SOME BASICS OF DIFFERENTIAL CALCULUS AND COMPLEX ANALYSIS

BRUCE K. DRIVER

Abstract. These are lecture notes from Math 240C on the basics of complex analysis.

CONTENTS

1. Calculus in Banach Spaces 1
   1.1. Basic Notation 1
   1.2. The Differential 3
   1.3. Product and Chain Rules 4
   1.4. The Riemann Integral 7
2. Contraction Mapping Principle 9
3. Inverse and Implicit Function Theorems 10
   3.1. Inverse Function Theorem 12
   3.2. Implicit Function Theorem 13
4. Basic Facts About Complex Numbers 14
5. Complex Differentiable Functions 14
   5.1. Contour integrals 19
   5.2. Weak characterizations of \( H(\Omega) \) 24
5.3. Homework #2 Due Friday April 20, 2001 27

1. CALCUlus IN BA Nach SPACES

1.1. Basic Notation. Let \( (X, \| \cdot \|_X) \) and \( (Y, \| \cdot \|_Y) \) be a Banach spaces\(^1\) Write \( U \subset_o X, U \subset X, \) and \( U \subset \subset X \) to denote the \( U \) is a open, close, or compact subset of \( X \) respectively. Let \( B_X(x, \epsilon) \equiv \{x' \in X | \|x' - x\| < \epsilon\} \) denote the open ball about \( x \) of radius \( \epsilon \). Again we will often write \( B(x, \epsilon) \) rather than \( B_X(x, \epsilon) \).

Notation 1.1 (\( \epsilon, O, \) and \( o \) notation). Let \( 0 \in U \subset_o X, \) and \( f : U \to Y \) be a function. We will write:

1. \( f(x) = \epsilon(x) \) if \( \lim_{\epsilon \to 0} \|f(x)\| = 0. \)

\(^1\)In the future we will often denote the norm on a Banach space simply by \( \| \cdot \| \) rather than \( \| \cdot \|_X \) or \( \| \cdot \|_Y \).

Date: April 9, 2001 File:COMPLEX2.tex.
Department of Mathematics, 0112.
University of California, San Diego.
La Jolla, CA 92093-0112.
2. $f(x) = O(x)$ if there are constants $C < \infty$ and $r > 0$ such that
\[ \|f(x)\| \leq C\|x\| \text{ for all } x \in B(0, r). \]
This is equivalent to the condition that
\[ \limsup_{x \to 0} \frac{\|f(x)\|}{\|x\|} < \infty, \]
where
\[ \limsup_{x \to 0} \frac{\|f(x)\|}{\|x\|} = \limsup_{r \to 0} \{\|f(x)\| : 0 < \|x\| \leq r\}. \]

3. $f(x) = o(x)$ if $f(x) = \epsilon(x)O(x)$, i.e. $\lim_{x \to 0} \|f(x)\|/\|x\| = 0$.

**Definition 1.2.** A function $f: U \subset \mathbb{R} \to Y$ is continuous at $x_0 \in U$ if $f(x) - f(x_0) = \epsilon(x - x_0)$ and $f$ is continuous on $U$ if $f$ is continuous at all points in $U$.

**Definition 1.3.** Let $L(X, Y)$ denote the set of linear operators $\Lambda : X \to Y$ which are continuous. It is well known that the linear space $L(X, Y)$ is also a Banach space with the “operator norm,”
\[ \|\Lambda\|_{L(X, Y)} = \sup_{x \neq 0} \frac{\|\Lambda x\|_Y}{\|x\|_X}. \]

Recall that if $Z$ is another Banach space and $\Gamma \in L(Y, Z)$ then $\Gamma \Lambda \in L(X, Z)$ and
\[ (1.1) \quad \|\Gamma \Lambda\| \leq \|\Gamma\| \|\Lambda\|. \]

**Proposition 1.4.** Let and $L^*(X, Y)$ denote those linear operators $\Lambda \in L(X, Y)$ which are also invertible. We will write the inverse operator as $\Lambda^{-1}$. (Recall that the open mapping theorem asserts that $\Lambda^{-1}$ is also a bounded operator, i.e. $\Lambda^{-1} \in L^*(Y, X)$ as well. Then $L^*(X, Y)$ is an open (possibly empty) subset of $L(X, Y)$. Moreover if $A \in L^*(X, Y)$ and $B \in L(X, Y)$ such that
\[ (1.2) \quad \|A - B\| < \|A^{-1}\| \]
then $B \in L^*(Y, X)$ and
\[ (1.3) \quad B^{-1} = \sum_{n=0}^{\infty} (I_X - A^{-1}B)\]
where the sum is convergent in $L^*(Y, X)$. (Here $I_X$ denotes the identity operator on $X$.)

Let us start first with the following special case.

**Lemma 1.5.** Suppose $E \in L(X)$ such that $\sum_{n=0}^{\infty} \|E^n\| < \infty$. Then $I - E \in L^*(X)$ and $(I - E)^{-1} = \sum_{n=0}^{\infty} E^n$, where sum converges in the operator norm. In particular, if $\|E\| < 1$, then $I - E \in L^*(X)$, i.e. $B_{L(X)}(I, 1) \subset L^*(X)$.

**Proof.** Define $C_N = \sum_{n=0}^{N} E^n \in L(X)$, Now for $M > N$,
\[ \|C_M - C_N\| = \| \sum_{n=N}^{M-1} E^n \| \leq \sum_{n=N}^{M-1} \|E^n\| \to 0 \text{ as } M, N \to \infty. \]
Therefore $C = \sum_{n=0}^{\infty} E^n \equiv \lim_{N \to \infty} C_N$ exists. Now clearly
\[ (I - E)C_N = C_N(I - E) = \sum_{n=0}^{N} E^n - \sum_{n=1}^{N+1} E^n = I - E^{N+1}. \]
Since $\|E^{N+1}\| \to 0$ as $N \to \infty$, we may let $N \to \infty$ in the last equation to find that $C(I - E) = (I - E)C = I$. □
Proof. Proof of Proposition 1.4. Let \( E \equiv A - B \), so \( B = A - E \). Working formally for the moment,

\[
B^{-1} = \frac{1}{B} = \frac{1}{A - E} = \frac{1}{A(I - A^{-1}E)} = (I - A^{-1}E)^{-1}A^{-1}.
\]

(The reason \( \frac{1}{A - E} \) was interpreted as \( (I - A^{-1}E)^{-1}A^{-1} \) rather than \( A^{-1}(I - A^{-1}E)^{-1} \) was to make sure the operator was a function from \( Y \to X \) rather than from \( X \to Y \).) Notice that if \( (I + A^{-1}E) \) is invertible then

\[
(I - A^{-1}E)^{-1}A^{-1} = [A(I - A^{-1}E)]^{-1} = [A - E]^{-1},
\]

so that \( B = A - E \) is invertible as well. Now by Lemma 1.5 above, we know that \( I - A^{-1}E \) is invertible if \( \|A^{-1}E\| < 1 \). Since \( \|A^{-1}E\| \leq \|A^{-1}\|\|E\| \) it suffices to assume that \( \|E\| < 1/\|A^{-1}\| \). In this case we also have by Lemma 1.5 that

\[
(I - A^{-1}E)^{-1} = \sum_{n=0}^{\infty} (A^{-1}E)^n.
\]

This equation combined with Eq. (1.4) and the fact that \( A^{-1}E = I - A^{-1}B \), shows that \( B^{-1} \) exists and is given by Eq. (1.3). \( \blacksquare \)

1.2. The Differential.

Definition 1.6. A function \( f : U \subset \subset \rightarrow Y \) is differentiable at \( x_0 \in U \) if there exists a linear transformation \( \Lambda \in L(X,Y) \) such that

\[
f(x) - f(x_0) - \Lambda(x-x_0) = o(x-x_0).
\]

As for continuity, \( f \) is differentiable on \( U \) if \( f \) is differentiable at all points in \( U \).

Remark 1.7. The linear transformation \( \Lambda \) in Definition 1.6 is necessarily unique. Indeed if \( \Lambda_1 \) is another linear transformation such that Eq. (1.5) holds with \( \Lambda \) replaced by \( \Lambda_1 \), then

\[
(\Lambda - \Lambda_1)(x-x_0) = o(x-x_0),
\]

i.e.

\[
\limsup_{x \to x_0} \frac{\| (\Lambda - \Lambda_1)(x-x_0) \|}{\| (x-x_0) \|} = 0.
\]

On the other hand, by definition of the operator norm,

\[
\limsup_{x \to x_0} \frac{\| (\Lambda - \Lambda_1)(x-x_0) \|}{\| (x-x_0) \|} = \| \Lambda - \Lambda_1 \|.
\]

The last two equations show that \( \Lambda = \Lambda_1 \).

If \( f : U \subset X \to Y \) is differentiable at \( x_0 \), we will denote the linear operator \( \Lambda \) in Eq. (1.5) by \( Df(x_0) \) of \( f'(x_0) \). \( Df(x_0) = f'(x_0) \) is called the differential of \( f \) at \( x_0 \).

Example 1.8. Assume that \( L^*(X,Y) \) is non-empty. Then \( f : L^*(X,Y) \to L^*(Y,X) \) defined by \( f(A) \equiv A^{-1} \) is differentiable and

\[
f'(A)B = -A^{-1}BA^{-1}.
\]
Indeed by Eq. (1.3),
\[
f(A + \mathcal{E}) - f(A) = \sum_{n=0}^{\infty} (-A^{-1}\mathcal{E})^n : A^{-1} - A^{-1} = -A^{-1}\mathcal{E}A^{-1} + \sum_{n=2}^{\infty} (-A^{-1}\mathcal{E})^n.\]

Since
\[
\|\sum_{n=2}^{\infty} (-A^{-1}\mathcal{E})^n\| \leq \sum_{n=2}^{\infty} \|A^{-1}\mathcal{E}\|^n \leq \frac{\|A^{-1}\|^2 \|\mathcal{E}\|^2}{1 - \|A^{-1}\mathcal{E}\|},
\]
we find that
\[
f(A + \mathcal{E}) - f(A) = -A^{-1}\mathcal{E}A^{-1} + o(\mathcal{E}).
\]

**Example 1.9.** Let \( X \equiv C([a, b] \to \mathbb{R}) \) the continuous functions from \([a, b]\) to \(\mathbb{R}\) with \(\| x \| \equiv \max_{t \in [a, b]} |x(t)|\) for all \(x \in X\). Suppose that \( f : \mathbb{R} \to \mathbb{R} \) is \(C^1\) and that \( F(x) \equiv \int_{a}^{b} g(x(t)) \, dt \). Then \( F : X \to \mathbb{R} \) is differentiable and \( F'(x)y = \int_{a}^{b} g'(x(t))y(t) \, dt \) \(= \Lambda y \).

To verify this last example, notice using the fundamental theorem of calculus,
\[
F(x + y) - F(x) - \Lambda y = \int_{a}^{b} [g(x(t) + y(t)) - g(x(t)) - g'(x(t))y(t)] \, dt
\]
\[
= \int_{a}^{b} \left( \int_{0}^{1} [g'(x(t) + ry(t)) - g'(x(t))] \, dr \right) y(t) \, dt.
\]

Therefore, \(|F(x + y) - F(x) - \Lambda y| \leq \|y\| \int_{a}^{b} \left( \int_{0}^{1} |g'(x(t) + ry(t)) - g'(x(t))| \, dr \right) dt\), which shows that \( F(x + y) - F(x) - \Lambda y = o(y) \) because of the dominated convergence theorem (or using uniform convergence),
\[
\int_{0}^{1} |g'(x(t) + ry(t)) - g'(x(t))| \, dr = \epsilon(y).
\]

**Exercise 1.10.** A function \( \sigma : (a, b) \to X \) is a differentiable at \( t \in (a, b) \) iff \( L := \lim_{h \to 0} \frac{\sigma(t + h) - \sigma(t)}{h} \) exists in \( X \). Moreover, \( D\sigma(t)a = La \) for all \( a \in \mathbb{R} \).

In the future, we will write \( L \) as \( \dot{\sigma}(t) \) or \( d\sigma(t)/dt \) and we will identify \( D\sigma(t) \) with \( \dot{\sigma}(t) \) in this context.

1.3. **Product and Chain Rules.** The following theorem summarizes some basic properties of the differential.

**Theorem 1.11.** The differential \( D \) has the following properties:

**Linearity:** \( D \) is linear.

**Product Rule:** If \( f : U \subset \mathbb{R} \to X \) and \( A : U \subset \mathbb{R} \to L(X, Z) \) are differentiable at \( x_0 \) then so is \( x \to (Af)(x) \equiv A(x)f(x) \) and \( D(Af)(x_0)x' = (DA(x_0)x')f(x_0) + A(x_0)Df(x_0)x' \).

**Chain Rule:** If \( f : U \subset \mathbb{R} \to V \subset \mathbb{R} \) is differentiable at \( x_0 \in U \), and \( g : V \subset \mathbb{R} \to Z \) is differentiable at \( y_0 \equiv f(x_0) \), then \( g \circ f \) is differentiable at \( x_0 \) and \( (g \circ f)'(x_0) = g'(y_0)f'(x_0) \).
Converse Chain Rule: Suppose that $f : U \subset_o X \to V \subset_o Y$ is continuous at $x_0 \in U$, $g : V \subset_o Y \to Z$ is differentiable $g(y_0) \equiv f(x_0)$, $g'(y_0)$ is invertible, and $g \circ f$ is differentiable at $x_0$, then $f$ is differentiable at $x_0$ and

\[
f'(x_0) \equiv [g'(y_0)]^{-1}(g \circ f)'(x_0).
\]

Proof. For the proof of linearity, let $f, g : U \subset_o X \to Y$ be two functions which are differentiable at $x_0 \in U$ and $c \in \mathbb{R}$, then

\[
(f + cg)(x_0 + x) = f(x_0) + Df(x_0)x + o(x) + c(g(x_0) + Dg(x_0)x + o(x))
\]

\[
= (f + cg)(x_0) + (Df(x_0) + cDg(x_0))x + o(x),
\]

which implies that $(f + cg)$ is differentiable at $x_0$ and that

\[
D(f + cg)(x_0) = Df(x_0) + cDg(x_0).
\]

For item 2, we have

\[
A(x_0 + x)f(x_0 + x) = (A(x_0) + DA(x_0)x + o(x))(f(x_0) + f'(x_0)x + o(x))
\]

\[
= A(x_0)f(x_0) + A(x_0)f'(x_0)x + [DA(x_0)x]f(x_0) + o(x),
\]

which proves item 2.

Similarly for item 3,

\[
(g \circ f)(x_0 + x) = g(f(x_0)) + g'(f(x_0))(f(x_0 + x) - f(x_0)) + o(f(x_0 + x) - f(x_0))
\]

\[
= g(f(x_0)) + g'(f(x_0))(Df(x_0)x + o(x)) + o(f(x_0 + x) - f(x_0))
\]

\[
= g(f(x_0)) + g'(f(x_0))Df(x_0)x + o(x),
\]

where in the last line we have used the fact that $f(x_0 + x) - f(x_0) = O(x)$ (see Eq. (1.5)) and $o(O(x)) = o(x)$.

Item 4. Since $g$ is differentiable at $y_0 = f(x_0)$,

\[
g(f(x_0 + x)) - g(f(x_0)) = g'(f(x_0))(f(x_0 + x) - f(x_0)) + o(f(x_0 + x) - f(x_0)).
\]

And since $g \circ f$ is differentiable at $x_0$,

\[
(g \circ f)(x_0 + x) - g(f(x_0)) = (g \circ f)'(x_0)x + o(x).
\]

Comparing these two equations shows that

\[
f(x_0 + x) - f(x_0) = g'(f(x_0))^{-1}\{(g \circ f)'(x_0)x + o(x) - o(f(x_0 + x) - f(x_0))\}
\]

\[
= g'(f(x_0))^{-1}(g \circ f)'(x_0)x + o(x)
\]

\[
- g'(f(x_0))^{-1}o(f(x_0 + x) - f(x_0)).
\]

Using the continuity of $f$, $f(x_0 + x) - f(x_0)$ is close to $0$ if $x$ is close to zero, and hence $\|o(f(x_0 + x) - f(x_0))\| \leq \frac{1}{2}\|f(x_0 + x) - f(x_0)\|$ for all $x$ sufficiently close to 0. (We may replace $\frac{1}{2}$ by any number $\alpha > 0$ above.) Using this remark, we may take the norm of both sides of equation (1.7) to find

\[
\|f(x_0 + x) - f(x_0)\| \leq \|g'(f(x_0))^{-1}(g \circ f)'(x_0)\||x|| + o(x) + \frac{1}{2}\|f(x_0 + x) - f(x_0)\|
\]

for $x$ close to 0. Solving for $\|f(x_0 + x) - f(x_0)\|$ in this last equation shows that

\[
f(x_0 + x) - f(x_0) = O(x).
\]
(This is an improvement, since the continuity of \( f \) only guaranteed that \( f(x_0 + x) - f(x_0) = \epsilon(x) \). Because of Eq. (1.8), we now know that \( o(f(x_0 + x) - f(x_0)) = o(x) \), which combined with Eq. (1.7) shows that

\[
f(x_0 + x) - f(x_0) = g'(f(x_0))^{-1}(g \circ f)'(x_0)x + o(x),
\]

i.e. \( f \) is differentiable at \( x_0 \) and \( f'(x_0) = g'(f(x_0))^{-1}(g \circ f)'(x_0) \). □

**Corollary 1.12.** Suppose that \( \sigma : (a, b) \to U \subset \sigma X \) is differentiable at \( t \in (a, b) \) and \( f : U \subset \sigma X \to Y \) is differentiable at \( \sigma(t) \in U \). Then \( f \circ \sigma \) is differentiable at \( t \) and

\[
d(f \circ \sigma)(t)/dt = f'(\sigma(t))\sigma'(t).
\]

**Definition 1.13** (Partial of Directional Derivative). Let \( f : U \subset \sigma X \to Y \) be a function, \( x_0 \in U \), and \( v \in X \). We say that \( f \) is differentiable at \( x_0 \) in the direction \( v \) iff \( \frac{d}{d\theta} f(x_0 + \theta v) = (\partial_v f)(x_0) \) exists. We call \( (\partial_v f)(x_0) \) the directional or partial derivative of \( f \) at \( x_0 \) in the direction \( v \).

Notice that if \( f \) is differentiable at \( x_0 \), then \( \partial_v f(x_0) \) exists and is equal to \( f'(x_0)v \), see Corollary 1.12.

**Theorem 1.14** (Mean Value Inequality). Suppose that \( f : [a, b] \to X \) is a continuous function such that \( \dot{f}(t) \) exists and is continuous in \( t \) for \( t \in (a, b) \). Then

\[
\|f(b) - f(a)\| \leq \int_a^b \|\dot{f}(t)\|dt \leq (b - a) \cdot \sup \left\{ \|\dot{f}(t)\| : a < t < b \right\}.
\]

**Proof.** We will give two proofs of this theorem. The first one use the Hahn Banach theorem while the second does not.

**First Proof.** Suppose that \( \lambda \in X^* \equiv L(X, \mathbb{R}) \). Then it is easily checked that \( \lambda(f(t)) \) is differentiable in \( t \) and the \( \frac{d}{dt} \lambda(f(t)) = \lambda(\dot{f}(t)) \). So by the fundamental theorem of calculus for real valued functions,

\[
\lambda(f(\beta) - f(\alpha)) = \lambda(f(\beta)) - \lambda(f(\alpha)) = \int_\alpha^\beta \lambda(\dot{f}(t))dt
\]

for all \( a < \alpha < \beta < b \). Taking absolute values of this last equation gives the estimate:

\[
|\lambda(f(\beta) - f(\alpha))| \leq \int_\alpha^\beta |\lambda(\dot{f}(t))|dt \leq \int_\alpha^\beta ||\lambda|||\dot{f}(t)||dt \leq ||\lambda|| \int_\alpha^\beta ||\dot{f}(t)||dt.
\]

Using the continuity of \( f \), we may let \( \alpha \downarrow a \) and \( \beta \uparrow b \) to find:

\[
\frac{|\lambda(f(b) - f(a))|}{||\lambda||} \leq \int_a^b ||\dot{f}(t)||dt.
\]

As a consequence of the Hahn Banach theorem,

\[
\sup_{\lambda \in X^*, \lambda \neq 0} \frac{|\lambda(f(b) - f(a))|}{||\lambda||} = ||f(b) - f(a)||
\]

which combined with the previous equation proves Eq. (1.9).

**Second Proof.** By a similar continuity argument used in the first proof it suffices to show that

\[
|f(\beta) - f(\alpha)| \leq \int_\alpha^\beta ||\dot{f}(t)||dt + \epsilon(\beta - \alpha)
\]
for any $\epsilon > 0$ and $a < \alpha < \beta < b$. Let $A$ denote the set of $T \in [a, \beta]$ such that

$$
\|f(T) - f(\alpha)\| \leq \int_a^T \|\dot{f}(t)\|dt + \epsilon(T - \alpha).
$$

Notice that $A$ is closed and non-empty since $\alpha \in A$. Let $t_0$ be the least upper bound for $A$. Since $A$ is closed $t_0 \in A$, i.e.

$$
\|f(t_0) - f(\alpha)\| \leq \int_a^{t_0} \|\dot{f}(t)\|dt + \epsilon(t_0 - \alpha).
$$

For sake of contradiction, suppose that $t_0 < \beta$. Then because $f$ is differentiable at $t_0$,

$$
\|f(t) - f(t_0)\| \leq \|\dot{f}(t_0)\|(t - t_0) + \frac{\epsilon}{2}(t - t_0),
$$

for all $t > t_0$ sufficiently close to $t_0$. Also the continuity of $\|\dot{f}\|$ implies that

$$
\left| \int_{t_0}^t \|\dot{f}(\tau)\|d\tau - \|\dot{f}(t_0)\|(t - t_0) \right| \leq \int_{t_0}^t \|\dot{f}(\tau)\|d\tau - \|\dot{f}(t_0)\|d\tau \leq \frac{\epsilon}{2}(t - t_0)
$$

for all $t > t_0$ sufficiently close to $t_0$. It now easily follows from Eq. (1.12) and (1.13) that

$$
\|f(t) - f(t_0)\| \leq \int_{t_0}^t \|\dot{f}(\tau)\|d\tau + \epsilon(t - t_0)
$$

for all $t > t_0$ sufficiently close to $t_0$. Using this equation, Eq. (1.11) and the triangle inequality, one shows that $\|f(t) - f(\alpha)\| \leq \int_a^t \|\dot{f}(\tau)\|d\tau + \epsilon(t - \alpha)$ for all $t > t_0$ sufficiently close to $t_0$. But this implies there exists $t \in A$ for which $t > t_0$ which contradicts the definition of $t_0$. ■

We have the following easy but useful corollary.

**Corollary 1.15.** Suppose that $f : [a, b] \rightarrow X$ is a continuous function such that $\dot{f}(t)$ exists and is equal to zero for $t \in (a, b)$. Then $f$ is a constant function.

1.4. The Riemann Integral. Let $J = [a, b]$ be a fixed interval. A function $S : J \rightarrow X$ is a step function of $S$ may be written in the form

$$
S(t) = x_0 1_{[a, t_1]}(t) + \sum_{i=1}^{n-1} x_i 1_{[t_i, t_{i+1}]}(t),
$$

where $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ and $x_i \in X$. Let $S$ denote the collection of step functions from $J$ to $X$. For $S$ as in Eq. (1.14), let

$$
I(S) = \sum_{i=0}^{n-1} (t_{i+1} - t_i)x_i.
$$

**Exercise 1.16.** Show that $I(S)$ is well defined, independent of how $S$ is represented as a step function. Also verify that $I : S \rightarrow X$ is a linear operator.

Taking the norm of Eq. (1.15) and using the triangle inequality we find,

$$
\|I(S)\| \leq \sum_{i=0}^{n-1} (t_{i+1} - t_i)\|x_i\| = \int_a^b \|S(t)\|dt \leq (b - a) \max_{t \in J} \|S(t)\|.
$$
Exercise 1.17. Let $\mathcal{F} \equiv \{ f : J \rightarrow X : \| f \|_\mathcal{F} \equiv \sup_{t \in [a,b]} \| f(t) \| < \infty \}$. Show that $(\mathcal{F}, \| \cdot \|_\mathcal{F})$ is a complete Banach space. Notice that $\overline{\mathcal{S}}$ denote the closure of $\mathcal{S}$ in $\mathcal{F}$. Also show that

$$C(J,X) \equiv \{ f : J \rightarrow X | f \text{ is continuous} \} \subset \overline{\mathcal{S}}.$$

Proposition 1.18. The linear function $I : \mathcal{S} \rightarrow X$ extends uniquely to a continuous linear map $\overline{I}$ from $\overline{\mathcal{S}}$ to $X$. Moreover, for all $f \in \overline{\mathcal{S}}$

$$\| \overline{I}(f) \| \leq \int_a^b \| f(t) \| \, dt \leq \| f \|_\mathcal{F}$$

Proof. (Sketch.) Let $f \in \overline{\mathcal{S}}$, and choose $S_n \in \mathcal{S}$ such that $\lim_{n \to \infty} \| f - S_n \|_\mathcal{F} = 0$. Notice that

$$\| I(S_n) - I(S_m) \| = \| I(S_n - S_m) \| \leq (b - a) \| S_n - S_m \|_\mathcal{F} \to 0 \text{ as } m, n \to \infty.$$  

Therefore, $\lim_{n \to \infty} I(S_n)$ exists. Moreover if $S'_n \in \mathcal{S}$ is another sequence such that $\lim_{n \to \infty} \| f - S'_n \|_\mathcal{F} = 0$, then

$$\| I(S'_n) - I(S_n) \| \leq (b - a) \| S'_n - S_n \|_\mathcal{F} \to 0 \text{ as } n \to \infty.$$  

Thus $\overline{I}(f) \equiv \lim_{n \to \infty} I(S_n)$ is well defined and is clearly the only possible continuous extension of $I$. The remainder of the proof, namely that $\overline{I}$ is linear and that the inequality in Eq. (1.17) holds, is left to the reader.

Notice that if $f \in C(J,X)$ and $\pi \equiv \{a = t_0 < t_1 < \cdots < t_n = b \}$ is a partition of $J$, $c_i \in [t_i, t_{i+1}]$ for $i = 0, 1, 2, \ldots, n - 1$ are chosen arbitrarily, then

$$f_\pi(t) \equiv f(c_0)I_{[a,c_1]}(t) + \sum_{i=1}^{n-1} f(c_i)I_{(t_i,t_{i+1}]}(t).$$

Moreover, by uniform continuity, $\| f - f_\pi \|_\mathcal{F} \to 0$ as $|\pi| \equiv \max\{|t_{i+1} - t_i| : i = 0, 1, 2, \ldots, n - 1 \} \to 0$. Therefore,

$$\overline{I}(f) = \lim_{|\pi| \to 0} I(f_\pi) = \lim_{|\pi| \to 0} \sum_{i=0}^{n-1} f(c_i)(t_{i+1} - t_i),$$

i.e. $\overline{I}(f)$ is a limit of Riemann sums. In the future, we will now denote $\overline{I}(f)$ by $\int_a^b f(t) \, dt$. Moreover, for $f \in \overline{\mathcal{S}}$ and $\alpha, \beta \in J$ such that $\alpha < \beta$, write $\int_\alpha^\beta f(t) \, dt$ for $\overline{I}(1_{[\alpha,\beta]} f)$. You should check that $1_{[\alpha,\beta]} f \in \overline{\mathcal{S}}$ provided that $f \in \overline{\mathcal{S}}$.

Theorem 1.19 (Fundamental Theorem of Calculus). Suppose that $f \in C(J,X)$. Then

1. $\frac{d}{dt} \int_a^t f(\tau) \, d\tau = f(t)$ for all $t \in (a,b)$.
2. Now assume that $F \in C(J,X)$, $F$ is continuously differentiable on $(a,b)$, and $F$ extends to a continuous function on $[a,b]$ which is still denoted by $F$. Then

$$\int_a^b \dot{F}(t) \, dt = F(b) - F(a).$$
Proof. (Sketch.) Let \( h > 0 \) be a small number and consider
\[
\| \int_a^{t+h} f(\tau)d\tau - \int_a^t f(\tau)d\tau - f(t)h \| = \| \int_t^{t+h} (f(\tau) - f(t))d\tau \|
\]
\[
\leq \int_t^{t+h} \| (f(\tau) - f(t)) \| d\tau
\]
\[
\leq h \epsilon(h),
\]
where \( \epsilon(h) \equiv \max_{\tau \in [t, t+h]} \| (f(\tau) - f(t)) \| \). Combining this with a similar computation when \( h < 0 \) shows, for all \( h \in \mathbb{R} \) sufficiently small, that
\[
\| \int_a^{t+h} f(\tau)d\tau - \int_a^t f(\tau)d\tau - f(t)h \| = |h| \epsilon(h),
\]
where now \( \epsilon(h) \equiv \max_{\tau \in [t-h, t]} \| (f(\tau) - f(t)) \| \). By continuity of \( f \) at \( t \), \( \epsilon(h) \to 0 \) and hence \( \frac{d}{dt} \int_a^t f(\tau)d\tau \) exists and is equal to \( f(t) \).

For the second item, set \( G(t) \equiv \int_a^t \hat{F}(\tau)d\tau - F(t) \). By what we have just proved, \( \hat{G}(t) = 0 \) for all \( t \in (a, b) \). You may also easily verify that \( G \) is continuous on \( J \). Hence we may apply Corollary 1.15 to conclude that \( G \) is a constant. In particular \( G(b) = G(a) \), i.e. \( \int_a^b \hat{F}(\tau)d\tau - F(b) = -F(a) \).

2. Contraction Mapping Principle

Theorem 2.1. Suppose that \((X, \rho)\) is a complete metric space and \( S : X \to X \) is a contraction, i.e. there exists \( \alpha \in (0, 1) \) such that \( \rho(S(x), S(y)) \leq \alpha \rho(x, y) \) for all \( x, y \in X \). Then \( S \) has a unique fixed point in \( X \), i.e. there exists a unique point \( x \in X \) such that \( S(x) = x \).

Proof. For uniqueness suppose that \( x \) and \( x' \) are two fixed points of \( S \), then
\[
\rho(x, x') = \rho(S(x), S(x')) \leq \alpha \rho(x, x').
\]
Therefore \((1 - \alpha)\rho(x, x') \leq 0\) which implies that \( \rho(x, x') = 0 \) since \( 1 - \alpha > 0 \). Thus \( x = x' \).

For existence, let \( x_0 \in X \) be any fixed point in \( X \) and define \( x_n \in X \) inductively by \( x_{n+1} = S(x_n) \) for \( n \geq 0 \). We will show that \( x \equiv \lim_{n \to \infty} x_n \) exists in \( X \). Assuming this for the moment, using the continuity of \( S \),
\[
x = \lim_{n \to \infty} x_{n+1} = \lim_{n \to \infty} S(x_n) = S( \lim_{n \to \infty} x_n ) = S(x),
\]
which shows that \( x \) is a fixed point of \( S \). So to finish the proof, because \( X \) is complete, it suffices to show that \( \{x_n\}_{n=1}^\infty \) is a Cauchy sequence in \( X \).

An easy inductive computation shows, for \( n \geq 0 \), that
\[
\rho(x_{n+1}, x_n) = \rho(S(x_n), S(x_{n-1})) \leq \alpha \rho(x_n, x_{n-1}) \leq \cdots \leq \alpha^n \rho(x_1, x_0).
\]
Another inductive argument using the triangle inequality shows, for \( m > n \), that,
\[
\rho(x_m, x_n) \leq \rho(x_m, x_{m-1}) + \rho(x_{m-1}, x_n) \leq \cdots \leq \sum_{k=n}^{m-1} \rho(x_{k+1}, x_k).
\]
Combining the last two inequalities gives (using again that \( \alpha \in (0, 1) \)),
\[
\rho(x_m, x_n) \leq \sum_{k=n}^{m-1} \alpha^k \rho(x_1, x_0) \leq \rho(x_1, x_0) \alpha^n \sum_{l=0}^{\infty} \alpha^l = \rho(x_1, x_0) \frac{\alpha^n}{1 - \alpha}.
\]
This last equation shows that \( \rho(x_m, x_n) \to 0 \) as \( m, n \to \infty \), i.e. \( \{x_n\}_{n=0}^{\infty} \) is a Cauchy sequence. \( \blacksquare \)

**Corollary 2.2 (Contraction Mapping Principle II).** Suppose that \((X, \rho)\) is a complete metric space and \( S : X \to X \) is a continuous map such that \( S^{(n)} \) is a contraction for some \( n \in \mathbb{N} \). Here

\[
S^{(n)} = S \circ S \circ \ldots \circ S
\]

and we are assuming there exists \( \alpha \in (0, 1) \) such that \( \rho(S^{(n)}(x), S^{(n)}(y)) \leq \alpha \rho(x, y) \) for all \( x, y \in X \). Then \( S \) has a unique fixed point in \( X \).

**Proof.** Let \( T \equiv S^{(n)} \), then \( T : X \to X \) is a contraction and hence \( T \) has a unique fixed point \( x \in X \). Since any fixed point of \( S \) is also a fixed point of \( T \), we see if \( S \) has a fixed point then it must be \( x \). Now

\[
T(S(x)) = S^{(n)}(S(x)) = S(S^{(n)}(x)) = S(T(x)) = S(x),
\]

which shows that \( S(x) \) is also a fixed point of \( T \). Since \( T \) has only one fixed point, we must have that \( S(x) = x \). So we have shown that \( x \) is a fixed point of \( S \) and this fixed point is unique. \( \blacksquare \)

**Lemma 2.3.** Suppose that \((X, \rho)\) is a complete metric space, \( n \in \mathbb{N} \), \( Z \) is a topological space, and \( \alpha \in (0, 1) \). Suppose for each \( z \in Z \) there is a map \( S_z : X \to X \) with the following properties:

- **Contraction property:** \( \rho(S_z^{(n)}(x), S_z^{(n)}(y)) \leq \alpha \rho(x, y) \) for all \( x, y \in X \) and \( z \in Z \).
- **Continuity in \( z \):** For each \( x \in X \) the map \( z \in Z \to S_z(x) \in X \) is continuous.

By Corollary 2.2 above, for each \( z \in Z \) there is a unique fixed point \( G(z) \in X \) of \( S_z \).

**Conclusion:** The map \( G : Z \to X \) is continuous.

**Proof.** Let \( T_z \equiv S_z^{(n)} \). If \( z, w \in Z \), then

\[
\rho(G(z), G(w)) = \rho(T_z(G(z)), T_w(G(w)))
\]

\[
\leq \rho(T_z(G(z)), T_w(G(z))) + \rho(T_w(G(z)), T_w(G(w)))
\]

\[
\leq \rho(T_z(G(z)), T_w(G(z))) + \alpha \rho(G(z), G(w)).
\]

Hence it follows that

\[
\rho(G(z), G(w)) \leq \frac{1}{1 - \alpha} \rho(T_z(G(z)), T_w(G(z))).
\]

Since \( w \to T_w(G(z)) \) is continuous it follows from the above equation that \( G(w) \to G(z) \) as \( w \to z \), i.e. \( G \) is continuous. \( \blacksquare \)

3. **Inverse and Implicit Function Theorems**

In this section, let \( X \) be a Banach spaces, \( U \subset X \) be an open set, and \( F : U \to X \) and \( \epsilon : U \to X \) be continuous functions. We now want to consider the following general problem. Suppose \( 0 \in U \) and that \( F(x) = x + \epsilon(x) \) where \( \epsilon \) is “small,” when is \( F \) restricted to small balls about 0 still a homeomorphism onto an open subset of \( X \)? To get an indication as to what we need to assume about \( \epsilon \) to make such a statement true, let’s look at the one dimensional case first. So for the moment assume that \( X = \mathbb{R}, U = (-1, 1), \) and \( \epsilon : U \to \mathbb{R} \) is \( C^1 \). Then \( F \) will be one to one.
iff $F$ is monotonic. This will be the case, for example, if $F' = 1 + \varepsilon' > 0$. This in turn is guaranteed by assuming that $|\varepsilon'| \leq \alpha < 1$. (This last condition makes sense on a Banach space whereas assuming $1 + \varepsilon' > 0$ is not as easily interpreted.)

**Lemma 3.1.** Suppose that $U = B = B(0, r)$ ($r > 0$) is a ball in $X$ and $\varepsilon : B \to X$ is a $C^1$ function such that $\|D\varepsilon\| \leq \alpha < \infty$ on $U$. Then for all $x, y \in U$ we have:

\[
\|\varepsilon(x) - \varepsilon(y)\| \leq \alpha\|x - y\|.
\]

**Proof.** By the fundamental theorem of calculus and the chain rule:

\[
\varepsilon(y) - \varepsilon(x) = \int_0^1 \frac{d}{dt} \varepsilon(x + t(y - x)) dt = \int_0^1 [D\varepsilon(x + t(y - x))] (y - x) dt.
\]

Therefore, by the triangle inequality and the assumption that $\|D\varepsilon(x)\| \leq \alpha$ on $B$,

\[
\|\varepsilon(y) - \varepsilon(x)\| \leq \int_0^1 \|D\varepsilon(x + t(y - x))\| dt \cdot \|y - x\| \leq \alpha \|y - x\|.
\]

\[
\]

**Remark 3.2.** It is easily checked that if $\varepsilon : B = B(0, r) \to X$ is $C^1$ and satisfies (3.1), then $\|D\varepsilon\| \leq \alpha$ on $B$.

Using the above remark and the analogy to the one dimensional example, one is lead to the following proposition.

**Proposition 3.3.** Suppose that $U = B = B(0, r)$ ($r > 0$) is a ball in $X$, $\alpha \in (0, 1)$, $\varepsilon : U \to X$ is continuous, $F(x) \equiv x + \varepsilon(x)$ for $x \in U$, and $\varepsilon$ satisfies

\[
\|\varepsilon(x) - \varepsilon(y)\| \leq \alpha\|x - y\| \quad \forall x, y \in U.
\]

Then $F(B)$ is open in $X$ and $F : B \to V := F(B)$ is a homeomorphism.

**Proof.** First notice from (3.2) that

\[
\|x - y\| = \|(F(x) - F(y)) - (\varepsilon(x) - \varepsilon(y))\|
\leq \|F(x) - F(y)\| + \|\varepsilon(x) - \varepsilon(y)\|
\leq \|F(x) - F(y)\| + \alpha\|x - y\|
\]

from which it follows that $\|x - y\| \leq (1 - \alpha)^{-1} \|F(x) - F(y)\|$. Thus $F$ is injective on $B$. Let $V = F(B)$ and $G = F^{-1} : V \to B$ denote the inverse function which exists since $F$ is injective.

We will now show that $V$ is open. For this let $x_0 \in B$ and $x_0 = F(x_0) = x_0 + \varepsilon(x_0) \in V$. We wish to show for $z$ close to $x_0$ that there is an $x \in B$ such that $F(x) = x + \varepsilon(x) = z$ or equivalently $x = z - \varepsilon(x)$. Set $S_z(x) = z - \varepsilon(x)$, then we are looking for $x \in B$ such that $x = S_z(x)$, i.e. we want to find a fixed point of $S_z$. We will show that such a fixed point exists by using the contraction mapping theorem.

**Step 1.** $S_z$ is contractive for all $z \in X$. In fact for $x, y \in B$,

\[
\|S_z(x) - S_z(y)\| = \|\varepsilon(x) - \varepsilon(y)\| \leq \alpha\|x - y\|.
\]
Step 2. For any $\delta > 0$ such the $C = \overline{B(x_0, \delta)} \subset B$ and $z \in X$ such that $\|z - z_0\| < (1 - \alpha)\delta$, we have $S_\delta(C) \subset C$. Indeed, let $x \in C$ and compute:

$$
\|S_\delta(x) - x_0\| = \|S_\delta(x) - S_\delta(x_0)\| = \|z - \epsilon(x) - (z_0 - \epsilon(x))\| = \|z - z_0 - (\epsilon(x) - \epsilon(x))\| \\
\leq \|z - z_0\| + \alpha\|x - x_0\| < (1 - \alpha)\delta + \alpha\delta = \delta.
$$

wherein we have used $z_0 = F(x_0)$ and (3.2).

Since $C$ is a closed subset of a Banach space $X$, we may apply the contraction mapping principle, Theorem 2.1 and Lemma 2.3, to $S_\delta$ to show that there is a continuous function $G : B(x_0, (1 - \alpha)\delta) \to C$ such that

$$
G(z) = S_\delta(G(z)) = z - \epsilon(G(z)) = z - F(G(z)) + G(z),
$$

i.e. $F(G(z)) = z$. This shows that $B(x_0, (1 - \alpha)\delta) \subset F(C) \subset F(B) = V$. That is $z_0$ is in the interior of $V$. Since $F^{-1}|_{B(x_0, (1 - \alpha)\delta)}$ is necessarily equal to $G$ which is continuous, we have also shown that $F^{-1}$ is continuous in a neighborhood of $z_0$. Since $z_0 \in V$ was arbitrary, we have shown that $V$ is open and that $F^{-1} : V \to U$ is continuous. \qed

3.1. Inverse Function Theorem.

Theorem 3.4 (Inverse Function Theorem). Suppose that $f : U \to X$ is a $C^1$ function, $x_0 \in U$, and $Df(x_0)$ is invertible. Then there is a ball $B = B(x_0, r)$ in $U$ centered at $x_0$ such that

1. $V = f(B)$ is open,
2. $f|_B : B \to V$ is a homeomorphism,
3. $g = (f|_B)^{-1}$ is $C^1$.

Proof. Define $F(x) \equiv [Df(x_0)]^{-1}f(x+x_0)$ and $\epsilon(x) \equiv x - F(x)$ for $x \in (U - x_0)$. Notice that $0 \in U - x_0$, $DF(0) = I$, and that $D\epsilon(0) = I - I = 0$. Choose $r > 0$ such that $\hat{B} \equiv B(0, r) \subset U - x_0$ and $\|D\epsilon(x)\| \leq \frac{1}{2}$ for $x \in \hat{B}$. By Lemma 3.1, $\epsilon$ satisfies (3.3) with $\alpha = 1/2$. By Proposition 3.3, $F(\hat{B})$ is open and $F|_B : B \to F(\hat{B})$ is a homeomorphism. Let $G \equiv F|_B^{-1}$ which we know to be a continuous map from $F(\hat{B}) \to \hat{B}$.

Since $\|D\epsilon(x)\| \leq 1/2$ for $x \in \hat{B}$, $DF(x) = I + D\epsilon(x)$ is invertible, see Proposition 1.4. Since $H(z) \equiv z$ is $C^1$ and $H = F \circ G$ on $F(\hat{B})$, it follows from the converse to the chain rule, Theorem 1.11, that $G$ is differentiable and

$$
DG(z) = [DF(G(z))]^{-1}DH(z) = [DF(G(z))]^{-1}.
$$

Since $G$, $DF$, and the map $A \in L^*(X) \to A^{-1} \in L^*(X)$ are all continuous maps, (see Example 1.8) the map $z \in F(\hat{B}) \to DG(z) \in L(X)$ is also continuous, i.e. $G$ is $C^1$.

Let $B = \hat{B} + x_0 = B(x_0, r) \subset U$. Since $f(x) = Df(x_0)[F(x - x_0) + Df(x_0) + x]$ is invertible (hence an open mapping), $f(B) = [Df(x_0)]F(\hat{B})$ is open in $X$. It is also easily checked that $f|_B^{-1}$ exists and is given by

$$
f|_B^{-1}(y) = x_0 + G([Df(x_0)]^{-1}y)
$$

(3.4)
for \( y \in f(B) \). This shows that \( f|_B \) is a homeomorphism of \( B \) onto \( f(B) \). It now follows from (3.4), the fact the \( G \) is \( C^1 \), and the chain rule that \( f|_B \) is \( C^1 \). □

3.2. Implicit Function Theorem.

**Theorem 3.5** (Implicit Function Theorem). Now suppose that \( X, Y, \) and \( W \) are three Banach spaces, \( A \subset X \times Y \) is an open set, \( (x_0, y_0) \) is a point in \( A \), and \( f : A \rightarrow W \) is a \( C^1 \) - map such \( f(x_0, y_0) = 0 \). Assume that \( D_2 f(x_0, y_0) \equiv D(f(x_0, \cdot))(y_0) : Y \rightarrow W \) is a bounded invertible linear transformation. Then there is an open neighborhood \( U_0 \) of \( x_0 \) in \( X \) such that for all connected open neighborhoods \( U \) of \( x_0 \) contained in \( U_0 \), there is a unique continuous function \( u : U \rightarrow Y \) such that \( u(x_0) = y_0, (x, u(x)) \in A \) and \( f(x, u(x)) = 0 \) for all \( x \in U \). Moreover \( u \) is necessarily \( C^1 \) and

\[
Du(x) = -D_2 f(x, u(x))^{-1} D_1 f(x, u(x)) \text{ for all } x \in U.
\]

**Proof.** Proof of 3.5. By replacing \( f \) by \( (x, y) \rightarrow D_2 f(x_0, y_0)^{-1} f(x, y) \) if necessary, we may assume with out loss of generality that \( W = Y \) and \( D_2 f(x_0, y_0) = I_Y \).

Define \( F : A \rightarrow X \times Y \) by \( F(x, y) \equiv (x, f(x, y)) \) for all \( (x, y) \in A \). Notice that

\[
DF(x, y) = \begin{bmatrix} I & D_1 f(x, y) \\ 0 & D_2 f(x, y) \end{bmatrix}
\]

which is invertible if \( D_2 f(x, y) \) is invertible and if \( D_2 f(x, y) \) is invertible then

\[
DF(x, y)^{-1} = \begin{bmatrix} I & -D_1 f(x, y)D_2 f(x, y)^{-1} \\ 0 & D_2 f(x, y)^{-1} \end{bmatrix}.
\]

Since \( D_2 f(x_0, y_0) = I \) is invertible, the implicit function theorem guarantees that there exists a neighborhood \( U_0 \) of \( x_0 \) and \( V_0 \) of \( y_0 \) such that \( U_0 \times V_0 \subset A \), \( F(U_0 \times V_0) \) is open in \( X \times Y \), \( F|_{U_0 \times V_0} \) has a \( C^1 \)-inverse which we call \( F^{-1} \). Let \( \pi_2(x, y) \equiv y \) for all \( (x, y) \in X \times Y \) and define \( C^1 \)-function \( u_0 \) on \( U_0 \) by \( u_0(x) \equiv \pi_2 \circ F^{-1}(x, 0) \).

Since \( F^{-1}(x, 0) = (\tilde{x}, u_0(\tilde{x})) \) iff \( (x, 0) = F(\tilde{x}, u_0(\tilde{x})) = (\tilde{x}, \tilde{f}(\tilde{x}, u_0(\tilde{x}))) \), it follows that \( x = \bar{x} \) and \( f(\bar{x}, u_0(\bar{x})) = 0 \). Thus \( (\bar{x}, u_0(\bar{x})) = F^{-1}(x, 0) \subset U_0 \times V_0 \subset A \) and \( f(\bar{x}, u_0(\bar{x})) = 0 \) for all \( x \in U_0 \). Moreover, \( u_0 \) is \( C^1 \) being the composition of the \( C^1 \)-functions, \( x \rightarrow (x, 0) \), \( F^{-1} \), and \( \pi_2 \). So if \( U \subset U_0 \) is a connected set containing \( x_0 \), we may define \( u \equiv u_0|_U \) to show the existence of the functions \( u \) as described in the statement of the theorem. The only statement left to prove is the uniqueness of such a function \( u \).

Suppose that \( u_1 : U \rightarrow Y \) is another continuous function such that \( u_1(x_0) = y_0 \), and \( (x, u_1(x)) \in A \) and \( f(x, u_1(x)) = 0 \) for all \( x \in U \). Let

\[
O \equiv \{ x \in U | u(x) = u_1(x) \} = \{ x \in U | u_0(x) = u_1(x) \}.
\]

Clearly \( O \) is a (relatively) closed subset of \( U \) which is not empty since \( x_0 \in O \).

Because \( U \) is connected, if we show that \( O \) is also an open set we will have shown that \( O = U \) or equivalently that \( u_1 = u_0 \) on \( U \). So suppose that \( x \in O \), i.e. \( u_0(x) = u_1(x) \). For \( \tilde{x} \) near \( x \in U \),

\[
0 = 0 - 0 = f(\tilde{x}, u_0(\tilde{x})) - f(\tilde{x}, u_1(\tilde{x})) = R(\tilde{x})(u_1(\tilde{x}) - u_0(\tilde{x}))
\]

where

\[
R(\tilde{x}) \equiv \int_0^1 D_2 f((\tilde{x}, u_0(\tilde{x}) + t(u_1(\tilde{x}) - u_0(\tilde{x})))dt.
\]
From Eq. (3.7) and the continuity of \( u_0 \) and \( u_1 \), \( \lim_{\tilde{x} \to x} R(\tilde{x}) = D_{\tilde{x}} f(x, u_0(x)) \) which is invertible\(^2\). Thus \( R(\tilde{x}) \) is invertible for all \( \tilde{x} \) sufficiently close to \( x \). Using Eq. (3.6), this last remark implies that \( u_1(\tilde{x}) = u_0(\tilde{x}) \) for all \( \tilde{x} \) sufficiently close to \( x \). Since \( x \in O \) was arbitrary, we have shown that \( O \) is open. 

4. BASIC FACTS ABOUT COMPLEX NUMBERS

**Definition 4.1.** \( \mathbb{C} = \mathbb{R}^2 \) and we write \( 1 = (1, 0) \) and \( i = (0, 1) \). As usual \( \mathbb{C} \) becomes a field with the multiplication rule determined by \( i^2 = 1 \) and \( i^2 = -1 \), i.e.

\[
(a + ib)(c + id) \equiv (ac - bd) + i(bc + ad).
\]

**Notation 4.2.** If \( z = a + ib \) with \( a, b \in \mathbb{R} \) let \( \bar{z} = a - ib \) and \( |z|^2 = z\bar{z} = a^2 + b^2 \).

Also notice that if \( z \neq 0 \), then \( z \) is invertible with inverse given by

\[
z^{-1} = \frac{1}{z} = \frac{\bar{z}}{|z|^2}.
\]

Given \( w = a + ib \in \mathbb{C} \), the map \( z \in \mathbb{C} \to wz \in \mathbb{C} \) is complex and hence real linear so we may view this a linear transformation \( M_w : \mathbb{R}^2 \to \mathbb{R}^2 \). To work out the matrix of this transformation, let \( z = c + id \), then the map is \( c + id \to wz = (ac - bd) + i(bc + ad) \) which written in terms of real and imaginary parts is equivalent to

\[
\begin{pmatrix}
  a & -b \\
  b & a
\end{pmatrix}
\begin{pmatrix}
  c \\
  d
\end{pmatrix}
= \begin{pmatrix}
  ac - bd \\
  bc + ad
\end{pmatrix}.
\]

Thus

\[
M_w = \begin{pmatrix}
  a & -b \\
  b & a
\end{pmatrix} = aI + bJ \quad \text{where } J = \begin{pmatrix}
  0 & -1 \\
  1 & 0
\end{pmatrix}.
\]

**Remark 4.3.** Continuing the notation above, \( M^*_w = M_w \), \( \det(M_w) = a^2 + b^2 = |w|^2 \), and \( M_w M_z = M_{wz} \) for all \( w, z \in \mathbb{C} \). Moreover the ready may easily check that a real \( 2 \times 2 \) matrix \( A \) is equal to \( M_w \) for some \( w \in \mathbb{C} \) iff \( 0 = [A, J] := AJ - JA \).

Hence \( \mathbb{C} \) and the set of real \( 2 \times 2 \) matrices \( A \) such that \( 0 = [A, J] \) are algebraically isomorphic objects.

5. COMPLEX DIFFERENTIABLE FUNCTIONS

**Definition 5.1.** A function \( F : \Omega \subset_o \mathbb{C} \to \mathbb{C} \) is complex differentiable at \( z_0 \in \Omega \) if

\[
\lim_{z \to z_0} \frac{F(z) - F(z_0)}{z - z_0} = w
\]

exists.

**Proposition 5.2.** A function \( F : \Omega \subset_o \mathbb{C} \to \mathbb{C} \) is complex differentiable iff \( F : \Omega \to \mathbb{C} \) is differentiable (in the real sense as a function from \( \Omega \subset_o \mathbb{R}^2 \to \mathbb{R}^2 \)) and \( [F'(z_0), J] = 0 \), i.e. by Remark 4.3,

\[
F'(z_0) = M_w = \begin{pmatrix}
  a & -b \\
  b & a
\end{pmatrix}
\]

for some \( w = a + ib \in \mathbb{C} \).

\(^2\)Notice that \( DF(x, u_0(x)) \) is invertible for all \( x \in U_0 \) since \( F|_{U_0 \times U_0} \) has a \( C^1 \) inverse. Therefore \( D_{\tilde{x}} f(x, u_0(x)) \) is also invertible for all \( x \in U_0 \).
Proof. Eq. (5.1) is equivalent to the equation:
\[ F(z) = F(z_0) + w(z - z_0) + o(z - z_0) \]
(5.2)
\[ = F(z_0) + M_w(z - z_0) + o(z - z_0) \]
and hence \( F \) is complex differentiable iff \( F \) is differentiable and the differential is of the form \( F'(z_0) = M_w \) for some \( w \in \mathbb{C} \). ■

**Corollary 5.3** (Cauchy Riemann Equations). \( F : \Omega \to \mathbb{C} \) is complex differentiable at \( z_0 \in \Omega \) iff \( F'(z_0) \) exists\(^3\) and, writing \( z_0 = x_0 + iy_0 \),
\[ i \frac{\partial F(x_0 + iy_0)}{\partial x} = \frac{\partial F}{\partial y}(x_0 + iy_0) \]
or in short we write \( \frac{\partial F}{\partial x} + i \frac{\partial F}{\partial y} = 0 \).

Proof. The differential \( F'(z_0) \) is, in general, an arbitrary matrix of the form
\[ F'(z_0) = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \]
where
\[ \frac{\partial F}{\partial x}(z_0) = a + ib \text{ and } \frac{\partial F}{\partial y}(z_0) = c + id. \]
Since \( F \) is complex differentiable at \( z_0 \) iff \( d = a \) and \( c = -b \) which is easily seen to be equivalent to Eq. (5.3) by Eq. (5.4) and comparing the real and imaginary parts of \( iF_x(z_0) \) and \( F_y(z_0) \). ■

**Notation 5.4.** Let
\[ \bar{\partial} = \frac{1}{2} \left( \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \text{ and } \partial = \frac{1}{2} \left( \frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right). \]
With this notation we have
\[ \partial f dz + \bar{\partial} f d\bar{z} = \frac{1}{2} \left( \frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) f (dx + idy) + \frac{1}{2} \left( \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) f (dx - idy) \]
\[ = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = df. \]
In particular if \( \sigma(s) \in \mathbb{C} \) is a smooth curve, then
\[ \frac{d}{ds} f(\sigma(s)) = df(\sigma(s)) \sigma'(s) + \bar{\partial} f(\sigma(s)) \bar{\sigma}'(s). \]

**Corollary 5.5.** Let \( \Omega \subset \mathbb{C} \) be a given open set and \( f : \Omega \to \mathbb{C} \) be a \( C^1 \) function in the real variable sense. Then the following are equivalent:
1. The complex derivative \( df(z)/dz \) exists for all \( z \in \Omega \).\(^4\)
2. The real differential \( f'(z) \) satisfies \( \left| f'(z), J \right| = 0 \) for all \( z \in \Omega \).
3. The function \( f \) satisfies the Cauchy Riemann equations \( \bar{\partial} f = 0 \) on \( \Omega \).

**Notation 5.6.** A function \( f : \Omega \to \mathbb{C} \) satisfying any and hence all of the conditions in Corollary 5.5 is said to be a holomorphic or an analytic function on \( \Omega \). We will let \( H(\Omega) \) denote the space of holomorphic functions on \( \Omega \).

\(^3\)For example this is satisfied if \( F : \Omega \to \mathbb{C} \) is continuous at \( z_0 \), \( F_x \) and \( F_y \) exists in a neighborhood of \( z_0 \) and are continuous near \( z_0 \).

\(^4\)As we will see later the assumption that \( f \) is \( C^1 \) in this condition is redundant. Complex differentiability already implies that \( f \) is \( C^\infty \).
Corollary 5.7. The chain rule holds for complex differentiable functions. In particular, \( \Omega \subset \mathbb{C} \xrightarrow{f} D \subset \mathbb{C} \) are functions, \( z_0 \in \Omega \) and \( u_0 = f(z_0) \in D \). Assume that \( f'(z_0) \) exists, \( g'(u_0) \) exists then \( (g \circ f)'(z_0) \) exists and is given by

\[
(g \circ f)'(z_0) = g'(f(z_0))f'(z_0)
\]

Proof. This is a consequence of the chain rule for \( F : \mathbb{R}^2 \to \mathbb{R}^2 \) when restricted to those functions whose differentials commute with \( J \). Alternatively, one can simply follow the usual proof in the complex category as follows:

\[
g \circ f(z) = g(f(z)) = g(u_0) + g'(u_0)(f(z) - f(z_0)) + o(f(z) - f(z_0))
\]

and hence

\[
\frac{g \circ f(z) - g(f(z_0))}{z - z_0} = g'(u_0)\frac{f(z) - f(z_0)}{z - z_0} + \frac{o(f(z) - f(z_0))}{z - z_0}.
\]

Since \( \frac{o(f(z) - f(z_0))}{z - z_0} \to 0 \) as \( z \to z_0 \) we may pass to the limit \( z \to z_0 \) in Eq. (5.6) to prove Eq. (5.5).

Lemma 5.8 (Converse to the Chain rule). Suppose \( f : \Omega \subset \mathbb{C} \to U \subset \mathbb{C} \) and \( g : U \subset \mathbb{C} \to \mathbb{C} \) are functions such that \( f \) is continuous, \( g \in H(U) \) and \( h := g \circ f \in H(\Omega) \), then \( f \in H(\Omega \setminus \{z : g'(f(z)) = 0\}) \). Moreover \( f'(z) = h'(z)/g'(f(z)) \) when \( z \in \Omega \) and \( g'(f(z)) \neq 0 \).

Proof. This follow from the previous converse to the chain rule or directly as follows\(^5\). Suppose that \( z_0 \in \Omega \) and \( g'(f(z_0)) \neq 0 \). On one hand

\[
h(z) = h(z_0) + h'(z_0)(z - z_0) + o(z - z_0)
\]

while on the other

\[
h(z) = g(f(z)) = g(f(z_0)) + g'(f(z_0))(f(z) - f(z_0)) + o(f(z) - f(z_0)).
\]

Combining these equations shows

\[
h'(z_0)(z - z_0) = g'(f(z_0))(f(z) - f(z_0)) + o(f(z) - f(z_0)) + o(z - z_0).
\]

Since \( g'(f(z_0)) \neq 0 \) we may conclude that

\[
f(z) - f(z_0) = o(f(z) - f(z_0)) + O(z - z_0),
\]

in particular it follow that

\[
|f(z) - f(z_0)| \leq \frac{1}{2} |f(z) - f(z_0)| + O(z - z_0) \text{ for } z \text{ near } z_0
\]

and hence that \( f(z) - f(z_0) = O(z - z_0) \). Using this back in Eq. (5.7) then shows that

\[
h'(z_0)(z - z_0) = g'(f(z_0))(f(z) - f(z_0)) + o(z - z_0)
\]

or equivalently that

\[
f(z) - f(z_0) = \frac{h'(z_0)}{g'(f(z_0))}(z - z_0) + o(z - z_0).
\]

Example 5.9. Here are some examples.

\(^5\)One could also apeal to the inverse function theorem here as well.
1. \( f(z) = z \) is analytic and more generally \( f(z) = \sum_{n=0}^{k} a_n z^n \) with \( a_n \in \mathbb{C} \) are analytic on \( \mathbb{C} \).
2. If \( f, g \in H(\Omega) \) then \( f \cdot g, f + g, cf \in H(\Omega) \) and \( f/g \in H(\Omega \setminus \{g = 0\}) \).
3. \( f(z) = \bar{z} \) is not analytic any function \( f : \mathbb{C} \to \mathbb{R} \) is not analytic unless \( f \) is constant.

The next theorem shows that analytic functions may be averaged to produce new analytic functions.

**Theorem 5.10.** Let \( g : \Omega \times X \to \mathbb{C} \) be a function such that

1. \( g(\cdot, x) \in H(\Omega) \) for all \( x \in X \) and write \( g'(z, x) \) for \( \frac{\partial}{\partial z} g(z, x) \).
2. There exists \( G \in L^1(X, \mu) \) such that \( |g'(z, x)| \leq G(x) \) on \( \Omega \times X \).
3. \( g(z, \cdot) \in L^1(X, \mu) \) for \( z \in \Omega \).

Then

\[
f(z) := \int_{X} g(z, \xi) d\mu(\xi)
\]

is holomorphic on \( \Omega \) and the complex derivative is given by

\[
f'(z) = \int_{X} g'(z, \xi) d\mu(\xi).
\]

**Exercise 5.11.** Prove Theorem 5.10 using the dominated convergence theorem along with the mean value inequality of Theorem 1.14.

As an application we will show that power series give example of complex differentiable functions.

**Corollary 5.12.** Suppose that \( \{a_n\}_{n=0}^{\infty} \subset \mathbb{C} \) is a sequence of complex numbers such that series

\[
f(z) := \sum_{n=0}^{\infty} a_n (z - z_0)^n
\]

is convergent for \( |z - z_0| < R \), where \( R \) is some positive number. Then \( f : D(z_0, R) \to \mathbb{C} \) is complex differentiable on \( D(z_0, R) \) and

\[
f'(z) = \sum_{n=0}^{\infty} na_n (z - z_0)^{n-1} = \sum_{n=1}^{\infty} na_n (z - z_0)^{n-1}.
\]

By induction it follows that \( f^{(k)} \) exists for all \( k \) and that

\[
f^{(k)}(z) = \sum_{n=0}^{\infty} n(n-1) \ldots (n-k+1) a_n (z - z_0)^{n-k}.
\]

**Proof.** Let \( \rho < R \) be given and choose \( r \in (\rho, R) \). Since \( z = z_0 + r \in D(z_0, R) \), by assumption the series \( \sum_{n=0}^{\infty} a_n r^n \) is convergent and in particular \( M := \sup_{n} |a_n r^n| < \infty \). We now apply Theorem 5.10 with \( X = \mathbb{N} \cup \{0\} \), \( \mu \) being counting measure, \( \Omega = D(z_0, \rho) \) and \( g(z, n) := a_n (z - z_0)^n \). Since

\[
|g'(z, n)| = |na_n (z - z_0)^{n-1}| \leq n |a_n| |\rho|^{n-1} \leq \frac{1}{r^n} \left( \frac{\rho}{r} \right)^{n-1} M
\]

is uniformly bounded on \( D(z_0, \rho) \), then \( f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \) is analytic on \( D(z_0, \rho) \) and

\[
f'(z) = \sum_{n=1}^{\infty} na_n (z - z_0)^{n-1}.
\]
and the function $G(n) := \frac{M}{r} n \left( \frac{r}{|z|} \right)^{n-1}$ is summable (by the Ratio test for example), we may use $G$ as our dominating function. It then follows from Theorem 5.10 that

$$f(z) = \int g(z, n) d\mu(n) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

is complex differentiable with the differential given as in Eq. (5.8).

**Example 5.13.** Let $w \in \mathbb{C}$, $\Omega := \mathbb{C} \setminus \{w\}$ and $f(z) = \frac{1}{z-w}$. Then $f \in H(\Omega)$. Let $z_0 \in \Omega$ and write $z = z_0 + h$, then

$$f(z) = \frac{1}{w-z} = \frac{1}{w-z_0 + h} = \frac{1}{w-z_0} \frac{1}{1 - h/(w-z_0)}$$

$$= \frac{1}{w-z_0} \sum_{n=0}^{\infty} \left( \frac{h}{w-z_0} \right)^n = \sum_{n=0}^{\infty} \left( \frac{1}{w-z_0} \right)^{n+1} (z-z_0)^n$$

which is valid for $|z-z_0| < |w-z_0|$. Summarizing this computation we have shown

$$\frac{1}{w-z} = \sum_{n=0}^{\infty} \left( \frac{1}{w-z_0} \right)^{n+1} (z-z_0)^n \text{ for } |z-z_0| < |w-z_0|. \quad (5.9)$$

**Proposition 5.14.** The exponential function $e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$ is holomorphic on $\mathbb{C}$ and $\frac{d}{dz} e^z = e^z$. Moreover,

1. $e^{z+w} = e^z e^w$ for all $z, w \in \mathbb{C}$.
2. (Euler’s Formula) $e^{i\theta} = \cos \theta + i \sin \theta$ for all $\theta \in \mathbb{R}$ and $|e^{i\theta}| = 1$ for all $\theta \in \mathbb{R}$.
3. $e^{x+iy} = e^x (\cos y + i \sin y)$ for all $x, y \in \mathbb{R}$.
4. $e^{x} = e^{\Re x}$.

**Proof.** By the chain rule for functions of a real variable,

$$\frac{d}{dt} [e^{-tw} e^{(z+tw)}] = -we^{-tw} e^{(z+tw)} + e^{-tw} we^{(z+tw)} = 0$$

and hence $e^{-tw} e^{(z+tw)}$ is constant in $t$. So by evaluating this expression at $t = 0$ and $t = 1$ we find

$$e^{-w} e^{(z+w)} = e^z \text{ for all } w, z \in \mathbb{C}. \quad (5.10)$$

Choose $z = 0$ in Eq. (5.10) implies $e^{-w} e^w = 1$, i.e. $e^{-w} = 1/e^w$ which used back in Eq. (5.10) proves item 1. Similarly,

$$\frac{d}{d\theta} [e^{-i\theta} (\cos \theta + i \sin \theta)] = -ie^{-i\theta} (\cos \theta + i \sin \theta) + e^{-i\theta} (-\sin \theta + i \cos \theta) = 0.$$

Hence $e^{-i\theta} (\cos \theta + i \sin \theta) = e^{-i\theta} (\cos \theta + i \sin \theta)|_{\theta=0} = 1$ which proves item 2. Item 3. is a consequence of items 1) and 2) and item 4) follows from item 3) or directly from the power series expansion.

**Remark 5.15.** One could define $e^z$ by $e^z = e^x (\cos(y) + i \sin(y))$ when $z = x + iy$ and then use the Cauchy Riemann equations to prove $e^z$ is complex differentiable.

**Exercise 5.16.** By comparing the real and imaginary parts of the equality $e^{i\theta} e^{i\alpha} = e^{i(\theta + \alpha)}$ prove the formulas:

$$\cos(\theta + \alpha) = \cos \theta \cos \alpha - \sin \theta \sin \alpha$$

$$\sin(\theta + \alpha) = \cos \theta \sin \alpha + \cos \alpha \sin \theta$$
for all \( \theta, \alpha \in \mathbb{R} \).

**Exercise 5.17.** Find all possible solutions to the equation \( e^z = w \) where \( z \) and \( w \) are complex numbers. Let \( \log(w) \equiv \{ z : e^z = w \} \). Note that \( \log : \mathbb{C} \to \) (subsets of \( \mathbb{C} \)). One often writes \( \log : \mathbb{C} \to \mathbb{C} \) and calls log a multi-valued function. A continuous function \( l \) defined on some open subset \( \Omega \) of \( \mathbb{C} \) is called a branch of log if \( l(w) \in \log(w) \) for all \( w \in \Omega \). Use the reverse chain rule to show that any branch of log is holomorphic on its domain of definition and that \( l'(z) = 1/z \) for all \( z \in \Omega \).

**Exercise 5.18.** Let \( \Omega = \{ w = re^{i\theta} \in \mathbb{C} : r > 0, \quad -\pi < \theta < \pi \} = \mathbb{C} \setminus (-\infty, 0] \), and define \( L_n : \Omega \to \mathbb{C} \) by \( L_n(re^{i\theta}) = \ln(r) + id\theta \) for \( r > 0 \) and \( |\theta| < \pi \). Show that \( L_n \) is a branch of log. This branch of the log function is often called the principle value branch of log. The line \( (-\infty, 0] \) where \( L_n \) is not defined is called a branch cut.

**Exercise 5.19.** Let \( \sqrt[n]{w} = \{ z \in \mathbb{C} : z^n = w \} \). The “function” \( w \to \sqrt[n]{w} \) is another example of a multi-valued function. Let \( h(w) \) be any branch of \( \sqrt[n]{w} \), that is \( h \) is a continuous function on an open subset \( \Omega \) of \( \mathbb{C} \) such that \( h(w) \in \sqrt[n]{w} \). Show that \( h \) is holomorphic away from \( w = 0 \) and that \( h'(w) = \frac{1}{n} h(w)/w \).

**Exercise 5.20.** Let \( l \) be any branch of the log function. Define \( w^z = e^{z(l(w))} \) for all \( z \in \mathbb{C} \) and \( w \in D(l) \) where \( D(l) \) denotes the domain of \( l \). Show that \( w^{1/n} \) is a branch of \( \sqrt[n]{w} \) and also show that \( \frac{d}{dw} w^z = zw^{z-1} \).

### 5.1. Contour integrals.

**Definition 5.21.** Suppose that \( \sigma : [a, b] \to \Omega \) is a Piecewise \( C^1 \) function and \( f : \Omega \to \mathbb{C} \) is continuous, we define the contour integral of \( f \) along \( \sigma \) (written \( \int_{\sigma} f(z)dz \)) by

\[
\int_{\sigma} f(z)dz := \int_{a}^{b} f(\sigma(t))\dot{\sigma}(t)dt.
\]

**Notation 5.22.** Given \( \Omega \subset_{o} \mathbb{C} \) and a \( C^1 \) map \( \sigma : [a, b] \times [0, 1] \to \Omega \), let \( \sigma_s := \sigma(\cdot, s) \in C^1([a, b] \to \Omega) \). In this way, the map \( \sigma \) may be viewed as a map

\[
s \in [0, 1] \to \sigma_s := \sigma(\cdot, s) \in C^1([a, b] \to \Omega),
\]

i.e. \( s \to \sigma_s \) is a path of contours in \( \Omega \).

**Definition 5.23.** Given a region \( \Omega \) and \( \alpha, \beta \in C^1([a, b] \to \Omega) \), we will write \( \alpha \simeq \beta \) in \( \Omega \) provided there exists a \( C^1 \) map \( \sigma : [a, b] \times [0, 1] \to \Omega \) such that \( \sigma_0 = \alpha, \sigma_1 = \beta \), and \( \sigma \) satisfies either of the following two conditions:

1. \( \frac{d}{ds}\sigma(a, s) = \frac{d}{ds}\sigma(b, s) = 0 \) for all \( s \in [0, 1] \), i.e. the end points of the paths \( \sigma_s \) for \( s \in [0, 1] \) are fixed.
2. \( \sigma(a, s) = \sigma(b, s) \) for all \( s \in [0, 1] \), i.e. \( \sigma_s \) is a loop in \( \Omega \) for all \( s \in [0, 1] \).

**Proposition 5.24.** Let \( \Omega \) be a region and \( \alpha, \beta \in C^1([a, b], \Omega) \) be two contours such that \( \alpha \simeq \beta \) in \( \Omega \). Then

\[
\int_{\alpha} f(z)dz = \int_{\beta} f(z)dz \quad \text{for all} \quad f \in H(\Omega).
\]
Proof. Let $\sigma : [a, b] \times [0, 1] \to \Omega$ be as in Definition 5.23, then it suffices to show the function
\[ F(s) := \int_{\sigma_s} f(z)dz \]
is constant for $s \in [0, 1]$. For this we compute:
\[ F'(s) = \frac{d}{ds} \int_a^b f(\sigma(t, s))\dot{\sigma}(t, s)dt = \int_a^b \frac{d}{ds} [f(\sigma(t, s))\dot{\sigma}(t, s)] dt \]
\[ = \int_a^b \left\{ f'(\sigma(t, s))\dot{\sigma}'(t, s) \dot{\sigma} + f(\sigma(t, s))\dot{\sigma}'(t, s) \right\} dt \]
\[ = \int_a^b \frac{d}{dt} [f(\sigma(t, s))\dot{\sigma}'(t, s)] dt \]
\[ = [f(\sigma(t, s))\dot{\sigma}'(t, s)]_{t=a}^{t=b} = 0 \]
where the last equality is a consequence of either of the two endpoint assumptions of Definition 5.23. 

Remark 5.25. For those who know about differential forms and such we may generalize the above computation to $f \in C^1(\Omega)$ using $df = \partial f dz + \tilde{\partial} f d\tilde{z}$. We then find
\[ F'(s) = \frac{d}{ds} \int_a^b f(\sigma(t, s))\dot{\sigma}(t, s)dt = \int_a^b \frac{d}{ds} [f(\sigma(t, s))\dot{\sigma}(t, s)] dt \]
\[ = \int_a^b \left\{ \left[ \partial f(\sigma(t, s))\dot{\sigma}'(t, s) + \tilde{\partial} f(\sigma(t, s))\sigma'(t, s) \right] \dot{\sigma}(t, s) + f(\sigma(t, s))\dot{\sigma}'(t, s) \right\} dt \]
\[ = \int_a^b \left\{ \left[ \partial f(\sigma(t, s))\dot{\sigma}'(t, s) + \tilde{\partial} f(\sigma(t, s))\sigma'(t, s) \dot{\sigma}(t, s) \right] + f(\sigma(t, s))\dot{\sigma}'(t, s) \right\} dt \]
\[ + \int_a^b \tilde{\partial} f(\sigma(t, s)) (\sigma'(t, s)\dot{\sigma}(t, s) - \sigma(t, s)\sigma'(t, s)) dt \]
\[ = \int_a^b \frac{d}{dt} [f(\sigma(t, s))\dot{\sigma}'(t, s)] dt + \int_a^b \tilde{\partial} f(\sigma(t, s)) (\sigma'(t, s)\dot{\sigma}(t, s) - \sigma(t, s)\sigma'(t, s)) dt \]
\[ = [f(\sigma(t, s))\dot{\sigma}'(t, s)]_{t=a}^{t=b} + \int_a^b \tilde{\partial} f(\sigma(t, s)) (\sigma'(t, s)\dot{\sigma}(t, s) - \sigma(t, s)\sigma'(t, s)) dt \]
\[ = \int_a^b \tilde{\partial} f(\sigma(t, s)) (\sigma'(t, s)\dot{\sigma}(t, s) - \sigma(t, s)\sigma'(t, s)) dt. \]
Integrating this expression on $s$ then shows that
\[ \int_{\sigma_1} f dz - \int_{\sigma_0} f dz = \int_0^1 ds \int_a^b dt \tilde{\partial} f(\sigma(t, s)) (\sigma'(t, s)\dot{\sigma}(t, s) - \sigma(t, s)\sigma'(t, s)) \]
\[ = \int_{\sigma} \tilde{\partial} (f dz) = \int_{\sigma} \tilde{\partial} f d\tilde{z} \wedge dz \]
We have just given a proof of Green’s theorem in this context.

The main point of this section is to prove the following theorem.

Theorem 5.26. Let $\Omega \subset \mathbb{C}$ be an open set and $f \in C^1(\Omega, \mathbb{C})$, then the following statements are equivalent:
1. \( f \in H(\Omega) \),
2. For all disks \( D = D(z_0, \rho) \) such that \( \bar{D} \subset \Omega \),
   \[
   f(z) = \frac{1}{2\pi i} \oint_{\partial D} \frac{f(w)}{w-z} dw \text{ for all } z \in D.
   \]
3. For all disks \( D = D(z_0, \rho) \) such that \( \bar{D} \subset \Omega \), \( f(z) \) may be represented as a
   convergent power series
   \[
   f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \text{ for all } z \in D.
   \]

   In particular \( f \in C^\infty(\Omega, \mathbb{C}) \),
   Moreover if \( D \) is as above, we have
   \[
   f^{(n)}(z) = \frac{n!}{2\pi i} \oint_{\partial D} \frac{f(w)}{(w-z)^{n+1}} dw \text{ for all } z \in D
   \]
   and the coefficients \( a_n \) in Eq. (5.12) are given by
   \[
   a_n = f^{(n)}(z_0)/n! = \frac{1}{2\pi i} \oint_{\partial D} \frac{f(w)}{(w-z_0)^{n+1}} d.w.
   \]

   Proof. 1) \( \implies \) 2) For \( s \in [0, 1] \), let \( z_s = (1-s)z_0 + sz \), \( \rho_s := \text{dist}(z_s, \partial D) = \rho - s|z - z_0| \)
   and \( \sigma_s(t) = z_s + \rho_s e^{it} \) for \( 0 \leq t \leq 2\pi \). Notice that \( \sigma_0 \) is a parametrization
   of \( \partial D \), \( \sigma_0 \simeq \sigma_1 \) in \( \Omega \setminus \{z\} \), \( w \rightarrow \frac{\{w\}}{w-z} \) is in \( H(\Omega \setminus \{z\}) \) and hence by Proposition
   5.24,
   \[
   \oint_{\partial D} \frac{f(w)}{w-z} dw = \int_{\sigma_0} f(w) \frac{dw}{w-z} = \int_{\sigma_1} f(w) \frac{dw}{w-z}.
   \]
   Now let \( \tau_s(t) = z + s\rho_1 e^{it} \) for \( 0 \leq t \leq 2\pi \) and \( s \in (0, 1] \). Then \( \tau_1 = \sigma_1 \) and \( \tau_1 \simeq \tau_s \) in
   \( \Omega \setminus \{z\} \) and so again by Proposition 5.24,
   \[
   \oint_{\partial D} \frac{f(w)}{w-z} dw = \int_{\tau_s} f(w) \frac{dw}{w-z} = \int_{\sigma_1} f(w) \frac{dw}{w-z} = \int_{\tau_s} f(w) \frac{dw}{w-z} = \int_0^{2\pi} f(z + s\rho_1 e^{it}) \frac{is\rho_1 e^{it} dt}{s\rho_1 e^{it}} dt
   \]
   \[
   = i \int_0^{2\pi} f(z + s\rho_1 e^{it}) dt \to 2\pi i f(z) \text{ as } s \downarrow 0.
   \]

2) \( \implies \) 3) By 2) and Eq. (5.9)
   \[
   f(z) = \frac{1}{2\pi i} \oint_{\partial D} \frac{f(w)}{w-z} dw
   \]
   \[
   = \frac{1}{2\pi i} \oint_{\partial D} f(w) \sum_{n=0}^{\infty} \left( \frac{1}{w-z_0} \right)^{n+1} (z-z_0)^n d.w
   \]
   \[
   = \frac{1}{2\pi i} \sum_{n=0}^{\infty} \left( \oint_{\partial D} f(w) \left( \frac{1}{w-z_0} \right)^{n+1} d.w \right) (z-z_0)^n.
   \]

   (The reader should justify the interchange of the sum and the integral.) The last equation proves Eq. (5.12) and shows that
   \[
   a_n = \frac{1}{2\pi i} \oint_{\partial D} \frac{f(w)}{(w-z_0)^{n+1}} d.w.
   \]
Also using Theorem 5.10 we may differentiate Eq. (5.11) repeatedly to find
\begin{equation}
\frac{f^{(n)}(z)}{2\pi i} = \frac{n!}{\partial D} \int_{\partial D} \frac{f(w)}{(w-z)^{n+1}} \, dw \text{ for all } z \in D
\end{equation}
which evaluated at \(z = z_0\) shows that \(a_n = f^{(n)}(z_0)/n!\).

3) \implies 1) This follows from Corollary 5.12 and the fact that being complex differentiable is a local property.

The proof of the theorem also reveals the following corollary.

**Corollary 5.27.** If \(f \in H(\Omega)\) then \(f' \in H(\Omega)\) and by induction \(f^{(n)} \in H(\Omega)\) with \(f^{(n)}\) defined as in Eq. (5.14).

**Corollary 5.28** (Cauchy Estimates). Suppose that \(f \in H(\Omega)\) where \(\Omega \subset \mathbb{C}\) and suppose that \(D(z_0, \rho) \subset \Omega\), then
\begin{equation}
\left| f^{(n)}(z_0) \right| \leq \left( \frac{n!}{2\pi} \sup_{|z - z_0| < \rho} |f(\xi)| \right) \frac{1}{\rho^n}.
\end{equation}

**Proof.** From Eq. (5.14) evaluated at \(z = z_0\) and letting \(\sigma(t) = z_0 + \rho e^{it}\) for \(0 \leq t < 2\pi\), we find
\begin{align*}
f^{(n)}(z_0) &= \frac{n!}{2\pi i} \int_{\partial D} \frac{f(w)}{(w-z_0)^{n+1}} \, dw = \frac{n!}{2\pi i} \int_{\sigma} \frac{f(w)}{(w-z_0)^{n+1}} \, dw \\
&= \frac{n!}{2\pi i} \int_{0}^{2\pi} f(z_0 + \rho e^{it}) \left( \frac{1}{\rho e^{it}} \right)^{n+1} i\rho e^{it} \, dt \\
&= \frac{n!}{2\pi i} \int_{0}^{2\pi} f(z_0 + \rho e^{it}) \frac{1}{e^{n+1}i\rho^2} \, dt.
\end{align*}
Therefore,
\begin{align*}
\left| f^{(n)}(z_0) \right| &\leq \frac{n!}{2\pi i} \int_{0}^{2\pi} \left| f(z_0 + \rho e^{it}) \frac{1}{e^{n+1}i\rho^2} \right| \, dt = \frac{n!}{2\pi i} \int_{0}^{2\pi} \left| f(z_0 + \rho e^{it}) \right| \, dt \\
&\leq \left( \frac{n!}{2\pi} \sup_{|z - z_0| < \rho} |f(\xi)| \right) \frac{1}{\rho^n}.
\end{align*}

**Exercise 5.29.** Show that Theorem 5.10 is still valid with conditions 2) and 3) in the hypothesis being replaced by:
- There exists \(G \in L^1(X, \mu)\) such that \(|g(z, x)| \leq G(x)\).

**Hint:** Use the Cauchy estimates.

**Corollary 5.30** (Liouville’s Theorem). If \(f \in H(\mathbb{C})\) and \(f\) is bounded then \(f\) is constant.

**Proof.** This follows from Eq. (5.15) with \(n = 1\) and the letting \(n \to \infty\) to find \(f'(z_0) = 0\) for all \(z_0 \in \mathbb{C}\).

**Corollary 5.31** (Fundamental theorem of algebra). Every polynomial \(p(z)\) of degree larger than 0 has a root in \(\mathbb{C}\).

**Proof.** Suppose that \(p(z)\) is polynomial with no roots in \(z\). Then \(f(z) = 1/p(z)\) is a bounded holomorphic function and hence constant. This shows that \(p(z)\) is a constant, i.e. \(p\) has degree zero.
Definition 5.32. We say that $\Omega$ is a region if $\Omega$ is a connected open subset of $\mathbb{C}$.

Corollary 5.33. Let $\Omega$ be a region and $f \in H(\Omega)$ and $Z(f) = f^{-1}(\{0\})$ denote the zero set of $f$. Then either $f \equiv 0$ or $Z(f)$ has no accumulation points in $\Omega$. More generally if $f, g \in H(\Omega)$ and the set $\{z \in \Omega : f(z) = g(z)\}$ has an accumulation point in $\Omega$, then $f \equiv g$.

Proof. The second statement follows from the first by considering the function $f - g$. For the proof of the first assertion we will work strictly in $\Omega$ with the relative topology.

Let $A$ denote the set of accumulation points of $Z(f)$ (in $\Omega$). By continuity of $f$, $A \subset Z(f)$ and $A$ is a closed subset of $\Omega$ with the relative topology. The proof is finished by showing that $A$ is open and thus $A = \emptyset$ or $A = \Omega$ because $\Omega$ is connected.

Suppose that $z_0 \in A$, and express $f(z)$ as its power series expansion

$$f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$$

for $z$ near $z_0$. Since $0 = f(z_0)$ it follows that $a_0 = 0$. Let $z_k \in Z(f) \setminus \{z_0\}$ such that

$$\lim_{k \to \infty} z_k = z_0.$$ 

Then

$$0 = \frac{f(z_k)}{z_k - z_0} = \sum_{n=1}^{\infty} a_n(z_k - z_0)^{n-1} \to a_1 \text{ as } k \to \infty$$

so that $f(z) = \sum_{n=2}^{\infty} a_n(z - z_0)^n$. Similarly

$$0 = \frac{f(z_k)}{(z_k - z_0)^2} = \sum_{n=2}^{\infty} a_n(z_k - z_0)^{n-2} \to a_2 \text{ as } k \to \infty$$

and continuing by induction, it follows that $a_n \equiv 0$, i.e. $f$ is zero in a neighborhood of $z_0$. $\blacksquare$

Definition 5.34. For $z \in \mathbb{C}$, let

$$\cos(z) = \frac{e^{iz} + e^{-iz}}{2} \text{ and } \sin(z) = \frac{e^{iz} - e^{-iz}}{2i}.$$ 

Exercise 5.35. Show the these formula are consistent with the usual definition of cos and sin when $z$ is real. Also shows that the addition formula in Exercise 5.16 are valid for $\theta, \alpha \in \mathbb{C}$. This can be done with no additional computations by making use of Corollary 5.33.

Exercise 5.36. Let

$$f(z) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}x^2 + zx\right)dm(x) \text{ for } z \in \mathbb{C}.$$ 

Show $f(z) = \exp\left(\frac{z^2}{4}\right)$ using the following outline:

1. Show $f \in H(\Omega)$.
2. Show $f(z) = \exp\left(\frac{z^2}{4}\right)$ for $z \in \mathbb{R}$ by completing the squares and using the translation invariance of $m$. Also recall that you have proved in the first quarter that $f(0) = 1$.

---

$^6$Recall that $x \in A$ iff $V_x^* \cap Z \neq \emptyset$ for all $x \in V_x \subset \mathbb{C}$ where $V_x^* := V_x \setminus \{x\}$. Hence $x \notin A$ iff there exists $x \in V_x \subset \mathbb{C}$ such that $V_x^* \cap Z = \emptyset$. Since $V_x^*$ is open, it follows that $V_x^* \subset A^c$ and thus $V_x \subset A^c$. So $A^c$ is open, i.e. $A$ is closed.
3. Conclude \( f(z) = \exp(\frac{1}{2}z^2) \) for all \( z \in \mathbb{C} \) using Corollary 5.33.

**Corollary 5.37 (Mean value property).** Let \( \Omega \subset \mathbb{C} \) and \( f \in H(\Omega) \), then \( f \) satisfies the mean value property

\[
f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + \rho e^{i\theta}) d\theta
\]

which holds for all \( z_0 \) and \( \rho \geq 0 \) such that \( \overline{D(z_0, \rho)} \subset \Omega \).

**Proof.** Take \( n = 0 \) in Eq. (5.16). ■

**Lemma 5.38.** Suppose that \( f \in H(D) \) where \( D = D(z_0, \rho) \) for some \( \rho > 0 \). If \( |f(z)| = k \) is constant on \( D \) then \( f \) is constant on \( D \).

**Proof.** If \( k = 0 \) we are done, so assume that \( k > 0 \). By assumption

\[
0 = \partial^2 f = \partial |f|^2 = \partial (\bar{f} f) = \partial \bar{f} \cdot f + \bar{f} \partial f
\]

wherein we have used

\[
\partial \bar{f} = \frac{1}{2} (\partial_x - i \partial_y) \bar{f} = \frac{1}{2} (\partial_x + i \partial_y) f(z) = \bar{\partial f} = 0
\]

by the Cauchy Riemann equations. Hence \( f' = 0 \) and \( f \) is constant. ■

**Corollary 5.39 (Maximum modulus principle).** Let \( \Omega \) be a bounded region and \( f \in C(\overline{\Omega}) \cap H(\Omega) \). Then for all \( z \in \Omega \), \( |f(z)| \leq \sup_{z \in \partial \Omega} |f(z)| \). Furthermore if there exists \( z_0 \in \Omega \) such that \( |f(z_0)| = \sup_{z \in \partial \Omega} |f(z)| \) then \( f \) is constant.

**Proof.** Assume there exists \( z_0 \in \Omega \) and \( |f(z_0)| = \sup_{z \in \partial \Omega} |f(z)| =: M \). Then for \( \rho \) small, Corollary 5.37 implies

\[
M = |f(z_0)| \leq \frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + \rho e^{i\theta})| d\theta \leq M
\]

from which we conclude that \( |f(z_0 + \rho e^{i\theta})| = M = |f(z_0)| \) for all \( \theta \). (Why?) Since this is true for all \( \rho > 0 \) such that \( \overline{D(z_0, \rho)} \subset \Omega \), Lemma 5.38 shows \( f \) is constant on any such \( D(z_0, \rho) \). Since \( \Omega \) is connected, Corollary 5.33 implies that \( f \) is constant on \( \Omega \). ■

5.2. **Weak characterizations of** \( H(\Omega) \). The next theorem is the deepest theorem of this section.

**Theorem 5.40.** Let \( \Omega \subset \mathbb{C} \). Then \( f \in H(\Omega) \Rightarrow \oint_{\partial \Omega} f(z) dz = 0 \) for all solid triangles \( T \subseteq \Omega \).

**Proof.** Write \( T = S_1 \cup S_2 \cup S_3 \cup S_4 \) as in Figure 5.2 below.
Spliting $T$ into four similar triangles of equal size.

Let $T_1 \in \{S_1, S_2, S_3, S_4\}$ such that $|\int_{\partial T_i} f(z)dz| = \max\{|\int_{\partial S_i} f(z)dz| : i = 1, 2, 3, 4\}$, then

$$|\int_{\partial T} f(z)dz| = \left| \sum_{i=1}^{4} \int_{\partial S_i} f(z)dz \right| \leq \sum_{i=1}^{4} \left| \int_{\partial S_i} f(z)dz \right| \leq 4 \int_{\partial T_1} f(z)dz.$$

Repeating the above argument with $T$ replaced by $T_1$ again and again, we find by induction there are triangles $\{T_i\}_{i=1}^{\infty}$ such that

1. $T \supseteq T_1 \supseteq T_2 \supseteq T_3 \supseteq \ldots$
2. $\ell(\partial T_n) = 2^{-n} \ell(\partial T)$ where $\ell(\partial T)$ denotes the length of the boundary of $T$,
3. $\text{diam}(T_n) = 2^{-n} \text{diam}(T)$ and

$$|\int_{\partial T} f(z)dz| \leq 4^n |\int_{\partial T_n} f(z)dz|.$$  \hspace{1cm} (5.17)

By finite intersection property of compact sets there exists $z_0 \in \bigcap_{n=1}^{\infty} T_n$. Because

$$f(z) = f(z_0) + f'(z_0)(z-z_0) + o(z-z_0)$$

we find

$$4^n \left| \int_{\partial T_n} f(z)dz \right| = 4^n \left| \int_{\partial T_n} f(z_0)dz + \int_{\partial T_n} f'(z_0)(z-z_0)dz + \int_{\partial T_n} o(z-z_0)dz \right|$$

$$= 4^n \left| \int_{\partial T_n} o(z-z_0)dz \right| \leq C\epsilon_n 4^n \int_{\partial T_n} |z-z_0|d|z|$$

where $\epsilon_n \to 0$ as $n \to \infty$. Since

$$\int_{\partial T_n} |z-z_0|d|z| \leq \text{diam}(T_n)\ell(\partial T_n) = 2^{-n} \text{diam}(T)2^{-n} \ell(\partial T) = 4^{-n} \text{diam}(T)\ell(\partial T)$$
we see
\[ 4^n \left| \int_{\partial T_n} f(z)dz \right| \leq C \epsilon_n 4^n 4^{-n} \text{diam}(T) \ell(\partial T) = C \epsilon_n \to 0 \text{ as } n \to \infty. \]

Hence by Eq. (5.17), \( \int_{\partial T} f(z)dz = 0. \)

**Theorem 5.41** (Morera’s Theorem). Suppose that \( \Omega \subset \mathbb{C} \) and \( f \in C(\Omega) \) is a complex function such that

(5.18) \[ \int_{\partial T} f(z)dz = 0 \text{ for all solid triangles } T \subset \Omega, \]

then \( f \in H(\Omega). \)

**Proof.** Let \( D = D(z_0, \rho) \) be a disk such that \( \bar{D} \subset \Omega \) and for \( z \in D \) let

\[ F(z) = \int_{[z_0, z]} f(\xi)d\xi \]

where \([z_0, z]\) is by definition the contour, \( \sigma(t) = (1 - t)z_0 + tz \) for \( 0 \leq t \leq 1 \). For \( z, w \in D \) we have, using Eq. (5.18),

\[ F(w) - F(z) = \int_{[z, w]} f(\xi)d\xi = \int_0^1 f(z + t(w - z))(w - z)dt \]

\[ = (w - z) \int_0^1 f(z + t(w - z))dt. \]

From this equation and the dominated convergence theorem we learn that

\[ \frac{F(w) - F(z)}{w - z} = \int_0^1 f(z + t(w - z))dt \to f(z) \text{ as } w \to z. \]

Hence \( F' = f \) so that \( F \in H(D) \). Corollary 5.27 now implies \( f = F' \in H(D) \). Since \( D \) was an arbitrary disk contained in \( \Omega \) and the condition for being in \( H(\Omega) \) is local we conclude that \( f \in H(\Omega). \)

The method of the proof above also gives the following corollary.

**Corollary 5.42.** Suppose that \( \Omega \subset \mathbb{C} \) is convex open set. Then for every \( f \in H(\Omega) \) there exists \( F \in H(\Omega) \) such that \( F' = f \). In fact fixing a point \( z_0 \in \Omega \), we may define \( F \) by

\[ F(z) = \int_{[z_0, z]} f(\xi)d\xi \text{ for all } z \in \Omega. \]

**Exercise 5.43.** Let \( \Omega \subset \mathbb{C} \) and \( \{f_n\} \subset H(\Omega) \) be a sequence of functions such that \( f(z) = \lim_{n \to \infty} f_n(z) \) exists for all \( z \in \Omega \) and the convergence is uniform on compact subsets of \( \Omega \). Show \( f \in H(\Omega) \) and \( f'(z) = \lim_{n \to \infty} f_n'(z) \).

**Hint:** Use Morera’s theorem to show \( f \in H(\Omega) \) and then use Eq. (5.13) with \( n = 1 \) to prove \( f'(z) = \lim_{n \to \infty} f'_n(z). \)
Theorem 5.44. Let $\Omega \subset \mathbb{C}$ be an open set. Then
\begin{equation}
H(\Omega) = \left\{ f : \Omega \to \mathbb{C} \text{ such that } \frac{df(z)}{dz} \text{ exists for all } z \in \Omega \right\}.
\end{equation}
In other words, if $f : \Omega \to \mathbb{C}$ is complex differentiable at all points of $\Omega$ then $f'$ is automatically continuous and hence $C^\infty$ by Theorem 5.26!!!

Proof. Combine Theorems 5.40 and 5.41. ■

Corollary 5.45 (Removable singularities). Let $\Omega \subset \mathbb{C}$, $z_0 \in \Omega$ and $f \in H(\Omega \setminus \{z_0\})$.
If $\limsup_{z \to z_0} |f(z)| < \infty$, i.e. $\sup_{0<|z-z_0|<\epsilon} |f(z)| < \infty$ for some $\epsilon > 0$, then
$\lim_{z \to z_0} f(z)$ exists. Moreover if we extend $f$ to $\Omega$ by setting $f(z_0) = \lim_{z \to z_0} f(z)$, then
$f \in H(\Omega)$.

Proof. Set
\[ g(z) = \begin{cases} (z-z_0)^2 f(z) & \text{for } z \in \Omega \setminus \{z_0\} \\ 0 & \text{for } z = z_0 \end{cases} . \]
Then $g'(z_0)$ exists and is equal to zero. Therefore $g'(z)$ exists for all $z \in \Omega$ and hence $g \in H(\Omega)$. We may now expand $g$ into a power series using $g(z_0) = g'(z_0) = 0$ to learn $g(z) = \sum_{n=2}^{\infty} a_n (z-z_0)^n$ which implies
\[ f(z) = \frac{g(z)}{(z-z_0)^2} = \sum_{n=0}^{\infty} a_n (z-z_0)^{n-2} \text{ for } 0 < |z-z_0| < \epsilon. \]
Therefore, $\lim_{z \to z_0} f(z) = a_2$ exists. Defining $f(z_0) = a_2$ we have $f(z) = \sum_{n=0}^{\infty} a_n (z-z_0)^{n-2}$ for $z$ near $z_0$. This shows that $f$ is holomorphic in a neighborhood of $z_0$ and since $f$ was already holomorphic away from $z_0$, $f \in H(\Omega)$. ■

5.3. Homework #2 Due Friday April 20, 2001. Do the exercises from Section 5 of these notes. Only hand in Exercises: 5.17, 5.20, 5.29, 5.36 and 5.43.