# MAT237 Multivariable Calculus
# Lecture Notes

Yuchen Wang, Tingfeng Xia

April 11, 2019

## Contents

# 1   Limits, Continuity, and Related Topics

## 1.1   Open, Closed, and other subsets of $\mathbb{R}^n$

**Definition - Open Ball**   Assume $\mathbf{a} \in \mathbb{R}^n$ and $r > 0$. The <u>open ball with centre $\mathbf{a}$and radius $r$</u> is the set denoted $B(r, \mathbf{a})$, defined by

$$B(r, \mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{a}| < r\}$$

**Definition - Sphere**   The <u>sphere with centre $\mathbf{a}$and radius $r$</u> is the set of points whose distance from $\mathbf{a}$ exactly equals $r$:

$$\{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{a}| = r\}$$

**Set Terminologies**

1. A set $S \subset \mathbb{R}^n$ is <u>bounded</u> if there exists some $r > 0$ such that $S \subset B(r, \mathbf{0})$

2. A set is <u>unbounded</u> if and only if it is not bounded

# 2   Taylor's Theorem

## 2.1   Review of Taylor's Theorem in 1 Dimension

**Definition of Taylor polynomials**   Assume $I \subset \mathbb{R}$ is an open interval and that $f : I \rightarrow \mathbb{R}$ is a function of class $C^k$ on $I$.

For a point $a \in I$, the *kth* order Taylor polynomial of $f$ at $a$ is the unique polynomial of order at most $k$, denoted$P_{a,k}(h)$ such that

$$f(a) = P_{a,k}(0)$$
$$f'(a) = P'_{a,k}(0)$$
$$\vdots$$
$$f^{(k)}(a) = P^{(k)}_{a,k}(0)$$

$$P_{a,k}(h) = f(a) + hf'(a) + \frac{h^2}{2}f"(a) + ... + \frac{h^k}{k!}f^{(k)}(a)$$

$$= \sum_{j=0}^{k} \frac{h^j}{j!}f^{(j)}(a)$$

**Remark** Taylor's Theorem guarantees that $P_{a,k}(h)$ is a very good approximation of $f(a+h)$, and that the quality of the approximation increases as $k$ increases.

**Taylor's Theorem in 1D** Assume $I \subset \mathbb{R}$ is an open interval and that $f : I \to \mathbb{R}$ is a function of class $C^k$ on $I$. For $a \in I$ and $h \in \mathbb{R}$ such that $a+h \in I$, let $P_{a,k}(h)$ denote the $k$th-order Taylor polynomial at $a$ and define the remainder

$$R_{a,k}(h) := f(a+h) - P_{a,k}(h)$$

Then

$$\lim_{h \to 0} \frac{R_{a,k}(h)}{h^k} = 0$$

## 2.2 Taylor's Theorem in higher dimensions

Assume $S \subset \mathbb{R}^n$ is an open set and that $f : S \to \mathbb{R}$ is a function of class $C^k$ on $S$. For a point $a \in S$, the *kth* order Taylor polynomial of $f$ at $a$ is the unique polynomial of order at most $k$, denoted $P_{a,k}(\mathbf{h})$ such that

$$f(\mathbf{a}) = P_{\mathbf{a},k}(\mathbf{0})$$
$$\partial^\alpha f(\mathbf{a}) = \partial^\alpha P_{\mathbf{a},k}(\mathbf{0})$$

for all partial derivatives of order up to $k$.

**Taylor's Theorem in nD** Assume $S \subset \mathbb{R}^n$ is an open interval and that $f : S \to \mathbb{R}$ is a function of class $C^k$ on $I$. For $a \in S$ and $h \in \mathbb{R}^n$ such that $a+h \in S$, let $P_{a,k}(h)$ denote the $k$th-order Taylor polynomial at $a$ and define the remainder

$$R_{a,k}(h) := f(a+h) - P_{a,k}(h)$$

Then

$$\lim_{h \to 0} \frac{R_{a,k}(h)}{|h|^k} = 0$$

**A Taylor polynomial formula for k = 2**

$$P_{\mathbf{a},2}(\mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{h} + \frac{1}{2}(H(\mathbf{a})\mathbf{h}) \cdot \mathbf{h}$$

where we remember that $\mathbf{h} = \mathbf{x} - \mathbf{a}$ if we want the result in terms of $x, y$.

# 3   Critical Points

**Definition**   A symmetric $n \times n$ matrix A is

1. **positive definite** if $\mathbf{x}^T A \mathbf{x} > 0$ for all $x \in \mathbb{R}^n \backslash \{\mathbf{0}\}$

2. **nonnegative definite** if $\mathbf{x}^T A \mathbf{x} \geq 0$ for all $x \in \mathbb{R}^n$

In addition, we say that A is

1. **negative definite** if -A is positive definite

2. **nonpositive definite** if -A is nonnegative definite

A matrix A is **indefinite** if none of the above holds. Equivalently, A is indefinite if there exist $\mathbf{x}, \mathbf{y} \in \mathbb{R}$ such that $\mathbf{x}^T A \mathbf{x} < 0 < \mathbf{y}^T A \mathbf{y}$

**Theorem 1**   Assume that A is a symmetric matrix. Then

1. A is positive definite $\iff$ all its eigenvalues are positive
   $\iff \exists \lambda_1 > 0$ such that $\mathbf{x}^T A \mathbf{x} \geq \lambda_1 |\mathbf{x}|^2$ for all $\mathbf{x} \in \mathbb{R}^n$

2. A is nonnegative definite $\iff$ all its eigenvalues are nonnegative

3. A is indefinite $\iff$ A has both positive and negative eigenvalues

**Remark**   If A is a symmetric matrix then
The smallest eigenvalue of A $= \min_{\{\mathbf{u} \in \mathbb{R}^n : |\mathbf{u}| = 1\}} \mathbf{u}^T A \mathbf{u}$

**Theorem 2**   For the matrix $A = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$,

1. if $det A < 0$, then A is indefinite

2. if $det A > 0$, then

   if $\alpha > 0$ then A is positive definite

   if $\alpha < 0$ then A is negative definite

3. if $det A = 0$ then at least one eigenvalue equals zero.

**Definition**   A critical point $\mathbf{a}$ of $C^2$ function $\mathbf{f}$ is <u>degenerate</u> if $\det(D_{\mathbf{H}}(\mathbf{a})) = 0$

**Theorem 3 - first derivative test**   If $\mathbf{f} : S \in \mathbb{R}^n \to \mathbb{R}$ is differentiable, then every local extremum is a critical point.

**Theorem 4 - second derivative test**

1. If $f : S \to \mathbb{R}$ is $C^2$ and $\mathbf{a}$ is a local minimum point for $f$, then $\mathbf{a}$ is a critical point of $f$ and $H(\mathbf{a})$ is nonnegative definite.

2. If $\mathbf{a}$ is a critical point and $H(\mathbf{a})$ is positive definite, then $\mathbf{a}$ is a local minimum point.

**Corollary**   Assume that $f$ is $C^2$ and $\nabla f(\mathbf{a}) = \mathbf{0}$

1. If H(a) is positive definite, then a is a local min;

2. If H(a) is negative definite, then a is a local max;

3. If H(a) is indefinite, then a is a saddle point;

4. If none of the above hold, then we cannot determine the character of the critical point without further thought.

**E.Knight's approach to critical points.**   In solving a question of $f : \mathbb{R}^2 \to \mathbb{R}$ we could use the following "quick check" approach:

1. Calculate the gradient of $F$, equating it to zero to find the critical points

2. Calculate the Hessian of $F$, find the correspnding matrices for each critical points, where the Hessian is defined as

$$H(f) = \begin{bmatrix} \partial_{xx}f & \partial_{xy}f = \partial_{yx}f \\ \partial_{xy}f = \partial_{yx}f & \partial_{yy}f \end{bmatrix}$$

3. Calculate the determinant of the hessian, and there are the following cases to consider

   (a) $\det H < 0$, then $sig(H) = (1,1)$ and the point is a saddle point
   (b) $\det H > 0$, then
       i. $tr(H) < 0 \implies sig(H) = (2,0)$ and the point is a local minimum
       ii. $tr(H) > 0 \implies sig(H) = (0,2)$ and the point is a local maximum
   (c) $\det H = 0$, then the test is inconclusive. We have to do this case by starring at it.

6

# 4 Lagrange Multipliers

## 4.1 Two constraints

Assume that $f, g_1$ and $g_2$ are functions $\mathbb{R}^n \to \mathbb{R}$ of class $C^1$. Assume also that $\{\nabla g_1(\mathbf{x}), \nabla g_2(\mathbf{x})\}$ are linearly independent at all $\mathbf{x}$ where $g_1(\mathbf{x}) = g_2(\mathbf{x}) = 0$ Then if $x$ is any solution to the optimization problem, there exists $\lambda_1, \lambda_2 \in \mathbb{R}$ such that the following system of equations is satisfies by $\mathbf{x}, \lambda_1$ and $\lambda_2$:

$$\begin{cases} \nabla f(\mathbf{x}) + \lambda_1 \nabla g_1(\mathbf{x}) + \lambda_2 \nabla g_2(\mathbf{x}) & = \mathbf{0} \\ g_1(\mathbf{x}) & = 0 \\ g_2(\mathbf{x}) & = 0 \end{cases}$$

## 4.2 Inequality Constraints

Consider the problem

$$\begin{cases} \text{minimize/maximize } f(\mathbf{x}) \\ \text{subject to the constraint: } g(\mathbf{x}) \leq 0 \end{cases}$$

where we assume that $g$ is $C^2$, say, and that $\nabla g(\mathbf{x} \neq 0$ on the set $\{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = 0\}$
We can reduce this to problems we already know how to solve. EVT guarantees that the problem has a solution.

**Case 1** The max or min occurs in the set $\{x \in \mathbb{R}^n : g(\mathbf{x}) < 0$
Then it is a critical point, which we know how to find and classify

**Case 2** The max or min occurs in the set $\{x \in \mathbb{R}^n : g(\mathbf{x}) = 0$
Then we can find it by the Lagrange Multiplier technique.
Finally we can choose the smallest/largest value of $f$ and the point where that value is attained from among all the candidates found in steps 1 and 2 above.

# 5 The Implicit Function Theorem

Assume that S is an open subset of $\mathbb{R}^{n+k}$ and that $F : S \to \mathbb{R}^k$ is a function of class $C^1$. Assume also that $(\mathbf{a}, \mathbf{b})$ is a point in S such that $\mathbf{F}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$ and $\det D_{\mathbf{y}} \mathbf{F}(\mathbf{a}, \mathbf{b}) \neq 0$

1. Then there exists $r_0, r_1 > 0$ such that for every $\mathbf{x} \in \mathbb{R}^n$ such that $|\mathbf{x} - \mathbf{a}| < r_0$, there exists a unique $\mathbf{y} \in \mathbb{R}^k$ such that $|\mathbf{y} - \mathbf{b}| < r_1$

$$\mathbf{F(x, y) = 0} \quad (1)$$

In other words, equation (1) implicitly defines a function $\mathbf{y} = \mathbf{f(x)}$ for $x \in \mathbb{R}^n$ near $\mathbf{a}$, with $\mathbf{y} = \mathbf{f(x)}$ close to $\mathbf{b}$. Note in particular that $\mathbf{b} = \mathbf{f(a)}$.

2. Moreover, the function $\mathbf{f} : B(r_0, \mathbf{a}) \to B(r_1, \mathbf{b}) \subset \mathbb{R}^k$ from part (1) above is of class $C^1$, and its derivatives may be determined by differentiating the identity

$$\mathbf{F(x, f(x)) = 0}$$

and solving to find the partial derivatives of $\mathbf{f}$.

**Remark**

$$D\mathbf{f(a)} = -[D_\mathbf{y}\mathbf{F(a, b)}]^{-1} D_\mathbf{x}\mathbf{F(a, b)}$$

# 6 The Inverse Function Theorem

Let U and V be open sets in $\mathbb{R}^n$, and assume that $\mathbf{f} : U \to V$ is a mapping of class $C^1$.

Assume that $\mathbf{a} \in U$ is a point such that $D\mathbf{f(a)}$ is invertible.

and let $\mathbf{b} := \mathbf{f(a)}$. Then there exist open sets $M \subset U$ and $N \subset V$ such that

1. $\mathbf{a} \in M$ and $\mathbf{b} \in N$

2. $\mathbf{f}$ is one-to-one from M onto N (hence invertible), and

3. the inverse function $f^{-1} : N \to M$ is of class $C^1$

Moreover, if $x \in M$ and $y = \mathbf{f(x)} \in N$, then

$$D(\mathbf{f}^{-1})(\mathbf{y}) = [D\mathbf{f(x)}]^{-1}$$

In particular,

$$D(\mathbf{f}^{-1})(\mathbf{b}) = [D\mathbf{f(a)}]^{-1}$$

# 7 Some Important Coordinate Systems

## 7.1 Polar Coordinates in $\mathbb{R}^2$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} = \mathbf{f}(r, \theta)$$

For $\mathbf{f}$ to be a bijection between open sets, we have to restrict its domain and range. A common choice is to specify that $\mathbf{f}$ is a function $U \to V$ where

$$U := \{(r, \theta) : r > 0, |\theta| < \pi\}, V := \mathbb{R}^2 \backslash \{(x, 0) : x \le 0\}$$

(Note that there is a half of the x-axis missing)

## 7.2 Spherical Coordinates in $\mathbb{R}^3$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cos \theta \sin \varphi \\ r \sin \theta \sin \varphi \\ r \cos \varphi \end{pmatrix} = \mathbf{f}(r, \theta, \varphi)$$

If we want $\mathbf{f}$ to be a bijection between open sets U and V, it is necessary to restrict the domain and range in some appropriate way.

## 7.3 Cylindrical Coordinates in $\mathbb{R}^3$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cos \theta \\ r \sin \theta \\ z \end{pmatrix} = \mathbf{f}(r, \theta, z)$$

# 8 k-Dimensional Manifolds in $\mathbb{R}^n$

## 8.1 The General Case

Fix $k < n$. For a k-dimensional manifold $M$ in $\mathbb{R}^n$, we say that $M$ has "degrees of freedom" $k$. There are 3 natural ways to represent $M$ (be careful with the dimensions!!! ):

**1. As a graph:**
$$\mathbf{f} : U \subset \mathbb{R}^k \to \mathbb{R}^{n-k}$$

where U is open.

$$S = \{(\mathbf{x}, \mathbf{f}(\mathbf{x})) \in \mathbb{R}^n : \mathbf{x} = \mathbf{f}(\mathbf{x}), \forall \mathbf{x} \in U\}$$

**2. As a level set:**
$$\mathbf{F} : U \in \mathbb{R}^n \to \mathbb{R}^{n-k}$$

where U is open.

$$S = \{\mathbf{x} \in U : \mathbf{F}(\mathbf{x}) = \mathbf{c}\}$$

for some $\mathbf{c} \in \mathbb{R}^{n-k}$.
This is also called the *"zero locus"* of $\mathbf{F}$ when $\mathbf{c} = \mathbf{0}$

**Remark 1**   The regularity conditions that guarantees that S is smooth is that

1. $\nabla F_1(\mathbf{x}), ..., \nabla F_{n-k}(\mathbf{x})$ are linearly independent at each $\mathbf{x} \in S$. Or equivalently,

2. the matrix $D\mathbf{F}(\mathbf{x})$ has rank $n - k$ at every $\mathbf{x} \in S$.

**Remark 2**   It can happen that the above conditions are satisfied but S is not smooth. Example: The square of a smooth F, c = 0

**3. Parametrically**
$$\mathbf{f} : U \subset \mathbb{R}^k \to \mathbb{R}^n$$

where U is open.
$$S = \{\mathbf{f}(\mathbf{u}) : \mathbf{u} \in U\}$$

**Remark**   The regularity conditions that guarantees that S is smooth is that

1. $\partial_{u_1}\mathbf{f}(\mathbf{u}), ..., \partial_{u_k}\mathbf{f}(\mathbf{u})$ are linearly independent at each $\mathbf{u} \in U$. Or equivalently,

2. the matrix $D\mathbf{f}(\mathbf{u})$ has rank $k$ at every $\mathbf{u} \in U$.

**Notes**   We can prove that if the above conditions are satisfied, then S is smooth. Construct $\mathbf{F} : \mathbb{R}^{2k} \to \mathbb{R}^k$, then use IFT (the proof is hard but worthwhile to think about since the general case implies every specific case).

## 8.2   The Specific Cases

**Theorem 1 - When is a curve regular?**   Assume that $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}$ is $C^1$, and let
$$S := \{\mathbf{x} \in \mathbb{R}^2 : F(\mathbf{x}) = 0\}$$

If $\mathbf{a} \in S$ and $\nabla F(\mathbf{a}) \neq 0$, then there exists some $r > 0$ such that $B(r, \mathbf{a}) \cap S$ is a $C^1$ graph.
(Prove directly using IFT)

**Theorem 2 - When is the parametrization regular?**   Assume that $\mathbf{f} : (a, b) \to \mathbb{R}^2$ is $C^1$, and let

$$S := \{\mathbf{f}(t) : t \in (a, b)\}$$

If $\mathbf{f}'(c) \neq 0$ for some $c \in (a, b)$, then there exists some $r > 0$ such that $\{\mathbf{f}(t) : |t - c| < r\}$ is a $C^1$ graph.

**Remark**   It says only that the parametrization is regular near $t = c$, it does not say that S is regular near $\mathbf{f}(c)$. What it means is that when increasing/decreasing t, we have no control over the path of $f(t)$.

**Theorem 3- When is a surface regular?**   conditions: $\mathbf{a} \in S$ and $\nabla F(\mathbf{a}) \neq 0$

**Theorem 4 - When is the parametrization regular?**   conditions: $D\mathbf{f}(\mathbf{c})$ has rank 2 at some $c$

## 8.3   Remarks on smoothness of a parametric smooth curve

**Definition.**   If $I \subseteq \mathbb{R}$ is an interval, a $\mathcal{C}^1$ map $\gamma : I \to \mathbb{R}^2$ is said to be

1. A *regular curve* if $\gamma'(t) \neq 0, \forall t \in I$

2. A *simple curve* if $\gamma$ is injective on the interior of $I$.

Hence if $\gamma$ is regular, then there is a neighbourhood if each point whose image looks like a graph of a $\mathcal{C}^1$-function. Simplicity guarantees that no funny overlaps can happen, and this is what is need for the curve to be smooth. **As a conclusion, we say a parametric curve is smooth if and only if it is *regular* and *simple*.** Equivalently we could also convert such parametrization into a zero locus of a $F$ to use the good old method of gradient directly.

# 9   Integration

## 9.1   Zero content

**Zero content in 1-D**   A set $S \subset \mathbb{R}$ is said to have zero content if

$$\forall \epsilon > 0, \exists \text{ intervals } I_1, ..., I_n \ s.t. \ S \subseteq \bigcup_{i=1}^{n} I_i \wedge \sum_{i=1}^{n} Len(I_i) < \epsilon$$

11

**Multidimensional zero content.** A set $S \subset \mathbb{R}^n$ is said to have zero content if

$$\forall \epsilon > 0, \exists \text{ boxes } B_1, ..., B_n \text{ s.t. } S \subseteq \bigcup_{i=1}^n B_i \wedge \sum_{i=1}^n Area(B_i) < \epsilon$$

**Consequence of zero content.** If a set $Z$ has zero content, then

$$\forall \epsilon > 0, \exists \text{ boxes } B_1, ..., B_n \text{ s.t. } S \subseteq \bigcup_{i=1}^n B_i^{int} \wedge \sum_{i=1}^n Area(B_i) < \epsilon$$

Notice the extra *int*.

**Proposition on zero content**

1. If $Z \subset \mathbb{R}^2$ has zero content and $U \subset Z$, then $U$ has zero content.

2. If $Z_1, ..., Z_k$ have zero content, then so does $\bigcup_1^k Z_j$

3. $\mathbf{f} : (a_0, b_0) \to \mathbb{R}^2$ is of class $C_1$, then $\mathbf{f}([a, b])$ has zero content whenever $a_0 < a < b < b_0$

## 9.2  Theorems of 1-D Integral Calculus

**Lemma: Refined partitions give better approximations** Let $P$ be some partition over an interval and let $P'$ be a refinement of $P$, then

$$LS_{P'} f \geq LS_P f \wedge US_{P'} \leq US_P f$$

Where LS and US stands for lower sum and upper sum respectively.

**Lemma: Lower sum is always less then or equal to upper sum** If $P$ and $Q$ are any partitions of $[a, b]$, then $LS_P f \leq US_Q f$. The essence of this proof is to consider the common refinement of these two partitions.

**Lemma. $\epsilon - \delta$ definition of integrability** If $f$ is a bounded function on $[a, b]$, the following conditions are equivalent:

1. $f$ is integrable on $[a, b]$

2. $\forall \epsilon > 0, \exists P$ of $[a, b]$ such that $US_P f - LS_P f < \epsilon$

**Theorem: Integration is "Linear"**

1. Suppose $a < b < c$. If $f$ is integrable on $[a, b]$ and on $[b, c]$, then $f$ is integrable on $[a, c]$, further more

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx$$

2. If $f$ and $g$ are integrable on $[a, b]$, then so is $f + g$, further more

$$\int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx$$

**Theorem.**   Suppose $f$ is integrable on $[a, b]$.

1. If $c \in \mathbb{R}$, the $cf$ is integrable on $[a, b]$, and $\int_a^b cf(x) = c\int_a^b f(x)dx$

2. Of $[c, d] \subset [a, b]$, then $f$ is integrable on $[c, d]$.

3. If $g$ is integrable on $[a, b]$ and $f(x) \le g(x), \forall x \in [a, b]$, then $\int_a^b f(x)dx \le \int_a^b g(x)dx$

4. $|f|$ is integrable on $[a, b]$, and $\left|\int_a^b f(x)dx\right| \le \int_a^b |f(x)|dx$

**Theorem: Bounded + monotone $\implies$ integrable**   If $f$ is bounded and monotone on $[a, b]$, then $f$ is integrable on $[a, b]$. The proof of this uses the $\epsilon - \delta$ definition of integrability

**Theorem: Continuous $\implies$ integrable**   If $f$ is continuous on $[a, b]$, then $f$ is integrable on $[a, b]$. Note that continuous is a sufficient but not necessary condition of integrability

**Theorem: discontinuous at only finite pts $\implies$ integrable**   If $f$ is bounded on $[a, b]$ and continuous at all except finitely many points in $[a, b]$, then $f$ is integrable on $[a, b]$. A easy example of this would be any $\mathbb{R}$ function that has a hole in it.

**Theorem: Discontinuous at only zero content $\implies$ integrable**   If $f$ is bounded on $[a, b]$ and the set of points in $[a, b]$ at which $f$ is discontinuous has zero content, then $f$ is integrable on $[a, b]$.

**Proposition.**   Suppose $f$ and $g$ are integrable on $[a, b]$ and $f(x) = g(x)$ for all except finitely many points $x \in [a, b]$. Then $\int_a^b f(x)dx = \int_a^b g(x)dx$.

### The Fundamental Theorem Of Calculus

1. Let $f$ be an integrable function on $[a, b]$. For $x \in [a, b]$, let $F(x) = \int_a^x f(t)dt$. Then $F$ is continuous on $[a, b]$; more-over, $F'(x)$ exists and equals $f(x)$ at every $x$ at which $f$ is continuous,

2. Let $F$ be a continuous function on $[a, b]$ that is differentiable except perhaps at finitely many points in $[a, b]$, and let $f$ be a function on $[a, b]$ that agrees with $F'$ at all points where the latter is defined. If $f$ is integrable on $[a, b]$, then $\int_a^b f(t)dt = F(b) - F(a)$

**Proposition.**   Suppose $f$ is integrable on $[a, b]$. Given $\epsilon > 0, \exists \delta > 0$ such that if $P = \{x_0, ..., x_J\}$ is any partition of $[a, b]$ satisfying

$$max\{x_j - x_{j-1} | 1 \le j \le J\} < \delta$$

the sums $LS_P f$ and $US_P f$ differ from $\int_a^b f(x)dx$ by at most $\epsilon$.

## 9.3   Generalized Integral Calculus

### Theorems of double integrals

1. If $f_1$ and $f_2$ are integrable on the bounded set $S$ and $c_1, c_2 \in \mathbb{R}$, then $c_1 f_1 + c_2 f_2$ is integrable on $S$, and

$$\iint_S [c_1 f_1 + c_2 f_2]dA = c_1 \iint_S f_1 dA + c_2 \iint_S f_2 dA$$

2. Let $S_1$ and $S_2$ be bounded sets with no points in common (intersection $= \emptyset$), and let $f$ be a bounded function. If $f$ is integrable on $S_1$ and on $S_2$, then $f$ is integrable on $S_1 \cup S_2$, in which case

$$\iint_{S_1 \cup S_2} f dA = \iint_{S_1} f dA + \iint_{S_2} f dA$$

3. If $f$ and $g$ are integrable on $S$ and $f(\mathbf{x}) \le g(\mathbf{x})$ for $\mathbf{x} \in S$, then $\iint_S f dA \le \iint_S g dA$

4. If $f$ is integrable on $S$, then so is $|f|$, and

$$\left| \iint_S f dA \right| \le \iint_S |f| dA$$

**Theorem.**   Suppose $f$ is a bounded function on the rectangle $R$. If the set of points in $R$ at which $f$ is discontinuous has zero content, then f is integrable on $R$.

**Discontinuity of characteristic function**   The function $\chi_S$ is discontinuous at $\mathbf{x}$ if and only if $\mathbf{x}$ is in the boundary of $S$.

**Theorem.**   Let $S$ be a measurable subset of $\mathbb{R}^2$. Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is bounded and the set of points in $S$ at which $f$ is discontinuous has zero content. Then $f$ is integrable on $S$.

**Remark on this theorem:**   The only points where $f_{\chi_S}$ can be discontinuous are those points in the closure of S where either $f$ or $\chi_S$ is discontinuous. Both of these cases are discontuinity on a set of zero content. And we can definitely fix $S$ inside of a rectangle, then by the previously stated theorem (The theorem directly above), such function is integrable.

**Proposition: Integration on a set of zero content evaluates to zero.**
Suppose $Z \subset \mathbb{R}^2$ has zero content. If $f : \mathbb{R}^2 \to \mathbb{R}$ is bounded, then $f$ is integrable on $Z$ and $\int_Z f dA = 0$

**Corollary**

1. Suppose that $f$ is integrable on the set $S \subset \mathbb{R}^2$. If $g(\mathbf{x}) = f(\mathbf{x})$ except for $\mathbf{x}$ in a set of zero content, then $g$ is integrable on $S$ and $\int_S g dA = \int_S f dA$

2. Suppose that $f$ is integrable on $S$ and on $T$, and $S \cap T$ has zero content. Then $f$ is integrable on $S \cup T$, and $\int_{S \cup T} f dA = \int_S fd + \int_T f dA$

**Fubini's Theorem**   Let $R = \{(x,y) : a \leq x \leq b, c \leq y \leq d\}$, and let $f$ be an integrable function on R. Suppose that, for each $y \in [c,d]$, the function $f_y$ defined by $f_y(x) = f(x,y)$ is integrable on $[a,b]$, and the function $g(y) = \int_a^b f(x,y)dx$ is integrable on $[c,d]$. Then

$$\iint_R f dA = \int_c^d \left[ \int_a^b f(x,y)dx \right] dy$$

15

Likewise, if $f^x(y) = f(x, y)$ is integrable on $[c, d]$ for each $x \in [a, b]$, and $h(x) = \int_c^d f(x, y)dy$ is integrable on $[a, b]$, then

$$\iint_R f\,dA = \int_a^b \left[ \int_c^d f(x, y)dy \right] dx$$

## 9.4   Change of Variables

**Change of Variable formula 1D**   If $g$ is a one-to-one function of class $C^1$ on the interval $[a, b]$, then for any continuous function $f$,

$$\int_{[a,b]} f(g(u))|g'(u)|du = \int_{g([a,b])} f(x)dx$$

In practice it is often more convenient to have all the $g$'s on one side of the equation. If we set $I = g([a, b])$, we have $[a, b] = g^{-1}(I)$, and

$$\int_I f(x)\,dx = \int_{g^{-1}(I)} f(g(u))|g'(u)|du$$

goal: find the analogous formula for multiple integrals. The questions is: How does the volume of a tiny piece of $n$-space change when one applies the transformation G?

**Theorem - Change of Variable for linear mappings**   Let A be an invertible $n \times n$ matrix, and let $\mathbf{G}(\mathbf{u}) = A\mathbf{u}$ be the corresponding linear transformation of $\mathbb{R}^n$. Suppose S is a measurable region in $\mathbb{R}^n$ and $f$ is an integrable function on S. Then $G^{-1}(S) = \{A^{-1}\mathbf{x} : \mathbf{x} \in S\}$ is measurable and $f \circ \mathbf{G}$ is integrable on $\mathbf{G}^{-1}(S)$, and

$$\int \cdots \int_S f(\mathbf{x})d^n\mathbf{x} = |detA| \int \cdots \int_{\mathbf{G}^{-1}(S)} f(A\mathbf{u})d^n\mathbf{u}$$

**Theorem - Change of Variable for general functions**   Given open sets U and V in $\mathbb{R}^n$, let $\mathbf{G} : U \to V$ be a one-to-one transformation of class $C^1$ whose derivative $D\mathbf{G}(\mathbf{u})$ is invertible for all $\mathbf{u} \in U$. Suppose that $T \subset U$ and $S \subset V$ are measurable sets such that $\mathbf{G}(T) = S$. If $f$ is an integrable function on S, then $f \circ \mathbf{G}$ is integrable on T, and

$$\int \cdots \int_S f(\mathbf{x})d^n\mathbf{x} = \int \cdots \int_T f(\mathbf{G}(\mathbf{u}))| \det D\mathbf{G}(\mathbf{u})|d^n\mathbf{u}$$

**Some important determinants**

1. polar coordinates: factor $= r$

2. cylindrical coordinates: factor $= |\det(Dg)| = r$

3. spherical coordinates: factor $= r^2 \sin \varphi$

# 10   Functions Defined by Integrals

**Question**

$$F(\mathbf{x}) = \int \cdots \int_S f(\mathbf{x}, \mathbf{y}) d^n \mathbf{y}$$

What condition on $f$ guarantee that F behaves well?

**Theorem 1 - Continuity of F**   Suppose S and T are compact subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and S is measurable. If $f(\mathbf{x}, \mathbf{y})$ is continuous on the set $T \times S = \{(\mathbf{x}, \mathbf{y}) : \mathbf{x} \in T, \mathbf{y} \in S\}$, then the function F defined by

$$F(\mathbf{x}) = \int \cdots \int_S f(\mathbf{x}, \mathbf{y}) d^n \mathbf{y}$$

is continuous on T.

**Theorem 2 - Differentiability of F**   Suppose $S \subset \mathbb{R}^n$ is compact and measurable, and $T \subset \mathbb{R}^m$ is open. If f is a continuous function on $T \times S$ that is of class $C^1$ as a function of $\mathbf{x} \in T$ for each $\mathbf{y} \in S$, then the function F defined by

$$F(\mathbf{x}) = \int \cdots \int_S f(\mathbf{x}, \mathbf{y}) d^n \mathbf{y}$$

is of class $C^1$ on T, and

$$\frac{\partial F}{\partial x_j}(\mathbf{x}) = \int \cdots \int_S \frac{\partial f}{\partial x_j}(\mathbf{x}, \mathbf{y}) d^n \mathbf{y} (\mathbf{x} \in T)$$

**Remark**   Situation often occur in which the variable $\mathbf{x}$ occurs in the limits of integration as well as the integrand. For simplicity we consider the case where $x$ and $y$ are scalar variables:

$$F(x) = \int_a^{\varphi(x)} f(x, y) dy (*)$$

We suppose that $f$ is continuous in $x$ and $y$ and of class $C^1$ in $x$ for each $y$, and that $\varphi$ is of class $C^1$. If $f$ does not depend on $x$, the derivative of F can be computed by the fundamental theorem of calculus together with the chain rule:

$$\frac{d}{dx}\int_a^{\varphi(x)} f(y)fy = f(\varphi(x))\varphi'(x)$$

For the more general case (*), we can differentiate F by combining this result with Theorem 2: Differentiate with respect to each $x$ in (*) while treating the others as constants, and add the results

$$F'(x) = f(\varphi(x))\varphi'(x) + \int_a^{\varphi(x)} \frac{\partial f}{\partial x}(x,y)dy$$

# 11   Improper Multiple Integrals

There are many situations where one needs to integrate functions over infinite intervals(e.g. half-space or the whole space) or functions that are unbounded near some point in the region of integration.

## 11.1   Bounded Functions on Unbounded Domains

Suppose, for example, that $f$ is a continuous function on $\mathbb{R}^2$ and we wish to define $\iint_{\mathbb{R}^2} f dA$. The obvious idea is to set

$$\iint_{\mathbb{R}^2} f dA = \lim_{r\to\infty} \iint_{S_r} f dA$$

where the $S_r$'s are a family of measurable sets that fill out $\mathbb{R}^2$ as $r \to \infty$. For example, we could take $S_r$ to be

1. the disc of radius $r$ about the origin

2. the square of side length $r$ centered at the origin

3. the rectangle of side lengths $r$ and $r^2$ centered at the origin

4. the disc of radius $r$ centered at $(15, -37)$

5. . . .

No rationale for choosing one over another and no guarantee that different families $S_r$ will yield the same limit.

**Definition 1**   We will say the **improper integral** $\int \ldots \int_{\mathbb{R}^n} f(\mathbf{x}) d^n\mathbf{x}$ is **absolutely convergent** (or sometimes just "**the improper integral exists**") if there exists $L \in \mathbb{R}$ s.t.

$$\forall \varepsilon > 0, \exists R > 0 \text{ such that } \forall S \subset \mathbb{R}^n$$

if $B(R, \mathbf{0}) \subset S$, then $|\int \ldots \int_S f(\mathbf{x}) d^n\mathbf{x} - L| < \varepsilon$

**Theorem 2**   Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is continuous, and that

$$\lim_{R \to \infty} \int \ldots \int_{B(R)} |f(\mathbf{x})| d^n\mathbf{x} =: M \in [0, \infty) \text{ exists.}$$

Then the improper integral $\int \ldots \int_{\mathbb{R}^n} f(\mathbf{x}) d^n\mathbf{x}$ is absolutely convergent.

**Corollary 1**   Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is bounded and integrable on any bounded set and satisfies $|f(\mathbf{x})| \leq C \cdot \frac{1}{||\mathbf{x}||^p}$ whenever $||\mathbf{x}|| > R$, for some constants $C, R > 0$.
If $p > n$ then $\int \ldots \int_{\mathbb{R}^n} f d^n\mathbf{x}$ exists.

*proof:*   (n=2)Suppose there exist a $p > 2$ such that the hypothesis of the theorem is satisfied. Consider $B(\mathbf{0}, R)$.
Then

$$\int \ldots \int_{\mathbb{R}^n} f d^n\mathbf{x} = \int \ldots \int_{B(\mathbf{0}, R)} f d^n\mathbf{x} + \int \ldots \int_{\mathbb{R}^2 \setminus B(\mathbf{0}, R)} f d^n\mathbf{x}$$

(if it exists)
It's sufficient to show both (1) and (2) exist for all $R > 0$.
By the hypothesis of the theorem (that $f$ is integrable on any bounded set), (1) exists.
Consider (2)

19

Convert to polar coordinates:

$$\int \ldots \int_{\mathbb{R}^2 \setminus B(\mathbf{0}, R)} f d^n \mathbf{x} = \int_0^{2\pi} \int_R^\infty f r \, dr \, d\theta$$

$$\leq \int_0^{2\pi} \int_R^\infty |f| r \, dr \, d\theta$$

$$= 2\pi \int_R^\infty |f| r \, dr \, d\theta$$

$$\leq 2\pi \int_R^\infty C \frac{1}{\|\mathbf{x}\|^p} r \, dr$$

$$= 2\pi \int_R^\infty C \frac{1}{r^p} r \, dr \qquad (\|\mathbf{x}\| = r)$$

$$= 2\pi C \int_R^\infty r^{1-p} \, dr$$

$$\lim_{d \to \infty} 2\pi C (d^{1-p} - R^{1-p})$$

exists iff

$$p - 1 > 1 \iff p > 2$$

■

## 11.2  Unbounded Functions on Bounded Domains

Now let S be a measurable subset of $\mathbb{R}^n$, and for a point $\mathbf{a} \in S^{int}$, and consider a continuous but unbounded function $f : S \setminus \{\mathbf{a}\} \to \mathbb{R}$.

**Example**
$$f(x) = |\mathbf{x} - \mathbf{a}|^{-p}$$

for some $p > 0$.

**Definition 2**  For continuous $f : S \setminus \{\mathbf{a}\} \to \mathbb{R}$, we say the **improper integral** $\int \ldots \int_{S \setminus \{\mathbf{a}\}} f(\mathbf{x}) d^n \mathbf{x}$ is **absolutely convergent** (or sometimes just **"the improper integral exists"**) if there exists $L \in \mathbb{R}$ s.t.

$$\forall \varepsilon > 0, \exists r > 0 \text{ such that } \forall U \subset S \text{ with } \mathbf{a} \in U^{int}$$

if $U \subset B(r, \mathbf{a})$, then $|\int \ldots \int_{S \setminus U} f(\mathbf{x}) d^n \mathbf{x} - L| < \varepsilon$

**Theorem 3**   Assume that $f : S \setminus \{\mathbf{a}\} \to \mathbb{R}$ is continuous, and that

$$\lim_{r \to 0} \int \ldots \int_{S \setminus B(r,\mathbf{a})} |f(\mathbf{x})| d^n \mathbf{x} \text{ exists (and is finite)}$$

Then the improper integral $\int \ldots \int_S f(\mathbf{x}) d^n \mathbf{x}$ is absolutely convergent.

**Corollary 2**   Assume that S is a bounded measurable subset of $\mathbb{R}^n$ that contains the origin, and that $f : S \setminus \{\mathbf{0}\}$ is continuous. If

$$\exists C > 0 \text{ and } p < n \text{ such that } |f(\mathbf{x})| \leq C|\mathbf{x}|^{-p}$$

for all $\mathbf{x} \in S \setminus \{\mathbf{0}\}$, then
the improper integral $\int \ldots \int_S f(\mathbf{x}) d^n \mathbf{x}$ is absolutely convergent.

**Proposition**   For $p > 0$, define $f_p$ on $\mathbb{R}^n \setminus \{\mathbf{0}\}$ by $f_p(\mathbf{x}) = |\mathbf{x}|^{-p}$. The integral of $f_p$ over a ball $\{\mathbf{x} : |\mathbf{x}| < a\}$ is finite if and only if $p < n$; the integral of $f_p$ over the complement of a ball, $\{\mathbf{x} : |\mathbf{x}| > a\}$, is finite if and only if $p > n$.

**Proposition**

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

# 12   Line Integrals

Let $\mathbf{F}$ be an $\mathbb{R}^n$-valued function defined on some subset of $\mathbb{R}^n$. In this chapter, we think of such an $\mathbf{F}$ as a function that assigns to each point $\mathbf{x}$ in its domain a vector $\mathbf{F}(\mathbf{x})$, represented pictorially as an arrow based at $\mathbf{x}$, and we therefore call it a **vector field**.

## 12.1   Arc Length and Line Integrals

line integrals: integrals over curves
Given a smooth curve C, hot to compute its length?
Based on the idea of cutting up the curve into many tiny pieces, forming appropriate Riemann sums, and passing to the limit

**Differentials on Curve; Arc Length**   Suppose C is a smooth curve in $\mathbb{R}^n$. We consider two nearby points $\mathbf{x}$ and $\mathbf{x} + d\mathbf{x}$ on the curve; here

$$d\mathbf{x} = (dx_1, \ldots, dx_n)$$

is the vector difference between the two points, and we imagine it as being infinitely small. We may, however, be more interested in the distance between the two points, traditionally denoted by $ds$, which is

$$ds = |d\mathbf{x}| = \sqrt{dx_1^2 + \ldots + dx_n^2}(*)$$

Arc length of C = result of "adding up all the $ds$'s" = $\int_C ds$

To give these differentials a precise meaning that can be used for calculations, the best procedure is to paramatrize the curve. Thus, we assume that C is given by parametric equations $\mathbf{x} = \mathbf{g}(t), a \leq t \leq b$, where $\mathbf{g}$ is of class $C^1$ and $\mathbf{g}'(t) \neq \mathbf{0}$. Then the neighboring points $\mathbf{x}$ and $\mathbf{x} + d\mathbf{x}$ are given by $\mathbf{g}(t)$ and $\mathbf{g}(t + dt)$, so

$$d\mathbf{x} = \mathbf{g}(t + dt) - \mathbf{g}(t) = \mathbf{g}'(t)dt = (\frac{dx_1}{dt}, \ldots, \frac{dx_n}{dt})dt$$

Moreover,

$$|d\mathbf{x}| = |\mathbf{g}'(t)|dt = \sqrt{(\frac{dx_1}{dt})^2 + \ldots + (\frac{dx_n}{dt})^2}dt$$

which is just what one gets by formally multiplying and dividing the expression on the right of (*) by $dt$

Integration of the vector increments $d\mathbf{x}$ just gives the total vector increment, that is, the vector difference between the initial and final points on the curve:

$$\int_C d\mathbf{x} = \int_a^b \mathbf{g}'(t)dt = \mathbf{g}(b) - \mathbf{g}(a)$$

**Definition of arc length for a smooth curve**   $ds$ is the *arc length* of the bit of curve between $d\mathbf{x}$ and $\mathbf{x} + d\mathbf{x}$. Adding these up gives the total arc length of the curve

$$\text{Arc length} = \int_C ds = \int_a^b |\mathbf{g}'(t)|dt$$

**Notes**   The arc length and the vector difference between the two endpoints of the curve should not depend on the particular parametrization we use. The issue here is that a parametrization $\mathbf{x} = \mathbf{g}(t)$ determines an **orientation** for the curve C, that is, a determination of which direction along the curve is "forward" (the direction in which the point $\mathbf{g}(t)$ moves as $t$ increases) and which direction is "backward".

**Definition of Piecewise Smoothness of a Function**   The function $\mathbf{g} : [a, b] \to \mathbb{R}^n$ is called **piecewise smooth** if

1. It is continuous

2. Its derivative exists and is continuous except perhaps at finitely many points $t_j$, at which the one-sided limits $\lim\limits_{t \to t_j \pm} \mathbf{g}'(t)$ exist.

**Line Integrals of Scalar Functions**   If $f$ is a continuous function whose domain includes a smooth (or piecewise smooth) curve C in $\mathbb{R}^n$, we can integrate $f$ over the curve, taking the differential in the integral to be the element of arc length $ds$. Thus, if C is parametrized by $\mathbf{x} = \mathbf{g}(t), a \leq t \leq b$, we define

$$\int_C f ds = \int_a^b f(\mathbf{g}(t))|\mathbf{g}'(t)| dt$$

This is independent of the parametrization and the orientation.

**Vector-Valued Line Integrals of Vector Fields**   We can define the integral of an $\mathbb{R}^m$-valued function over a curve in $\mathbb{R}^n$, simply by integrating each component separately; that is, if $\mathbf{F} = (F_1, \ldots, F_m)$, then $\int_C \mathbf{F} ds = (\int_C F_1 ds, \ldots, \int_C F_m ds)$

**Scalar-Valued Line Integrals for Vector Fields**   If C is a smooth (or piecewise smooth) curve in $\mathbb{R}^n$ and $\mathbf{F}$ is a continuous vector field defined on some neighborhood of C in $\mathbb{R}^n$, the **line integral** of $\mathbf{F}$ over C is

$$\int_C \mathbf{F} \cdot d\mathbf{x} = \int_C (F_1 dx_1 + F_2 dx_2 + \ldots + F_n dx_n)$$

That is, if C is described parametrically by $\mathbf{x} = \mathbf{g}(t), a \leq t \leq b$, then

$$\int_C \mathbf{F} \cdot d\mathbf{x} = \int_a^b \mathbf{F}(\mathbf{g}(t)) \cdot \mathbf{g}'(t) dt$$

**Fundamental Theorem of Line Integrals**   If $C$ is a curve that starts at $\mathbf{p} \in \mathbb{R}^n$ and ends at $\mathbf{q} \in \mathbb{R}^n$, then

$$\int_C \nabla \mathbf{f} \cdot d\mathbf{x} = \mathbf{f}(\mathbf{g}(b)) - \mathbf{f}(\mathbf{g}(a))$$

where $\mathbf{g}(a) = \mathbf{p}$ and $\mathbf{g}(b) = \mathbf{q}$ and the line integral around any closed path is 0.

## 12.2  Green's Theorem

"The integral of something over the boundary of a region equals the integral of something else over the region itself."

Let $R \subset \mathbb{R}^2$ be a region such that $R = \overline{R^{int}}$. Assume that $R$ is bounded. We say that $\partial R$ is piecewise smooth if:

$$\partial R = \bigcup_{i=1}^{n} C_i \text{ with } C_i \text{ being smooth, } \forall i$$

**Simple-Closed curve.**   A Jordan Simple-Closed curve is a curve in $\mathbb{R}^2$ that is closed and non-self-overlapping. Such curve divides the region of $\mathbb{R}^2$ into to portions: those that are inside this curve, and those that are outside. More precisely, a simple closed curve is one that can be parametrized by a continuous map $\mathbf{x} = \mathbf{g}(t), a \leq t \leq b$, such that $\mathbf{g}(a) = \mathbf{g}(b)$ but $\mathbf{g}(s) \neq \mathbf{g}(t)$ unless $\{s, t\} = \{a, b\}$

**Positive/ Natural Orientation**   The **positive orientation** on $\partial S$ is the orientation on each of the closed curves that make up the boundary such that the region S is on the *left* with respect to the positive direction on the curve

**Regular Region**   A compact set in $\mathbb{R}^n$ that is the closure of its interior.

**Statement of the Green's Theorem - Textbook Version**   Suppose S is a regular region in $\mathbb{R}^2$ with piecewise smooth boundary $\partial S$. Suppose also that $\mathbf{F}$ is a vector field of class $C^1$ on $\bar{S}$. Then

$$\int_{\partial S} \mathbf{F} \cdot d\mathbf{x} = \iint_S \left( \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dA$$

In the more common notation, if we set $\mathbf{F} = (P, Q)$ and $\mathbf{x} = (x, y)$,

$$\int_{\partial S} P \, dx + Q \, dy = \iint_S (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}) dA$$

The sum of line integrals over all curves that make up $\partial S$

**Statement of the Green's Theorem - E.Knight's Version**   Let $C$ be a positively oriented, piece-wise smooth, simple closed curve in a plane, and let $D$ be the region bounded by $C$. If $L(\cdot)$ and $M(\cdot)$ are functions of $(x, y)$ defined on an open region containing $D$ and have continuous partial derivatives there, then

$$\oint_C (L dx + M dy) = \iint_D \left( \frac{\partial M}{\partial x} - \frac{\partial L}{\partial y} \right) dx dy$$

where the path of integration along $C$ is anticlockwise.

**Notes**

1. Green's Theorem implies that

$$\int_{\partial S} x \, dy = -\int_{\partial S} y \, dx = \int_{\partial S} \frac{1}{2} (x \, dy - y \, dx) = \iint_S 1 \, dA = area(S)$$

**Reformulation of Green's Theorem**   Let $S \subset \mathbb{R}^2$ be a regular domain with piecewise smooth boundary. If $\mathbf{F}$ is a $C^1$ vector field defined on an open set that contained S, then

$$\iint_S (\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y}) \, dA = \int_{\partial S} \mathbf{F} \cdot \mathbf{n} \, ds$$

where $\mathbf{n}(t) = (t_2, -t_1) \perp \partial S$.

# 13   Surface Integrals

On a curve the orientation is a matter of deciding which direction along a curve is "positive"; on a surface it is a matter of deciding which side of the surface is the "positive" side.
Not every surface can be oriented. Counterexample: Mobius band
However, if a surface forms part of the boundary of a regular region in $\mathbb{R}^3$, it is always orientable, and the positive normal vector is defined to be the one pointing out of the region.

**Element of Area**

$$dA = |\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}| \, du \, dv$$

**Area for a parametrized surface**  If R is a measurable subset of W in the $uv$-plane and $\mathbf{G}(R)$ is the corresponding region in the surface S,

$$\text{Area of } \mathbf{G}(\text{R}) \;=\; \iint_R |\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}| \, du \, dv$$

**Surface Integrals of scalar Functions**  If S admits a parametrization $\mathbf{x} = \mathbf{G}(u,v)$ with $(u,v) \in W$, where W is tacitly assumed to be measurable,

$$\iint_S f \, dA = \iint_W f(\mathbf{G}(u,v))|\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}| \, du \, dv$$

If S is the graph of a function $z = \varphi(x,y), (x,y) \in W$, the result is

$$\iint_S f \, dA = \iint_W f(x,y,\varphi(x,y))\sqrt{1 + (\partial_x \varphi)^2 + (\partial_y \varphi)^2} \, dx \, dy$$

**Surface Integrals of Vector Fields**  It is natural to regard the vector $(\partial_u \mathbf{G} \times \partial_v \mathbf{G}) \, du \, dv$ itself as a "vector element of area" for S: its magnitude gives the area of a small bit of S, and its direction, namely the normal direction to S, specifies how that bit is oriented in space. That is, we have

$$(\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}) \, du \, dv = \mathbf{n} \, dA$$

where $\mathbf{n}$ is a unit normal vector to the surface S.

$dA$ is independent of the parametrization, and clearly so is $\mathbf{n}$ up to a factor of $\pm 1$ (using a different parametrization might result in replacing $\mathbf{n}$ by $-\mathbf{n}$. Suppose S is a surface with a specified orientation, and $\mathbf{F}$ is a continuous vector field defined on a neighborhood of S. The **surface integral** of $\mathbf{F}$ over S is defined to be

$$\iint_S \mathbf{F} \cdot \mathbf{n} \, dA$$

Thus, if S is parametrized by $\mathbf{x} = \mathbf{G}(u,v), (u,v) \in W$, we have

$$\iint_S \mathbf{F} \cdot \mathbf{n} \, dA = \iint_W \mathbf{F}(\mathbf{G}(u,v)) \cdot (\frac{\partial \mathbf{G}}{\partial u} \times \frac{\partial \mathbf{G}}{\partial v}) \, du \, dv$$

## 14   Vector Derivatives

$$\nabla = (\partial_1, \ldots, \partial_n)$$
$$grad\, f = \nabla f = (\partial_1 f, \ldots, \partial_n f)$$
$$div\, \mathbf{F} = \nabla \cdot \mathbf{F} = \partial_1 F_1 + \ldots + \partial_n F_n$$

Suppose $n = 3$. If $\mathbf{F}$ is a $C^1$ vector field on an open subset of $\mathbb{R}^3$, the **curl** of $\mathbf{F}$ is the vector field defined by

$$curl\, \mathbf{F} = \nabla \times \mathbf{F} = (\partial_2 F_3 - \partial_3 F_2)\mathbf{i} + (\partial_3 F_1 - \partial_1 F_3)\mathbf{j} + (\partial_1 F_2 - \partial_2 F_1)\mathbf{k}$$

**Formulas**

$$grad(fg) = f\, grad\, g + g\, grad\, f \tag{1}$$
$$grad(\mathbf{F} \cdot \mathbf{G}) = (\mathbf{F} \cdot \nabla)\mathbf{G} \times (curl\, \mathbf{G}) + (curl\, \mathbf{G}) + (\mathbf{G} \cdot \nabla)\mathbf{F} + \mathbf{G} \times (curl\, \mathbf{F}) \tag{2}$$

$$curl(f\mathbf{G}) = f\, curl\, \mathbf{G} + (grad\, f) \times \mathbf{G} \tag{3}$$
$$curl(\mathbf{F} \times \mathbf{G}) = (\mathbf{G} \cdot \nabla)\mathbf{F} + (div\, \mathbf{G})\mathbf{F} - (\mathbf{F} \cdot \nabla)\mathbf{G} - (div\, \mathbf{F})\mathbf{G} \tag{4}$$
$$div(f\mathbf{G}) = f\, div\, \mathbf{G} + (grad\, f) \cdot \mathbf{G} \tag{5}$$
$$div(\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot (curl\, \mathbf{F}) - \mathbf{F} \cdot (curl\, \mathbf{G}) \tag{6}$$

**Combined Operations**

$$curl(grad\, f) = (\partial_2 \partial_3 f - \partial_3 \partial_2 f)\mathbf{i} + (\partial_3 \partial_1 f - \partial_1 \partial_3 f)\mathbf{j} + (\partial_1 \partial_2 f - \partial_2 \partial_1 f)\mathbf{k} = \mathbf{0} \tag{7}$$
$$div(curl\, \mathbf{F}) = \partial_1(\partial_2 F_3 - \partial_3 F_2) + \partial_2(\partial_3 F_1 - \partial_1 F_3) + \partial_3(\partial_1 F_2 - \partial_2 F_1) = 0 \tag{8}$$
$$\tag{9}$$

**Laplacian**

$$\nabla^2 f = \Delta f = div(grad\, f) = \partial_1^2 f + \ldots + \partial_n^2 f$$

Other two combinations together yield the Laplacian for vector fields in $\mathbb{R}^3$:

$$grad(div\, \mathbf{F}) - curl(curl\, \mathbf{F}) = \nabla^2 \mathbf{F} = (\nabla^2 F_1)\mathbf{i} + (\nabla^2 F_2)\mathbf{j} + (\nabla^2 F_3)\mathbf{k}$$

## 15   The Divergence Theorem

The 3-dimensional analogue of Green's theorem; Relates surface integrals over the boundary of a regular region in $\mathbb{R}^3$ to volume integrals over the region itself; valid for regions with piece-wise smooth boundaries.

**The Divergence Theorem**   Suppose R is a regular region in $\mathbb{R}^3$ with piecewise smooth boundary $\partial R$, oriented so that the positive normal points out of R. Suppose also that $\mathbf{F}$ is a vector field of class $C^1$ on R. Then

$$\iint_{\partial R} \mathbf{F} \cdot \mathbf{n} \, dA = \iiint_R div \, \mathbf{F} \, dV$$

*proof:*
Begin by considering a class of simple regions.
We say that R is $xy$-simple if it has the form

$$R = \{(x, y, z) : (x, y) \in W, \varphi_1(x, y) \le z \le \varphi_2(x, y)\}$$

where W is a regular region in the $xy$-plane and $\varphi_1$ and $\varphi_2$ are piecewise smooth functions on W. We define the notions of $yz$-simple and $xz$-simple similarly, and we say that R is $\underline{\text{simple}}$ if it is $xy$-simple, $yz$-simple and $xz$-simple. Next, show $\iint_{\partial R} F_3 \mathbf{k} \cdot \mathbf{n} \, dA = \iint_R \partial_3 F_3 \, dV$
The proof for $F_1 i$ and $F_2 j$ is the same.
It now follows that the divergence theorem is valid for regions that can be cut up into finitely many simple regions $R_1, \ldots, R_k$

## 15.1   Geometric Interpretation of div

Suppose $\mathbf{F}$ is a $C^1$ vector field on some open set containing the point $\mathbf{a}$.
For $r > 0$, let $B_r$ be the ball of radius $r$ about $\mathbf{a}$. If $r$ is very small, the average value of $div\mathbf{F}(\mathbf{x})$ on the ball $B_r$ is very nearly equal to $div\mathbf{F}(\mathbf{a})$. Therefore, by the divergence theorem,

$$div\mathbf{F}(\mathbf{a}) \approx \frac{3}{4\pi r^3} \iiint_{B_r} div\mathbf{F} \, dV = \frac{3}{4\pi r^3} \iint_{\partial B_r} \mathbf{F} \cdot \mathbf{n} \, dA$$

This approximation becomes better and better as $r \to 0$, and hence

$$div\mathbf{F}(\mathbf{a}) = \lim_{r \to 0} \frac{3}{4\pi r^3} \iint_{|\mathbf{x} - \mathbf{a}| = r} \mathbf{F} \cdot \mathbf{n} \, dA$$

The flux of $\mathbf{F}$ across $\partial B_r$ from the inside $(B_r)$ to the outside (the complement of $B_r$)

**Corollary (Green's Formulas)**   Suppose R is a regular region in $\mathbb{R}^3$ with piecewise smooth boundary, and $f$ and $g$ are functions of class $C^2$ on $\bar{R}$. Then

$$\iint_{\partial R} f \nabla g \cdot \mathbf{n} \, dA = \iint_R (\nabla f \cdot \nabla g + f \nabla^2 g) dV$$

$$\iint_{\partial R} (f\nabla g - g\nabla f) \cdot \mathbf{n}\, dA = \iint_{R} (f\nabla^2 g - g\nabla^2 f)\, dV$$

The directional derivative $\nabla f \cdot \mathbf{n}$ that occurs in these formulas is called the **outward normal derivative** of $f$ on $\partial R$ and is often denoted by $\frac{\partial f}{\partial n}$.

# 16    Stokes's Theorem

"the generalization of Green's theorem in which the plane is replaced by a curved surface"

**Statement of Stokes's Theorem**    Suppose $S_0$ is a smooth surface in $\mathbb{R}^3$, and S is a region in $S_0$ bounded by a piecewise curve $\partial S$. Assume S has an orientation given by $\mathbf{n}$ and $\partial S$ is positively oriented with respect to this orientation.
Let $\mathbf{F}$ be a $C^1$ vector field defined on some neighborhood of S in $\mathbb{R}^3$. Then

$$\int_{\partial S} \mathbf{F} \cdot d\mathbf{x} = \iint_{S} (curl\, \mathbf{F}) \cdot \mathbf{n}\, dA$$

**Special case**    If S is a region in the $xy$-plane, then $\mathbf{n} = \mathbf{k} = (0,0,1)$; moreover, $\mathbf{F} \cdot d\mathbf{x}$ involves only the $x$- and $y$-components of $\mathbf{F}$, i.e., $F_1$ and $F_2$, and $(curl\mathbf{F}) \cdot \mathbf{n} = \partial_1 F_2 - \partial_2 F_1$. Hence Stokes's theorem reduces to Green's theorem in this case.

**Note**    A closed curve in $\mathbb{R}^3$ is the boundary of infinitely many surfaces in $\mathbb{R}^3$. Stokes's theorem says that if C is a closed curve in $\mathbb{R}^3$ and S is any oriented surfaced bounded by C, then

$$\int_{\partial S} \mathbf{F} \cdot d\mathbf{x} = \iint_{S} (curl\, \mathbf{F}) \cdot \mathbf{n}\, dA$$

for any $C^1$ vector field $\mathbf{F}$, provided that the orientations on C and S are compatible.

**Example**    Use Stokes's Theorem twice to let the complicated integral vanish

29

**Corollary**   If S is a closed surface (i.e., a surface with no boundary) in $\mathbb{R}^3$ with unit outward normal $\mathbf{n}$, and $\mathbf{F}$ is a $C^1$ vector field on S, then $\iint_S (curl\mathbf{F}) \cdot \mathbf{n} \, dA = 0$

*proof:*

If $\mathbf{F}$ is differentiable on the region R inside S, this follows from the divergence theorem since $div(curl\mathbf{F}) = 0$ for any $\mathbf{F}$ More generally, draw a simple closed curve C in S that divides S into two regular regions $S_1$ and $S_2$, and we have

$$\iint_S (curl\mathbf{F}) \cdot \mathbf{n} \, dA = \iint_{S_1} (curl\mathbf{F}) \cdot \mathbf{n} \, dA + \iint_{S_2} (curl\mathbf{F}) \cdot \mathbf{n} \, dA$$

On the other hand, if we give C the orientation compatible with $S_1$, Stokes's theorem gives

$$\iint_S (curl\mathbf{F}) \cdot \mathbf{n} \, dA = \int_C \mathbf{F} \cdot d\mathbf{x} = -\iint_{S_2} (curl\mathbf{F}) \cdot \mathbf{n} \, dA$$

Hence the terms on the right cancel.

## 16.1   Geometric Interpretation of Curl

Suppose $\mathbf{F}$ is a $C^1$ vector field on some open set containing the point $\mathbf{a}$.

Fix a unit vector $\mathbf{u}$.

Let $D_\epsilon$ be the disc of radius $\epsilon$ centered at $\mathbf{a}$ in the plane perpendicular to $\mathbf{u}$, oriented so that $\mathbf{u}$ is the positive normal for $D_\epsilon$.

As $\epsilon \to 0$, the average value of $(curl\mathbf{F}) \cdot \mathbf{u}$ over $D_\epsilon$ approaches its value at $\mathbf{a}$:

$$(curl\mathbf{F}(\mathbf{a})) \cdot \mathbf{u} = \lim_{\epsilon \to 0} \frac{1}{\pi\epsilon^2} \iint_{D_\epsilon} (curl\mathbf{F}) \cdot \mathbf{u} \, dA$$

Since $\mathbf{u}$ is the normal to $D_\epsilon$, Stokes's theorem gives

$$(curl\mathbf{F}(\mathbf{a})) \cdot \mathbf{u} = \lim_{\epsilon \to 0} \frac{1}{\pi\epsilon^2} \int_{C_\epsilon} \mathbf{F} \cdot d\mathbf{x}$$

where $C_\epsilon$ is the circle of radius $\epsilon$ about $\mathbf{a}$ in the plane perpendicular to $\mathbf{u}$, traversed counterclockwise as viewed from the side on which $\mathbf{u}$ lies.

Think of $\mathbf{F}$ as a force field, $\int_{C_\epsilon} \mathbf{F} \cdot d\mathbf{x}$ is the work done by $\mathbf{F}$ on a particle that moves around $C_\epsilon$. Thus $(curl\mathbf{F}(\mathbf{a})) \cdot \mathbf{u}$ represents the tendency of the force $\mathbf{F}$ to push the particle around $C_\epsilon$.

# 17   Integrating Vector Derivatives

Consider the equation $\nabla f = \mathbf{G}$

**Proposition**   Suppose $\mathbf{G}$ is a continuous vector field on an open set R in $\mathbb{R}^n$. The following two conditions are equivalent:

1. If $C_1$ and $C_2$ are any two oriented curves in R with the same initial point and the same final point, then $\int_{C_1} \mathbf{G} \cdot d\mathbf{x} = \int_{C_2} \mathbf{G} \cdot d\mathbf{x}$

2. If C is any closed curve in R, $\int_C \mathbf{G} \cdot d\mathbf{x} = 0$

*proof:* see textbook p258

**Conservative Vector Field**   A vector field $\mathbf{G}$ is called <u>conservative</u> in the region R if it satisfies the above conditions.

**Geometric Interpretation**   $\mathbf{G}$ a force field, the force does no net work on a particle that returns to its starting point.

**Proposition**   A continuous vector field $\mathbf{G}$ in an open set $R \subset \mathbb{R}^n$ is conservative $\iff$ there is a $C^1$ function $f$ on R such that $\mathbf{G} = \nabla f$ *proof:*
$\leftarrow$: If $\mathbf{G} = \nabla f$ and C is a closed curve parametrized by $\mathbf{x} = \mathbf{g}(t), a \leq t \leq b$, by the chain rule we have

$$\int_C \nabla f \cdot d\mathbf{x} = \int_a^b \nabla f(\mathbf{g}(t)) \cdot \mathbf{g}'(t) \, dt = \int_a^b \frac{d}{dt} f(\mathbf{g}(t)) dt = f(\mathbf{g}(b)) - f(\mathbf{g}(a)) = 0$$

$\rightarrow$: If $\mathbf{G}$ is conservative in R. Need to construct a function $f$ such that $\nabla f = \mathbf{G}$. Assume R is (piecewise) connected, pick $\mathbf{a} \in R$. For any $\mathbf{x} \in R$, let C be a curve in R from $\mathbf{a}$ to $\mathbf{x}$. We shall show $G_i = \partial_j f$ for each $j$.
For case $j = 1$:
Let $\mathbf{h} = (h, 0, \ldots, 0)$. Suppose h is small enough so that the line segment L from $\mathbf{x}$ to $\mathbf{x} + \mathbf{h}$ lies entirely in R. We define $f(x) = \int_C \mathbf{G} \cdot d\mathbf{x}$ where C is a curve from $\mathbf{a}$ to $\mathbf{x}$.
Make a curve from $\mathbf{a}$ to $\mathbf{x} + \mathbf{h}$ by joining L onto the end of C, so that

$$f(\mathbf{x} + \mathbf{h}) = \int_C \mathbf{G} \cdot d\mathbf{x} + \int_L \mathbf{G} \cdot d\mathbf{x}$$

But then

$$\frac{f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})}{h} = \frac{1}{h} \int_L \mathbf{G} \cdot d\mathbf{x} = \frac{1}{h} \int_0^h G_1(x_1 + t, x_2, \ldots, x_n) \, dt$$

By letting $h \to 0$ we obtain $\partial_1 f(\mathbf{x}) = G_1(\mathbf{x})$                    ∎

**Theorem**   Suppose R is a convex open set in $\mathbb{R}^n$ and $\mathbf{G}$ is a $C^1$ vector field on R. $curl(\mathbf{G}) = \mathbf{0} \implies \mathbf{G} = \nabla f \implies \mathbf{G}$ is conservative. *proof:*
Similar as previous proof.
By Stokes's Theorem,

$$\int_{\partial S} \mathbf{G} \cdot d\mathbf{x} = \iint_S (curl\,\mathbf{G})\mathbf{n}\,dA = 0$$

(since $curl\,\mathbf{G} = 0$)

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = \int_{L(\mathbf{a},\mathbf{x})} \mathbf{G} \cdot d\mathbf{x} - \int_{L(\mathbf{a},\mathbf{x}+\mathbf{h})} \mathbf{G} \cdot d\mathbf{x} = \int_{L(\mathbf{x},\mathbf{x}+\mathbf{h})} \mathbf{G} \cdot d\mathbf{x}$$

Now the same argument applies.

**Theorem**   Suppose R is a convex open set in $\mathbb{R}^3$ and $\mathbf{G}$ is a $C^1$ vector field on R. If $\mathbf{G}$ satisfies $div\mathbf{G} = 0$ on R, then $\mathbf{G}$ is the curl of a $C^2$ vector field on R.

**Definition**   If $div\mathbf{G} = 0$, a vector field $\mathbf{F}$ such that $curl\mathbf{F} = \mathbf{G}$ is called a vector potential for $\mathbf{G}$.

**Theorem**   Let R be a bounded convex open set in $\mathbb{R}^3$. For any $C^1$ function $g$ on $\bar{R}$ and any $C^2$ vector field $\mathbf{G}$ on $\bar{R}$ such that $div\mathbf{G} = 0$, there is a $C^2$ vector field $\mathbf{F}$ on $\bar{R}$ such that $curl\mathbf{F} = \mathbf{G}$ and $div\mathbf{F} = g$ on $R$

# 18   Fourier Series

**Definition**   Infinite series that use the trigonometric functions $\cos n\theta$ and $\sin n\theta$, or, equivalently, $e^{in\theta}$ and $e^{-in\theta}$, as the basic building blocks.

**Piecewise Continuity of Functions**   A function $f$ defined on an interval $[a, b]$ is piecewise continuous on $[a, b]$ if it is continuous except at finitely many points in $[a, b]$, and at each such point the one-sided limits

$$f(x+) = \lim_{\epsilon \to 0+} f(\mathbf{x} + \epsilon), \quad f(x-) = \lim_{\epsilon \to 0+} f(\mathbf{x} - \epsilon)$$

exist (and are finite)
By making the change of variable $\theta = \frac{2\pi x}{P}$, we can convert any P-periodic function into a $2\pi$-periodic function.

**Basic Idea**  An arbitrary piecewise continuous $2\pi$-periodic function $f(\theta)$ can be expanded as an infinite linear combination of the functions $e^{in\theta}(n = 0, \pm 1, \pm 2, \ldots)$, of the form

$$f(\theta) = \sum_{-\infty}^{\infty} c_n e^{in\theta}$$

(advantage: exponentials tend to be easier to manipulate than trig functions)

Since $e^{\pm in\theta} = \cos n\theta \pm i \sin n\theta$, then $c_n e^{in\theta} + c_{-n} e^{-in\theta} = (c_n + c_{-n}) \cos n\theta + i(c_n - c_{-n}) \sin n\theta = a_n \cos n\theta + b_n \sin n\theta$

$$f(\theta) = \frac{1}{2} a_0 + \sum_{1}^{\infty} (a_n \cos n\theta + b_n \sin n\theta)$$

(advantage: $\cos n\theta$ and $\sin n\theta$ are real-valued and are respectively even and odd)

f may be either real-valued or complex-valued

**Question**  Given a $2\pi$-periodic function $f$, can it be expanded in a series of the form as above? If so, how do we find the coefficients $c_n$ in this series?

**Fourier coefficients**  It can be derived that (see textbook p358)

$$a_n = \frac{1}{\pi} \int_{\pi}^{\pi} f(\theta) \cos n\theta \, d\theta$$

$$b_n = \frac{1}{\pi} \int_{\pi}^{\pi} f(\theta) \sin n\theta \, d\theta$$

$$c_n = \frac{1}{2\pi} \int_{\pi}^{\pi} f(\theta) e^{-in\theta} \, d\theta$$

**Fourier Series**

$$\sum_{-\infty}^{\infty} c_n e^{in\theta} = \frac{1}{2} a_0 + \sum_{1}^{\infty} (a_n \cos n\theta + b_n \sin n\theta)$$

f may be either real-valued or complex-valued

**E.Knight's Version of Fourier Series, A Reformulation**   To simplify our calculations, let's reformat the above mentioned form of Fourier Series into the following:

$$f(\theta) = a_0 + \sum_{n \geq 1} b_n \cos n\theta + \sum_{m \geq 1} c_m \sin m\theta$$

where the Fourier coefficients are to be calculated by

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) d\theta$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(\theta) \cos n\theta d\theta$$

$$c_m = \frac{1}{\pi} \int_0^{2\pi} f(\theta) \sin m\theta d\theta$$

notice that we here are integrating over the region $[0, 2\pi]$, but in fact any region of length $2\pi$ would do, since we are assuming a $2\pi$-periodicity. We simplify our calculation in this reformulation in the way that $a_0$ now is just the average value of the original function $f(\theta)$ in some interval of length $2\pi$.

**Convergence of Fourier Series Decomposition.**   We shall study the convergence of such series since all of these would be pointless if we can not write down an converging series at the end of the day. We shall first look at the following question: "What can we say about the Fourier coefficients $b_n, c_n$ as $n \to \infty$"? To answer, we look at the following theorem:

$$\text{Theorem:} \qquad \sum_{n=1}^{\infty} b_n^2 < \infty \wedge \sum_{n=1}^{\infty} c_n^2 < \infty$$

In paticular, we have $\lim_{n\to\infty} b_n = \lim_{n\to\infty} c_n = 0$.
<u>proof:</u> Let $N \in \mathbb{N}$, and define $f_N(x) = \sum_{n=1}^{N} (b_n \cos(nx) + c_n \sin(nx))$, then we consider

$$\int_0^{2\pi} f^2(x)dx = \int_0^{2\pi} f_N(x)dx$$

$$+ 2 \int_0^{2\pi} f_N(x)(f(x) - f_N(x))dx + \int_0^{2\pi} (f(x) - f_N(x))^2 \, dx \quad (10)$$

using brute force, we will find that $\int_0^{2\pi} f_N(x)dx = \sum_{n=1}^{N} (b_n^2 \pi + c_n^2 \pi)$, and that $b_n = 0$. We look at our equation (10) above, and see that

$$\int_0^{2\pi} (f(x) - f_N(x))^2 \, dx \geq 0$$

since we are intergrating an non-negative function. So rearranging what we have yields us

$$\sum_{n=1}^{N} \left(b_n^2 \pi + c_n^2 \pi\right) = \pi \sum_{n=1}^{N} \left(b_n^2 + c_n^2\right) \leq \frac{1}{\pi} \int_0^{2\pi} f^2(x)dx$$

but this inequality holds for any choice of $n$, and we can break the series into two as we please, so indeed we have found an upperbound and both of the two serieses are convergent. ∎

**Piece-wise smooth functions have exact Fourier Decompositions**
Suppose that $f(x)$ is $2-\pi$ periodic piece-wise smooth, then $\forall x \in [0, 2\pi)$, we have

$$\frac{f(x+) + f(x-)}{2} = a_0 + \sum_{n=1}^{N} \left(b_n \cos(nx) + c_n \sin(nx)\right)$$