int_{x_0}^{x} M dt\right| = M|x_0 - x| \leq M\delta \leq M \cdot \frac{b}{M} = b.$$

   Thus, each integral exists and so each term is well-defined.

2. *Proof.* Recall the following homework problem:

   Lemma. Let $\phi_n \colon X \to \mathbb{R}$ be a sequence of functions defined on some set X, and suppose that some sequence c_n of non-negative reals is given such that for every $x \in X$, $|\phi_n(x) - \phi_{n+1}(x)| \leq c_n$. Suppose also that $\sum_{n=1}^{\infty} c_n$ is finite. Then the sequence ϕ_n is uniformly convergent.

   Proof. Let $\sigma = \sum_{n=1}^{\infty} c_n$. By the Vanishing Condition, $\lim_{n\to\infty} c_n = 0$. Let $\epsilon > 0$ be arbitrarily chosen. Then $\exists N \in \mathbb{Z}^+$ such that $\sum_{n=1}^{N} c_n > S - \epsilon$. Choose such an N. Then $\sum_{n=N}^{\infty} c_n < \epsilon$. Then $\forall x \in X, \forall m > N, n > N$, we have $|\phi_n(x) - \phi_m(x)| < \epsilon$.
   Now we need such a sequence of constants corresponding to the sequence of ϕ_n.

$$|\phi_n(x) - \phi_{n-1}(x)| = \left| \int_{x_0}^{x} f(t, \phi_{n-1}(t))dt - \int_{x_0}^{x} f(t, \phi_{n-2}(t))dt \right|$$
$$\leq \left| \int_{x_0}^{x} |f(t, \phi_{n-1}(t)) - f(t, \phi_{n-2}(t))|dt \right| \leq \left| \int_{x_0}^{x} K|\phi_{n-1}(t) - \phi_{n-2}(t)|dt \right|$$
$$\leq | \int_{x_0}^{x} K \frac{Mk^{n-2}}{(n-1)!} |t - x_0|^{n-1} dt| = \frac{MK^{n-1}}{(n-1)!} \int_{0}^{|x-x_0|} t^{n-1} dt$$
$$= \frac{MK^{n-1}}{n!} |x - x_0|^n \leq \frac{M\delta^n K^{n-1}}{n!}.$$

   But $\sum_{n=1}^{\infty} \frac{M\delta^n K^{n-1}}{n!} = \frac{M}{K} \cdot \sum_{n=1}^{\infty} \frac{(\delta K)^n}{n!}$ converges, so the sequence of ϕ_n converges by the Lemma.

3. *Proof.* Since $\{\phi_n\}_{n\in\mathbb{N}}$ univermly converges, we can say that

$$\phi(x) = \lim_{n\to\infty} \phi_n(x) = y_0 + \lim_{n\to\infty} \int_{x_0}^{x} f(t, \phi_n(t))dt = y_0 + \int_{x_0}^{x} f(t, \lim_{n\to\infty} \phi_n(t))dt = \int_{x_0}^{x} f(t, \phi(t))dt$$
$$\implies \phi'(x) = fx, \phi(x), \text{ as desired.}$$

   Finally

$$\phi(x_0) = y_0 \int_{x_0}^{x^0} f(t, \phi(t))dt = y_0, \text{ so we have constructed a solution.}$$

4. *Proof.* We know that ϕ is a solution. Suppose that ψ is another solution. Then $\forall x \in (x_0, x_0 + \delta]$, define

$$\chi(x) = |\phi(x) - \psi(x)| = \left| \int_{x_0}^{x} f(t, \phi(t))dt - K \cdot \int_{x_0}^{x} f(t, \psi(t))dt \right|$$
$$\leq \int_{x_0}^{x} |f(t, \phi(t)) - f(t, psi(t))|dt \leq \int_{x_0}^{x} |\phi(x) - psi(x)|dx = K \cdot \int_{x_0}^{x} \chi(x)dx$$
$$\implies \chi' = K \cdot \chi.$$

   This is a Type 1 equation with solutions $\chi(x) = 0$ or $c \cdot Kx$ for some nonzero $c \in \mathbb{R}$. But $\chi(x_0) = 0$, which the latter cannot achieve. Thus

$$\chi = 0 \implies \psi = psi, \text{ so there is a unique solution.}$$

■

4 Numerical Approximation Methods

As usual, we are given a function $f : \mathbb{R}^2 \to \mathbb{R}$ and fixed $x_0, y_0 \in \mathbb{R}$, and wish to find a sequence $\{\phi_n\}_{n\in\mathbb{N}}$ of approximations of a function $y : \mathbb{R}_{\geq x+0} \to \mathbb{R}$ such that

$$y'(x) = f(x, y(x)) \qquad \text{and} \qquad y(x_0) = y_0.$$

For each of the following recursively-defined sequences, let $\phi_0 = y_0$. Now $\forall n \in \mathbb{N}$,

1. **Picard's Method.**

$$\boxed{\phi_{n+1}(x) = \phi_0 + \int_a^b f(x, \phi_n(x))dx}.$$

2. **Euler Method.**
Let $\{x_n\}_{n\in\mathbb{N}}$ be a monotonically increasing sequence of real numbers with first term x_0. For each $n \in \mathbb{N}$, let $f_n = f(x_n, \phi_n)$. Then

$$\phi_{n+1} = \phi_n + f(x_n, \phi_n)(x_{n+1} - x_n) = \boxed{\phi_n + f_n \cdot (x_{n+1} - x_n)}.$$

If the sequences is increasing by the same increment h every time, then the above is $\boxed{\phi_{n+1} = \phi_n + hf_n}$.

3. **Improved Euler Method.**
Let $\{x_n\}_{n\in\mathbb{N}}$ be a monotonically increasing sequence of real numbers with first term x_0, such that $\forall n \in \mathbb{N}, x_{n+1} = x_n + h$ for some fixed $h \in \mathbb{R}^+$. Then

$$\phi_{n+1} = \phi_n + \frac{f(x_n, \phi_n) + f(x_n + h, \phi_n + hf(x_n, \phi_n))}{2} = \boxed{\phi_n + \frac{f_n + f(x_n + h, \phi_n + hf_n)}{2}}.$$

4. **Runge-Kutta Method.**

$$\boxed{\phi_{n+1} = \phi_n + \frac{\phi_n + k_{n1} + 2k_{n2} + 2k_{n3} + k_{n4}}{2} \cdot h} \text{ where } \begin{cases} k_{1n} = f\left(x_n, \phi_n\right) \\ k_{2n} = f\left(x_n + \dfrac{h}{2}, \phi_n + \dfrac{hk_{n2}}{2}\right) \\ k_{3n} = f\left(x_n + \dfrac{h}{2}, \phi_n + \dfrac{hk_{n3}}{2}\right) \\ k_{4n} = f(x_n + h, \phi_n + hk_{n3}) \end{cases}$$

5 Linear Constant-coefficient Equations

The general linear ODE of degree $n \in \mathbb{N}$ is $\displaystyle\sum_{k=0}^{n} p_k y^{(k)} = g(x)$, where $p_0, \ldots, p_n, g : [a, b] \to \mathbb{R}$ are given functions.

Theorem. If $p_0, p_1, \ldots, p_n$ are continuous on $I = [a, b]$, then there exists a unique solution $y = \phi(x)$ that satisfies the following n initial conditions for fixed $\alpha, x_0, x_1, \ldots, x_n \in \mathbb{R}$

$$y(\alpha) = x_0, y^{(1)}(\alpha) = x_1, \ldots, y^{(n-1)}(\alpha) = x_{n-1}.$$

Proof. Omitted because apparently, according to the textbook, it is difficult. The point is that if we find a solution, it's the only one and we are done. ∎

We will be looking specifically at **constant-coefficient** linear equations: $\sum_{k=0}^{n} c_k y^{(k)} = g(x)$.

First we solve constant-coefficient **homogeneous** linear equations: $\sum_{k=0}^{n} c_k y^{(k)} = 0$.

Definition. The **characteristic** equation of $\sum_{k=0}^{n} c_k y^{(k)} = 0$ is $\sum_{k=0}^{n} c_k r^k = 0$ where the variable is r.

Let the n complex roots of the characteristic equation be $r_1, \ldots, r_n$. Define a set of n linearly independent functions of x on $[a, b]$ as follows. For each $1 \leq k \leq n$:

1. If $r_k \in \mathbb{R}$ and has multiplicity $m \geq 1$, then add $e^{r_k x}, xe^{r_k x}, \ldots, x^{m-1}e^{r_k x}$ to the set.
2. If $r_k = u \pm iv \in \mathbb{C} \backslash \mathbb{R}$ and has multiplicity $m \geq 1$, then add

$$e^{ux}\cos(vx), e^{ux}\sin(vx), xe^{ux}\cos(vx), xe^{ux}\sin(vx), \ldots, x^{m-1}e^{ux}\cos(vx), x^{m-1}e^{ux}\sin(vx)$$

to the set.

Definition. Let the n elements of the above set be $y_1, \ldots, y_n$. Define the **Wronskian** as

$$W(y_1, \ldots, y_n) = \begin{vmatrix} y_1 & y_2 & \cdots & y_n \\ y_1^{(1)} & y_2^{(1)} & \cdots & y_3^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{vmatrix}$$

Theorem. $\exists c_1, \cdots, c_n \in \mathbb{R}$ such that $y = \sum_{k=1}^{n} c_k y_k$ satisfies $\begin{cases} x_0 = y(\alpha) = (c_1 y_1(\alpha) + \cdots + c_n y_n(\alpha)) \\ \vdots \\ x_{n-1} = y^{(n-1)}(\alpha) = (c_1 y_1(\alpha) + \cdots + c_n y_n(\alpha))^{(n-1)} \end{cases}$ if and only if $W(y_1, \ldots, y_n) \neq 0$.

Proof. Note that

$$\begin{cases} x_0 = y(\alpha) = (c_1 y_1(\alpha) + \cdots + c_n y_n(\alpha)) \\ \vdots \\ x_{n-1} = y^{(n-1)}(\alpha) = (c_1 y_1(\alpha) + \cdots + c_n y_n(\alpha))^{(n-1)} \end{cases}$$

$$\iff \begin{cases} x_0 = y(\alpha) = c_1 y_1(\alpha) + \cdots + c_n y_n(\alpha) \\ \vdots \\ x_{n-1} = y^{(n-1)}(\alpha) = c_1 y_1^{(n-1)}(\alpha) + \cdots + c_n y_n^{(n-1)}(\alpha) \end{cases}$$

The rest follows directly from Cramer's Rule (see Appendix). ■

Theorem. Every solution is a linear combination of the n elements of the set.

Proof. Omitted.

Reasoning for (2): Change of basis of the vector space of solutions to real linear combinations:

$$\text{span}\left(e^{(u+iv)x}, xe^{(u+iv)x}, \ldots, x^{m-1}e^{(u+iv)x}\right) \cup \text{span}\left(e^{(u-iv)x}, xe^{(u-iv)x}, \ldots, x^{m-1}e^{(u-iv)x}\right)$$

$$
\begin{aligned}
&= \operatorname{span}\left(e^u \cdot \frac{e^{ivx}+e^{-ivx}}{2}, xe^u \cdot \frac{e^{ivx}+e^{-ivx}}{2}, \dots, x^{m-1}e^u \cdot \frac{e^{ivx}+e^{-ivx}}{2}\right) \\
&\cup \left(e^u \cdot \frac{e^{ivx}-e^{-ivx}}{2i}, xe^u \cdot \frac{e^{ivx}-e^{-ivx}}{2i}, \dots, x^{m-1}e^u \cdot \frac{e^{ivx}-e^{-ivx}}{2i}\right) \\
&= \operatorname{span}\left(e^{ux}\cos(vx), xe^{ux}\cos(vx), \dots, x^{m-1}e^{ux}\cos(vx)\right) \\
&\cup \operatorname{span}\left(e^{ux}\sin(vx), xe^{ux}\sin(vx), \dots, x^{m-1}e^{ux}\sin(vx)\right).
\end{aligned}
$$

Now we solve constant-coefficient **non-homogeneous** linear equations: $\sum_{k=0}^{n} c_k y^{(k)} = g(x)$. ■

Theorem. Given the equation $\sum_{k=0}^{n} c_k y^{(k)} = g(x)$, let ϕ be a solution. Then every solution ψ is such that $\psi = \phi + \chi$, where χ is a solution of the **auxiliary** equation $\sum_{k=0}^{n} c_k y^{(k)} = 0$.

Proof. Let ϕ be a fixed solution of $\sum_{k=0}^{n} c_k y^{(k)} = g(x)$. Let ψ be any other solution. Then

$$
\sum_{k=0}^{n} c_k \phi^{(k)} = g(x) = \sum_{k=0}^{n} c_k \psi^{(k)} \implies \sum_{k=0}^{n} c_k (\phi - \psi)^{(k)}.
$$

Let $\chi = \phi - \psi$, which is a solution of the auxiliary equation. The theorem follows. ■

Note. It is finding ϕ that is the difficulty. This is done by guessing a general form, based on $g(x)$ and subbing it in to find the coefficients of the terms (Method of Undetermined Coefficients).

6 Matrix Exponentiation

Definition 1. For $A \in M_{n\times n}(\mathbb{R})$, and $t \in \mathbb{R}$, define $e^{tA} = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}$. (matrix exponentiation)

Property 1. Matrix exponentiation is **well-defined** because this series always converges.

Property 2. If I_0 is the $n \times n$ matrix of all 0 entries, then $e^{I_0} = I_n$, where I_n is the $n \times n$ identity matrix.

Property 3. Suppose that $D \in M_{n\times n}(\mathbb{R})$ is diagonal. Then

$$
D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \implies D^k = \begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix} \quad \text{and} \quad e^{tD} = \begin{bmatrix} e^{t\lambda_1} & 0 & \dots & 0 \\ 0 & e^{t\lambda_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{t\lambda_n} \end{bmatrix}.
$$

Property 4. If $A, B \in M_{n\times n}(\mathbb{R})$ such that $AB = BA$, then $e^{A+B} = e^A \cdot e^B$.

Proof.

$$
e^{A+B} = \sum_{k=0}^{\infty} \frac{(A+B)^k}{k!} = \sum_{k=0}^{\infty} \left(\frac{1}{k!} \sum_{j=0}^{\infty} \binom{k}{j} A^j B^{k-j} \right).
$$

Change variables: $k = i + j$. Then the above expression is

$$
\sum_{i,j=0}^{\infty} \frac{1}{(i+k)!} \binom{i+j}{j} A^j B^i = \sum_{i,j=0}^{\infty} \frac{A^j B^i}{i!j!} = \left(\sum_{j=0}^{\infty} \frac{A^j}{j!} \right) \cdot \left(\sum_{i=0}^{\infty} \frac{B^i}{i!} \right) = e^A \cdot e^B.
$$

Corollary. For $t, s \in \mathbb{R}$ and $A \in M_{n\times n}(\mathbb{R})$, $e^{(t+s)A} = e^{tA} \cdot e^{sA}$.

Property 5. Time to relate matrix exponentiation to differential equations. We have that $\frac{d}{dt}e^{tA} = A \cdot e^{tA}$.

Proof.

$$\frac{d}{dt}e^{tA} = \frac{d}{dt}\left(\sum_{k=0}^{\infty}\frac{t^k A^k}{k!}\right) = \sum_{k=1}^{\infty}\frac{kt^{k-1}A^k}{k!} = A \cdot \sum_{k=1}^{\infty}\frac{kt^{k-1}A^{k-1}}{(k-1)!} = A \cdot \sum_{k=0}^{\infty}\frac{t^k A^k}{k!} = A \cdot e^{tA}.$$

Property 6. If $C, D \in M_{n\times n}(\mathbb{R})$ such that D is diagonal and $A = C^{-1}DC$, then $e^A = e^{C^{-1}DC} = C^{-1}e^D C$.

Proof.

$$e^A = e^{C^{-1}DC} = \sum_{k=0}^{\infty}\frac{(C^{-1}DC)^k}{k!} = \sum_{k=0}^{\infty}\frac{C^{-1}D^kC}{k!} = C^{-1}\left(\sum_{k=0}^{\infty}\frac{D^k}{k!}\right)C = C^{-1}e^D C.$$

7 Systems of Linear Differential Equations

System of Linear **Homogeneous** Constant-coefficient Equations:

Suppose we are given a system of linear homogeneous constant-coefficient ODEs:

$$\begin{cases} y_1' = a_{11}y_1 + \cdots + a_{1n}y_n \\ \vdots \\ y_n' = a_{n1}y_1 + \cdots + a_{nn}y_n \end{cases} \quad \text{and } \; y(0) = y_0 \; \text{ is given, where } \forall x, y(x) = \begin{pmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{pmatrix}.$$

We can state the problem as finding $y(x)$. Let $y(x) = e^{Ax} \cdot y_0$ where $A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$.

The problem is solved since

$$y(0) = e^{A0} \cdot y_0 = I_n \cdot y_0 = y_0 \qquad \text{and} \qquad y'(x) = \left(e^{Ax} \cdot y_0\right)' = Ae^{Ax} \cdot y_0 = Ay.$$

System of Linear **Non-homogeneous** Constant-coefficient Equations:

8 Power Series

Definition. A **power series** $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ is said to converge at x if $\lim_{x\to\infty}\sum_{n=0}^{m} a_n(x-x_0)^n$ exists at x.

Index shift. $\forall m \in \mathbb{Z}, \sum_{n=k}^{\infty} f(k) = \sum_{n=k+m}^{\infty} f(k-m)$.

Theorem. There exists a **radius of convergence** R of $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ which is a non-negative real number such that the series converges absolutely for $|x - x_0| < R$ and diverges for $|x - x_0| > R$.

Proof. Omitted. ∎

The following is a practical method of finding R which works in many, but not all, cases.

Ratio Test. Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence of real numbers. Let $\sigma = \sum_{n=0}^{\infty} a_n$ and $L = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$. Then

1. $L < 1 \implies \sigma$ converges absolutely.
2. $L > 1 \implies \sigma$ diverges.
3. $L = 1 \implies$ the test is inconclusive as either of convergence or divergence is possible.

Proof. See Spivak's *Calculus*, p.476. ■

Definition. If $f \in C^\infty$ for all $x \in \mathbb{R}$ such that $|x - x_0| < R$, then the **Taylor series** of about $x = x_0$ is

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} \cdot (x - x_0)^n,$$

and f is said to be **analytic** at x_0.

In section 5, we solved the general constant-coefficient linear equation. Now we will solve the case where the coefficients depend on the independent variable, via series solutions.

We first tackle the homogeneous second order ODE

$$P(x) \cdot y'' + Q(x) \cdot y' + R(x) \cdot y = 0.$$

Definition. If $x_0 \in \mathbb{R}$ such that $P(x) \neq 0$, then x_0 is called an **ordinary** point for this ODE. If $P(x_0) = 0$, then x_0 is called a **singular** point for this ODE.

If we are looking for solutions around an ordinary point x_0, then we can divide by $P(x)$ and let $q(x) = \dfrac{Q(x)}{P(x)}$ and $r(x) = \dfrac{R(x)}{P(x)}$ to get $y'' + q(x) \cdot y' + r(x) \cdot y = 0$.

Algorithm for Solving.

1. Assume that a series solution $y = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ exists in a neighbourhood of an ordinary point x_0.
2. Calculate
$$\begin{cases} y' = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} (x - x_0)^n \\ y'' = \sum_{n=2}^{\infty} n(n-1) a_n (x - x_0)^{n-2} = \sum_{n=0}^{\infty} (n+2)(n+1) a_{n+2} (x - x_0)^n. \end{cases}$$
3. Summing,
$$\begin{aligned} 0 &= y'' + q(x) \cdot y' + r(x) \cdot y \\ &= \sum_{n=0}^{\infty} (n+2)(n+1) a_{n+2} (x - x_0)^n + q(x) \cdot \sum_{n=0}^{\infty} (n+1) a_{n+1} (x - x_0)^n + r(x) \cdot \sum_{n=0}^{\infty} a_n (x - x_0)^n \end{aligned}$$
4. In the final sum, the coefficient of each $(x - x_0)^n$ equals 0 and this should establish a **recurrence relation** among the a_n.
5. Given sufficient initial conditions on the coefficients a_n, it should be possible to find a unique $\{a_n\}_{n \in \mathbb{N}}$.

9 Calculus of Variations

Definition. A **functional** is a function mapping from a set of functions to $\mathbb{R}$.

Fundamental Lemma. Let $f : [a,b] \to \mathbb{R}$ be of class C^k. Suppose that for every C^k function $h : [a,b] \to \mathbb{R}$ such that $h(a) = h(b) = 0$,

$$\int_a^b f(x)h(x)dx = 0.$$

Then $\forall x \in [a,b], f(x) = 0$.

Proof. Let f satisfy the hypotheses.
Define $r : [a,b] \to \mathbb{R}$ by $\forall x \in [a,b], r(x) = (x-a)(b-x)$. Then $r \in C^\infty \implies r \in C^k$. Note that $r(a) = r(b) = 0$ and $\forall x \in (a,b), r(x) > 0$.
Define $h : [a,b] \to \mathbb{R}$ by $\forall x \in [a,b], h(x) = r(x)f(x)$. Then $h \in C^k$ and $h(a) = h(b) = 0$, so

$$\int_a^b r(x)[f(x)]^2 dx = \int_a^b f(x)h(x)dx = 0.$$

So $r(x)[f(x)]^2 = 0$ for every $x \in [a,b]$ except possibly a set of measure zero (Munkres' *AOM*, p.96). Suppose that $\exists x_0 \in [a,b] : r(x_0)[f(x_0)]^2 \neq 0$. Then $r(x_0)[f(x_0)]^2 > 0$ and by continuity, there exists a neighbourhood around x_0 where $r(x)[f(x)]^2 > 0$. But then $\int_a^b r(x)[f(x)]^2 dx > 0$, which is a contradiction.
Thus, $\forall x \in [a,b], r(x)[f(x)]^2 = 0 \implies \forall x \in [a,b], f(x) = 0$. ∎

Euler-Lagrange. Let $f : [a,b] \to \mathbb{R}$ be a given function n-times differentiable, for some fixed $n \in \mathbb{N}$, with fixed values of $f^{(k)}(a), f^{(k)}(b)$ for $k \in (0, 1, \ldots, n-1)$. Define $\phi : [a,b] \to \mathbb{R}^{n+2}$ as $\forall x \in [a,b], \phi(x) = (x, f(x), f^{(1)}(x), \ldots, f^{(n)})(x)$. Let $F : \phi([a,b]) \to \mathbb{R}$ be a given C^1 function. If f extremizes the functional

$$J(f) = \int_a^b F(\phi(x))dx.$$

then f satisfies the differential equation

$$\frac{\partial F}{\partial f} - \frac{d}{dx}\left(\frac{\partial F}{\partial f^{(1)}}\right) + \frac{d^2}{dx^2}\left(\frac{\partial F}{\partial f^{(2)}}\right) - \cdots + (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial F}{\partial f^{(n)}}\right) = 0.$$

∎

Note. This is a necessary, but not sufficient in general, condition that can be solved to reduce the number of potential functions where extrema are achieved.

We will prove the more concrete 3-variable case and the general method will be implicit.

Theorem. Let $f : [a,b] \to \mathbb{R}$ be a given differentiable function, with fixed values of $f(a) = A, f(b) = B$. Define $\phi : [a,b] \to \mathbb{R}^3$ as $\forall x \in [a,b], \phi(x) = (x, f(x), f'(x))$. Let $F : \phi([a,b]) \to \mathbb{R}$ be a given C^1 function. If f extremizes the functional

$$J(f) = \int_a^b F(\phi(x))dx,$$

then f satisfies the differential equation

$$\frac{\partial F}{\partial f} - \frac{d}{dx}\left(\frac{\partial F}{\partial f'}\right) = 0.$$

Proof. Let the hypotheses be satisfied. Suppose f is an extremum. Then any small perturbation on f that preserves the boundary values will increase or decrease the value of the functional. Let $\eta : [a,b] \to \mathbb{R}$ be an arbitrary C^1 function satisfying $\eta(a) = \eta(b) = 0$.

For each $\epsilon > 0$, define:

1. $g_\epsilon : [a,b] : \mathbb{R}$ by $\forall x \in [a,b], g_\epsilon(x) = f(x) + \epsilon\eta(x)$,
2. $\phi_\epsilon : g_\epsilon([a,b]) : \mathbb{R}^3$ by $\forall x \in g_\epsilon([a,b]), g_\epsilon(x) = (x, g_\epsilon(x), g'_\epsilon(x))$,
3. $J_\epsilon = \displaystyle\int_a^b F(g_\epsilon(x))dx = \int_a^b F_\epsilon$, where $F_\epsilon = F(g_\epsilon$.

Then by the Fundamental Theorem of Calculus,

$$\frac{dJ_\epsilon}{d\epsilon} = \frac{d}{d\epsilon}\int_a^b F_\epsilon = \frac{d}{d\epsilon}\int_a^b F(g_\epsilon(x))dx = \int_a^b \frac{d}{d\epsilon}F(g_\epsilon(x))dx.$$

Now by the Corollary to the Chain Rule (see Appendix),

$$\frac{d}{d\epsilon}F(g_\epsilon(x)) = \frac{dx}{d\epsilon}\left(\frac{\partial F_\epsilon}{\partial x}\right) + \frac{dg_\epsilon}{d\epsilon}\left(\frac{\partial F_\epsilon}{\partial g_\epsilon}\right) + \frac{dg_{\epsilon'}}{d\epsilon}\left(\frac{\partial F_\epsilon}{\partial g_{\epsilon'}}\right) = \eta(x)\left(\frac{\partial F_\epsilon}{\partial g_\epsilon}\right) + \eta'(x)\left(\frac{\partial F_\epsilon}{\partial g'_\epsilon}\right).$$

Then

$$\frac{dJ_\epsilon}{d\epsilon} = \int_a^b \eta(x)\left(\frac{\partial F_\epsilon}{\partial g_\epsilon}\right) + \eta'(x)\left(\frac{\partial F_\epsilon}{\partial g'_\epsilon}\right)dx.$$

Since f is an extremum, the above evaluated at $\epsilon = 0$ is

$$\begin{aligned}
0 = \left.\frac{dJ_\epsilon}{d\epsilon}\right|_{\epsilon=0} &= \int_a^b \left[\eta(x)\left(\frac{\partial F}{\partial f}\right) + \eta'(x)\left(\frac{\partial F}{\partial f'}\right)\right]dx \\
&= \int_a^b \left[\frac{\partial F}{\partial f} - \frac{d}{dx}\left(\frac{\partial F}{\partial f'}\right)\right]\eta(x)dx + \left[\eta(x)\left(\frac{\partial F}{\partial f'}\right)\right]_a^b \quad \text{(Integration by Parts)} \\
&= \int_a^b \left[\frac{\partial F}{\partial f} - \frac{d}{dx}\left(\frac{\partial F}{\partial f'}\right)\right]\eta(x)dx \quad (\eta(a) = \eta(b) = 0) \\
&\implies \frac{\partial F}{\partial f} - \frac{d}{dx}\left(\frac{\partial F}{\partial f'}\right) = 0 \quad \text{(Fundamental Lemma).}
\end{aligned}$$

■

Euler-Lagrange Reductions. We can state the Euler-Lagrange equations as $F_y - \frac{d}{dx}F_{y'} = 0$. This is a second order ODE. Cases that should be easy to solve:

1. F does not depend on $y' \iff F_{y'} = 0 \implies$ EL equation is $F_y = 0$, which is an algebraic equation.
2. F does not depend on $y \iff F_y = 0 \implies$ EL equation is $F_{y'} = c$ for any fixed $c \in \mathbb{R}$, which is a first order ODE.
3. F does not depend on $x \iff F_x = 0 \implies$ EL equation is

$$\begin{aligned}
0 &= F_y - \frac{d}{dx}F_{y'} \\
&= F_y - (1 \cdot 0 + F_{y'y} \cdot y' + y_{y'y'} \cdot y'') \quad \text{(Chain Rule)} \\
&= \frac{d}{dx}(F - y'F_{y'}) \quad \text{(Chain Rule)} \\
&\iff F - y'F_{y'} = c \ \text{ for any fixed } \ c \in \mathbb{R}, \ \text{ which is a first order ODE.}
\end{aligned}$$

Lagrange Multipliers. We are given $f : \mathbb{R}^n \to R$ and a restriction $g : \mathbb{R}^n \to \mathbb{R}$ where $\forall x \in \mathbb{R}^n, g(x) = 0$. We wish to find the extrema of f given the restraint g. The method, which I will not prove, is:

1. Define $F : \mathbb{R}^{n+1} \to \mathbb{R}$ by $\forall x \in \mathbb{R}^n, F(x, \lambda) = f(x) - \lambda \cdot g(x)$.
2. Calculate the gradient of F (the $(n+1)$-dimensional vector of partials of F) and set it equal to 0.
3. Solve the $n+1$ equations for $n+1$ unknowns, and the vector of the non-λ values is the point of extrema.
4. This may not be necessary, but substituting the point of extrema back into f give the extreme value.

10 Appendix

Chain Rule. Let $\begin{cases} A \subset \mathbb{R}^m \\ B \subset \mathbb{R}^n \end{cases}$ and $\begin{cases} f : A \to \mathbb{R}^n \\ g : B \to \mathbb{R}^p \end{cases}$ be functions with $f(A) \subset B$. Then

$$\begin{cases} f \text{ is differentiable at } a \\ g \text{ is differentiable at } f(a) = b \end{cases} \implies \begin{cases} g \circ f \text{ is differentiable at } a \\ D(g \circ f)(a) = Dg(b) \cdot Df(a) \end{cases}.$$

Corollary. Let $\begin{cases} A \subset \mathbb{R} \\ B \subset \mathbb{R}^n \end{cases}$ and $\begin{cases} f : A \to \mathbb{R}^n \\ g : B \to \mathbb{R} \end{cases}$ be functions with $f(A) \subset B$. Then the composite function $F = g \circ f$ maps from a subset of $\mathbb{R}$ to $\mathbb{R}$. Thus $\forall x \in A$,

$$DF(x) = Dg(f(x)) \cdot Df(x) = [D_1 g(f(x)) \cdots D_n g(f(x))] \cdot \begin{bmatrix} Df_1(x) \\ \vdots \\ Df_n(x) \end{bmatrix} = \sum_{k=1}^{n} [Df_k(x) \cdot D_k g(f(x))].$$

Cramer's Rule. Let $Ax = b$ be the matrix form of a system of n equations in n unknowns. If $\det(A) \neq 0$, then this system has a unique solution.

Proof. See Friedberg's *Linear Algebra*, p.224. ■

Euler's Formula. $e^{ix} = \cos x + i \sin x$.

Corollary. $\sin(nx) = \dfrac{e^{inx} - e^{-inx}}{2i}$ and $\cos(nx) = \dfrac{e^{inx} + e^{-inx}}{2}$.

Definition. A sequence $\{a_n\}_{n \in \mathbb{N}}$ of real numbers is said to **converge** if $\lim_{n \to \infty} a_n = L \in \mathbb{R}$, meaning

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \implies |a_n - L| \leq \epsilon.$$

Otherwise, it is said to **diverge**. It is said to converge **absolutely** if $\{|a_n|\}_{n \in \mathbb{N}}$ converges.

Definition. A series $\sum_{n=0}^{\infty} a_n$ is said to converge if the sequence of partial sums $\left\{ \sum_{k=0}^{n} a_k \right\}_{n \in \mathbb{N}}$ converges. It is said to converge **absolutely** if $\sum_{n=0}^{\infty} |a_n|$ converges.

Theorem. If a series absolutely converges, then it converges. The converse is not true in general.

Proof. See Spivak's *Calculus*, p.480. ■

Arc Length. For a differentiable function $f : [a, b] \to \mathbb{R}$, the length of the arc over this interval is

$$s = \int_a^b \sqrt{1 + [f'(x)]} dx.$$

∞