Analysis II HS21

by plaffranchi

2 ODE (ordinary differential equation)

Theorem 2.1.6

Let $F : \mathbb{R}^2 \to \mathbb{R}$ be differentiable. Let $x_0 \in \mathbb{R}$ and $y_0 \in \mathbb{R}^2$. Then the ODE $y' = F(x, y)$ has a **unique solution** f defined on a "largest" open interval I containing x_0 such that $f(x_0) = y_0$.

Definition 2.2.1

Let $I \subset \mathbb{R}$ be an open interval and $k \in \mathbb{N}_0$. An **homogeneous linear ODE of order** k on I is of the form $y^{(k)}+a_{k-1}y^{(k-1)}+\cdots+a_1y'+a_0y = 0$ where the coefficients $a_0, \ldots, a_{k-1}$ are complex-valued functions on I, and the unknown is a function $I \to \mathbb{C}$ that is k-times differentiable on I. An equation of the form $y^{(k)} + a_{k-1}y^{(k-1)} + \cdots + a_1y' + a_0y = b$, where $b : I \to \mathbb{C}$ is another function, is called an **inhomogeneous linear ODE**.

Recognize an ODE

1. no coefficients before the highest derivative
2. all coefficients are continuous
3. no products of y or their derivatives
4. no powers of y or their derivatives
5. no functions depending on y or their derivatives

Proposition 2.3.1

Any solution of $y' + ay = 0$ is of the form $f(x) = z\exp(-A(x))$ where A is a primitive of a. The unique solution with $f(x_0) = y_0$ is $f(x) = y_0 \exp(A(x_0) - A(x))$.

Solving inhomogeneous equations

Case 1: Make a guess. For example $y' = y+x^2$ guess $f(x) = ax^2+bx+c$, and solve the equation.
Case 2: Use the variation of the constant. Assume $f_p = z(x)\exp(-A(x))$ for $z : I \to \mathbb{C}$. Then $z'(x) = b(x)\exp(A(x)) \implies k(x) = \int b(x)\exp(A(x))dx$.

Definition Linear differential equations with constant coefficients

Let $k \in \mathbb{N}_0$, $a_0, \ldots, a_{k-1} \in \mathbb{C}$ fixed and b a general continuous function, then $y^{(k)} + a_{k-1}y^{(k-1)} + \cdots + a_1y' + a_0y = b$ is such equation.

Solution of hom. diff. eq. with constant coefficients

Look for solutions of the form $f(x) = e^{\alpha x}$ for $\alpha \in \mathbb{C}$. Then we have $f^{(j)}(x) = \alpha^j e^{\alpha x}$ for all $j \geqslant 0$ and for all x, which means that

$$\begin{aligned} &f^{(k)}(x) + a_{k-1}f^{(k-1)}(x) + \cdots + a_1f'(x) + a_0f(x) \\ =&e^{\alpha x}\left(\alpha^k + a_{k-1}\alpha^{k-1} + \cdots + a_1\alpha + a_0\right). \end{aligned}$$

This translates into finding the zeros of the characteristic polynomial:

$$\begin{aligned} P(X) =&X^k + a_{k-1}X^k + \cdots + a_1X + a_0 \\ =&(X - \alpha_1)\cdots(X - \alpha_k) = 0 \end{aligned}$$

Imaginary roots

If a root is not real i.e. $\alpha = \beta + i\gamma$, the solution $f(x) = e^{\alpha x}$ does not take real values, but $\overline{\alpha} = \beta - i\gamma$ is also a root, hence we can write $\tilde{f}_1(x) = e^{\beta x}\cos(\gamma x)$, $\tilde{f}_2(x) = e^{\beta x}\sin(\gamma x)$ instead of $f_1(x) = e^{\alpha x}$, $f_2(x) = e^{\bar{\alpha} x}$

Multiple roots

Case 1: no multiple roots. Any solution of the equation is of the form $f(x) = z_1e^{a_1x} + \cdots + z_ke^{a_kx}$.
Case 2: multiple roots. Suppose that α is a multiple root of order j with $2 \leqslant j \leqslant k$. Then the k functions $f_{\alpha,0}(x) = e^{\alpha x}$, $f_{\alpha,1}(x) = xe^{\alpha x}$, $\cdots$, $f_{\alpha,j-1}(x) = x^{j-1}e^{\alpha x}$ are linearly independent solutions. Taking the union of the functions $f_{\alpha,j}$ for all roots of P, each with its multiplicity, gives a basis of the space of solutions.

3 Differential calculus in $\mathbb{R}^n$

Definition 3.2.1.

Let $(x_k)_{k\in\mathbb{N}}$ where $x_k \in \mathbb{R}^n$. Write $x_k = (x_{k,1}, \ldots, x_{k,n})$. Let $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. The sequence (x_k) **converges to** $(\to)$ y as $k \to +\infty$ if $\forall\varepsilon > 0$, if $\exists N \geqslant 1$ such that $\forall n \geqslant N$, we have $\|x_k - y\| < \varepsilon$.

Lemma 3.2.2.

$(x_k) \to y$ as $k \to +\infty \iff$ either: (1) $\forall i, 1 \leqslant i \leqslant n$, the sequence of real numbers $(x_{k,i}) \to y_i$. (2) The sequence of real numbers $\|x_k - y\| \to 0$ as $k \to +\infty$.

Definition 3.2.3.

Let $X \subset \mathbb{R}^n$ and $f : X \to \mathbb{R}^m$. (1) Let $x_0 \in X$. f is **continuous at** x_0 if $\forall\varepsilon > 0\ \exists\delta > 0$ s.t. $\|x - x_0\| < \delta \implies \|f(x) - f(x_0)\| < \varepsilon, \forall x \in X$. (2) f is **continuous** on X if it is continuous at $x_0\ \forall x_0 \in X$.

Proposition 3.2.4.

Let $X \subset \mathbb{R}^n$ and $f : X \to \mathbb{R}^m$. Let $x_0 \in X$. The function f is continuous at $x_0 \iff \forall (x_k)_k \geqslant 1$ in X s.t. $x_k \to x_0$ as $k \to +\infty$, the sequence $(f(x_k))_k \geqslant 1$ in $\mathbb{R}^m$ converges to $f(x)$.

Definition 3.2.5.

Let $X \subset \mathbb{R}^n$ and $f : X \to \mathbb{R}^m$. Let $x_0 \in X$ and $y \in \mathbb{R}^m$. We say that f has the **limit** y as $x \to x_0$ with $x \neq x_0$ if for every $\varepsilon > 0$, there exists $\delta > 0$, s.t. $\forall x \in X, x \neq x_0$, s.t. $\|x - x_0\| < \delta$, we have $\|f(x) - y\| < \varepsilon$. We then write $\lim_{\substack{x\to x_0 \\ x\neq x_0}} f(x) = y$.

Proposition 3.2.7.

Let $X \subset \mathbb{R}^n$ and $f : X \to \mathbb{R}^m$. Let $x_0 \in X$ and $y \in \mathbb{R}^m$. We have $\lim_{\substack{x\to x_0 \\ x\neq x_0}} f(x) = y \iff \forall (x_k) \in X$ s.t. $x_k \to x$ as $k \to +\infty$, and $x_k \neq x_0$, the sequence $(f(x_k))$ in $\mathbb{R}^m$ converges to y.

Proposition 3.2.9.

Let $X \subset \mathbb{R}^n, Y \subset \mathbb{R}^m$ and $p \geqslant 1$ an integer. Let $f : X \to Y$ and $g : Y \to \mathbb{R}^p$ be continuous functions. Then the composite $g \circ f$ is continuous.

Definition 3.2.11.

(1) A subset $X \subset \mathbb{R}^n$ is **bounded** if the set of $\|x\|$ for $x \in X$ is bounded in $\mathbb{R}$. (2) A subset $X \subset \mathbb{R}^n$ is **closed** if for every sequence (x_k) in X that converges in $\mathbb{R}^n$ to some vector $y \in \mathbb{R}^n$, we have $y \in X$. (3) A subset $X \subset \mathbb{R}^n$ is **compact** if it is bounded and closed.

Proposition 3.2.13.

Let $g : \mathbb{R}^n \to \mathbb{R}^m$ be a continuous map. For any closed set $Y \subset \mathbb{R}^m$, $f^{-1}(Y) = x \in \mathbb{R}^n : f(x) \in Y \subset \mathbb{R}^n$ is closed.

Theorem 3.2.15.

Let $X \subset \mathbb{R}^n$ be a non-empty compact set and $f : X \to \mathbb{R}$ a continuous function. Then f is bounded and achieves its max and min. I.e. $\exists x_+, x_- \in X$ s.t. $f(x_+) = \sup_{x\in X} f(x)$, $f(x_-) = \inf_{x\in X} f(x)$.

Definition 3.3.1.

A subset $X \subset \mathbb{R}^n$ is **open** if, for any $x = (x_1, \ldots, x_n) \in X$, there exists $\delta > 0$ such that the set

$$\{y = (y_1, \ldots, y_n) \in \mathbb{R}^n : |x_i - y_i| < \delta \text{ for all } i\}$$

is contained in X. In other words: any point of $\mathbb{R}^n$ obtained by changing any coordinate of x by at most δ is still in X.

Proposition 3.3.2.

A set $X \subset \mathbb{R}^n$ is open if and only if the complement $Y = \{x \in \mathbb{R}^n : x \notin X\}$ is closed.

Corollary 3.3.3.

If $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous and $Y \subset \mathbb{R}^m$ is open, then $f^{-1}(Y)$ is open in $\mathbb{R}^n$.

Definition 3.3.5

Let $X \subset \mathbb{R}^n$ be an open set. Let $f : X \to \mathbb{R}^m$ be a function. Let $1 \leqslant i \leqslant n$. We say that f has a **partial derivative** on X with respect to the i-th variable, or coordinate, if for all $x_0 = (x_{0,1}, \ldots, x_{0,n}) \in X$, the function defined by

$$g(t) = f(x_{0,1}, \ldots, x_{0,i-1}, t, x_{0,i+1}, \ldots, x_{0,n})$$

on the set $I = \{t \in \mathbb{R} : (x_{0,1}, \ldots, x_{0,i-1}, t, x_{0,i+1}, \ldots, x_{0,n}) \in X\}$ is differentiable at $t = x_{0,i}$. Its **derivative** $g'(x_{0,i})$ at $x_{0,i}$ is denoted

$$\frac{\partial f}{\partial x_i}(x_0), \quad \partial_{x_i} f(x_0), \quad \partial_i f(x_0)$$

Proposition 3.3.7.

Consider $X \subset \mathbb{R}^n$ open and f, g functions from X to $\mathbb{R}^m$. Let $1 \leqslant i \leqslant n$. (1) If f and g have partial derivatives with respect to the i-th coordinate on X, then $f + g$ also does, and

$$\partial_{x_i}(f + g) = \partial_{x_i}(f) + \partial_{x_i}(g).$$

(2) If $m = 1$, and if f and g have partial derivatives with respect to the i-th coordinate on X, then fg also does and

$$\partial_{x_i}(fg) = \partial_{x_i}(f)g + f\partial_{x_i}(g).$$

Furthermore, if $g(x) \neq 0$ for all $x \in X$, then f/g has a partial derivative with respect to the i-th coordinate on X, with

$$\partial_{x_i}(f/g) = (\partial_{x_i}(f)g - f\partial_{x_i}(g))/g^2.$$

Definition 3.3.9. (Jacobi matrix)

Let $X \subset \mathbb{R}^n$ open and $f : X \to \mathbb{R}^m$ a function with partial derivatives on X. Write

$$f(x) = (f_1(x), \ldots, f_m(x)).$$

For any $x \in X$, the matrix

$$J_f(x) = \left(\partial_{x_j} f_i(x)\right)_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$$

with m rows and n columns is called the **Jacobi matrix** of f at x.

Definition 3.3.11 (Gradient, Divergence)

Let $X \subset \mathbb{R}^n$ be open. (1) Let $f : X \to \mathbb{R}$ be a function. If all partial derivatives of f exist at $x_0 \in X$, then the column vector

$$\begin{pmatrix} \partial_{x_1} f(x_0) \\ \cdots \\ \partial_{x_n} f(x_0) \end{pmatrix}$$

is called the **gradient** at x_0, and is denoted $\nabla f(x_0)$. (2) Let $f = (f_1, \ldots, f_n) : X \to \mathbb{R}^n$ be a function with values in $\mathbb{R}^n$ such that all partial derivatives of all coordinates f_i of f exist at $x_0 \in X$. Then the real number

$$\operatorname{Tr}(J_f(x_0)) = \sum_{i=1}^{n} \partial_{x_i} f_i(x_0),$$

the trace of the Jacobi matrix, is called the **divergence** of f at x_0, and is denoted $\operatorname{div}(f)(x_0)$.

Definition 3.4.2.

Let $X \subset \mathbb{R}^n$ be open and $f : X \to \mathbb{R}^m$ be a function. Let u be a linear map $\mathbb{R}^n \to \mathbb{R}^m$ and $x_0 \in X$. We say that f **is differentiable at** x_0 **with differential** u if

$$\lim_{x \to x_0} \frac{1}{\|x - x_0\|}(f(x) - f(x_0) - u(x - x_0)) = 0$$

where the limit is in $\mathbb{R}^m$. We then denote $df(x_0) = u$. If f is differentiable at every $x_0 \in X$, then we say that f is differentiable on X.

Proposition 3.4.4.

Let $X \subset \mathbb{R}^n$ be open and $f : X \to \mathbb{R}^m$ be a function that is differentiable on X. (1) The function f is continuous on X. (2) The function f admits partial derivatives on X with respect to each variable. (3) Assume that $m = 1$. Let $x_0 \in X$, and let $u(x_1, \ldots, x_n) = a_1 x_1 + \cdots + a_n x_n$ be the differential of f at x_0. We then have $\partial_{x_i} f(x_0) = a_i$ for $1 \leqslant i \leqslant n$.

Proposition 3.4.6.

Let $X \subset \mathbb{R}^n$ be open, $f : X \to \mathbb{R}^m$ and $g : X \to \mathbb{R}^m$ differentiable functions on X. (1) The function $f + g$ is differentiable with differential $d(f + g) = df + dg$, and if $m = 1$, then fg is differentiable. (2) If $m = 1$ and if $g(x) \neq 0$ for all $x \in X$, then f/g is differentiable.

Proposition 3.4.7.

Let $X \subset \mathbb{R}^n$ be open, $f : X \to \mathbb{R}^m$ a function on X. If f has all partial derivatives on X, and if the partial derivatives of f are continuous on X, then f is differentiable on X, with differential determined by its partial derivatives, in the sense that the matrix of the differential $df(x_0)$, with respect to the canonical basis of $\mathbb{R}^n$ and $\mathbb{R}^m$, is the Jacobi matrix of f at x_0.

Proposition 3.4.9 (Chain rule)

Let $X \subset \mathbb{R}^n$ be open, $Y \subset \mathbb{R}^m$ be open, and let $f : X \to Y$ and $g : Y \to \mathbb{R}^p$ be differentiable functions. Then $g \circ f : X \to \mathbb{R}^p$ is differentiable on X, and for any $x \in X$, its differential is given by the composition

$$d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0).$$

In particular, the Jacobi matrix satisfies

$$J_{g \circ f}(x_0) = J_g(f(x_0)) J_f(x_0)$$

where the right-hand side is a matrix product.

Definition 3.4.11.

Let $X \subset \mathbb{R}^n$ be open and $f : X \to \mathbb{R}^m$ a function that is differentiable. Let $x_0 \in X$ and $u = df(x_0)$ be the differential of f at x_0. The graph of the affine linear approximation

$$g(x) = f(x_0) + u(x - x_0)$$

from $\mathbb{R}^n$ to $\mathbb{R}^m$, or in other words the set

$$\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : y = f(x_0) + u(x - x_0)$$

is called the **tangent space** at x_0 to the graph of f.

Definition 3.4.13.

Let $X \subset \mathbb{R}^n$ be an open set and let $f : X \to \mathbb{R}^m$ be a function. Let $v \in \mathbb{R}^n$ be a non-zero vector and $x_0 \in X$. We say that f has **directional derivative** $w \in \mathbb{R}^m$ **in the direction** v, if the function g defined on the set

$$I = \{t \in \mathbf{R} : x_0 + tv \in X\}$$

by

$$g(t) = f(x_0 + tv)$$

has a derivative at $t = 0$, and this is equal to w. In other words, this means that the limit

$$\lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(x_0 + tv) - f(x_0)}{t}$$

exists and is equal to w.

Proposition 3.4.15.

Let $X \subset \mathbb{R}^n$ be an open set and let $f : X \to \mathbb{R}^m$ be a differentiable function. Then for any $x \in X$ and non-zero $v \in \mathbb{R}^n$, the function f has a directional derivative at x_0 in the direction v, equal to $\mathrm{df}(x_0)(v)$.

Definition 3.5.1.

Let $X \subset \mathbb{R}^n$ be open and $f : X \to \mathbb{R}^m$. We say that f is of class C^1 if f is differentiable on X and all its partial derivatives are continuous. The set of functions of class C^1 from X to $\mathbb{R}^m$ is denoted $C^1(X; \mathbb{R}^m)$. Let $k \geqslant 2$. We say, by induction, that f is of class C^k if it is differentiable and each partial derivative $\partial_{x_i} f : X \to \mathbb{R}^m$ is of class C^{k-1}. The set of functions of class C^k from X to $\mathbb{R}^m$ is denoted $C^k(X; \mathbb{R}^m)$.
If $f \in C^k(X; \mathbb{R}^m)$ for all $k \geqslant 1$, then we say that f is of class C^∞. The set of such functions is denoted $C^\infty(X; \mathbb{R}^m)$.

Proposition 3.5.4 (Mixed derivatives commute)

$k \geqslant 2$. Let $X \subset \mathbf{R}^n$ be open and let $f : X \to \mathbf{R}^m$ be a function of class C^k. Then the partial derivatives of order k are independent of the order in which the partial derivatives are taken: for any variables x and y, we have $\partial_{x,y} f = \partial_{y,x} f$ and for any variables x, y, z, we have

$$\partial_{x,y,z} f = \partial_{x,z,y} f = \partial_{y,z,x} f = \partial_{z,x,y} f = \cdots$$

Definition 3.5.9 (Hessian).

Let $X \subset \mathbf{R}^n$ be open and $f : X \to \mathbf{R}$ a C^2 function. For $x \in X$, the **Hessian matrix** of f at x is the symmetric square matrix

$$\mathrm{Hess}_f(x) = \left(\partial_{x_i, x_j} f\right)_{1 \leqslant i,j \leqslant n}.$$

We also sometimes write simply $H_f(x)$.