Review of Multi-Calculus

(Study Guide for Spivak’s CHAPTER ONE TO THREE)

This material is for June 9 to 16 (Monday to Monday)

Chapter I: Functions on $\mathbb{R}^n$

- **Dot product and norm for vectors in $\mathbb{R}^n$:** Let $X = (x_1, \ldots, x_n), Y = (y_1, \ldots, y_n)$ be two vectors in $\mathbb{R}^n$, the dot product $X \cdot Y$ is defined by
  $$X \cdot Y = x_1y_1 + \cdots + x_ny_n.$$  
  The norm (length) of $X$ is
  $$\|X\|^2 = X \cdot X = x_1^2 + \cdots + x_n^2.$$

- **Open, close and compact subsets.** A subset $A \subset \mathbb{R}^n$ is open if for every point $X_0 \in A$, there exists $\epsilon > 0$ such that
  $$\{X \in \mathbb{R}^n \mid \|X - X_0\| < \epsilon\} \subset A.$$  
  A subset $A \subset \mathbb{R}^n$ is closed if its complement is open. A subset $A$ is compact if every open cover of $A$ contains a finite subcollections of open sets which also covers $A$. A subset $A \subset \mathbb{R}^n$ is compact iff it is closed and bounded (Hein-Borel Theorem).

- **Continuous functions.** Recall that in Calculus, a function $f$ is continuous at $x_0$ means that $\lim_{x \to x_0} f(x) = f(x_0)$, i.e. for every $\epsilon > 0$, there exists a $\delta > 0$ such that for all $0 < |x - x_0| < \delta$, we have
  $$|f(x) - f(x_0)| < \epsilon.$$  
  $f$ is continuous on $(a, b)$ if $f$ is continuous at every point in $(a, b)$. Similarly, a function $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is said to be continuous on $U$ if for every $X_0 \in U$ and every $\epsilon > 0$, there exists a $\delta > 0$ such that for all $0 < \|X - X_0\| < \delta$, we have
  $$\|F(X) - F(X_0)\| < \epsilon.$$  
  One of the pleasant surprise about the concept of continuity is that it can be described (defined) without using the limits: $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous iff $f^{-1}(U)$ is open whenever $U \subset \mathbb{R}^m$ is open.
Part II (Chapter 2): Differentiation

- **Differentiability and (total) Derivatives:** One of the basic ideas of differential calculus is to approximate differentiable maps by linear maps as to reduce analytic (hard) problems to linear-algebraic (easy) problems whenever possible. Consider a function \( f : U \subset \mathbb{R}^n \to \mathbb{R}^m \). The linear approximation of \( f \), locally at \( x \in U \), is the derivative \( Df(x) : \mathbb{R}^n \to \mathbb{R}^m \) of \( f \) at \( x \). The differential is characterized by \( f(x+h) = f(x) + Df(x)(h) + \phi(h) \) where \( \lim_{h \to 0} \phi(h)/\|h\| = 0 \). This leads to the following definition:

**Definition:** A function \( f : U \subset \mathbb{R}^n \to \mathbb{R}^m \) is said to be differentiable at \( a \in U \) if there exists a linear transformation \( \lambda : \mathbb{R}^n \to \mathbb{R}^m \) such that

\[
\lim_{h \to 0} \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} = 0.
\]

It is proved that if \( f \) is differentiable at \( a \), then \( \lambda \) is unique! We denote \( \lambda \) by \( Df(a) \) and call it the derivative of \( f \) at \( a \). So \( Df(a) \), the derivative of \( f \) at \( a \), is a linear transformation from \( \mathbb{R}^n \) to \( \mathbb{R}^m \).

- **Directional and Partial Derivatives:** Let \( f : U \subset \mathbb{R}^n \to \mathbb{R}^m \). Let \( a \in U \) and let \( v \in \mathbb{R}^n \). The **directional derivative** of \( f \) at \( a \) in the direction \( v \) is defined by

\[
D_v f(a) = \lim_{t \to 0} \frac{f(a + tv) - f(a)}{t}.
\]

Note here \( t \in \mathbb{R} \) so \( \frac{f(a + tv) - f(a)}{t} \) makes sense. \( D_{e_i} f \), where \( e_1 = (1, \ldots, 0) \), \ldots, \( e_n = (0, \ldots, 1) \), is called the **partial derivatives** of \( f \), and is denoted by \( D_i f(a) \) (or \( \frac{\partial f}{\partial x_i} |_a \)).

**Remark:** \( f \) is differentiable at \( a \) implies the partials \( \frac{\partial f}{\partial x_i} |_a \) \( 1 \leq i \leq n, \ 1 \leq j \leq m \), exist. However, unlike the \( n = 1 \) case, the existence of partials \( \frac{\partial f}{\partial x_i} |_a \) does not imply that \( f \) is differentiable at \( a \).

- According to the linear algebra, any linear transformation \( T : V \to W \) can be associated to a matrix \( A \) after a basis of \( V \) and a basis \( W \) are fixed. Now assume that a function \( f = (f_1, \ldots, f_m) : U \subset \mathbb{R}^n \to \mathbb{R}^m \) is differentiable at \( a \in U \), so \( Df(a) : \mathbb{R}^n \to \mathbb{R}^m \) exists. Choose the standard basis \( \{e_1, \ldots, e_n\} \) for \( \mathbb{R}^n \) and \( \{e_1, \ldots, e_m\} \) for \( \mathbb{R}^m \), then the matrix of \( Df(a) \) with respect to these two standard basis is the \( m \times n \) matrix (called the **Jacobian matrix**).
by $f'(a)$, given by (see page 17 on the textbook)

$$f'(a) = \begin{pmatrix}
\frac{\partial f_1}{\partial x_1} |_{a} & \cdots & \frac{\partial f_1}{\partial x_n} |_{a} \\
\vdots & \ddots & \vdots \\
\frac{\partial f_m}{\partial x_1} |_{a} & \cdots & \frac{\partial f_m}{\partial x_n} |_{a}
\end{pmatrix}. $$

With this matrix, we have,

$$Df(a) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = f'(a) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix}
\frac{\partial f_1}{\partial x_1} |_{a} & \cdots & \frac{\partial f_1}{\partial x_n} |_{a} \\
\vdots & \ddots & \vdots \\
\frac{\partial f_m}{\partial x_1} |_{a} & \cdots & \frac{\partial f_m}{\partial x_n} |_{a}
\end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. $$

- **Nations for this book:** $Df(a)$ is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, $f'(a)$ is an $m \times n$ matrix, $D_i f(a)$ are partial derivatives of $f$ at $a$.

- **Chain Rule:** The main computational tool for derivatives is:

**Theorem (Chain Rule)** If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a$ and if $g : \mathbb{R}^m \to \mathbb{R}^p$ is differentiable at $b = f(a)$, then $g \circ f : \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at $a$ and $D(g \circ f)(a) = D(g(b)) Df(a)$.

**Theorem (Chain Rule for partials)** If $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$ are continuously differentiable at $a$ and if $f : \mathbb{R}^m \to \mathbb{R}$ is differentiable at $(g_1(a), \ldots, g_m(a))$, then $F = f(g_1, \ldots, g_m)$ has partials

$$D_i F(a) = \sum_{j=1}^{m} D_j f(g_1(a), \ldots, g_m(a)) D_i g_j(a).$$

**Theorem:** If $f : D \to \mathbb{R}$ is a function defined on an open subset $D \subset \mathbb{R}^n$ and if $f$ and its first and second order partials are exist throughout $D$ and are continuous there, then $D_i D_j f(a) = D_j D_i f(a)$ for all $a \in D$.

- **The Inverse Function Theorem and Implicit Function Theorem:** Again, the basic ideas of differential calculus is to approximate differentiable maps by linear maps as to reduce analytic (hard) problems to linear-algebraic (easy) problems whenever possible. Consider a function $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$. The linear approximation of $f$, locally at $a \in U$, is the derivative $Df(a) : \mathbb{R}^n \to \mathbb{R}^m$ of $f$ at $a$. In the case when $n = m$, one expects that the inversability of the linear approximation $Df(a)$ (which can be easily determined, according to the linear
algebra, by looking at its determinant) implies the locally inversability of $f$ at $a$. This is exactly the meaning of the following “Inverse Function Theorem”.

**The Inverse Function Theorem (see Page 35):** Let $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable. Assume that $a \in U$ and $\det f'(a) \neq 0$ (i.e. $Df(a)$ is invertible). Then there is a neighborhood $V \subset U$ of $a$ and $W$ of $f(a)$ and a differentiable function $f^{-1} : W \to V$ (i.e. $f \circ f^{-1}$ is an identity map) such that, for every $y \in W$,

$$(f^{-1})'(y) = [f'(f^{-1}(y))]^{-1}.$$ 

**Exercise:** Show that the theorem is true for linear functions.

To prove the Inverse function theorem, we first prove the following lemma:

**Lemma:** Let $f : A \to \mathbb{R}^m$ be defined and differentiable on a convex open set $A$. Then for all $x, y \in A$, one has

$$|f(x) - f(y)| \leq Mmn|x - y|$$

where $M$ is chosen to be an upper bound on all the $D^j f(b)$ for all $i, j$ and $b \in A$.

**Proof.** Let $g(x) = f(x + t(y - x))$. For each $j$, the mean-value theorem says that there is a $c_j$ with $1 \leq c_j \leq 1$ such that

$$f^j(x) - f^j(y) = g^j(1) - g^j(0) = (g^j)'(c_j).$$

By the chain rule,

$$|(g^j)'(c_j)| = \left| \sum_{j=1}^{n} D_j f(x + c_j(y - x))(y^j - x^j) \right| \leq Mn|x - y|.$$ 

Hence

$$|f(x) - f(y)| \leq \sum_{j=1}^{m} |f^j(x) - f^j(y)| \leq Mmn|x - y|.$$ 

**The Proof of the Inverse Function Theorem:** By replacing $f$ with $(Df(a))^{-1}f$, we can assume that $Df(a)$ is the identity map (Why?).
Apply the Lemma above to \( g(x) = f(x) - x \), we get \(|f(x) - x - (f(y) - y)| \leq Mn^2|x - y|\) for \( x, y \) in an open rectangle containing \( a \). But then
\[
|x - y| - |f(x) - f(y)| \leq |f(x) - x - (f(y) - y)| \leq Mn^2|x - y|.
\]
Rearranging gives
\[
|x - y|(1 - Mn^2) \leq |f(x) - f(y)|.
\]
Since \( g(x) = f(x) - x \), \( g'(a) = 0 \) and so by choosing the rectangle small enough, we can assume that \( Mn^2 < 1/2 \) and so
\[
|x - y| \leq 2|f(x) - f(y)|.
\]
In particular, it follows that \( f \) is one-to-one when restricted to this rectangle, and the inverse will be continuous if it exists. Replacing the rectangle with a smaller one, we can assume the same is true when \( f \) is restricted to the closure of the rectangle. Now the boundary \( B \) of the rectangle is compact and so \( f(B) \) is also compact and does not contain \( f(a) \). Let \( d = \min_{x \in B} |f(x) - f(a)| \), then \( d > 0 \). Let \( V \) be the set \( V = \{ y \mid |y - f(a)| < d/2 \} \).

We now prove that \( f^{-1} \) exists on \( V \). For every \( y \in V \), we claim that there is at least one \( x \) in the rectangle with \( f(x) = y \). In fact, consider the function \( h(x) = |f(x) - y|^2 \). The image of \( h \) under the closure of the rectangle. The minimum of this function does not occur on \( B \) because \( y \in V \). So it must occur where the derivative is zero, i.e. one has:
\[
\sum_{i=1}^{n} -2(f^i(x) - y^i)D_j f^i(x) = 0
\]
for \( j = 1, \ldots, n \). But by taking the rectangle sufficiently small, we can assume that \( \det f'(x) \neq 0 \) for all \( x \) in the rectangle. But then the only solution of this system of linear equations is \( f(x) - y = 0 \). This proves the claim. If \( U = f^{-1}(V) \), then \( f \) maps the open set \( U \) one-to-one and onto the open set \( V \).

It remains to check differentiability of the inverse. Since \( f \) is differentiable, one has for \( x \in U \),
\[
f(y) - f(x) - f'(x)(y - x) = \phi(y - x)
\]
where \( \lim_{h \to 0} \frac{\phi(h)}{|h|} = 0 \). Letting \( u = f(x) \) and \( v = f(y) \), we get after substitution:
\[
v - u - f'(f^{-1}(u))(f^{-1}(v) - f^{-1}(u)) = \phi(y - x).
\]
Rearranging gives:

\[ f^{-1}(v) - f^{-1}(u) - (f'(f^{-1}(u)))^{-1}(v - u) = - (f'(f^{-1}(u)))^{-1} \phi(y - x). \]

It remains to show that

\[ \lim_{v \to u} \frac{|-(f'(f^{-1}(u)))^{-1} \phi(y - x)|}{|v - u|} = 0. \]

Since the derivative is just a linear function, it is enough to show that

\[ \lim_{v \to u} \frac{|\phi(y - x)|}{|v - u|} = 0. \]

As \( v \to u \), we have \( y \to x \) because \( f^{-1} \) is continuous. So \( \lim_{v \to u} \frac{|\phi(y - x)|}{|v - u|} = 0. \) But we know that \( \frac{|y - x|}{|v - u|} \leq 2 \). So, the limit of the product is zero, as desired.

Note: It follows from the formula for the derivative of the inverse that the inverse is also continuously differentiable.

**Implicit Function Theorem:** Suppose that \( f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m \) is continuously differentiable on an open set containing \((a, b)\) and \( f(a, b) = 0 \). Let \( M \) be the \( m \times m \) matrix

\[ (D_{n+j}f(a, b)) \quad 1 \leq i, j \leq m. \]

If \( \det M \neq 0 \), then there is an open subset \( A \subset \mathbb{R}^n \) containing \( a \) and an open subset \( B \subset \mathbb{R}^m \) containing \( b \) with the following property: for each \( x \in A \) there is a unique \( g(x) \in B \) such that \( f(x, g(x)) = 0 \). The function \( g \) is differentiable.

Proof. We can extend \( f \) to a function \( F : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m} \) by \( F(x, y) = (x, f(x, y)) \) and apply the Inverse Function Theorem to get an inverse \( G \) defined in an open subset containing \((a, 0)\) and mapping onto an open subset containing \((a, b)\). Let \( g(x) = \pi \circ G(x, 0) \) where \( \pi \) is the projection onto the last \( m \) coordinates of \( \mathbb{R}^{n+m} \). Then one has \( f(x, g(x)) = f(x, \pi \circ G(x, 0)) = (\pi \circ F)(G(x, 0)) = \pi(x, 0) = 0. \)

**Exercise 2:** Show that the Inverse Function Theorem is a Corollary of the Implicit Function Theorem.

**Special cases when** \( n = 1 \) **and** \( n = 2 \): Let \( f(x) \) be a one-variable function. If a function \( f \) is differentiable at \( x_0 \) then can write, by letting \( h = x - x_0 \),

\[ f(x_0 + h) = f(x_0) + f'(x_0)h + \phi(h) \]
where \( \lim_{h \to 0} \frac{\phi(h)}{h} = 0 \). In other words, up to an "error term", we have \( f(x_0 + h) \sim f(x_0) + f'(x_0)h \), i.e. we can use the linear function \( f(x_0) + f'(x_0)h \) to approximate the value \( f(x_0 + h) \) the the "error" is small as \( h \to 0 \) (or \( x \to x_0 \)). Geometrically, this means that we can use the straight line \( y = f(x_0) + f'(x_0)h \) to approach the graph of \( f \) near \((x_0, f(x_0))\). In other words, the "derivative property" (the linear approximation) of \( f \) are expected to tell us some information about the function itself. For example, we have (in calculus) that if \( f'(x_0) \neq 0 \), then \( f \) is locally (near \( x_0 \)) is one-to-one (since the linear function is always one-to-one).

**Theorem (Inverse function theorem):** Let \( f : (a, b) \to \mathbb{R} \) be a function. Let \( x_0 \in (a, b) \). Assume that \( f'(x_0) \neq 0 \), then there is a neighborhood \( W \) of \( x_0 \) and a neighborhood \( V \) of \( f(x_0) \) and a differentiable function \( g : V \to W \) such that \( f(g(y)) = y \) for every \( y \in V \) (i.e. \( g = f^{-1} \)).

It is important (which will help you to understand the (general) implicit function theorem below) to note that we can re-write the above formula as follows: let \( y_0 = f(x_0) \), and consider \( F(x, y) = y - f(x) \). Then \( F(x_0, y_0) = 0 \), and \( \frac{\partial f}{\partial x} = -f'(x_0) \). So \( F : U \subset \mathbb{R}^2 \to \mathbb{R} \) with rank of the matrix \( F'(x_0, y_0) \) as 1 (assume that \( f'(x_0) \neq 0 \)). Then the above Inverse function theorem can be restated as the following (special) implicit theorem:

**(Special) Implicit Function Theorem.** Assume that \( f'(x_0) \neq 0 \). Consider \( F(x, y) = y - f(x) \). Let \( y_0 = f(x_0) \) such that \( F(x_0, y_0) = 0 \). Then there exists a neighborhood \( V \) of \( x_0 \) and a neighborhood \( W \) of \( y_0 \) and a differentiable function \( g : W \to V \) such that \( F(g(y), y) = 0 \) for every \( y \in W \) (i.e \( y = f(g(y)) \)).

We can consider a more general function \( F \) of two variables (not necessarily \( F(x, y) = f(x) - y \)), and, in stead of considering \( F(x, y) = 0 \), we can consider \( F(x, y) = c \) where \( c \) is any real number. We have

**Implicit Function Theorem for functions of two variables.** Consider \( F(x, y) : U \subset \mathbb{R}^2 \to \mathbb{R} \). Let \( (x_0, y_0) \in U \) such that \( F(x_0, y_0) = c \). Assume that the rank of the matrix \( F'(x_0, y_0) \) is 1. Then there exists a neighborhood \( V \) of \( x_0 \) and a neighborhood \( W \) of \( y_0 \) and a differentiable function \( \phi : W \to V \) such that \( F(\phi(y), y) = c \) for all \( y \in W \).

Or, by switching \( x \) with \( y \), we re-state it as follows:
Implicit Function Theorem for functions of two variables. Consider $F(x, y) : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. Let $(x_0, y_0) \in U$ such that $F(x_0, y_0) = c$. Assume that the rank of the matrix $F'(x_0, y_0)$ is 1. Then there exists a neighborhood $V$ of $x_0$ and a neighborhood $W$ of $y_0$ and a differentiable function $\phi : V \rightarrow W$ such that $F(x, \phi(x)) = c$ for all $x \in V$.

For example, for $F(x, y) = x^2 + y^2$. Consider the equation $F(x, y) = 1$ (i.e. with $c = 1$). Take $x_0 = 1, y_0 = 0$. Then $\phi(y) = \sqrt{1 - y^2}$ satisfying $F(\phi(y), y) = c$ for every $y \in W = (-1, 1)$.

Part III (Chapter 3): Integration

In order to simplify the presentation, we will develop integration using the Riemann (or Jordan) approach rather than the more general theory of Lebesgue integration.

- Integration Over a Rectangle.

**Basic Definition** Let $A \subset \mathbb{R}^n$ be a rectangular, and $f : A \rightarrow \mathbb{R}$ be a bounded function. The integral (or integrability) of $\int_A f$ is defined, through the partition $P$ of $A$, by comparing the upper sums and lower sums of $f$ for $P$ (see Page 47 on the textbook). The lower integral (respectively upper integral of on is defined to be $L \int_A f = \sup_P \{L(f, P)\}$ (respectively $U \int_A f = \inf_P \{U(f, P)\}$). If $L \int_A f = U \int_A f$, then $f$ is called integrable over $A$ with integral equal to this common value. Another notation $\int_A f(x^1, \ldots, x^n)dx^1 \cdots dx^n$.

The following result is very useful to decide whether $f$ is integrable over $A$.

**Theorem 3-3** (Page 48) $f$ is integrable over $A$ iff for every $\epsilon > 0$, there is a partition $P$ of $A$ such that $U(f, P) - L(f, P) < \epsilon$.

**Evaluating integration.** The key to evaluate a double integral $\int \int_\Omega f(x, y)dxdy$ is to reduce it to repeated integrations one-variable integration which we learnt in Calculus I) as justified by the following theorem. To do so, you need to break the region $\Omega$ into several pieces of type I and type II regions.

**Fubini’s Theorem, see Pgae 58**. Let $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^m$ be closed rectangles and $f : A \times B \rightarrow \mathbb{R}$ be integrable. For $x \in A$, let $f_x : B \rightarrow \mathbb{R}$ be defined by $f_x(y) = f(x, y)$ and let $\mathcal{L} : A \rightarrow \mathbb{R}$ and $\mathcal{U} : B \rightarrow \mathbb{R}$ be defined by:

$$\mathcal{L}(x) = \mathcal{L} \int_B f_x,$$

$$\mathcal{U}(x) = \mathcal{U} \int_B f_x.$$
\[ \mathcal{U}(x) = U \int_B f_x. \]

Then \( \mathcal{L} \) and \( \mathcal{U} \) are both integrable on \( A \) and

\[
\int_{A \times B} f = \int_A \mathcal{L} = \int_A L \int_B f_x, \\
\int_{A \times B} f = \int_A \mathcal{U} = \int_A U \int_B f_x.
\]

In other words, the double integral can be calculated as either of two iterated integrals:

\[
\int_{A \times B} f = \int_A (L \int_B f(x, y) dy) dx = \int_A (U \int_B f(x, y) dy) dx.
\]

**Extending the Integral: Characteristic Functions.** We have only defined the integral over a rectangle. However, integral over other sets (bounded sets only) can be easily defined by using the characteristic function.

**Exercise:** State and prove a result which would make the following definition legitimate and reasonable.

**Definition:** If \( C \subset \mathbb{R}^n \), then its characteristic function \( \chi_C \) is defined by \( \chi_C(x) = 0 \) if \( x \notin C \) and \( \chi_C(x) = 1 \) if \( x \in C \). If \( C \subset R \) for some closed rectangle, then a bounded function \( f : R \to \mathbb{R} \) is said to be integrable with integral

\[
\int_C f := \int_R f \cdot \chi_C
\]

provided that this last quantity is defined.

**Criterion of integrability:** Let \( f : R \to \mathbb{R} \) be a bounded function defined on a closed rectangle \( R \). Then \( f \) is integrable if and only if the set of discontinuity points of \( f \) has measure zero (see Page 53, Theorem 3-8).

**Corollary:** Let \( C \subset \mathbb{R}^n \) be a bounded set. Then its characteristic function \( \chi_C : \mathbb{R}^n \to \mathbb{R} \) is integrable if and only if the boundary of \( C \) is of measure zero.

**Extending the Integral: Partitions of Unity:** The goal of this section is to define an integral for functions defined on open sets. This is based on a technical result:
Theorem (Partitions of the Unity) (see Theorem 3-11 on page 63). Let \( \mathcal{O} \) be an open cover of a subset \( A \subset \mathbb{R}^n \). There is a set \( \Phi \) of \( C^\infty \) function \( \phi \) defined in an open set containing \( A \) and satisfying:

1. For each \( x \in A \), one has \( 0 \leq \phi(x) \leq 1 \) and \( \sum_{\phi \in \Phi} \phi(x) = 1 \) where the sum is defined because for every such \( x \) there is an open set \( V \) containing \( x \) where there are but finitely many \( \phi \in \Phi \) such that \( \phi|_V \neq 0 \).

2. For each \( \phi \in \Phi \), there is a \( U \in \mathcal{O} \) such that \( \phi = 0 \) outside of some compact set contained in \( U \).

A set \( \Phi \) of functions satisfying condition (1) of the above Theorem is called a \( C^\infty \) partition of unity of \( A \). When both conditions are satisfied, we call \( \Phi \) a partition of unity of \( A \) subordinate to the cover \( \mathcal{O} \).

The result about the ”Partition of the Unity” is Extremely important, it allows to piece things together from local to global one! It is used in the theory of integration (see page 65): The integration \( \int_A f \) previously is only defined for a bounded function over a rectangular \( A \). By using the ”Partition of the Unit”, we can define it for general \( A \) by reducing it to the rectangular case (see Page 65). Note that the ”Partition of the Unity” is also useful in doing ”local-to-global” properties about ”manifolds”.

The prove of above theorem (Partition of the Unity) replies on the following lemma which concerns the existence of so-called the bump-function. See the appendix below about the bump function (also see Exercise 1-22 and 2-26):

**Lemma 1:** (i) If \( C \subset U \) is a compact subset of an open set \( U \). Then there is a compact subset \( D \) of \( U \) such that \( C \) is contained in the interior of \( D \).

(ii) If is \( C \) a compact subset of an open set \( U \), then there is a \( C^\infty \) function \( f : U \to \mathbb{R} \) such that \( f(x) > 0 \) on all of \( C \) and such that \( f = 0 \) outside of a compact subset of \( U \). In fact, \( f \) can be chosen so it maps into \([0, 1]\) and such that \( f(x) = 1 \) for all \( x \in C \).

**Definition (see Page 65):** An open cover \( \mathcal{O} \) of a set \( A \subset \mathbb{R}^n \) is called admissible if each element of \( \mathcal{O} \) is contained in \( A \). Let \( f : A \to \mathbb{R} \) be a function and \( \mathcal{O} \) be admissible for \( A \) (and so \( A \) is open). Then \( f \) is said to be integrable in the extended sense with integral

\[
\int_A f := \sum_{\phi \in \Phi} \int_A \phi \cdot f
\]
provided that the terms of the series are defined and $\sum_{\phi \in \Phi} \int_A \phi \cdot |f|$ is convergent.

By Theorem 3-12, this integral is independent of the choice of $O$ and $\Phi$.

- **The Change of Variables formula**: (See Page 67-72).

**APPENDIX: The bump function and the partition of unity**:

The Bump Functions:

**Proposition 1** (see exercise 2-25 on page 29). Define

$$f(x) = \begin{cases} 
  e^{-\frac{1}{x}} & \text{for } x > 0 \\
  0 & \text{for } x \leq 0.
\end{cases}$$

$f(x)$ is smooth. In particular, $f^{(n)}(0) = 0$ for all $n \geq 0$.

**Proof.** It is clear that $f(x)$ is smooth except at $x = 0$. So we just need to verify that $f^{(n)}(0) = 0$ for all $n \geq 0$. (Note that if we prove that $f^{(n)}(0) = 0$, then that implies already that $f^{(n-1)}(x)$ is continuous at $x = 0$. It is obvious that $f^{(n-1)}(x)$ is continuous away from $x = 0$. So if we prove all the derivatives exist, we automatically get that they are continuous, which is part of the definition of smoothness.)

First of all, we check that $f'(0) = 0$. It is clear that

$$\lim_{x \to 0^-} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^-} \frac{0}{x} = 0.$$

The other one-sided limit

$$\lim_{x \to 0^+} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^+} \frac{e^{-\frac{1}{x}}}{x} = \lim_{x \to 0^+} \frac{\left(\frac{1}{x}\right)^{\frac{1}{e^{\frac{1}{x}}}}} = \lim_{y \to \infty} \frac{y}{e^y} \lim_{y \to \infty} e^y = 0.$$

Here we’ve substituted $y = \frac{1}{x}$, and $\frac{e^y}{e^y}$ indicates the use of l’Hôpital’s rule. Since the left and right hand limits are equal, we have $f'(0) = 0$.

For higher derivatives, compute for $x > 0$, $f'(x) = \frac{1}{x^2} e^{-\frac{1}{x}}$. It is easy to show by induction that all higher derivatives have a similar form: for $x > 0$,

$$f^{(n)}(x) = P_n \left(\frac{1}{x}\right) e^{-\frac{1}{x}}$$
for $P_n$ a polynomial function.

**Proof.** We’ve checked that for $n = 1$, $x > 0$, $f'(x)$ is of the form $P_1(\frac{1}{x})e^{\frac{1}{x}}$ for $P_1(y) = y^2$. Now assume that $f^{(n)}(x) = P_n(\frac{1}{x}) e^{-\frac{1}{x}}$ for $x > 0$. Then compute

$$f^{(n+1)}(x) = P'_n(\frac{1}{x}) \left(-\frac{1}{x^2}\right) e^{-\frac{1}{x}} + P_n(\frac{1}{x}) e^{-\frac{1}{x}} \left(\frac{1}{x^2}\right)$$

for $P_{n+1}(y) = -y^2P'_n(y) + y^2P_n(y)$, which is a polynomial whenever $P_n$ is. So we’ve checked that $f'(x) = f^{(1)}(x)$ satisfies the hypothesis, and also that if the hypothesis is true for $f^{(n)}(x)$, then it’s true for $f^{(n+1)}(x)$. By induction, the hypothesis is true for all positive integers $n$. This finishes the proof.

We use induction again to show that $f^{(n)}(0) = 0$. We’ve checked the case $n = 1$ above. Now assume that $f^{(n)}(0) = 0$. To compute $f^{(n+1)}(0)$, the left-hand limit is 0 as above, and the right-hand limit

$$\lim_{x \to 0^+} \frac{f^{(n)}(x) - f^{(n)}(0)}{x - 0} = \lim_{x \to 0^+} \frac{f^{(n)}(x) - 0}{x} = \lim_{x \to 0^+} \frac{1}{x} \int P_n \left(\frac{1}{x}\right) e^{-\frac{1}{x}} = \lim_{x \to 0^+} \frac{\tilde{P}_n(\frac{1}{x})}{e^{\frac{1}{x}}} = \lim_{y \to \infty} \frac{\tilde{P}_n(y)}{e^y} = 0.$$

Here $y = \frac{1}{x}$, and $\tilde{P}_n(y) = yP_n(y)$ is another polynomial. The last limit is a standard fact that can be proved by l’Hôpital’s rule and induction. Since the left and right hand limits are both 0, we have $f^{(n+1)}(0) = 0$, which completes the induction. This finishes the proof of Proposition 1.

**Proposition 2** (2-26 on page 29 of the textbook). Given $a < b$, there is a smooth function $h(x)$ so that $h(x) = 0$ for $x \leq a$, $0 < h(x) < 1$ for $x \in (a, b)$, and $h(x) = 1$ for $x \geq b$.

**Proof.** First let $g(x) = f(x - a)f(b - x)$ for $f$ defined above. It is clear $g$ is smooth, nonnegative, and $g(x) > 0 \iff x \in (a, b)$. Then define

$$h(x) = \frac{\int^x_{-\infty} g(t) \, dt}{\int_{-\infty}^{\infty} g(t) \, dt}.$$
It is easy to check that $h(x)$ has the relevant properties.

**Theorem (existence of bump function)** (2-26 on page 29 of the textbook). *Let $0 < a < b$. Then there is a smooth real-valued function $\lambda(x)$ on $\mathbb{R}$ so that $\lambda(x) = 1$ if $|x| \leq a$, $0 < \lambda(x) < 1$ for $x \in (a,b)$, and $\lambda(x) = 0$ for $|x| \geq b$.*

**Proof.** Let $\lambda(x) = 1 - h(x)$ for $h(x)$ above. Then it is easy to it satisfies the relevant properties. In particular, the chain rule shows that $\lambda$ is smooth except possibly at $x = 0$. But in a neighborhood of $x = 0$, $\lambda(x)$ is identically 1, so it is smooth there as well.

Notice that the graph of $\lambda(z)$ looks like a bump. It is called a **bump function.**