

MATH 358 Notes : Honours Advanced Calculus

Samy Lahlou

These notes are based on lectures given by Professor John Toth at McGill University in Winter 2025. The subject of these lectures is to extend the theorems from the usual Real Analysis in one variable to several variables. This is not a calculus course so most of the results will be proved rigorously. The only prerequisite would be Analysis 1 and Analysis 2.

As a disclaimer, it is more than possible that I made some mistakes. Feel free to correct me or ask me anything about the content of this document at the following address : samy.lahloukamal@mcgill.ca

Contents

1 Preliminaries	2
2 Differentiation	2
2.1 Definition	2
2.2 Basic Properties	5
2.3 Partial Derivatives	10
2.4 Inverse Function Theorem and Implicit Function Theorem	13
2.5 Taylor Expansions	14
2.6 Lagrange Multipliers	15
3 Integration	15
3.1 Riemann Integral over Box Domains	15
3.2 Riemann Integral over Generalized Domains	17
3.3 Fubini's Theorem	18
3.4 Change of Variables Formula	20
4 Toward Green's and Stokes' Theorem	21
4.1 Integration over Paths	21
4.2 Integration over Surfaces	23
4.3 Green's Theorem	25
4.4 Stokes' Theorem	26
4.5 Divergence Theorem	26

1 Preliminaries

Historically, Analysis can be characterized by three very useful tools: Limits, Derivatives and Integrals. The goal is to generalize these three notions to \mathbb{R}^n . Notice that we already know how to generalize the concept of limit to any metric space from Analysis 2. Let's recall some definitions.

Definition (Euclidean Norm). In \mathbb{R}^n , the function $\|\cdot\|_n : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$\|(x_1, \dots, x_n)\|_n = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}$$

is called the Euclidean Norm on \mathbb{R}^n . When it is clear from the context, we can make the notation lighter by simply writing $\|x\|$ instead of $\|x\|_n$.

We will not prove it here but the Euclidean Norm is indeed norm. This means that for all $x \in \mathbb{R}^n$, the norm of x is zero if and only if x is the zero vector, for all $\lambda \in \mathbb{R}$, we have $\|\lambda x\| = |\lambda| \|x\|$, and for all $x, y \in \mathbb{R}^n$, we have the Triangle Inequality: $\|x + y\| \leq \|x\| + \|y\|$. Since every norm induces a metric, we can define Limits in \mathbb{R}^n as follows.

Definition (Limits of Functions). Given a function $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, a $x_0 \in \overline{\Omega}$ and a vector $L \in \mathbb{R}^m$, we say that f converges to L as x approaches x_0 if for all $\epsilon > 0$, there is a $\delta > 0$ such that $\|f(x) - f(x_0)\|_m < \epsilon$ whenever $0 < \|x - x_0\|_n < \delta$.

If f converges to some value when x approaches x_0 , then we can prove that this value must be unique. Hence, we can write $\lim_{x \rightarrow x_0} f(x) = L$ or sometimes $f(x) \rightarrow L$ as $x \rightarrow x_0$. As we could expect from this definition, it turns out that given two functions $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ and a $x_0 \in \overline{\Omega}$ such that both f and g converge as x approaches x_0 , we have

$$\lim_{x \rightarrow x_0} (f + g)(x) = \lim_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} g(x), \quad \lim_{x \rightarrow x_0} (fg)(x) = \left(\lim_{x \rightarrow x_0} f(x) \right) \left(\lim_{x \rightarrow x_0} g(x) \right)$$

$$\lim_{x \rightarrow x_0} \lambda f(x) = \lambda \lim_{x \rightarrow x_0} f(x)$$

for all $\lambda \in \mathbb{R}$. Beside these properties, we should build everything else in the following sections.

2 Differentiation

2.1 Definition

Let $\Omega \subset \mathbb{R}^n$ be an open and connected set and consider a function $f : \Omega \rightarrow \mathbb{R}^m$. Let's define what it means for f to be differentiable at some $x_0 \in \Omega$. Recall that for a function $g : (a, b) \rightarrow \mathbb{R}$ and a $x_0 \in (a, b)$, we say that g is differentiable at x_0 if

$$\lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0} = L \tag{1}$$

for some real number L . We want an equivalent definition of differentiability in the simple one-variable case that can be generalized to the multivariable case. For the moment, our problem is that the expression

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

is undefined since we cannot take the quotient of two vectors. Our strategy here will be to notice that even though we cannot take the quotient of two vectors, we can take the quotient of their respective norm since the norm of a vector is a real number. Hence, equation (1) can be rewritten as

$$\lim_{x \rightarrow x_0} \left(\frac{g(x) - g(x_0)}{x - x_0} - L \right) = 0$$

which is equivalent to

$$\lim_{x \rightarrow x_0} \left| \frac{g(x) - g(x_0)}{x - x_0} - L \right| = 0.$$

Putting everything on the same denominator and distributing the absolute value gives

$$\lim_{x \rightarrow x_0} \frac{|g(x) - g(x_0) - L(x - x_0)|}{|x - x_0|} = 0 \quad (2)$$

Here, if we replace the absolute values by Euclidean Norms in \mathbb{R}^n and \mathbb{R}^m , the expression

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0) - L(x - x_0)\|_m}{\|x - x_0\|_n} \quad (3)$$

is still undefined since $f(x) - f(x_0)$ represents a vector in \mathbb{R}^m and $L(x - x_0)$ is a vector in \mathbb{R}^n and so we cannot subtract vectors of different dimensions. We can fix this by thinking of L as an $m \times n$ matrix instead of a real number. With this modification, the expression at (3) is now well-defined provided the limit exists. Therefore, our definition of differentiability will be the following.

Definition. Given an open and connected set $\Omega \subset \mathbb{R}^n$ and a function $f : \Omega \rightarrow \mathbb{R}^m$, we say that f is differentiable at $x_0 \in \Omega$ if there exists a linear map $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0) - L(x - x_0)\|_m}{\|x - x_0\|_n} = 0.$$

Equivalently, f is differentiable at x_0 if there exists a linear map $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that for all $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\|f(x) - f(x_0) - L(x - x_0)\|_m \leq \epsilon \|x - x_0\|_n$$

for all $x \in \Omega$ satisfying $\|x - x_0\|_n < \delta$.

In the one variable case, we call L the derivative of g at x_0 . For the moment, in the definition above, we cannot define L to be *the* derivative of f at x_0 . This comes from the fact that in the one variable case, L is defined as a limit and so it is automatically unique. Here, it is not clear from the definition that the linear map L is unique.

Proposition 2.1.1. *Let $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ and suppose that there exist linear maps $L_1, L_2 : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that the expression at (3) is equal to 0 for both $L = L_1$ and $L = L_2$, then $L_1 = L_2$.*

Proof. Let $\epsilon > 0$, then there exist $\delta_1, \delta_2 > 0$ such that

$$\|f(x) - f(x_0) - L_1(x - x_0)\|_m \leq \frac{\epsilon}{2} \|x - x_0\|_n$$

for all $x \in \Omega$ satisfying $\|x - x_0\|_n < \delta_1$ and

$$\|f(x) - f(x_0) - L_2(x - x_0)\|_m \leq \frac{\epsilon}{2} \|x - x_0\|_n$$

for all $x \in \Omega$ satisfying $\|x - x_0\|_n < \delta_2$. Thus, if we let $\delta = \min(\delta_1, \delta_2)$ and $x \in \Omega$ such that $\|x - x_0\|_n < \delta$, then

$$\begin{aligned} & \|(L_1 - L_2)(x - x_0)\|_m \\ &= \|f(x) - f(x_0) - L_2(x - x_0) - (f(x) - f(x_0) - L_2(x - x_0))\|_m \\ &\leq \|f(x) - f(x_0) - L_2(x - x_0)\|_m + \|(f(x) - f(x_0) - L_2(x - x_0))\|_m \\ &\leq \frac{\epsilon}{2} \|x - x_0\|_n + \frac{\epsilon}{2} \|x - x_0\|_n \\ &= \epsilon \|x - x_0\|_n. \end{aligned}$$

Let $i \in \llbracket 1, n \rrbracket$ and $\epsilon > 0$, then by what we showed previously, there is a $\delta > 0$ such that $\|(L_1 - L_2)(x - x_0)\|_m \leq \epsilon \|x - x_0\|_n$ whenever $\|x - x_0\|_n < \delta$. Take $x = x_0 + \frac{\delta}{2}e_i$ and notice that

$$\|x - x_0\|_n = \left\| x_0 + \frac{\delta}{2}e_i - x_0 \right\|_n = \frac{\delta}{2} \|e_i\|_n < \delta.$$

Thus, we get $\|(L_1 - L_2)e_i\|_m \leq \epsilon$. Since it holds for all $\epsilon > 0$, then taking $\epsilon \rightarrow 0$ gives us $\|(L_1 - L_2)e_i\|_m = 0$. By properties of norms, it follows that $(L_1 - L_2)e_i = 0$. Since it holds for all $i \in \llbracket 1, n \rrbracket$ and L is linear, then $L_1 - L_2$ is the zero map and so $L_1 = L_2$. ■

Thus, if f is differentiable at $x_0 \in \Omega$, then L is unique. We denote L by $Df(x_0)$ and call it the Jacobian of f at x_0 . From the definition of differentiability, we can prove the following useful differentiability criterion.

Proposition 2.1.2. *If we write the function $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ as $f = (f_1, \dots, f_m)$ with $f_i : \Omega \rightarrow \mathbb{R}$, then f is differentiable at $x_0 \in \Omega$ if and only if f_j is differentiable at x_0 for all $j \in \llbracket 1, m \rrbracket$.*

Proof. Suppose that f is differentiable at x_0 , then there exists a linear map $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that for all $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\|f(x) - f(x_0) - L(x - x_0)\| \leq \epsilon \|x - x_0\|$$

for all $x \in \Omega$ satisfying $\|x - x_0\| < \delta$. Let $j \in \llbracket 1, m \rrbracket$ and write L as (L_1, \dots, L_m) , then

$$\|f(x) - f(x_0) - L(x - x_0)\| = \left(\sum_{i=1}^m (f_i(x) - f_i(x_0) - L_i(x - x_0))^2 \right)^{1/2}$$

which implies that

$$|f_j(x) - f_j(x_0) - L_j(x - x_0)| \leq \|f(x) - f(x_0) - L(x - x_0)\|.$$

Thus, for all $\epsilon > 0$, there exists a $\delta > 0$ such that

$$|f_j(x) - f_j(x_0) - L_j(x - x_0)| \leq \epsilon \|x - x_0\|$$

for all $\|x - x_0\| < \delta$ with $x \in \Omega$. It follows that f_j is differentiable at x_0 for all $j \in \llbracket 1, m \rrbracket$. Conversely, suppose that f_j is differentiable at x_0 for all $j \in \llbracket 1, m \rrbracket$, then for all $\epsilon > 0$, there exists $\delta_j > 0$ such that

$$|f_j(x) - f_j(x_0) - L_j(x - x_0)| \leq \epsilon \|x - x_0\|$$

for all $\|x - x_0\| < \delta_j$ where L_j is a linear map from \mathbb{R}^n to \mathbb{R} . Define the linear map $L = (L_1, \dots, L_m)$ and let $\epsilon > 0$, then for all $j \in \llbracket 1, m \rrbracket$, there exists a $\delta_j > 0$ such that

$$|f_j(x) - f_j(x_0) - L_j(x - x_0)| \leq \frac{\epsilon}{m} \|x - x_0\|.$$

Define $\delta = \min(\delta_1, \dots, \delta_m)$ and notice that for all $x \in B_\delta(x_0) \cap \Omega$,

$$\begin{aligned} \|f(x) - f(x_0) - L(x - x_0)\| &= \left(\sum_{j=1}^m (f_j(x) - f_j(x_0) - L_j(x - x_0))^2 \right)^{1/2} \\ &\leq \sum_{j=1}^m |f_j(x) - f_j(x_0) - L_j(x - x_0)| \\ &\leq \sum_{j=1}^m \frac{\epsilon}{m} \|x - x_0\| \\ &= \epsilon \|x - x_0\|. \end{aligned}$$

Therefore, f is differentiable at x_0 . ■

2.2 Basic Properties

Now that we defined the notion of differentiability in a more general case, let's prove the basic and usual properties that we know from the one variable case. In the one variable case, when we know that differentiability at a point implies continuity at that point. The same implication holds in the general case.

Proposition 2.2.1. *Let $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $x_0 \in \Omega$ be differentiable at $x_0 \in \Omega$, then f is continuous at x_0 .*

Proof. Let $\epsilon = 1$, then there exists a $\delta > 0$ such that for all $x \in \Omega$ with $\|x - x_0\| < \delta$, we have $\|f(x) - f(x_0) - L(x - x_0)\| \leq \|x - x_0\|$. Thus, using properties of the matrix-norm, we obtain

$$\begin{aligned} \|f(x) - f(x_0)\| &\leq \|f(x) - f(x_0) - L(x - x_0) + L(x - x_0)\| \\ &\leq \|x - x_0\| + \|L\| \|x - x_0\| \\ &= (1 + \|L\|) \|x - x_0\|. \end{aligned}$$

Thus, f is Lipschitz continuous at x_0 and so f is continuous at x_0 . ■

Some of the most useful properties of the derivative are its algebraic properties which lets us compute the derivative of a complicated function by splitting it into an algebraic expression involving elementary functions.

Theorem 2.2.2. *Let $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ be differentiable at $x_0 \in \Omega$, then $f + g$ is differentiable at x_0 with*

$$D(f + g)(x_0) = Df(x_0) + Dg(x_0).$$

Proof. Let $\epsilon > 0$, then by differentiability, there exist δ_f and δ_g strictly positive such that

$$\forall x \in B_{\delta_f}(x_0) \cap \Omega : \quad \|f(x) - f(x_0) - Df(x_0)(x - x_0)\| \leq \frac{\epsilon}{2} \|x - x_0\|,$$

$$\forall x \in B_{\delta_g}(x_0) \cap \Omega : \quad \|g(x) - g(x_0) - Dg(x_0)(x - x_0)\| \leq \frac{\epsilon}{2} \|x - x_0\|.$$

If we let $\delta = \min(\delta_f, \delta_g)$, then for any $x \in B_\delta \cap \Omega$:

$$\begin{aligned} & \|f(x) + g(x) - f(x_0) - g(x_0) - (Df(x_0) + Dg(x_0))(x - x_0)\| \\ & \leq \|f(x) - f(x_0) - Df(x_0)(x - x_0)\| + \|g(x) - g(x_0) - Dg(x_0)(x - x_0)\| \\ & \leq \epsilon \|x - x_0\|. \end{aligned}$$

Therefore, $f + g$ is differentiable at x_0 with $D(f + g)(x_0) = Df(x_0) + Dg(x_0)$. ■

Theorem 2.2.3. *Let $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable at $x_0 \in \Omega$, then fg is differentiable at x_0 with*

$$D(fg)(x_0) = Df(x_0)g(x_0) + f(x_0)Dg(x_0).$$

Proof. Let $\epsilon \in (0, 1]$, then by differentiability, there exists $\delta_1, \delta_2 > 0$ such that

$$\forall x \in B_{\delta_1}(x_0) \cap \Omega : \quad |f(x) - f(x_0) - Df(x_0)(x - x_0)| \leq \epsilon \|x - x_0\|,$$

$$\forall x \in B_{\delta_2}(x_0) \cap \Omega : \quad |g(x) - g(x_0) - Dg(x_0)(x - x_0)| \leq \epsilon \|x - x_0\|.$$

Moreover, since f is differentiable at x_0 , then f is continuous at x_0 . Hence, there is a $\delta_3 > 0$ such that $|f(x) - f(x_0)| \leq \epsilon$ whenever $\|x - x_0\| < \delta_3$. Define $\delta = \min(\delta_1, \delta_2, \delta_3)$

and take an arbitrary $x \in \Omega$ such that $\|x - x_0\| < \delta$, then

$$\begin{aligned}
& |f(x)g(x) - f(x_0)g(x_0) - [Df(x_0)g(x_0) + f(x_0)Dg(x_0)](x - x_0)| \\
&= |f(x)g(x) + f(x)g(x_0) - f(x)g(x_0) - f(x_0)g(x_0) \\
&\quad - Df(x_0)g(x_0)(x - x_0) - f(x_0)Dg(x_0)(x - x_0)| \\
&\leq |g(x_0)||f(x) - f(x_0) - Df(x_0)(x - x_0)| \\
&\quad + |f(x)g(x) - f(x)g(x_0) - f(x_0)Dg(x_0)(x - x_0)| \\
&\leq \epsilon|g(x_0)|\|x - x_0\| + |f(x)g(x) - f(x)g(x_0) - g(x_0)Df(x_0)(x - x_0)| \\
&\quad + |g(x_0)Df(x_0)(x - x_0) - f(x_0)Dg(x_0)(x - x_0)| \\
&\leq \epsilon|g(x_0)|\|x - x_0\| + |f(x)||g(x) - g(x_0) - Dg(x_0)(x - x_0)| \\
&\quad + |f(x) - f(x_0)|Dg(x_0)(x - x_0)| \\
&\leq \epsilon|g(x_0)|\|x - x_0\| + \epsilon|f(x)|\|x - x_0\| + \epsilon\|Dg(x_0)\|\|x - x_0\| \\
&\leq \epsilon|g(x_0)|\|x - x_0\| + \epsilon|f(x) - f(x_0)|\|x - x_0\| \\
&\quad + \epsilon|f(x_0)|\|x - x_0\| + \epsilon\|Dg(x_0)\|\|x - x_0\| \\
&\leq \epsilon(|g(x_0)| + \epsilon + |f(x_0)| + \|Dg(x_0)\|)\|x - x_0\| \\
&\leq \epsilon(1 + |g(x_0)| + |f(x_0)| + \|Dg(x_0)\|)\|x - x_0\|.
\end{aligned}$$

Since $(1 + |g(x_0)| + |f(x_0)| + \|Dg(x_0)\|)$ is constant, then fg is differentiable at x_0 and the Jacobian at x_0 has the desired form. \blacksquare

Theorem 2.2.4. *Let $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable at $x_0 \in \Omega$ with $g(x_0) \neq 0$, then f/g is differentiable at x_0 with*

$$D\left(\frac{f}{g}\right)(x_0) = \frac{Df(x_0)g(x_0) - f(x_0)Dg(x_0)}{g(x_0)^2}$$

Proof. First, let's show that $1/g$ is differentiable at $x_0 \in \Omega$ with

$$D\left(\frac{1}{g}\right)(x_0) = -\frac{1}{g(x_0)^2}Dg(x_0).$$

To do so, let $\epsilon > 0$ and assume that $\epsilon \leq 1$, then by differentiability of g , there is a $\delta_1 > 0$ such that

$$|g(x) - g(x_0) - Dg(x_0)(x - x_0)| \leq \epsilon\|x - x_0\|$$

for all $x \in B_{\delta_1}(x_0) \cap \Omega$. Moreover, g is continuous at x_0 and $g(x_0) \neq 0$ so

$$\lim_{x \rightarrow x_0} g(x) = g(x_0) \implies \lim_{x \rightarrow x_0} \frac{1}{g(x)} = \frac{1}{g(x_0)}.$$

Hence, there is a $\delta_2 > 0$ such that $|\frac{1}{g(x)} - \frac{1}{g(x_0)}| \leq \epsilon$ for all $x \in B_{\delta_2}(x_0) \cap \Omega$. Let

$\delta = \min(\delta_1, \delta_2)$ and take $x \in B_\delta(x_0) \cap \Omega$, then

$$\begin{aligned}
& \left| \frac{1}{g(x)} - \frac{1}{g(x_0)} - \left(-\frac{1}{g(x_0)^2} Dg(x_0) \right) (x - x_0) \right| \\
&= \left| \frac{g(x_0) - g(x)}{g(x)g(x_0)} + \frac{Dg(x_0)(x - x_0)}{g(x_0)^2} \right| \\
&= \left| \frac{g(x) - g(x_0)}{g(x)g(x_0)} - \frac{Dg(x_0)(x - x_0)}{g(x_0)^2} \right| \\
&= \left| \frac{g(x) - g(x_0) - Dg(x_0)(x - x_0)}{g(x)g(x_0)} + \frac{Dg(x_0)(x - x_0)}{g(x_0)g(x)} - \frac{Dg(x_0)(x - x_0)}{g(x_0)^2} \right| \\
&\leq \left| \frac{1}{g(x)} \right| \frac{1}{|g(x_0)|} |g(x) - g(x_0) - Dg(x_0)(x - x_0)| \\
&\quad + \frac{1}{|g(x_0)|} \left| \frac{1}{g(x)} - \frac{1}{g(x_0)} \right| |Dg(x_0)(x - x_0)| \\
&\leq \left| \frac{1}{g(x)} - \frac{1}{g(x_0)} \right| \frac{1}{|g(x_0)|} \epsilon \|x - x_0\| + \frac{1}{|g(x_0)^2|} \epsilon \|x - x_0\| \\
&\quad + \frac{1}{|g(x_0)|} \epsilon \|Dg(x_0)\| \|x - x_0\| \\
&= \epsilon \left[\frac{1}{|g(x_0)|} + \frac{1}{|g(x_0)^2|} + \frac{1}{|g(x_0)|} \|Dg(x_0)\| \right] \|x - x_0\| \\
&= C\epsilon \|x - x_0\|
\end{aligned}$$

which proves that $D(\frac{1}{g})(x_0)$ exists and is equal to $-\frac{1}{g(x_0)^2} Dg(x_0)$. Now, using the previous theorem, we get that f/g is differentiable at x_0 and

$$\begin{aligned}
D\left(\frac{f}{g}\right)(x_0) &= D\left(f \cdot \frac{1}{g}\right)(x_0) \\
&= Df(x_0) \frac{1}{g(x_0)} + f(x_0) \left(-\frac{1}{g(x_0)^2} Dg(x_0) \right) \\
&= \frac{Df(x_0)g(x_0) - f(x_0)Dg(x_0)}{g(x_0)^2}.
\end{aligned}$$

■

Theorem 2.2.5 (Chain Rule). *Let $f : \Omega \subset \mathbb{R}^n \rightarrow \tilde{\Omega} \subset \mathbb{R}^m$ and $g : \tilde{\Omega} \rightarrow \mathbb{R}^k$ be differentiable at $x_0 \in \Omega$ and $f(x_0) \in \tilde{\Omega}$ respectively, then $g \circ f$ is differentiable at x_0 with*

$$D(g \circ f)(x_0) = Dg(f(x_0))Df(x_0).$$

Proof. Let $\epsilon > 0$ and suppose that $\epsilon \leq 1$, then by differentiability of f at x_0 ,

$$x \in B_{\delta_1}(x_0) \cap \Omega \implies \|f(x) - f(x_0) - Df(x_0)(x - x_0)\| \leq \epsilon \|x - x_0\|$$

and similarly, by differentiability of g at $f(x_0)$,

$$y \in B_{\delta_2}(f(x_0)) \cap \tilde{\Omega} \implies \|g(y) - g(f(x_0)) - Dg(f(x_0))(y - f(x_0))\| \leq \epsilon \|y - f(x_0)\|$$

for some δ_1 and δ_2 depending on ϵ . Moreover, since f is differentiable at x_0 , then f is continuous at x_0 . This implies that there exists a $\delta_3 > 0$ such that for all $x \in B_{\delta_3}(x_0) \cap \Omega$:

$$\|x - x_0\| < \delta_3 \implies \|f(x) - f(x_0)\| < \delta_2.$$

Thus, if we let $y = f(x)$ and $x \in B_{\delta_3}(x_0) \cap \Omega$, then

$$\|g(f(x)) - g(f(x_0)) - Dg(f(x_0))(f(x) - f(x_0))\| \leq \epsilon \|f(x) - f(x_0)\|.$$

Define $\delta = \min(\delta_1, \delta_3)$ and let $x \in B_\delta(x_0) \cap \Omega$, then

$$\begin{aligned} & \|(g \circ f)(x_0) - (g \circ f)(x_0) - Dg(f(x_0))Df(x_0)(x - x_0)\| \\ &= \|(g \circ f)(x_0) - (g \circ f)(x_0) - Dg(f(x_0))(f(x) - f(x_0)) \\ &\quad + Dg(f(x_0))(f(x) - f(x_0)) - Dg(f(x_0))Df(x_0)(x - x_0)\| \\ &\leq \|g(f(x_0)) - g(f(x_0)) - Dg(f(x_0))(f(x) - f(x_0))\| \\ &\quad + \|Dg(f(x_0))(f(x) - f(x_0)) - Dg(f(x_0))Df(x_0)(x - x_0)\| \\ &\leq \epsilon \|f(x) - f(x_0)\| + \|Dg(f(x_0))\| \|f(x) - f(x_0) - Df(x_0)(x - x_0)\| \\ &\leq \epsilon \|f(x) - f(x_0) - Df(x_0)(x - x_0)\| + \epsilon \|Df(x_0)(x - x_0)\| \\ &\quad + \epsilon \|Dg(f(x_0))\| \|x - x_0\| \\ &\leq \epsilon^2 \|x - x_0\| + \epsilon \|Df(x_0)\| \|x - x_0\| + \epsilon \|Dg(f(x_0))\| \|x - x_0\| \\ &\leq \epsilon(1 + \|Df(x_0)\| + \|Dg(f(x_0))\|) \|x - x_0\| \\ &= C\epsilon \|x - x_0\| \end{aligned}$$

which proves that $g \circ f$ is differentiable at x_0 with $D(g \circ f)(x_0) = Dg(f(x_0))Df(x_0)$. \blacksquare

Theorem 2.2.6 (Mean Value Theorem). *Let $B = B_r(c_0) \subset \mathbb{R}^n$ be a ball and $f : B \rightarrow \mathbb{R}$ be differentiable on B . For any $x, y \in B$, there is a $z \in B$ such that*

$$f(x) - f(y) = Df(z)(x - y).$$

Proof. Let $x, y \in B$ and define $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ by $\gamma(t) = tx + (1 - t)y$. Notice that for $t \in [0, 1]$, we have

$$\begin{aligned} \|\gamma(t) - c_0\| &= \|tx + (1 - t)y - c_0\| \\ &= \|tx + (1 - t)y - tc_0 - (1 - t)c_0\| \\ &\leq \|tx - tc_0\| + \|(1 - t)y - (1 - t)c_0\| \\ &= t\|x - c_0\| + (1 - t)\|y - c_0\| \\ &< tr + (1 - t)r \\ &= r. \end{aligned}$$

It follows that $\text{Im } \gamma \subset B$. Now, set $F = f \circ \gamma : [0, 1] \rightarrow \mathbb{R}$ and notice that F is differentiable. By the Mean Value Theorem in one variable, there exists a $t^* \in [0, 1]$ such that $F(1) - F(0) = F'(t^*)$. Using the Chain Rule we get that

$$F'(t^*) = D(f \circ \gamma)(t^*) = Df(\gamma(t^*))D\gamma(t^*).$$

Since $D\gamma(t^*) = x - y$, then if we define $z = \gamma(t^*) \in B$, we get

$$f(x) - f(y) = F(1) - F(0) = Df(z)(x - y).$$

\blacksquare

2.3 Partial Derivatives

Usually, in a multivariable Calculus class, when we think about generalizing the concept of derivatives, we probably first think of partial derivatives instead of the Jacobian. This probably comes from the fact that partial derivatives really are the usual one dimensional derivative, and hence, it closer to what we are used to with usual single variable derivatives. We will explore in this subsection the link between these two notions.

Definition. Let $f = (f_1, \dots, f_m) : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, then for all $x \in \Omega$, $i \in \llbracket 1, m \rrbracket$ and $j \in \llbracket 1, n \rrbracket$, define

$$\frac{\partial f_i}{\partial x_j}(x) = \lim_{h \rightarrow 0} \frac{f_i(x + he_j) - f_i(x)}{h}$$

provided the limit exists. In that case, we call it the i th partial derivative of f with respect to x_j at $x \in \Omega$.

The first proposition links the Jacobian of a function to its partial derivatives in a way that makes it easier to determine in general if a function is differentiable.

Proposition 2.3.1. *Let $f = (f_1, \dots, f_m) : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ be differentiable at $x_0 \in \Omega$, then, all the partial derivatives of f exist and*

$$Df(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \dots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}$$

such that the entry in the i th row and j th column is equal to the i th partial derivative of f with respect to x_j at $x \in \Omega$.

Proof. First, write $Df(x_0) = L$ as the matrix with entries $L_{i,j}$ and denote by L_1, \dots, L_m its columns. Let $i \in \llbracket 1, m \rrbracket$, $j \in \llbracket 1, n \rrbracket$ and $\epsilon > 0$, then there is a $\delta > 0$ such that

$$\|f(x) - f(x_0) - L(x - x_0)\| \leq \epsilon \|x - x_0\|$$

for all $x \in B_\delta(x_0) \cap \Omega$. Let $0 < |h| < \delta$ and define $x = x_0 + he_j$, then $\|x - x_0\| = \|he_j\| = |h| < \delta$. Thus,

$$\begin{aligned} |f_i(x) - f_i(x_0) - L_i(x - x_0)| &\leq \left(\sum_{k=1}^m (f_k(x) - f_k(x_0) - L_k(x - x_0))^2 \right)^{1/2} \\ &= \|f(x) - f(x_0) - L(x - x_0)\| \\ &\leq \epsilon \|x - x_0\| \end{aligned}$$

which implies

$$|f_i(x_0 + he_j) - f_i(x_0) - L_i(he_j)| \leq h\epsilon$$

and is equivalent to

$$\left| \frac{f_i(x_0 + he_j) - f_i(x_0)}{h} - L_{i,j} \right| \leq \epsilon.$$

Therefore, by definition of the limit, we have

$$\frac{\partial f_i}{\partial x_j}(x_0) = \lim_{h \rightarrow 0} \frac{f_i(x_0 + he_j) - f_i(x_0)}{h} = L_{i,j}$$

which proves that all the partial derivatives exist and that they are the entries of the Jacobian. \blacksquare

Notice that the converse need not be true. Take for example the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} \frac{xy}{\sqrt{x^2+y^2}} & (x, y) \neq (0, 0), \\ 0 & (x, y) = (0, 0). \end{cases}$$

To see why it is a counterexample, consider the point $(0, 0)$.

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = 0$$

and

$$\frac{\partial f}{\partial y}(0, 0) = \lim_{h \rightarrow 0} \frac{f(0, h) - f(0, 0)}{h} = 0.$$

Therefore, by the previous theorem, if f is differentiable at $(0, 0)$, then the Jacobian would be

$$Df(0, 0) = \begin{pmatrix} \frac{\partial f}{\partial x}(0, 0) & \frac{\partial f}{\partial y}(0, 0) \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$$

which would imply that the following equation holds

$$\lim_{(x,y) \rightarrow (0,0)} \frac{|f(x, y) - f(0, 0) - Df(0, 0)(x, y)|}{\sqrt{x^2 + y^2}} = 0.$$

After simplifying the previous expression, we obtain the following equivalent form:

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{x^2 + y^2} = 0.$$

But this is impossible because if we consider the sequence $\{(\frac{1}{n}, \frac{1}{n})\}_n$ which converges to $(0, 0)$, we obtain

$$\lim_{n \rightarrow \infty} \frac{(\frac{1}{n})(\frac{1}{n})}{(\frac{1}{n})^2 + (\frac{1}{n})^2} = \lim_{n \rightarrow \infty} \frac{n^2}{2n^2} = \frac{1}{2}.$$

Therefore, this example shows why the converse is false and may not hold. However, by adding a little extra assumption on the partial derivatives, we obtain the following partial converse.

Theorem 2.3.2. *Let $f = (f_1, \dots, f_m) : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. Suppose that all the partial derivatives exist and are continuous on a neighborhood $U \subset \Omega$ of $x_0 \in \Omega$, then f is differentiable at $x_0 \in \Omega$.*

Proof. First, notice that it we can assume without loss of generality that $U = \Omega$ since otherwise, we can just restrict ourself to $f|_U$ instead of f . Moreover, we can assume that $m = 1$ because if we prove that each f_i is differentiable at x_0 , then it automatically follows that f is differentiable at x_0 . Let's prove that $f : \Omega \rightarrow \mathbb{R}$ is differentiable at $x_0 \in \Omega$. Let $\epsilon > 0$, then by continuity, for all $i \in \llbracket 1, n \rrbracket$, there exists a $\delta_i > 0$ such that

$$\left| \frac{\partial f}{\partial x_i}(y) - \frac{\partial f}{\partial x_i}(x_0) \right| \leq \frac{\epsilon}{n}$$

for all $y \in B_{\delta_j}(x_0) \cap \Omega$. Let δ be the minimum of the δ_i 's, take an arbitrary $x = (x^{(1)}, \dots, x^{(n)}) \in B_{\delta_i}(x_0) \cap \Omega$ and write $x_0 = (x_0^{(1)}, \dots, x_0^{(n)})$. Notice that by the Mean Value Theorem in one variable, for all $i \in \llbracket 1, n \rrbracket$, there exists a z_i between $x^{(i)}$ and $x_0^{(i)}$ such that

$$\begin{aligned} f(x) - f(x_0) &= \sum_{i=1}^n f(x_0^{(1)}, \dots, x_0^{(i-1)}, x^{(i)}, \dots, x^{(n)}) - f(x_0^{(1)}, \dots, x_0^{(i)}, x^{(i+1)}, \dots, x^{(n)}) \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0^{(1)}, \dots, x_0^{(i-1)}, z_i, x^{(i+1)}, \dots, x^{(n)}). \end{aligned}$$

Now, define $L = \left(\frac{\partial f}{\partial x_1}(x_0) \quad \dots \quad \frac{\partial f}{\partial x_n}(x_0) \right)$ and notice that by continuity of the partial derivatives,

$$\begin{aligned} |f(x) - f(x_0) - L(x - x_0)| &= \left| \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x_0^{(1)}, \dots, x_0^{(i-1)}, z_i, x^{(i+1)}, \dots, x^{(n)}) - \frac{\partial f}{\partial x_i}(x_0) \right) (x^{(i)} - x_0^{(i)}) \right| \\ &\leq \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(x_0^{(1)}, \dots, x_0^{(i-1)}, z_i, x^{(i+1)}, \dots, x^{(n)}) - \frac{\partial f}{\partial x_i}(x_0) \right| |x^{(i)} - x_0^{(i)}| \\ &\leq \sum_{i=1}^n \frac{\epsilon}{n} |x^{(i)} - x_0^{(i)}| \\ &\leq \frac{\epsilon}{n} \sum_{i=1}^n \|x - x_0\| \\ &= \epsilon \|x - x_0\|. \end{aligned}$$

Therefore, f is differentiable at x_0 . ■

Notice that the continuity of the partial derivatives is sufficient but not necessary. Consider the following example:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \quad f(x) = \begin{cases} \frac{x^2 y^2}{x^2 + y^4} & (x, y) \neq (0, 0), \\ 0 & (x, y) = (0, 0). \end{cases}$$

The function f is differentiable at $(0, 0)$ and its partial derivatives exist but they are not continuous at $(0, 0)$.

Definition. Let $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ be differentiable on Ω , then we say that f is twice differentiable at x_0 if the map $Df : \Omega \rightarrow \mathbb{R}^{mn}$ is differentiable at x_0 . In that case, we write $D^2f(x_0) = D(Df)(x_0)$. Recursively, we define $D^k f(x_0) = D(D^{k-1}f)(x_0)$ provided $D^{k-1}f$ is defined on Ω .

Definition. Let $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, we say that $f \in C^k(\Omega)$ for some natural k if all the partial derivatives to order k are continuous on Ω .

Theorem 2.3.3 (Clairault's Theorem). *Given $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, if its second-order partial derivatives exist and are continuous on Ω , then*

$$\frac{\partial f}{\partial x_i \partial x_j} = \frac{\partial f}{\partial x_j \partial x_i}$$

on Ω for all $i, j \in \llbracket 1, n \rrbracket$.

Proof. **TODO** ■

2.4 Inverse Function Theorem and Implicit Function Theorem

Given a differentiable function $f : (a, b) \rightarrow (c, d)$ such that $f' > 0$, then we know that there exists a function $g : (c, d) \rightarrow (a, b)$ which is differentiable and such that $g = f^{-1}$ because f is strictly increasing so it is both injective and surjective. Moreover, if a function is differentiable, then its inverse must be differentiable as well (provided the inverse exists). Let's generalize this result for multivariable functions. But first, we need some preliminary results.

Theorem 2.4.1 (Contraction Mapping Theorem). *Let (X, d) be a non-empty metric space and $f : X \rightarrow X$ be a function such that*

$$d(fx, fy) \leq \alpha d(x, y)$$

for all $x, y \in X$ and for some $0 < \alpha < 1$, then there exists a unique $x_0 \in X$ such that $fx_0 = x_0$.

Proof. **TODO** ■

Lemma 2.4.2. *The map $f : GL(n) \rightarrow GL(n)$ defined by $f(A) = A^{-1}$ is in C^∞ .*

Proof. **TODO** ■

Theorem 2.4.3 (Inverse Function Theorem). *Let $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ be C^1 and $x_0 \in \Omega$ with $Df(x_0) \in GL(n)$, then there exist neighborhoods U and V of x_0 and $f(x_0) = y_0$ respectively such that $f(U) = V$ and there is a C^1 inverse map $f^{-1} : V \rightarrow U$ such that*

$$D(f^{-1})(y) = [Df(f^{-1}y)]^{-1}$$

for all $y \in V$. Moreover, if f is C^k , then f^{-1} is C^k as well.

Proof. **TODO** ■

The Inverse Function Theorem is local in the sense that it can be applied even if the function is not invertible. For example, consider the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by

$$f(x, y) = (e^y \cos x, e^y \sin x),$$

its Jacobian is given by

$$Df(x, y) = \begin{pmatrix} -e^y \sin x & e^y \cos x \\ e^y \cos x & e^y \sin x \end{pmatrix}.$$

Thus,

$$\det Df(x, y) = -(\cos^2 x + \sin^2 x)e^y = -e^y \neq 0$$

for all $x, y \in \mathbb{R}$. However, f is not injective since its x component is 2π -periodic.

Theorem 2.4.4 (Implicit Function Theorem). *Let $F : \Omega \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ be a C^k map for some $k \geq 1$. Write $X = (x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ and $F = (F_1, \dots, F_m)$. Suppose there is a $X_0 = (x_0, y_0)$ such that $F(X_0) = 0$ and*

$$D_y F(X_0) = \begin{pmatrix} \frac{\partial F_1}{\partial y_1}(X_0) & \dots & \frac{\partial F_1}{\partial y_m}(X_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1}(X_0) & \dots & \frac{\partial F_m}{\partial y_m}(X_0) \end{pmatrix} \in GL(m),$$

then there exist neighborhoods U and V of $x_0 \in \mathbb{R}^n$ and $y_0 \in \mathbb{R}^m$ and a unique C^k map $f : U \rightarrow V$ such that $F(x, f(x)) = 0$ for all $x \in U$.

Proof. **TODO** ■

Corollary. *Given a C^k function $F : \Omega \subset \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$, if there is $(x_0, y_0) \in \Omega$ such that*

$$\frac{\partial F}{\partial y}(x_0, y_0) \neq 0,$$

then there exist neighborhoods U and V of x_0 and y_0 respectively and a unique C^k function $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ such that $F(x, f(x)) = 0$ for all $x \in U$.

Proof. **TODO** ■

2.5 Taylor Expansions

Theorem 2.5.1. *Let $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ with $f \in C^{k-1}(\Omega)$ and assume $D^{k-1}f$ is differentiable at $x_0 \in \Omega$, then*

$$f(x) = f(x_0) + \sum_{j=1}^k \frac{1}{j!} \left[\sum_{i_1, \dots, i_j=1}^n \frac{\partial^j f}{\partial x_{i_1} \dots \partial x_{i_j}}(x_0) (x_{i_1} - x_{i_1}^0) \dots (x_{i_j} - x_{i_j}^0) \right] + R_k(f)(x)$$

with

$$\frac{|R_k(f)(x)|}{\|x - x_0\|^k} \rightarrow 0$$

as $x \rightarrow x_0$.

Proof. **TODO** ■

For example, if we let $f(x, y) = e^x + \sin(xy)$, then we can compute the second order Taylor approximation. **TODO**

2.6 Lagrange Multipliers

Suppose we have two functions $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ which are C^1 . If we let $\Sigma = g^{-1}(\{0\})$, how can we find the maximum and minimum values of f on Σ ? We can see g as a constraint.

Theorem 2.6.1 (Lagrange Multipliers Theorem). *Let $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be C^1 functions and $\Sigma = g^{-1}(\{0\})$, then if $x_0 \in \Sigma$ is a maximum or minimum of f on Σ and $\nabla g(x_0) \neq 0$, then*

$$\nabla f(x_0) = \lambda \nabla g(x_0)$$

for some $\lambda \in \mathbb{R}$.

Proof. **TODO** ■

Theorem 2.6.2 (Max / Min Theorem). *Let $\Omega \subsetneq \tilde{\Omega} \subset \mathbb{R}^n$ be domains and assume $f, g \in C^1(\tilde{\Omega})$ satisfy*

$$\partial\Omega = g^{-1}(\{0\}), \quad \text{and} \quad 0 \notin g(\partial\Omega),$$

then the following are true:

1. *If x_0 is a local extremum of $f|_{\partial\Omega}$, then $\nabla f(x_0) = \lambda \nabla g(x_0)$ for some $\lambda \in \mathbb{R}$.*
2. *If $x_0 \in \text{Int}(\Omega)$ is a local extremum of f , then $\nabla f(x_0) = 0$.*
3. *The global extremum of f are either from the first or the second case.*

Proof. **TODO** ■

Theorem 2.6.3 (Lagrange Multipliers Theorem - Multiple Constraints). *Suppose that the functions $f, g_1, \dots, g_k : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ are all C^1 , let $\Gamma_0 = \cap_{i=1}^k g_i^{-1}(\{0\})$ and assume that the $\nabla g_i(x_0)$'s are linearly independent, then if $f|_{\Gamma_0}$ has an extremum at $x_0 \in \Gamma_0$, there exist $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ such that*

$$\nabla f(x_0) = \sum_{i=1}^k \lambda_i \nabla g_i(x_0).$$

Proof. **TODO** ■

3 Integration

3.1 Riemann Integral over Box Domains

In this section, we fix a natural number n .

Definition. A box is the cartesian product of n closed and bounded intervals of \mathbb{R} .

Definition. Given a box $B = \prod_{i=1}^n [a_i, b_i]$, we define its volume by

$$\text{Vol}(B) = \prod_{i=1}^n (b_i - a_i).$$

Definition. Let $B = \prod_{i=1}^n [a_i, b_i]$ be a box, then a partition P of B is a set $\mathcal{P} = \{P_1, \dots, P_n\}$ where each P_i is a partition of $[a_i, b_i]$.

Definition. Given a box B and two partitions $\mathcal{P}_1 = \{P_1^{(1)}, \dots, P_n^{(1)}\}$, $\mathcal{P}_2 = \{P_1^{(2)}, \dots, P_n^{(2)}\}$ of B , we say that \mathcal{P}_2 is a refinement of \mathcal{P}_1 if $P_i^{(2)}$ is a refinement of $P_i^{(1)}$ for all $i \in \llbracket 1, n \rrbracket$.

Each partition of a box B induces a set of subboxes of B . More precisely, Given a box $B = \prod_{i=1}^n [a_i, b_i]$ and a partition $\mathcal{P} = \{P_1, \dots, P_n\}$ where $P_i = \{a_i = x_i^0 < x_i^1 < \dots < x_i^{m_i}\}$ for all $i \in \llbracket 1, n \rrbracket$, then \mathcal{P} splits B into the subboxes

$$B_{i_1, \dots, i_n} = \prod_{i=1}^n [x_i^{i_1}, x_i^{i_1+1}]$$

where $i_k \in \llbracket 1, m_i - 1 \rrbracket$ for all $k \in \llbracket 1, n \rrbracket$. To make the notation lighter, denote the set of these subboxes by $B(\mathcal{P})$.

Definition. Let B be a box, $f : B \rightarrow \mathbb{R}$ be a bounded function and \mathcal{P} be a partition of B , define the lower sum with respect to \mathcal{P} by

$$L(f, \mathcal{P}) = \sum_{\beta \in B(\mathcal{P})} \inf_{\beta} (f) \text{Vol}(\beta)$$

and the upper sum with respect to \mathcal{P} by

$$U(f, \mathcal{P}) = \sum_{\beta \in B(\mathcal{P})} \sup_{\beta} (f) \text{Vol}(\beta).$$

Lemma 3.1.1. *Let B be a box, $f : B \rightarrow \mathbb{R}$ be a bounded function, \mathcal{P} and $\tilde{\mathcal{P}}$ be partitions of B such that $\tilde{\mathcal{P}}$ is a refinement of \mathcal{P} , then*

$$L(f, \mathcal{P}) \leq L(f, \tilde{\mathcal{P}}) \leq U(f, \tilde{\mathcal{P}}) \leq U(f, \mathcal{P}).$$

Proof. **TODO** ■

Definition. Let B be a box and $f : B \rightarrow \mathbb{R}$ be a bounded function, then we define its upper and lower integral by

$$U(f) = \inf_{\mathcal{P}} U(f, \mathcal{P}) \quad \text{and} \quad L(f) = \sup_{\mathcal{P}} L(f, \mathcal{P}).$$

We say that f is Riemann integrable on B if $U(f) = L(f)$. In that case, we write

$$\int_B f(x) dV = U(f) = L(f).$$

Theorem 3.1.2 (Integrability Criterion). *Let $f : B \rightarrow \mathbb{R}$ be bounded, then f is Riemann integrable over B if and only if for all $\epsilon > 0$, there exists a partition \mathcal{P} of B such that*

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) \leq \epsilon.$$

Proof. **TODO** ■

Theorem 3.1.3. Let $f : B \rightarrow \mathbb{R}$ be in $C^0(B)$, then it is Riemann integrable.

Proof. **TODO** ■

Theorem 3.1.4 (Mean Value Theorem). Let $f : B \rightarrow \mathbb{R}$ be continuous, then there exist a $x_0 \in B$ such that

$$\int_B f dV = f(x_0) \text{Vol}(B).$$

Proof. **TODO** ■

3.2 Riemann Integral over Generalized Domains

How to define $\int_\Omega f dV$ when Ω is not necessarily a box ?

Definition. Let $f : \Omega \rightarrow \mathbb{R}$ be a bounded function over a bounded domain, we define

$$\int_\Omega f dV = \int_B f \chi_\Omega dV$$

where B is a box containing Ω and where the right hand side exists.

Definition. We say that $\Gamma \subset \mathbb{R}^n$ is of content zero if for all $\epsilon > 0$, there exist open box domains B_1, \dots, B_n such that $\Gamma \subset \cup_{i=1}^n B_i$ and $\sum_{i=1}^n \text{Vol}(B_i) < \epsilon$.

Notice that a finite union of content zero sets has content zero as well.

Theorem 3.2.1. Let $f : B \rightarrow \mathbb{R}$ be a bounded function and $\Gamma(f) \subset B$ be the set of points x for which f is discontinuous at x . If $\Gamma(f)$ has content zero, then f is Riemann integrable over B .

Proof. **TODO** ■

Proposition 3.2.2. If $\Omega \subset \mathbb{R}^n$ is compact and $g : \Omega \rightarrow \mathbb{R}$ is continuous, then the graph of g , $\Gamma = \{(x, g(x)) : x \in \Omega\}$, has content zero.

Proof. **TODO** ■

Corollary. Suppose $f : B \rightarrow \mathbb{R}$ is bounded and its set of discontinuity $\Gamma(f)$ is the union of finitely many graphs of continuous functions over closed domains in \mathbb{R}^{n-1} , then f is Riemann integrable over B .

Proof. **TODO** ■

Corollary. Let $\Omega \subset \mathbb{R}^n$ be a compact set such that $\partial\Omega$ the union of finitely many graphs of continuous functions over closed domains. Suppose $f : \Omega \rightarrow \mathbb{R}$ is bounded and $\Gamma(f) \cap \Omega$ has content zero, then $\int_\Omega f dV$ exists.

Proof. **TODO** ■

Corollary. Let $G : \mathbb{R}^n \rightarrow \mathbb{R}$ be a differentiable function on \mathbb{R}^n such that $\Omega = \{G \leq 0\}$ is bounded. Suppose that $\nabla G \neq 0$ on $\{G = 0\}$. If $f : \Omega \rightarrow \mathbb{R}$ is a bounded function with the property that $\Gamma(f) \cap \Omega$ has content zero, then $\int_\Omega f dV$ exists.

Proof. **TODO** ■

3.3 Fubini's Theorem

To compute integrals in one variable, we use integration techniques and the Fundamental Theorem of Calculus. Usually, we learn in multivariable calculus that computing an integral over several variables is the same as computing the one-variable integral for each variable iteratively. However, from our definition of the multivariable integral, this may not be obvious. More precisely, why does the following formula holds ?

$$\int_{[a,b] \times [c,d]} f dV = \int_a^b \int_c^d f(x,y) dy dx.$$

The goal of this subsection is to understand why we can make such manipulations.

Theorem 3.3.1 (Fubini for Box Domains). *Suppose $f : B \rightarrow \mathbb{R}$ is Riemann integrable over $B = [a, b] \times [c, d]$ and for each $x \in [a, b]$, the integral $\int_c^d f(x, y) dy$ exists, then the function $x \mapsto \int_c^d f(x, y) dy$ is integrable over $[a, b]$ and its integral satisfies*

$$\int_B f dV = \int_a^b \int_c^d f(x, y) dy.$$

Similarly, if for each $y \in [c, d]$, the integral $\int_a^b f(x, y) dx$ exists, then the function $y \mapsto \int_a^b f(x, y) dx$ is integrable over $[c, d]$ and its integral satisfies

$$\int_B f dV = \int_c^d \int_a^b f(x, y) dx.$$

Proof. **TODO** ■

What about more general regions ?

Definition. A domain $\Omega \subset \mathbb{R}^2$ is called y -simple if it is of the form

$$\Omega = \{a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\}$$

where $\varphi_1, \varphi_2 : [a, b] \rightarrow \mathbb{R}$ are continuous. Similarly, $\Omega \subset \mathbb{R}^2$ is called x -simple if it is of the form

$$\Omega = \{c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\}$$

where $\psi_1, \psi_2 : [c, d] \rightarrow \mathbb{R}$ are continuous. Finally, Ω is called simple if it is both x -simple and y -simple. These three type of regions are called elementary regions.

Theorem 3.3.2 (Fubini for Elementary Regions). *Let $\Omega \subset \mathbb{R}^2$ be a bounded domain and assume that $f : \Omega \rightarrow \mathbb{R}$ is such that $\int_{\Omega} f dV$ exists.*

1. *If Ω is y -simple and $\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy$ exists for all $x \in [a, b]$, then*

$$\int_{\Omega} f dV = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy dx.$$

2. If Ω is x -simple and $\int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx$ exists for all $y \in [c, d]$, then

$$\int_{\Omega} f dV = \int_c^d \int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx dy.$$

3. If Ω is simple, $\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy$ exists for all $x \in [a, b]$ and $\int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx$ exists for all $y \in [c, d]$, then

$$\int_B f dV = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy dx = \int_c^d \int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx dy.$$

Proof. **TODO** ■

With these theorems, we can now compute integrals over elementary regions of \mathbb{R}^2 . Let's extend this to higher dimensions.

Theorem 3.3.3. Let $B = B_1 \times B_2$ be a box domain where $B_1 \subset \mathbb{R}^k$ and $B_2 \subset \mathbb{R}^{n-k}$. Suppose that $f : B \rightarrow \mathbb{R}$ is a Riemann integrable function. Set the functions

$$L(x) = \sup_{\mathcal{P}} L(f, B_2, \mathcal{P}), \quad U(x) = \inf_{\mathcal{P}} U(f, B_2, \mathcal{P})$$

where $x \in B_1$ and

$$\tilde{L}(y) = \sup_{\mathcal{P}} L(f, B_1, \mathcal{P}), \quad \tilde{U}(y) = \inf_{\mathcal{P}} U(f, B_1, \mathcal{P})$$

where $x \in B_2$, then L and U are Riemann integrable over B_1 and \tilde{L} and \tilde{U} are Riemann integrable over B_2 . Moreover,

$$\int_B f dV = \int_{B_1} L dV(x) = \int_{B_1} U dV(x)$$

and

$$\int_B f dV = \int_{B_2} \tilde{L} dV(y) = \int_{B_2} \tilde{U} dV(y).$$

Corollary. Let $B = B_1 \times B_2$ be a box domain where $B_1 \subset \mathbb{R}^k$ and $B_2 \subset \mathbb{R}^{n-k}$. Suppose that $f : B \rightarrow \mathbb{R}$ is a Riemann integrable function. If $f(x, y)$ is integrable on B_2 for all $x \in B_1$, then the function $x \mapsto \int_{B_2} f(x, y) dV(y)$ is integrable over B_1 and its integral satisfies

$$\int_B f dV = \int_{B_1} \left[\int_{B_2} f(x, y) dV(y) \right] dV(x).$$

Similarly, if $f(x, y)$ is integrable on B_1 for all $y \in B_2$, then the function $y \mapsto \int_{B_1} f(x, y) dV(x)$ is integrable over B_2 and its integral satisfies

$$\int_B f dV = \int_{B_2} \left[\int_{B_1} f(x, y) dV(x) \right] dV(y).$$

Definition. A domain $W \subset \mathbb{R}^3$ is called z -simple if there exists an elementary region $\Omega \subset \mathbb{R}^2$ and two continuous functions $\eta_1, \eta_2 : \Omega \rightarrow \mathbb{R}$ such that

$$W = \{(x, y, z) : (x, y) \in \Omega, \eta_1(x, y) \leq z \leq \eta_2(x, y)\}.$$

We can define x -simple and y -simple domains in \mathbb{R}^3 in the same way. These regions are called the elementary in \mathbb{R}^3 .

Theorem 3.3.4. Suppose $W \subset \mathbb{R}^3$ is z -simple and $f : W \rightarrow \mathbb{R}$ is Riemann integrable over W . If $\int_{\eta_1(x,y)}^{\eta_2(x,y)} f(x, y, z) dz$ exists for all $(x, y) \in \Omega$, then the function $(x, y) \mapsto \int_{\eta_1(x,y)}^{\eta_2(x,y)} f(x, y, z) dz$ is Riemann integrable over Ω . Moreover,

$$\int_W f dV = \int_{\Omega} \left[\int_{\eta_1(x,y)}^{\eta_2(x,y)} f(x, y, z) dz \right] dV(x, y).$$

3.4 Change of Variables Formula

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. We define $\int_{\Omega} f dV$ by looking at $L(f, \mathcal{P})$ and $U(f, \mathcal{P})$ where \mathcal{P} denotes a rectangular partition. What if we consider partitions that may not be rectangular?

Definition. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with a content zero boundary. A regular partition $\mathcal{P}_{\text{reg}} = \{D_k\}_{k=1}^N$ is a collection of subdomains $D_k \subset \Omega$ such that $\Omega = \cup_{k=1}^N D_k$, each D_k has a content zero boundary and $D_k \cap D_l$ has content zero whenever $k \neq l$.

Definition. Suppose $f : \Omega \rightarrow \mathbb{R}$ is bounded where $\Omega \subset \mathbb{R}^n$ is a bounded domain. Let $\mathcal{P}_{\text{reg}} = \{D_k\}_{k=1}^N$ be a regular partition of Ω . Define

$$U(f, \mathcal{P}_{\text{reg}}) = \sum_{k=1}^N \sup_{D_k} f \text{Vol}(D_k) \quad \text{and} \quad L(f, \mathcal{P}_{\text{reg}}) = \sum_{k=1}^N \inf_{D_k} f \text{Vol}(D_k).$$

Lemma 3.4.1. Let $f : B \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be a bounded function, then f is Riemann integrable over the box domain B if and only if for all $\epsilon > 0$, there is a regular partition \mathcal{P}_{reg} such that

$$U(f, \mathcal{P}_{\text{reg}}) - L(f, \mathcal{P}_{\text{reg}}) < \epsilon.$$

Proof. **TODO** ■

Theorem 3.4.2 (Change of Variables Formula). Let $\Omega, \tilde{\Omega} \subset \mathbb{R}^n$ be bounded domains. Let $T : \tilde{\Omega} \rightarrow \Omega$ be a C^1 bijection, then for any integrable function $f : \Omega \rightarrow \mathbb{R}$, we have

$$\int_{\Omega} f dV = \int_{\tilde{\Omega}} (f \circ T) |\det(DT)| dV$$

Proof. **TODO** ■

4 Toward Green's and Stokes' Theorem

4.1 Integration over Paths

A function of the form $\gamma : [a, b] \rightarrow \mathbb{R}^n$ is called a path. The graph of a path forms a line in \mathbb{R}^n . For any $t \in [a, b]$, the vector $\gamma'(t)$ represents the vector tangent to the graph of γ at $\gamma(t)$.

Definition. The length of a C^1 path $\gamma : [a, b] \rightarrow \mathbb{R}^n$ is defined by

$$L(\gamma) = \int_a^b \|\gamma'(t)\| dt.$$

Definition. The length of a piecewise C^1 path $\gamma : [a, b] \rightarrow \mathbb{R}^n$ is defined by

$$L(\gamma) = \sum_{i=1}^m \int_{a_{i-1}}^{a_i} \|\gamma'(t)\| dt$$

where γ is C^1 on $[a_{i-1}, a_i]$ for all i .

Definition. A path is said to be regular if it is C^1 and if its derivative is never 0. A path is said to be piecewise smooth if it is piecewise regular.

Definition. Let $\gamma : [a, b] \rightarrow \mathbb{R}^n$ and $\tilde{\gamma} : [c, d] \rightarrow \mathbb{R}^n$ be piecewise C^1 , then we say that $\tilde{\gamma}$ is a reparametrization of γ if there exists a C^1 bijective function $h : [c, d] \rightarrow [a, b]$ such that

$$\tilde{\gamma}(t) = \gamma(h(t))$$

for all $t \in [c, d]$. If $\tilde{\gamma}(c) = \gamma(a)$ and $\tilde{\gamma}(d) = \gamma(b)$, we say that the reparametrization is orientation-preserving. Otherwise, we say that the orientation is reversing.

Proposition 4.1.1. Suppose $\gamma : [a, b] \rightarrow \mathbb{R}^n$ is a piecewise C^1 curve and $\tilde{\gamma}$ is a reparametrization of γ , then $L(\gamma) = L(\tilde{\gamma})$.

Proof. **TODO** ■

Given a regular path $\gamma : [a, b] \rightarrow \mathbb{R}^n$, the function defined by

$$s(t) = \int_a^t \|\gamma'(t)\| dt$$

is called the arclength function. By the Fundamental Theorem of Calculus, its derivative is $\|\gamma'(t)\|$. Since γ is regular, then s has a non-zero derivative and so must be one-to-one from $[a, b]$ onto $[0, L(\gamma)]$. its inverse $s \mapsto t(s)$ is called the arclength parametrization of γ . Since $s'(t) = \|\gamma'(t)\|$, then

$$\frac{d}{dt} s^{-1}(t) = \frac{1}{\|\gamma'(s^{-1}(t))\|}.$$

It follows that if we consider the reparametrization $\tilde{\gamma} = \gamma \circ s^{-1} : [0, L] \rightarrow \mathbb{R}^n$, we obtain

$$\|\tilde{\gamma}'(t)\| = \left\| \frac{\gamma'(s^{-1}(t))}{\|\gamma'(s^{-1}(t))\|} \right\| = 1.$$

Definition. Suppose $\gamma : [a, b] \rightarrow \mathbb{R}^n$ is piecewise C^1 and $f : \text{Im}(\gamma) \rightarrow \mathbb{R}$ is bounded. The path integral of f along γ is defined by

$$\int_{\gamma} f ds = \sum_{i=1}^m \int_{a_{i-1}}^{a_i} f(\gamma(t)) \|\gamma'(t)\| dt$$

provided the integrals on the right hand side exist.

We can easily show that $\int_{\gamma} f ds$ is independent of the parametrization. By our previous observation, if we reparametrize by arclength, the formula becomes

$$\int_{\gamma} f ds = \int_0^L f(\gamma(t)) dt$$

where $L = L(\gamma)$.

Definition. A vector field in $\Omega \subset \mathbb{R}^n$ is a map $\vec{F} : \Omega \rightarrow \mathbb{R}^n$.

How to integrate a vector field over a path γ ? Notice that this is exactly the kind of quantity we are interested in when computing the work done by a force \vec{F} represented by a vector field over a path. Intuitively, the work done by a force over a path is simply the dot product between the force and the tangent of the path. This leads to the following definition.

Definition. The total work done by a vector field \vec{F} along a path γ is defined by

$$\int_{\gamma} \vec{F} \cdot d\vec{s} = \int_a^b \vec{F}(\gamma(t)) \cdot \gamma'(t) dt$$

provided the function is integrable. If we denote by $\vec{T}(t)$ the normalized tangent vector $\gamma' / \|\gamma'\|$, we obtain

$$\int_{\gamma} \vec{F} \cdot d\vec{s} = \int_{\gamma} \vec{F}(\gamma(t)) \cdot \vec{T}(t) ds$$

where ds denotes the arclength element $\|\gamma'(t)\| dt$. With this last equation, we can define the work done by \vec{F} over γ by the path integral of $\vec{F} \cdot \vec{T}$ over γ .

Proposition 4.1.2. Let $\gamma : [a, b] \rightarrow \mathbb{R}^n$ be a piecewise smooth path and $\tilde{\gamma} : [c, d] \rightarrow \mathbb{R}^n$ be a reparametrization of γ . Let \vec{F} be a vector field, then

$$\int_{\gamma} \vec{F} \cdot d\vec{s} = \int_{\tilde{\gamma}} \vec{F} \cdot d\vec{s}$$

is $\tilde{\gamma}$ is orientation-preserving and

$$\int_{\gamma} \vec{F} \cdot d\vec{s} = - \int_{\tilde{\gamma}} \vec{F} \cdot d\vec{s}$$

is $\tilde{\gamma}$ is orientation-reversing.

Proof. **TODO** ■

Definition. A path $\gamma : [a, b] \rightarrow \mathbb{R}^n$ is called simple if it is one-to-one.

Definition. A vector field \vec{F} is called conservative if it can be written as the gradient of a function f which we call the potential function.

Theorem 4.1.3. Suppose $\vec{F} = \nabla f$ for some C^1 function f in a neighborhood of a piecewise smooth path γ , then

$$\int_{\gamma} \vec{F} \cdot d\vec{s} = f(\gamma(b)) - f(\gamma(a)).$$

Proof. **TODO**(FTC) ■

Corollary. Suppose $\vec{F} = \nabla f$ for some C^1 function f in a neighborhood of a piecewise smooth closed path γ , then

$$\int_{\gamma} \vec{F} \cdot d\vec{s} = 0.$$

Let's find a way to determine quickly if a given vector field is conservative.

Lemma 4.1.4 (Differentiating under the Integral). Suppose $g(x, y) \in C^1(\Omega \times A)$ where $\Omega \subset \mathbb{R}^n$ is open and $A \subset \mathbb{R}^m$ is compact, then

$$\frac{\partial}{\partial x_i} \int_A g(x, y) dV(y) = \int_A \frac{\partial g}{\partial x_i}(x, y) dV(y)$$

for all $x \in \Omega$ and $i \in \llbracket 1, n \rrbracket$.

Proof. **TODO** ■

Definition. A domain $\Omega \subset \mathbb{R}^n$ is star-shaped if there exists a $s_0 \in \Omega$ such that for all $s \in \Omega$, the line segment connecting s to s_0 lies completely in Ω .

Theorem 4.1.5 (Poincaré). Suppose $\Omega \subset \mathbb{R}^n$ is star-shaped and \vec{F} is a C^1 vector field in Ω , then \vec{F} is conservative if and only if

$$\frac{\partial F_i}{\partial x_j}(x) = \frac{\partial F_j}{\partial x_i}(x)$$

for all $x \in \Omega$ and $i, j \in \llbracket 1, n \rrbracket$.

Proof. **TODO** ■

4.2 Integration over Surfaces

A surface is the graph of a C^1 function $g : D \rightarrow \mathbb{R}$ where $D \subset \mathbb{R}^2$ is a bounded domain. We define the unit normal vector to a surface at $g(t)$ by

$$\vec{n}(t) = \frac{\vec{T}_x(t) \times \vec{T}_y(t)}{\|\vec{T}_x(t) \times \vec{T}_y(t)\|} = \frac{(-g_x(t), -g_y(t), 1)}{\sqrt{1 + g_x(t)^2 + g_y(t)^2}}$$

where $\vec{T}_x(t)$ and $\vec{T}_y(t)$ are the tangent of S in the x and y direction respectively. More precisely:

$$\vec{T}_x(t) = (1, 0, g_x(t)) \quad \text{and} \quad \vec{T}_y(t) = (0, 1, g_y(t)).$$

Definition. The area of a surface S defined by $g : D \rightarrow \mathbb{R}$ is defined by

$$\text{area}(S) = \int_D \sqrt{1 + g_x^2 + g_y^2} dA$$

Definition. A parametrization of a surface S is the image of a map $\phi : D \rightarrow S \subset \mathbb{R}^3$. If ϕ is C^1 , we call S a C^1 surface. We say that S is regular if $\vec{\phi}_u \times \vec{\phi}_v \neq 0$ for all $(u, v) \in D$.

The tangent to the surface at any point $\phi(u, v)$ is spanned by $\vec{T}_u = \partial\phi/\partial u$ and $\vec{T}_v = \partial\phi/\partial v$. The tangent plane at $X_0 = (x(u_0, v_0), y(u_0, v_0), z(u_0, v_0))$ is given by the equation

$$(X - X_0) \cdot \vec{T}_u \times \vec{T}_v = 0.$$

It follows that the area of the surface S parametrized by ϕ is given by

$$\text{area}(S) = \int_D \sqrt{\left| \frac{\partial(x, y)}{\partial(u, v)} \right|^2 + \left| \frac{\partial(y, z)}{\partial(u, v)} \right|^2 + \left| \frac{\partial(x, z)}{\partial(u, v)} \right|^2} dA$$

Definition. Let S be a surface parametrized by ϕ , and let $f : S \rightarrow \mathbb{R}$ be a function, then we define the integral of f over S as

$$\int_S f dS = \int_D f(\phi(u, v)) \left\| \vec{T}_u \times \vec{T}_v \right\| dV.$$

In the special case when S is the graph of a function g , we get

$$\int_S f dS = \int_D f(x, y, g(x, y)) \sqrt{1 + g_x^2 + g_y^2} dV.$$

Definition. A parametrized C^1 regular surface is orientable if there is a unit normal $\vec{n}(p)$ at each $p \in S$ such that the function $n : S \rightarrow \mathbb{R}^3$ is a continuous map. If S is orientable with a unit map \vec{n} , then $-\vec{n}$ gives the opposite orientation.

Definition. Let \vec{F} be a vector field over an orientable regular surface S , we define its integral by

$$\int_S \vec{F} \cdot d\vec{s} = \int_S \vec{F} \cdot \vec{n} ds = \int_D \vec{F} \cdot (\phi_u \times \phi_v) dV$$

Definition. Let $\phi : D \subset \mathbb{R}^2 \rightarrow S \subset \mathbb{R}^3$ be a regular parametrization of S , then the function $\tilde{\phi} : \tilde{D} \subset \mathbb{R}^2 \rightarrow S$ is a reparametrization of S if there exists a C^1 bijection $h : \tilde{D} \rightarrow D$ with C^1 inverse that satisfies

$$\tilde{\phi}(s, t) = \phi(h(s, t))$$

for all $(s, t) \in \tilde{D}$. The reparametrization $\tilde{\phi}$ is orientation-preserving if the unit normal vector at each point of S is the same for both parametrizations. It is orientation-reversing otherwise.

Lemma 4.2.1. Let ϕ and $\tilde{\phi}$ be parametrizations of S such that $\tilde{\phi}$ is a reparametrization of ϕ , then if we write $h(s, t) = (u(s, t), v(s, t))$, we get that $\tilde{\phi}$ is orientation-preserving if and only if the determinant of the Jacobian of h is positive and orientation-reversing if and only if the determinant of the Jacobian of h is negative.

Proposition 4.2.2. *Let S be an oriented C^1 regular surface and $\vec{F} : S \rightarrow \mathbb{R}^3$ a continuous vector field, then*

$$\int_S \vec{F} \cdot d\vec{s}$$

is well-defined and independent of orientation-preserving reparametrizations.

Proof. **TODO** ■

4.3 Green's Theorem

The goal is to relate line integrals of vector fields and double integrals of planar domains. We say that a bounded planar domain is positively oriented if it has a normal $(0, 0, 1)$. The positive orientation of ∂D is defined to be the one induced by the Right Hand Rule.

Theorem 4.3.1 (Green's Theorem Part 1). *Let $D \subset \mathbb{R}^2$ be a piecewise smooth domain and $\vec{F} = (P, Q) : D \rightarrow \mathbb{R}^2$ be a C^1 vector field on D , then*

$$\int_{\partial D} \vec{F} \cdot d\vec{s} = \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA.$$

Proof. **TODO** ■

Corollary. *Let $D = D_1 \setminus D_2$ and $\partial D_1, \partial D_2$ be piecewise C^1 . If $\vec{F} = (P, Q)$ is C^1 with*

$$\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$$

on D , then

$$\int_{\partial D_1} \vec{F} \cdot d\vec{s} = \int_{\partial D_2} \vec{F} \cdot d\vec{s}.$$

Theorem 4.3.2 (Green's Theorem Part 2). *Let $D \subset \mathbb{R}^2$ be a piecewise smooth domain and $\vec{F} = (P, Q) : D \rightarrow \mathbb{R}^2$ be a C^1 vector field on D , then*

$$\int_{\partial D} \vec{F} \cdot \vec{n} ds = \int_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} \right) dA.$$

Proof. **TODO** Use Green part 1 ■

Notice that if $u \in C^2(D)$ is harmonic and $\vec{F} = \nabla u$, then

$$\int_{\partial D} \nabla u \cdot \vec{n} ds = \int_{\partial D} \vec{F} \cdot \vec{n} ds = \int_D \Delta u dA = 0.$$

Theorem 4.3.3 (Green's Identities). *If $\varphi, \psi \in C^2(D)$ and ∂D is piecewise C^1 , then*

$$\int_D \varphi \Delta \psi dA = - \int_D \nabla \varphi \cdot \nabla \psi dA + \int_{\partial D} \varphi \nabla \psi \cdot \vec{n} ds$$

and

$$\int_D [\varphi \Delta \psi - \psi \Delta \varphi] dA = \int_{\partial D} [\varphi \nabla \psi \cdot \vec{n} - \psi \nabla \varphi \cdot \vec{n}] ds.$$

Proof. **TODO** ■

4.4 Stokes' Theorem

Definition. Let $\vec{F} = (F_1, F_2, F_3) : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a C^1 vector field. We define the curl of F as

$$\text{curl } \vec{F} = \nabla \times \vec{F}.$$

Notice that if $\vec{F} = (P, Q, 0)$ and the $\vec{n} = (0, 0, 1)$, then

$$\text{curl } \vec{F} \cdot \vec{n} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}.$$

Theorem 4.4.1 (Stokes' Theorem). *Let S be a compact regular oriented surface with positively oriented piecewise C^1 oriented boundary ∂S , and let $\vec{F} : S \rightarrow \mathbb{R}^3$ be a C^1 vector field, then*

$$\int_S \text{curl } \vec{F} \cdot d\vec{S} = \int_{\partial S} \vec{F} \cdot d\vec{s}.$$

Proof. **TODO** ■

Corollary. *Suppose S_1 and S_2 are two oriented surfaces with $\partial S_1 = \partial S_2 = \gamma$. Suppose S_1 and S_2 induce opposite orientations on γ , then for any C^1 vector field \vec{F} ,*

$$\int_{S_1} (\nabla \times \vec{F}) \cdot d\vec{S} = - \int_{S_2} (\nabla \times \vec{F}) \cdot d\vec{S}.$$

Proof. **TODO** ■

Corollary. *If S is a compact oriented surface with $\partial S = \emptyset$ and $\vec{F} \in C^1(S)$, then*

$$\int_S (\nabla \times \vec{F}) \cdot d\vec{S} = 0.$$

Proof. **TODO** ■

Corollary. *Suppose $\vec{F} = \nabla f$ for some $f \in C^2$ and S is an oriented surface with $\partial S = \gamma_1 - \gamma_2$, then*

$$\int_{\gamma_1} \vec{F} \cdot d\vec{s} = \int_{\gamma_2} \vec{F} \cdot d\vec{s}.$$

Proof. **TODO** the curl is zero. ■

4.5 Divergence Theorem

Theorem 4.5.1 (Divergence Theorem). *Let $W \subset \mathbb{R}^3$ be a bounded solid region in \mathbb{R}^3 with smooth boundary $S = \partial W$ with positive orientation. Let $\vec{F} : W \rightarrow \mathbb{R}^3$ be a C^1 vector field, then*

$$\int_W (\nabla \cdot \vec{F}) dV = \int_S \vec{F} \cdot d\vec{s}.$$