Theorems and Definitions

Kyle Champley

Abstract

The following theorems and definitions were taken from lectures and textbooks on Real Analysis[1], Linear Algebra[2], Complex Analysis[3], Topology[4], and Probability[5]

[1]Lecture by Bob Higdon; Textbook by Royden
[2]Lecture by Tom Schmidt; Textbook by Friedberg, Insel, and Spence
[3]Lecture by Bent Petersen; Textbook by Bak and Newman
[4]Lecture by Bill Bogley; Textbook by Munkres
[5]Lecture by Ossiander; Textbook by Billingsley

Contents

1 Real Analysis

1.1 Set Theory and Basics

Completeness Axiom: Every nonempty set S of real numbers which has an upper bound has a least upper bound (but not necessarily in S).

Axiom of Archimedes: Given any real number x, there is an integer such that $x < n$.

Algebra: A collection of sets $\mathcal{A}$ of subsets of a space X is called an algebra of sets if (i) $A \cup B \in \mathcal{A}$, whenever $A, B \in \mathcal{A}$ and (ii) $A^c \in \mathcal{A}$, whenever $A \in \mathcal{A}$.

σ-algebra: An algebra $\mathcal{A}$ is a σ-algebra if every union of a countable collection of sets in $\mathcal{A}$ is again in $\mathcal{A}$. Note that from De Morgan's laws countable intersections of elements in $\mathcal{A}$ are also in $\mathcal{A}$.

Limits: L is a limit of $< x_n >$ iff $\forall \varepsilon > 0\ \exists N \in \mathbb{N}$ such that $|x_n - L| < \varepsilon\ \forall n \geq N$.

Cauchy Sequence: Given $\varepsilon > 0\ \exists N \in \mathbb{N}$ such that $\forall n \geq N$ and $\forall m \geq N$ $|x_n - x_m| < \varepsilon$.

Cauchy Criterion: A sequence of real numbers converges iff it is a Cauchy Sequence.

Supremum (least upper bound): The supremum of a set S is $\sup_{x \in S} = \sup\{x : x \in S\}$.

Infinum (greatest lower bound): The infinum of a set S is $\inf_{x \in S} = \inf\{x : x \in S\}$. Note that $\inf_{x \in S}\{x\} = -\sup_{x \in S}\{-x\}$.

Limit Superior: $\overline{\lim} x_n = \lim_{n \to \infty} \sup\{x_k : k \geq n\} = \inf_n \sup_{k \geq n} x_k$.

Limit Inferior: $\underline{\lim} x_n = \lim_{n \to \infty} \inf\{x_k : k \geq n\} = \sup_n \inf_{k \geq n} x_k$.

Open Set: A set O of real numbers is open iff $\forall x \in O\ \exists \delta > 0$ such that $\forall y$ with $|x - y| < \delta$ belongs to O.

Point of Closure: A real number x is a point of closure of a set E if $\forall \delta > 0\ \exists y \in E$ such that $|x - y| < \delta$. The set of points of closure of E is denoted $\overline{E}$ and $E \subset \overline{E}$. If $E = \overline{E}$, then E is a closed set.

Open Covering: A collection of open sets C is an open covering of a set F if $F \subset \cup\{O : O \in C$ and O is an open set$\}$.

Compact Set: A set K is compact iff every open covering of K has a finite subcovering. Also, $K \subset \mathbb{R}^n$ is compact iff K is closed and bounded.

Continuity: Suppose $E \subset \mathbb{R}$ and $f : E \to \mathbb{R}$. Then f is continuous at a point $x \in E$ iff $\forall \varepsilon > 0$ $\exists \delta > 0$ such that $|f(x) - f(y)| < \varepsilon\ \forall y \in E$ such that $|x - y| < \delta$.

Uniform Continuity: Suppose $E \subset \mathbb{R}$ and $f : E \to \mathbb{R}$. Then f is uniformly continuous iff $\forall \varepsilon > 0\ \exists \delta > 0$ such that $\forall x, y \in E$ with $|x - y| < \delta$ implies that $|f(x) - f(y)| < \varepsilon$.

Pointwise Convergence: The sequence $< f_n(x) >$ converges pointwise to $f(x)$ on E iff $\forall x \in E$ and $\forall \varepsilon > 0$ $\exists N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \varepsilon$ $\forall n \geq N$.

Uniform Convergence: The sequence $< f_n(x) >$ converges uniformly to $f(x)$ on E iff $\forall \varepsilon > 0$ $\exists N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \varepsilon$ $\forall n \geq N$ and $\forall x \in E$.

Borel Set: The collection $\mathcal{B}$ of Borel sets is the smallest σ-algebra which contains all the open sets.

Proposition 18: Let $f : \mathbb{R} \to \mathbb{R}$. Then f is continuous iff for each open set $O \subset \mathbb{R}$, $f^{-1}[O]$ is an open set.

Proposition 20: If a real-valued function f is defined and continuous on a compact set, F, of real numbers, then it is uniformly continuous on F.

Theorem: Every open set of real numbers is the union of a countable collection of disjoint of open intervals.

1.2 Lebesgue Measure

Measurability Axioms:
1) For an interval, I, $mI = l(I)$.
2) If $< E_n >$ is a sequence of disjoint sets (for which m is defined), then $m(\cup E_n) = \sum E_n$.
3) The measure is translation invariant, that is, $m(E + y) = m(E)$.

Outer Measure: Let $\{I_n\}$ be a countable collection of open intervals such that $A \subset \cup I_n$. Then the outer measure of A, denoted m^*A is: $m^*A = \inf_{A \subset \cup I_n} \sum l(I_n)$.

Properties of Outer Measure:
1) $m^*\emptyset = 0$
2) If $A \subset B$, then $m^*A \leq m^*B$.
3) If $\{A_n\}$ is a countable collection of sets of real numbers, then $m^*(\cup A_n) \leq \sum m^*A_n$.
4) If A is countable, then $m^*A = 0$.

Measurable: A set E is said to be measurable if for each set A we have $m^*A = m^*(A \cap E) + m^*(A \cap E^c)$. Note that we always have $m^*A \leq m^*(A \cap E) + m^*(A \cap E^c)$.

Properties of Measurable Sets:
1) If E_1 and E_2 are measurable, then so is $E_1 \cup E_2$ and E_i^c.
2) The family $\mathcal{M}$ of measurable sets is a σ-algebra.
3) The interval (a, ∞) is measurable.
4) Every Borel set is measurable. In particular each open set and each closed set is measurable.
5) Let $< E_i >$ be a sequence of measurable sets. Then $m(\cup E_i) \leq \sum mE_i$. If each E_i are disjoint, then $m(\cup E_i) = \sum m(E_i)$.
6) If a set E is measurable, then $m^*E = mE$.

Theorem: Suppose $E \subset F$ and E, F are measurable sets with finite measure. Then $m(F - E) = mF - mE$.

Proposition 14: Let $< E_n >$ be an infinite decreasing sequence of measurable sets, that is, a sequence with $E_{n+1} \subset E_n$ $\forall n$. Let $mE_1 < \infty$. Then $m(\cap E_i) = \lim_{n\to\infty} mE_n$.

Proposition 15: Let E be a given set, then the following are equivalent:
1) E is measurable.
2) Given $\varepsilon > 0$, there is an open set $O \supset E$ with $m^*(O - E) < \varepsilon$.
3) Given $\varepsilon > 0$, there is a closed set $F \subset E$ with $m^*(E - F) < \varepsilon$.
4) There is a G in G_δ with $E \subset G$ and $m^*(G - E) = 0$.
5) There is a F is F_σ with $F \subset E$ and $m^*(E - F) = 0$.
6) If $m^*E < \infty$, then the above are equivalent to: Given $\varepsilon > 0$, $\exists$ a finite union U of open intervals such that $m^*(U \Delta E) < \varepsilon$.

Proposition 18: Let f be an extended real-valued function whose domain is measurable. Then the following a equivalent.
1) $\forall \alpha \in \mathbb{R}$, $\{x : f(x) > \alpha\}$ is measurable.
2) $\forall \alpha \in \mathbb{R}$, $\{x : f(x) \geq \alpha\}$ is measurable.
3) $\forall \alpha \in \mathbb{R}$, $\{x : f(x) < \alpha\}$ is measurable.
4) $\forall \alpha \in \mathbb{R}$, $\{x : f(x) \leq \alpha\}$ is measurable.
If any of the above hold, then f is said to be measurable.

Proposition 19: Let $c \in \mathbb{R}$ and f and g be two measurable real-valued functions defined on the same domain. Then $f + c$, cf, $f + g$, $g - f$, and fg are also measurable.

Almost Everywhere: A property is said to hold almost everywhere, denoted a.e., if the set of points where it fails to hold is a set of measure zero.

Characteristic Function: If A is any set, we define the characteristic function χ_A of the set A to be the function given by $\chi_A(x) = \begin{cases} 1, \text{ if } x \in A \\ 0, \text{ if } x \notin A. \end{cases}$

Simple Function: A real-valued function φ, is called a simple if it is measurable and assumes only a finite number of values $\{\alpha_1, \ldots, \alpha_n\}$ i.e. $\varphi = \sum \alpha_i \chi_{A_i}$, where $A_i = \{x : \varphi(x) = \alpha_i\}$.

Theorem 20: Let $< f_n >$ be a sequence of measurable functions. Then the functions $\sup_n f_n$, $\inf_n f_n$, $\underline{\lim} f_n$, and $\overline{\lim} f_n$ are all measurable.

Proposition 21: If f is a measurable function and $f = g$ a.e., then g is measurable.

Proposition 22: Let f be a measurable function on $[a, b]$ and $m\{x : f(x) = \pm\infty\} = 0$. Then given $\varepsilon > 0$ we can find a step function φ and a continuous function h such that $|f - \varphi| < \varepsilon$ and $|f - h| < \varepsilon$ expect on a set of measure less than ε i.e. $m\{x : |f - h| \geq \varepsilon\} < \varepsilon$.

Egoroff's Theorem: If $< f_n >$ is a sequence of measurable functions such that $f_n \to f$ a.e. on a measurable set E of finite measure, then $\forall \delta > 0$ $\exists A \subset E$ with $mA < \delta$ such that f_n converges uniformly to f on $E - A$.

1.3 Lebesgue Integration

Riemann Integral: $R\int_a^b f(x)\,dx = \sup \sum_{i=1}^n (\xi_i - \xi_{i-1})m_i = \inf \sum_{i=1}^n (\xi_i - \xi_{i-1})M_i$, where $m_i = \inf_{\xi_{i-1}<x<\xi_i} f(x)$ and $M_i = \sup_{\xi_{i-1}<x<\xi_i} f(x)$

Proposition 3: Let f be defined and bounded on a measurable set E and $mE < \infty$. Let ψ and φ be simple functions. Then $\inf_{f\le\psi} \int_E \psi(x)\,dx = \sup_{f\ge\varphi} \int_E \varphi(x)\,dx$ iff f is measurable.

Lebesgue Integral(1): If f is a bounded measurable function defined on a measurable set E with $mE < \infty$, we define the Lebesgue integral of f over E by $\int_E f = \inf \int_E \psi$, where ψ is simple and $\psi \ge f$.

Proposition 4: Let f be bounded and defined on $[a,b]$. If f is Riemann integrable on $[a,b]$, then it is measurable and $R\int_a^b f = \int_a^b f$.

Proposition 5: If f and g are bounded, measurable, and defined on a measurable set E and $mE < \infty$, then
1) $\int_E af + bg = a\int_E f + b\int_E g$
2) If $f = g$ a.e., then $\int_E f = \int_E g$.
3) If $f \le g$ a.e., then $\int_E f \le \int_E g$.
4) If $A \le f(x) \le B$, then $AmE \le \int_E f \le BmE$.
5) If $A \cap B = \emptyset$ and $mA < \infty$, $mB < \infty$, then $\int_{A\cup B} f = \int_A f + \int_B f$.

Bounded Convergence Theorem: Let $< f_n >$ be a sequence of measurable functions defined on E where $mE < \infty$ and suppose $\exists M$ such that $|f_n(x)| \le M$ $\forall n$, $\forall x \in E$. If $\lim f_n(x) = f(x)$ $\forall x \in E$, then $\int_E f = \lim \int_E f_n$.

Proposition 7: A bounded function f on $[a,b]$ is Riemann integrable iff the set of points at which f is discontinuous has measure zero.

Lebesgue Integral(2): If $f \ge 0$ and is defined on a measurable set E, we define $\int_E f = \sup_{h\le f} \int_E h$, where h is a bounded measurable function such that $m\{x : h(x) \ne 0\} < \infty$.

Proposition 8: If f and g are nonnegative, measurable functions and a, b are nonnegative constants, then (1) $\int_E af + bg = a\int_E f + b\int_E g$ and (2) If $f \le g$, then $\int_E f \le \int_E g$.

Fatou's Lemma: If $< f_n >$ is a sequence of nonnegative measurable functions and $\lim f_n(x) = f(x)$ a.e. on a measurable set E, then $\int_E f \le \underline{\lim} \int_E f_n$.

Monotone Convergence Theorem: Let $< f_n >$ be a sequence of nonnegative measurable functions such that $\lim f_n(x) = f(x)$ and $f_n(x) \le f(x)$ a.e., then $\int f = \lim \int f_n$.

Corollary 11: Let u_n be a sequence of nonnegative measurable functions, and let $f = \sum_{n=1}^\infty u_n$. Then $\int f = \sum_{n=1}^\infty \int u_n$.

Proposition 12: Let f be a nonnegative function and $< E_i >$ be a disjoint sequence of measurable sets. Let $E = \cup E_i$. Then $\int_E f = \sum \int_{E_i} f$.

Integrable: A nonnegative measurable function f is called integrable over the measurable set

E if $\int_E f < \infty$.

Proposition 14: Let f be a nonnegative measurable function which is integrable over a set E. Then given $\varepsilon > 0$ $\exists \delta > 0$ such that for every set $A \subset E$ with $mA < \delta$ we have $\int_A f < \epsilon$.

Lebesgue Dominated Convergence Theorem: Let g be integrable over E and $< f_n >$ be a measurable sequence such that $|f_n| \leq g$ and $f_n \to f$ a.e. on E. Then $\int_E f = \lim \int_E f_n$.

Theorem 17: Let $< g_n >$ be integrable functions that converge to an integrable function g. Let $< f_n >$ be a measurable sequence such that $|f_n| \leq g_n$ and $f_n \to f$ a.e. If $\int g = \lim \int g_n$, then $\int f = \lim \int f_n$.

Convergence in Measure: A sequence $< f_n >$ of measurable functions is said to converge to f in measure if, given $\varepsilon > 0$, there is an N such that for all $n \geq N$ we have $m\{x : |f(x) - f_n(x)| \geq \varepsilon\} < \varepsilon$.

Proposition 18: Let $< f_n >$ be a sequence of measurable functions that converges in measure to f. Then there is a subsequence $< f_{n_k} >$ that converges to f almost everywhere.

Corollary 19: Let $< f_n >$ be a sequence of measurable functions defined on a measurable set E of finite measure. Then $< f_n >$ converges to f in measure iff every subsequence of $< f_n >$ has in turn a subsequence that converges almost everywhere to f.

Proposition 20: Fatou's Lemma and the Monotone and Lebesgue Convergence Theorem's remain valid if 'convergence a.e.' is replaced by 'convergence in measure'.

1.4 Differentiation of the Integral

Theorem 3: Let f be an increasing function on $[a, b]$. Then f is differentiable a.e., the derivative f' is measurable, and $\int_a^b f' \leq f(b) - f(a)$.

Bounded Variation: Let f be a function defined on $[a, b]$ and let $a = x_0 < x_1 < \cdots < x_k = b$ be any finite subdivision of $[a, b]$. Define $p = \sum_{i=1}^k [f(x_i) - f(x_{i-1})]^+$, $n = \sum_{i=1}^k [f(x_i) - f(x_{i-1})]^-$, and $t = n + p = \sum_{i=1}^k |f(x_i) - f(x_{i-1})|$. Also let $P = \sup p$, $N = \sup n$, and $T = \sup t$. If $T < \infty$, then we say that f is of bounded variation, denoted $f \in BV$.

Lemma 4: If $f \in BV([a, b])$, then $T_a^b = P_a^b + N_a^b$ and $f(b) - f(a) = P_a^b - N_a^b$.

Theorem 5: A function $f \in BV([a, b])$ iff f is the difference of two monotone increasing real-valued functions on $[a, b]$.

Corollary 6: If $f \in BV([a, b])$, then $f'(x)$ exists a.e. on $[a, b]$.

Absolute Continuity: A function f is absolutely continuous if $\forall \varepsilon > 0$ $\exists \delta > 0$ such that $\sum_{i=1}^n |f(x_i') - f(x_i)| < \varepsilon$ for every finite collection $\{(x_i, x_i')\}$ of non-overlapping intervals with $\sum_{i=1}^n |x_i' - x_i| < \delta$.

Lemma 7: If f is integrable on $[a,b]$, then F, defined by $F(x) = \int_a^x f(t)\,dt$ is a continuous function of bounded variation on $[a,b]$.

Lemma 8: If f is integrable on $[a,b]$ and $\int_a^x f(t)\,dt = 0\ \forall x \in [a,b]$, then $f(t) = 0$ a.e. on $[a,b]$.

Lemma 9: If f is bounded and measurable on $[a,b]$ and $F(x) = \int_a^x f(t)\,dt + F(a)$, then $F' = f$ a.e. on $[a,b]$.

Theorem 10: Let f be an integrable function on $[a,b]$, and suppose that $F(x) = \int_a^x f(t)\,dt + F(a)$. Then $F' = f$ a.e. on $[a,b]$.

Lemma 11: If f is absolutely continuous on $[a,b]$, then $f \in BV([a,b])$.

Lemma 13: If f is absolutely continuous on $[a,b]$ and $f' = 0$ a.e., then f is constant.

Theorem 14: A function F is an indefinite integral iff it is absolutely continuous.

Corollary 15: Every absolutely continuous function is the indefinite integral of its derivative.

Convex Function: Let $0 \le \lambda \le 1$. Then $\varphi(\lambda x + (1-\lambda)y) \le \lambda\varphi(x) + (1-\lambda)\varphi(y)\ \forall x, y \in (a,b)$ iff φ is convex.

Jensen's Inequality: Let φ be a convex function on $\mathbb{R}$ and f be an integrable function on $[0,1]$. Then $\int \varphi(f(t))\,dt \ge \varphi[\int f]$.

1.5 L^p Spaces

L^p Norms: Let E be a measurable set. Then $||f||_{L^p(E)} = \left(\int_E |f|^p\right)^{1/p}$, for $0 < p < \infty$ and $||f||_{L^\infty(E)} = \inf\{M : m\{t : f(t) > M\} = 0\}$.

Properties of L^p Norms: Let $0 < p \le \infty$. If $\alpha \in \mathbb{R}$, then $||\alpha f|| = |\alpha| \cdot ||f||$, $||f|| \ge 0$, and $||f|| = 0$ iff f $= 0$ a.e.

Minkowski Inequality: If $f, g \in L^p$ for $1 \le p \le \infty$, then $||f + g|| \le ||f|| + ||g||$.

Holder Inequality: Let $0 \le p, q$, $\frac{1}{p} + \frac{1}{q} = 1$, and $f \in L^p$, $g \in L^q$, then $fg \in L^1$ and $\int |fg| \le ||f||_{L^p} \cdot ||g||_{L^q}$.

1.6 Banach Spaces

Linear Space: A set of elements, X, is a linear space iff $\forall f, g \in X$ and $\forall \alpha, \beta \in \mathbb{R}$ we have $\alpha f + \beta g \in X$.

Normed Linear Space: A linear space, X, is a normed linear space iff $\exists$ a norm $||\cdot|| : X \to \mathbb{R}$ such that $\forall f \in X$ (1) $||f|| \ge 0$, (2) $||f|| = 0$ iff $f = 0$, (3) $||\alpha f|| = |\alpha| \cdot ||f||\ \forall \alpha \in \mathbb{R}$, and (4) $||f + g|| \le ||f|| + ||g||$.

Convergence in a Normed Linear Space: A sequence $< f_n >$ in a normed linear space is said to converge to an element f in the space if $\forall \varepsilon > 0 \ \exists N \in \mathbb{N}$ such that $\forall n \geq N$ we have $||f - f_n|| < \varepsilon$.

Completeness: A normed linear space is complete if every Cauchy Sequence in the space converges.

Banach Space: A Banach Space is a complete normed linear space.

Proposition 5: A normed linear space X is complete iff every absolutely summable series is summable (f_n is absolutely summable iff $\sum_{n=0}^{\infty} ||f_n|| < \infty$).

Riesz-Fischer Theorem: The L^p spaces are complete.

Compact Support: The support of a function is the set $\{x : f(x) \neq 0\}$, denoted supp$f$. We say that f has compact support if the closure of supp$f$ is a compact set.

Density in L^p: Functions that are bounded, bounded with compact support, simple with compact support, step with compact support, continuous, smooth, and smooth with compact support are all dense in $L^p(\mathbb{R})$ for $1 \leq p < \infty$.

Linear Functional: A linear functional on a normed linear space X is a function $F : X \to \mathbb{R}$ such that $F(\alpha f + \beta g) = \alpha F(f) + \beta F(g) \ \forall f, g \in X$ and $\forall \alpha, \beta \in \mathbb{R}$.

Bounded Linear Functional: A linear functional $F : X \to \mathbb{R}$ is bounded iff $\exists M \in \mathbb{R}$ such that $|F(f)| \leq M||f|| \ \forall f \in X$.

Norm of a Linear Functional: We define a norm on $\mathcal{L}(X, \mathbb{R})$ by $||F|| = \sup_{f \neq 0} \frac{|F(f)|}{||f||} = \sup_{||f||=1} |F(f)|$. Note that $||F||$ and $||f||$ are different norms and $||F||$ exists only if F is a bounded linear functional.

Proposition 11: Let $p, q \geq 0$ and $\frac{1}{p} + \frac{1}{q} = 1$. Each $g \in L^q$ defines a bounded linear functional, F, on L^p by $F(f) = \int fg$. We also have $||F|| = ||g||_{L^q}$.

Lemma 12: Let $g \in L^1$ and suppose that $\exists M$ such that $|\int fg| \leq M||f||_p$ for all bounded measurable functions f. Then $g \in L^q$ and $||g||_{L^q} \leq M$.

Riesz Representation Theorem: Let F be a bounded linear functional on L^p for $1 \leq p < \infty$. Then $\exists g \in L^q$ such that $F(f) = \int fg$. We also have $||F|| = ||g||_{L^q}$.

1.7 Metric Spaces

Metric Space: A metric space (X, ρ) is a nonempty set X of elements with a real-valued function, $\rho : X \times X \to \mathbb{R}$ such that $\forall x, y, z \in X$ (1) $\rho(x, y) \geq 0$, (2) $\rho(x, y) = 0$ iff $x = y$, (3) $\rho(x, y) = \rho(y, x)$, and (4) $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$.

Open Sets: A set O is open iff $\forall x \in O \ \exists \delta > 0$ such that $\rho(x, y) < \delta$ implies that $y \in O$.

Point of Closure: A point $x \in X$ is called a point of closure of the set E iff $\forall \delta > 0\ \exists y \in E$ such that $\rho(x, y) < \delta$.

Proposition 2: If $A \subset B$, then $\overline{A} \subset \overline{B}$. Also, $\overline{(A \cup B)} = \overline{A} \cup \overline{B}$ and $\overline{(A \cap B)} \subset \overline{A} \cap \overline{B}$.

Dense: A subset S of a metric space (X, ρ) is dense in X iff $\overline{S} = X$ iff $\forall x \in X\ \forall \varepsilon > 0\ \exists s \in S$ such that $\rho(x, s) < \varepsilon$ iff $S \cap O \neq \emptyset$ for all open sets $O \subset X$.

Separable: A metric space is separable if it has a subset D which has a countable number of points and is dense in X.

Proposition 6: A metric space X is separable iff $\exists \{O_i\}_{i=1}^{\infty}$ of open sets such that for any open set $O \subset X$, $O = \cup_{O_i \subset O} O_i$.

Proposition 8: If $f : X \to Y$ and $g : Y \to Z$ are continuous, then $g \circ f : X \to Z$ is continuous.

Homeomorphism: Suppose that (X, ρ) and (Y, σ) are metric spaces. A homeomorphism between X and Y is a function $f : X \to Y$ such that (1) f is a bijection, (2) f is continuous on X, and (3) f^{-1} is continuous on Y.

Isometry: With f as above, if we also have $\sigma[f(x_1), f(x_2)] = \rho(x_1, x_2)\ \forall x_1, x_2 \in X$, then f is an isometry between X and Y.

Equivalent Metrics: Metrics ρ and σ on a set X are equivalent iff ρ and σ define the same open sets.

Equivalent Norms: Norms $|| \cdot ||$ and $|| \cdot ||^*$ on a normed linear space X are equivalent iff $\exists c_1, c_2 > 0$ such that $c_1 ||x|| \leq ||x||^* \leq c_2 ||x||\ \forall x \in X$.

Theorem 9: If (X, ρ) is an incomplete metric space, it is possible to find a complete metric space X^* in which X is isometrically embedded as a dense subset. If X is contained in an arbitrary complete metric space Y, then X^* is isometric with the closure of X in Y.

Proposition 12: Let X be a metric space and S a subspace of it. Then the closure of E relative to S is $\overline{E} \cap S$, where $\overline{E}$ denotes the closure of E in X. A set $A \subset S$ is closed relative to S iff $A = S \cap F$ with F closed in X. A set $A \subset S$ is open relative to S iff $A = S \cap O$ with O open in X.

Proposition 13: Every subspace of a separable metric space is separable.

Proposition 14: If a subset A of a metric space X is complete, then it is closed. Also, a closed subset of a complete metric space is itself complete.

Sequentially Compact: A space X is sequentially compact iff every sequence $< x_n >$ in X has a subsequence $< x_{n_k} >$ which converges to an element of X.

Bolzano-Weierstrauss Property: A space X is said to have the Bolzano-Weierstrauss property if every infinite sequence $< x_n >$ in X has at least one cluster point.

Theorem 18: Let f be a continuous real-valued function on a compact space. Then f is bounded and assumes its maximum and minimum.

Totally Bounded: A metric space is totally bounded if, for each $\varepsilon > 0$, there is a finite collection of points $\{x_1, \dots, x_n\}$ such that each $x \in X$ is within a distance of ε of one of the x_k.

Lemma 19: A sequentially compact metric space is totally bounded.

Theorem 21 (Borel-Lebesgue): Let X be a metric space. Then the following are equivalent: (1) X is compact, (2) X has the Bolzano-Weierstrauss property, and (3) X is sequentially compact.

Proposition 22: A closed subset of a compact space is compact. A compact subset of a metric space is closed and bounded.

Proposition 24: The continuous image of a compact set is compact.

Proposition 25: A metric space X is compact iff it is both complete and totally bounded.

Proposition 26: Let f be a continuous mapping of a compact metric space X into a metric space Y. Then f is uniformly continuous.

Theorem 27 (Baire): Let X be a complete metric space and $\{O_k\}$ a countable collection of dense open subsets of X. Then $\cap O_n$ is dense.

Corollary: Suppose (X, ρ) is a complete metric space and $X = \cup F_n$, where F_n is closed $\forall n$. Then $\exists N$ such that F_N contains a nonempty open set.

Nowhere Dense: Suppose (X, ρ) is a complete metric space. A set $E \subset X$ is nowhere dense in X iff $X - \overline{E}$ is dense in X.

Proposition: A subset E of a complete metric space is nowhere dense iff $\overline{E}$ contains no nonempty open sets.

Baire Category: A subset E of a complete metric space X is of first category (or meager) iff E is the union of countable many nowhere dense sets, otherwise E is said to be of second category.

Baire Category Theory: Let (X, ρ) be a complete metric space. Then every nonempty open subset of X is of second category.

Theorem 32 (Uniform Boundedness Principle): Let (X, ρ) be a complete metric space and $\mathcal{F}$ a family of continuous real-valued functions on X. Suppose $\forall x \in X\ \exists M_x \in \mathbb{R}$ such that $|f(x)| \leq M_x\ \forall f \in \mathcal{F}$. Then there exists a nonempty open subset O of X and $\exists M \in \mathbb{R}$ such that $|f(x)| \leq M\ \forall x \in O$ and $\forall f \in \mathcal{F}$.

Theorem: Suppose X is a Banach space and $\mathcal{F}$ a family of bounded linear functionals on X. Suppose $\forall x \in X\ \exists M_x \in R$ such that $|F(x)| \leq M_x\ \forall F \in \mathcal{F}$. Then the functionals in $\mathcal{F}$ are uniformly bounded i.e. $\exists M \in \mathbb{R}$ such that $||F|| \leq M\ \forall F \in \mathcal{F}$.

Equicontinuity: Suppose that (X,ρ) and (Y,σ) are metric spaces and $\mathcal{F}$ is a family of functions from X into Y. Then $\mathcal{F}$ is equicontinuous at a point $x \in X$ iff $\forall \varepsilon > 0\ \exists \delta > 0$ such that $\forall y \in X$ with $\rho(x,y) < \delta$ and $\forall f \in \mathcal{F}$, $\sigma(f(x), f(y)) < \varepsilon$.

Ascoli-Arzelá Theorem: Let $\mathcal{F}$ be an equicontinuous family of functions from X into Y, where (X,ρ) is a separable metric space and (Y,σ) is a metric space. Let $< f_n >$ be a sequence in $\mathcal{F}$ such that $\forall x \in X$ the closure of $\{f_n(x) : n \geq 1\}$ is compact. Then there exists a subsequence $< f_{n_k} >$ which converges pointwise to a continuous function f on X and $f_n \to f$ uniformly on any compact subset of X.

Contraction Mapping: Let (X,ρ) be a metric space. A function $g : X \to X$ is a contraction mapping iff $\exists \lambda$ with $0 < \lambda < 1$ such that $\rho(g(x), g(y)) \leq \lambda\rho(x,y)\ \forall x, y \in X$.

Fixed Point Theorem: Let (X,ρ) be a complete metric space and let $g : X \to X$ be a contraction mapping. Then (1) exists a unique $\alpha \in X$ such that $g(\alpha) = \alpha$ and (2) let $< x_n >$ be a sequence in X such that $x_{n+1} = g(x_n)\ \forall n \geq 1$, then $x_n \to \alpha$.

1.8 Measure and Integration

Measure Space: A measure space is a pair $(X, \mathcal{B})$, where X is a set and $\mathcal{B}$ is a σ-algebra of subsets of X. A set $A \subset X$ is measurable iff $A \in \mathcal{B}$.

Measure: A measure is a nonnegative extended real-valued function on $\mathcal{B}$, μ, such that $\mu(\emptyset) = 0$ and $\mu\left(\cup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i)$ for all collections $\{E_i\}$ of disjoint sets in $\mathcal{B}$. The triple $(X, \mathcal{B}, \mu)$ is a measure space.

Proposition 1: If $A \in \mathcal{B}$, $B \in \mathcal{B}$, and $A \subset B$, then $\mu A \leq \mu B$.

Proposition 2: If $E_i \in \mathcal{B}$, $\mu E_1 < \infty$, and $E_i \supset E_{i+1}$, then $\mu\left(\cap_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} \mu E_n$.

Proposition 3: If $E_i \in \mathcal{B}$, then $\mu\left(\cup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \mu E_i$.

Finite Measure Space: The triple $(X, \mathcal{B}, \mu)$ is a finite measure space if $\mu(X) < \infty$.

σ-Finite Measure Space: Let the triple $(X, \mathcal{B}, \mu)$ be a measure space. If $\exists < E_n >$ with $E_n \in \mathcal{B}$ such that $X = \cup_{i=1}^{\infty} E_n$ and $\mu E_n < \infty\ \forall n$, then $(X, \mathcal{B}, \mu)$ is a σ-finite measure space.

Semifinite Measure: If each measurable set of infinite measure contains measurable sets of arbitrarily large finite measure, then the measure is said to be a semifinite measure.

Complete Measure Space: Let $(X, \mathcal{B}, \mu)$ be a measure space. If $\mathcal{B}$ contains all subsets of sets of measure zero, then $(X, \mathcal{B}, \mu)$ is a complete measure space.

Proposition: Let $(X, \mathcal{B}, \mu)$ be a measure space. A function $f : X \to X$ is a measurable function iff $f^{-1}(I) \in \mathcal{B}$ for all intervals $I \subset X$.

Product Measure: Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces. We define the product measure on $X \times Y$ as follows. If $A \in \mathcal{A}$ and $B \in \mathcal{B}$, then $\lambda(A \times B) = \mu A \cdot \nu B$.

Fubini's Theorem: Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be complete measure spaces. Suppose $f : X \times Y \to \mathbb{R}$ is integrable. Then
(1) for any $x \in X$, define $f_x : Y \to \mathbb{R}$ by $f_x(y) = f(x, y)$ $\forall y \in Y$. Then for a.e. $x \in X$, f_x is integrable on Y.
(2) The function defined a.e. in X by $\int_Y f(x, y)\, d\nu(y)$ is integrable on X.
(3) We have $\int_{X\times Y} f\, d(\mu \times \nu) = \int_X \left[\int_Y f(x, y)\, d\nu(y)\right] d\mu(x) = \int_Y \left[\int_X f(x, y)\, d\mu(x)\right] d\nu(y)$.
Note that in (1) and (2) we may interchange the roles of x and y.

Tonelli's Theorem: If we replace complete with σ-finite and integrable with nonnegative and measurable in Fubini's Theorem, then we have Tonelli's Theorem.

1.9 Hilbert Spaces

Inner Product Space: An inner product space H is a vector space with a function $(\cdot, \cdot) : H \times H \to \mathbb{C}$ such that:
(1) $(x, x) \geq 0$ $\forall x \in H$ and $(x, x) = 0$ iff $x = 0_H$
(2) $(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1(x_1, y) + \alpha_2(x_2, y)$ $\forall x_1, x_2, y \in H$ and $\forall \alpha_1, \alpha_2 \in \mathbb{C}$
(3) $(x, y) = \overline{(y, x)}$ $\forall x, y \in H$
Define $||x|| = \sqrt{(x, x)}$ $\forall x \in H$.

Cauchy-Schwarz Inequality: $|(x, y)| \leq ||x|| \cdot ||y||$ $\forall x, y \in H$

Hilbert Space: An inner product space which is complete with respect to the norm induced by the inner product is called a Hilbert Space.

Orthogonal: Elements $x, y \in H$ are orthogonal iff $(x, y) = 0$.

Orthonormal System: A subset $\mathcal{S}$ of H is called an orthonormal system iff $(\varphi, \psi) = 0$ and $||\varphi|| = ||\psi|| = 1$ $\forall \varphi, \psi \in \mathcal{S}$ with $\varphi \neq \psi$.

Proposition: Let H be an inner product space and $\mathcal{S}$ an orthonormal system in H. Then any two distinct elements of $\mathcal{S}$ are a distance $\sqrt{2}$ from each other. Also, if H is separable, then $\mathcal{S}$ is countable.

Complete Orthonormal System: An orthonormal system $\mathcal{S}$ in an inner product space H is complete if $(z, \varphi) = 0$ $\forall \varphi \in \mathcal{S}$ implies $z \equiv 0 \in H$.

Proposition: Let $\mathcal{S} = \{\varphi_\nu\}$ be an orthonormal system in an inner product space H, and let $\mathcal{S}_N = \text{span}\{\varphi_1, \ldots, \varphi_N\}$. Let $x \in H$. Let $y = \sum_{i=1}^{N}(x, \varphi_\nu)\varphi_\nu$. Then $||x - y|| < ||x - z||$ $\forall z \in \mathcal{S}_N \backslash \{y\}$.

Bessel's Inequality: Let $\mathcal{S} = \{\varphi_\nu\}$ be an countably infinite orthonormal system in an inner product space H, $x \in H$, and $a_\nu = (x, \varphi_\nu)$ $\forall \nu$. Then $\sum_{\nu=1}^{\infty} |a_\nu|^2 \leq ||x||^2$.

Parseval's Theorem: Suppose H is a Hilbert Space and $\mathcal{S} = \{\varphi_\nu\}$ is a complete orthonormal system in H. Then $\forall x \in H$, $x = \sum_{i=1}^{\infty} a_\nu \varphi_\nu$, where $a_\nu = (x, \varphi_\nu)$ $\forall \nu$ and $||x||^2 = \sum_{i=1}^{\infty} |a_\nu|^2$.

Complex Fourier Series in $L^2([-\pi, \pi])$: $f(x) = \sum_{-\infty}^{\infty} \widehat{f}(n)e^{inx}$, where $\widehat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\, dx$ $\forall n \in \mathbb{Z}$.

2 Linear Algebra

2.1 Vector Spaces

Vector Space (linear space): A set V is a vector space over a field $\mathbb{F}$ if
(1) V is an abelian group under $+$
(2) $1 \cdot x = x$ for $1 \in \mathbb{F}$ and $x \in V$
(3) $(ab)x = a(bx)$ for $a, b \in \mathbb{F}$ and $x \in V$
(4) $a(x + y) = ax + ay$ for $a \in \mathbb{F}$ and $x, y \in V$
(5) $(a + b)x = ax + bx$ for $a, b \in \mathbb{F}$ and $x \in V$

Subspace: A subset W of a vector space V over of field $\mathbb{F}$ is called a subspace of V if (1) $0 \in W$ (2) $x + y \in W$, whenever $x, y \in W$, and (3) $cx \in W$ whenever $c \in \mathbb{F}$ and $x \in W$.

Proposition: Let $W \subset V$ as vector spaces. If $\dim(W) = \dim(V)$, then $W = V$.

2.2 Linear Transformations and Matrices

Linear Transformation: Let V be a vector space over $\mathbb{F}$ and let W be vector space. Then a function $T : V \to W$ is a linear transformation if $T(ax + by) = aT(x) + bT(y)$ $\forall a, b \in \mathbb{F}$ and $\forall x, y \in V$.

Null Space: Let V and W be vector spaces, and let $T : V \to W$ be linear. We define the null space (or kernel) of T by $N_T = \{x \in V : T(x) = 0\}$.

Range Space: Let V and W be vector spaces, and let $T : V \to W$ be linear. We define the range space of T by $R_T = \{T(x) : x \in V\}$.

Theorem 2.2: Let V and W be vector spaces and let $T : V \to W$ be linear. If $\beta = \{v_1, \ldots, v_n\}$ is a basis for V, then $R_T = \text{span}(\{T(v_1), \ldots, T(v_n)\})$.

Rank-Nullity Theorem: Let V and W be vector spaces and let $T : V \to W$ be linear. If V is finite-dimensional, then $\text{nullity}(T) + \text{rank}(T) = \dim(V)$.

Theorem 2.4: Let V and W be vector spaces, and let $T : V \to W$ be linear. Then T is one-to-one iff $N_T = \{0\}$ iff $T(x) = T(y)$ implies $x = y$.

Theorem 2.5: Let V and W be vector spaces of equal (finite) dimension, and let $T : V \to W$ be linear. Then T is one-to-one iff T is onto.

Space of Linear Transformations: Let V and W be vector spaces. We denote the vector space of all linear transformations from V into W by $\mathcal{L}(V, W)$. Then $\dim(\mathcal{L}(V, W)) = \dim(V) \cdot \dim(W)$.

Theorem 2.11: Let V, W, Z be finite-dimensional vector spaces with ordered bases α, β, γ, respectively. Let $T : V \to W$ and $U : W \to Z$ be linear. Then $[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta$.

Theorem 2.14: Let V and W be finite-dimensional vector spaces with ordered bases β and γ, respectively and let $T : V \to W$ be linear. Then for each $u \in V$ we have $[T(u)]_\gamma = [T]_\beta^\gamma [u]_\beta$.

Left-Multiplication Transformation: Let A be an m $\times$ n matrix with entries from a field $\mathbb{F}$. We denote L_A the mapping $L_A : \mathbb{F}^n \to \mathbb{F}^m$ defined by $L_A(x) = Ax$ for each $x \in \mathbb{F}^n$. We call L_A a left-multiplication transformation. Let β, γ be the standard ordered basis for $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively. Then (1) $[L_A]_\beta^\gamma = A$, (2) $L_A = L_B$ iff $A = B$, (3) $L_{A+B} = L_A + L_B$ and $L_{aA} = aL_A$, $a \in \mathbb{F}$, and (4) $L_{AB} = L_A L_B$.

Inverses: Let V and W be vector spaces and let $T : V \to W$ be linear. A function $U : W \to V$ is said to be an inverse of T if $TU = I_W$ and $UT = I_V$. If T has an inverse, then T is invertible. If T and U are invertible functions, then $(TU)^{-1} = U^{-1}T^{-1}$.

Lemma: Let $T : V \to W$ be linear, where V and W are finite-dimensional vector spaces. If T is invertible, then $\dim(V) = \dim(W)$.

Lemma: Let $T : V \to W$ be linear, where V and W are finite-dimensional vector spaces. Then T is invertible iff T is one-to-one and onto.

Isomorphisms: Let V and W be vector spaces. We say that V is isomorphic to W if there exists a linear transformation $T : V \to W$ that is invertible. Here, T is an isomorphism from V to W.

Theorem 2.22: Let β and β' be two ordered bases for a finite-dimensional vector space V, and let $Q = [I_V]_{\beta'}^\beta$. Then (1) Q is invertible and (2) $[v]_\beta = Q[v]_{\beta'}$ $\forall v \in V$. Here, Q is called the change of coordinates matrix.

Linear Operator: A linear transformation that maps a vector space V into itself is called a linear operator on V.

Theorem 2.23: Let T be a linear operator on a finite-dimensional vector space V, and let β and β' be ordered basis for V. Let $Q = [I_V]_{\beta'}^\beta$. Then $[T]_{\beta'} = Q^{-1}[T]_\beta Q$.

Similar Matricies: Let $A, B \in \mathcal{M}_{n\times n}(\mathbb{F})$. We say that B is similar to A if there exsits an invertible matrix Q such that $B = Q^{-1}AQ$.

Linear Functional: A linear transformation from a vector space V into its field of scalars $\mathbb{F}$, which is itself a vector space of dimension 1 over itself, is called a linear functional on V.

Dual Space: For a vector space V over $\mathbb{F}$, we define the dual space of V to be the vector space $\mathcal{L}(V, \mathbb{F})$ denoted V^*. Note that $\dim(V) = \dim(V^*)$.

Theorem 2.24: Suppose that V is a finite-dimensional vector space with ordered basis $\beta = \{x_1, \ldots, x_n\}$. Let f_i for $1 \le i \le n$ be the i-th coordinate function with respect to β i.e. $f_i(x_j) = \delta_{ij}$. Let $\beta^* = \{f_1, \ldots, f_n\}$. Then β^* is an ordered basis for V^*, and $\forall f \in V^*$ we have $f = \sum_i f(x_i)f_i$. We call β^* the dual basis of β.

Theorem 2.25: Let V and W be finite-dimensional vector spaces over $\mathbb{F}$ with ordered bases β, γ, respectively. For any linear transformation $T : V \to W$, the mapping $T^t : W^* \to V^*$ defined by $T^t(g) = gT$ $\forall g \in W^*$ is a linear transformation with the property $[T^t]_{\gamma^*}^{\beta^*} = ([T]_\beta^\gamma)^t$.

Double Dual: For a vector $x \in V$, we define $\widehat{x} : V^* \to \mathbb{F}$ by $\widehat{x}(f) = f(x)$, $\forall f \in V^*$. Then $\widehat{x}$ is a linear functional on V^*, so $\widehat{x} \in V^{**}$, the double dual of V.

Lemma: Let V be a finite-dimensional vector space, and let $x \in V$. If $\widehat{x}(f) = 0\, \forall f \in V^*$, then $x = 0$.

Theorem 2.26: Let V be a finite-dimensional vector space and define $\psi : V \to V^{**}$ by $\psi(x) = \widehat{x}$. Then ψ is an isomorphism.

Corollary: Let V be a finite-dimensional vector space with dual space V^*. Then every ordered basis of V^* is the dual basis for some basis for V.

2.3 Diagonalization

Properties of Determinants: If E is an elementary matrix obtained by interchanging two rows of I, then $\det(E) = -1$. If E is an elementary matrix obtained by scalar multiplication (by k) of a row, then $\det(E) = k$. If E is an elementary matrix obtained by adding a multiple of one row to another, then $\det(E) = 1$. For $A \in \mathcal{M}_{n\times n}$ we have $\det(A^{-1}) = [\det(A)]^{-1}$, $\det(A^t) = \det(A)$, and $\det(A) = (-1)^n p(x)$, where $A \in \mathcal{M}_{n\times n}$ and $p(x)$ is monic. We also have that A is invertible iff $\det(A) \neq 0$.

Diagonalizable: If a linear operator T on a finite-dimensional vector space V is diagonalizable if there exists a basis β for V such that $[T]_\beta$ is a diagonal matrix.

Eigenvectors and Eigenvalues: Let $T : V \to V$ be linear, where V is a vector space over $\mathbb{F}$. Suppose that there exists $v \in V\backslash\{0\}$ such that $Tv = \lambda v$, where $\lambda \in \mathbb{F}$. Then v is called an eigenvector of T with associated eigenvalue λ.

Theorem 5.4: A linear operator T on a finite dimensional vector space V is diagonalizable iff there exists an ordered basis β for V consisting of eigenvectors of T. Furthermore, if T is diagonalizable, $\beta = \{v_1, \ldots, v_n\}$ is an ordered basis of eigenvectors of T, and $D = [T]_\beta$, then D is a diagonal matrix and D_{ii} is the eigenvalue corresponding to v_i for $1 \leq i \leq n$.

Theorem 5.11: The characteristic polynomial of any diagonalizable linear operator splits.

Eigenspace: The eigenspace of a linear operator: $E_\lambda = \{x \in V : T(x) = \lambda x\} = N(T - \lambda I)$. Note that $E_{\lambda_i} \cap E_{\lambda_j} = \{0\}$ for $i \neq j$.

Diagonalization Criteria: Let T be a linear operator on V over $\mathbb{F}$. If $c_T(x)$ splits over $\mathbb{F}$ and for each eigenvalue, λ, of T the algebraic multiplicity of λ is equal to the dimension of $E_\lambda = n - \text{rank}(T - \lambda I)$, then T is diagonalizable.

T-invariant Subspace: Let V be a vector space and let $T : V \to V$ be linear. Then W is a T-invariant subspace iff W is a subspace of V and $T(W) \subseteq W$. Note that $\{0\}, V, R_T, N_T, E_\lambda$ are all T-invaiant subspaces.

Smallest T-invariant Subspace: Let $x \in V\backslash\{0\}$ and $W = \text{span}(\{x, Tx, T^2x, \ldots\})$. Then W is the T-cyclic subspace of V generated by x. Here, $W \subset V$ and W is the "smallest" T-invariant subspace of V containing x.

Theorem 5.26: Let V be a finite-dimensional vector space, $T : V \to V$ be linear, and W be a T-invariant subspace of V. Then the characteristic polynomial of $T_W \equiv T|_W$ divides $c_T(x)$.

Theorem 5.27: Let V be a finite dimensional vector space, $T : V \to V$ be linear, and W be a T-invariant subspace of V generated by some $x \in V\backslash\{0\}$. Let $k = \dim(W)$. Then $\{x, Tx, \ldots, T^{k-1}x\}$ is a basis for W and if $a_0x + a_1Tx + \cdots + a_{k-1}T^{k-1}x + T^kx = 0$, then $c_{T_W}(t) = (-1)^k(a_0 + \cdots + a_{k-1}t^{k-1} + t^k)$.

Cayley-Hamilton Theorem: Let V be a finite-dimensional vector space and let $T : V \to V$ be linear. Then $c_T(T) = T_0$, the zero operator.

Theorem 5.29: Let V be a finite-dimensional vector space and let $T : V \to V$ be linear. Suppose that $V = W_1 \oplus \cdots \oplus W_k$, where W_i is a T-invariant subspace of V. Let f be the characteristic polynomial of T and f_i be the characteristic polynomials of T_{W_i}. Then $f(x) = f_1(x) \cdots f_k(x)$.

2.4 Inner Product Spaces

Inner Product: Let V be a vector space over $\mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$). An inner product on V is a function that assigns to every ordered pair of vectors $x, y \in V$ a scalar in $\mathbb{F}$, denoted $< x, y >$. We also define $\sqrt{< x, x >} = ||x||$. Then $\forall x, y, z \in V$ and $\forall c \in \mathbb{F}$ we have
(1) $< cx + z, y >= c < x, y > + < z, y >$
(2) $\overline{< x, y >} =< y, x >$
(3) $< x, x > \quad > 0$ if $x \neq 0$
(4) if $< x, y >=< x, z > \forall x \in V$, then $y = z$
(5) $\sqrt{< x, x >} = ||x||$
(6) $||x + y|| \leq ||x|| + ||y||$ and $| < x, y > | \leq ||x|| \cdot ||y||$

Inner Product Space: An inner product space is a vector space equipped with an inner product.

Orthogonal: Let V be an inner product space. Vectors $x, y \in V$ are orthogonal if $< x, y >= 0$. A subset S of V is orthogonal if any two distinct elements of S are orthogonal.

Theorem 6.3: Let V be an inner product space, and let $S = \{v_1, \ldots, v_k\}$ be an orthogonal set of nonzero vectors. If $y = \sum_i a_iv_i$, then $a_j =< y, v_j > /||v_j||^2 \; 1 \leq j \leq k$.

Corollary: Let V be an inner product space, and let S be an orthogonal set of nonzero vectors. Then S is linearly independent.

Gram-Schmidt Orthogonalization Process: Let V be an inner product space, and let $S = \{w_1, \ldots, w_n\}$ be a linearly independent subset of V. Define $S' = \{v_1, \ldots, v_n\}$, where $v_1 = w_1$ and $v_k = w_k - \sum_{j=1}^{k-1} \frac{<w_k, v_j>}{||v_j||^2} v_j$ for $2 \leq k \leq n$. Then S' is an orthogonal set such that $\text{span}(S') = \text{span}(S)$.

Corollary: Let V be a finite-dimensional inner product space with an orthonormal basis $\beta = \{v_1, \ldots, v_n\}$. Let T be a linear operator on V, and let $A = [T]_\beta$. Then $(A)_{ij} =< T(v_j), v_i >$ and $T(v_j) = \sum_i < T(v_j), v_i > v_i$.

Orthogonal Complement: Let S be a subset of a vector space V. Then the orthogonal complement of S is the set of all vectors that are orthogonal to every element in S, that is, $S^\perp = \{x \in V :< x, y >= 0 \, \forall y \in S\}$.

Proposition 6.6: Let W be a finite-dimensional subspace of an inner product space V, and let $y \in V$. Then there exists a unique $u \in W$ and $z \in W^\perp$ such that $y = u + z$. Furthermore, if $\{v_i\}$ is an orthonormal basis of W, then $u = \sum_{i=1}^k < y, v_i > v_i$.

Theorem 6.7: Suppose that $S = \{v_1, \ldots, v_k\}$ is an orthonormal set in an n-dimensional inner product space V. Then (1) S can be extended to an orthonormal basis $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ for V, (2) if $W = \text{span}(S)$, then $\text{span}(\{v_{k+1}, \ldots, v_n\}) = W^\perp$, and (3) $V = W \oplus W^\perp$.

Corollary: In the notation above, the vector u is the unique vector in W that is "closest" to y, that is, $\forall x \in W$, $||y - x|| \geq ||y - u||$ (with equality iff $x = u$). This vector u is called the orthogonal projection of y on W.

Theorem 6.8: Let V be a finite-dimensional inner product space over $\mathbb{F}$, and let $g : V \to \mathbb{F}$ be linear. Then there exists a unique $y \in V$ such that $g(x) =< x, y > \forall x \in V$. Moreover, $y = \sum_{i=1}^n \overline{g(v_i)} v_i$, where $\{v_1, \ldots, v_n\}$ is an orthonormal basis for V.

Theorem 6.9: Let V be a finite-dimensional inner product space, and let T be a linear operator on V. Then there exists a unique function $T^* : V \to V$ such that $< T(x), y >=< x, T^*(y) >$ $\forall x, y \in V$. Furthermore T^* is linear and $[T^*]_\beta = [T]_\beta^*$, where β is an orthonormal basis. We also have $(AB)^* = B^* A^*$ and $\text{rank}(A^* A) = \text{rank}(A)$ for matricies A and B.

Theorem 6.13: Let $A \in \mathcal{M}_{m \times n}(\mathbb{F})$ and $y \in \mathbb{F}^m$. Then $\exists x_0 \in \mathbb{F}^n$ such that $(A^* A) x_0 = A^* y$ and $||A x_0 - y|| \leq ||Ax - y|| \; \forall x \in \mathbb{F}^n$. Furthermore, if $\text{rank}(A) = n$, then $x_0 = (A^* A)^{-1} A^* y$.

Theorem 6.14: Let T be a linear operator on a finite-dimensional inner product space. Suppose that the characteristic polynomial splits. Then there exists an orthonormal basis β for V such that $[T]_\beta$ is upper triangular.

Normal Operator: Let V be an inner product space, and let T be a linear operator on V. We say that T is normal if $TT^* = T^* T$.

Theorem 6.15: Let V be an inner product space, and let T be a normal operator on V. Then
(1) $||T(x)|| = ||T^*(x)|| \; \forall x \in V$
(2) $T - cI$ is normal $\forall c \in \mathbb{F}$
(3) if $Tx = \lambda x$, then $T^* x = \overline{\lambda} x$
(4) if $x_1 \in E_{\lambda_1}$, $x_2 \in E_{\lambda_2}$, and $\lambda_1 \neq \lambda_2$, then $x_1 \perp x_2$

Theorem 6.16: Let T be a linear operator on a finite-dimensional complex inner product space V. Then T is normal iff there exists an orthonormal basis for V consisting of eigenvectors of T.

Self-Adjoint: Let T be a linear operator on an inner product space V. We say that T is self-adjoint (or Hermitian) if $T^* = T$.

Lemma: Let T be a self-adjoint operator on a finite dimensional inner product space V. Then every eigenvalue of T is real. If in addition $\mathbb{F} = \mathbb{R}$, then the characteristic polynomial splits.

Theorem 6.17: Let T be a linear operator on a finite-dimensional real inner product space V. Then T is self-adjoint iff there exists an orthonormal basis for V consisting of eigenvectors of T.

Unitary and Orthogonal Operators: Let T be a linear operator on an inner product space V (over $\mathbb{F}$). If $||T(x)|| = ||x|| \; \forall x \in V$, we call T a unitary operator for $\mathbb{F} = \mathbb{C}$ and an orthogonal operator for $\mathbb{F} = \mathbb{R}$. If V is infinite dimensional, we call T is isometry.

Theorem 6.18: Let T be a linear operator on a finite-dimensional inner product space V. Then the following are equivalent: (1) $TT^* = T^*T = I$, (2) $< T(x), T(y) >=< x, y > \forall x, y \in V$, (3) if β is an orthonormal basis for V, then so is $T(\beta)$, and (4) $||T(x)|| = ||x|| \; \forall x \in V$.

Lemma: Let U be a self-adjoint operator on a finite-dimensional inner product space V. If $< x, U(x) >= 0 \; \forall x \in V$, then $U(x) = 0 \; \forall x \in V$.

Corollary 1: Let T be a linear operator on a finite-dimensional real inner product space V. Then V has an orthonormal basis of eigenvectors of T with corresponding eigenvalues ± 1 iff T is both self-adjoint and orthogonal.

Theorem 6.19: Let $A \in \mathcal{M}_{n\times n}(\mathbb{C})$. Then A is normal iff A is unitarily equivalent to a diagonal matrix i.e. there exists a diagonal matrix D and a unitary matrix P such that $A = P^*DP$.

Theorem 6.20: Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Then A is symmetric iff A is orthogonally equivalent to a real diagonal matrix.

Theorem 6.21 (Shur): Let $A \in \mathcal{M}_{n\times n}(\mathbb{F})$ be a matrix whose characteristic polynomial splits over $\mathbb{F}$. If $\mathbb{F} = \mathbb{C}$, then A is unitarily equivalent to a complex upper triangular matrix. If $\mathbb{F} = \mathbb{R}$, then A is orthogonally equivalent to a real upper triangular matrix.

Orthogonal Projection: Let V be an inner product space and let $T : V \to V$ be a projection. We say that T is an orthogonal projection if $R_T^\perp = N_T$ and $N_T^\perp = R_T$.

Theorem 6.23: Let V be an inner product space and let T be a linear operator on V. Then T is an orthogonal projection iff T has an adjoint T^* and $T^2 = T = T^*$.

Theorem 6.24 (Spectral Theorem): Suppose that T is a linear operator on a finite-dimensional inner product space V over $\mathbb{F}$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$. Assume that T is normal if $\mathbb{F} = \mathbb{C}$ and self-adjoint if $\mathbb{F} = \mathbb{R}$. For each i $(1 \le i \le k)$ let W_i be the eigenspace of T corresponding to the eigenvalue λ_i and let T_i be the orthogonal projection onto W_i. Then the following are true.
(1) $V = W_1 \oplus \cdots \oplus W_k$
(2) if W_i' denotes the direct sum of W_j, $j \ne i$, then $W_i^\perp = W_i'$
(3) $T_iT_j = \delta_{ij}T_i$ for $1 \le i, j \le k$

(4) $I = T_1 + \cdots + T_k$ and $T = \lambda_1 T_1 + \cdots + \lambda_k T_k$

Corollary 1: If $\mathbb{F} = \mathbb{C}$, then T is normal iff $T^* = g(T)$ for some polynomial g.

Corollary 2: If $\mathbb{F} = \mathbb{C}$, then T is unitary iff T is normal and $|\lambda| = 1$ for every eigenvalue λ of T.

Corollary 3: If $\mathbb{F} = \mathbb{C}$ and T is normal, then T is self-adjoint iff every eigenvalue of T is real.

2.5 Canonical Forms

Generalized Eigenvectors: Let V be a finite-dimensional vector space and $T : V \to V$ be linear. Then $x \in V \backslash \{0\}$ is a generalized eigenvector of T corresponding to $\lambda \in \mathbb{R}$ if $(T - \lambda I)^p x = 0$ for some $p > 0$.

Generalized Eigenspace: Let V be a finite-dimensional vector space, $T : V \to V$ be linear, and λ be an eigenvalue of T. The generalized eigenspace of T corresponding to λ, denoted $K_\lambda(T)$, is the subset of V defined by $K_\lambda(T) = \{x \in V : (T - \lambda I)^p x = 0 \text{ for some } p > 0\}$.

Theorem 7.4: Let V be a finite-dimensional vector space and $T : V \to V$ be linear. Suppose that $c_T(x)$ splits over $\mathbb{F}$. Let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of T with algebraic multiplicities $m_1, \ldots, m_k$ and β_i be bases for K_{λ_i}. Then $\beta_i \cap \beta_j = \emptyset$ for $i \neq j$, $\beta = \cup_{i=1}^k \beta_i$ is a basis for V, $\dim(K_{\lambda_i}(T)) = m_i$, and $E_{\lambda_i} = K_{\lambda_i}$ iff T is diagonalizable.

Cycle of Generalized Eigenvectors: The set $\{(T - \lambda I)^{p-1} x, \ldots, (T - \lambda I)x, x\}$ is a cycle of generalized eigenvectors, also known as a Jordan Chain. If two cycles use different λ's, then their union is linearly independent.

Theorem 7.7: Let V be a finite-dimensional vector space, $T : V \to V$ be linear, and λ be an eigenvalue of T. Then $K_\lambda(T)$ has an ordered basis consisting of a union of disjoint cycles of generalized eigenvectors corresponding to λ.

Corollary: Let V be a finite-dimensional vector space and $T : V \to V$ be linear. Then $c_T(x)$ splits over $\mathbb{F}$ iff T has a Jordan Canonical Form.

Minimal Polynomial: Let V be a finite-dimensional vector space and $T : V \to V$ be linear. A polynomial $p(t)$ is called the minimal polynomial of T if $p(t)$ is the monic polynomial of least degree for which $p(T) = T_0$.

Theorem 7.12: Let V be a finite-dimensional vector space, $T : V \to V$ be linear, and $m_T(x)$ be the minimal polynomial of T. Then for any polynomial $g(t)$ such that $g(T) = T_0$, then $m_T(x) \,|\, g(x)$; in particular $m_T(x) \,|\, c_T(x)$. The minimal polynomial is unique. If β is an ordered basis for V and $A = [T]_\beta$, then $m_T(A) = 0$.

Theorem 7.14: Let T, V be as usual. Then $\lambda \in \mathbb{R}$ is an eigenvalue of T iff $m_T(\lambda) = 0$ i.e. $c_T(x)$ and $m_T(x)$ have the same roots.

Corollary: Let T, V be as usual. If $c_T(t) = (\lambda_1 - t)^{n_1} (\lambda_2 - t)^{n_2} \cdots (\lambda_k - t)^{n_k}$ where λ_i are distinct, then $m_T(t) = (\lambda_1 - t)^{m_1} (\lambda_2 - t)^{m_2} \cdots (\lambda_k - t)^{m_k}$, where $1 \leq m_i \leq n_i$. If V is a T-cyclic subspace of itself, then $c_T(x) = (-1)^n m_T(x)$. Also, $m_i = 1$ $\forall i$ iff T is diagonalizable.

3 Complex Analysis

3.1 Topology and Basics

Fundamental Theorem of Algebra: Every non-constant polynomial with complex coefficients has a zero in $\mathbb{C}$.

Region: An open connected set.

Simply Connected: A region D is simply connected if its complement is "connected within ε to ∞". That is, if for any $z_0 \in D^c$ and $\varepsilon > 0$, there exists a continuous curve $\gamma(t)$ for $0 \leq t < \infty$ such that $d(\gamma(t), D^c) < \varepsilon \ \forall t \leq 0$, $\gamma(0) = z_0$, and $\lim \gamma(t) = \infty$.

Star-like Region: A set S is called star-like if $\exists \alpha \in S$ such that the line segment connecting α and z is contained in $S \ \forall z \in S$.

Proposition: We have convex $\subset$ star-like $\subset$ simply connected.

Theorem: Suppose that $\overline{\lim} |c_k|^{1/k} = 1/R$. Then $\sum_{k=0}^{\infty} c_k z^k$ converges uniformly for $\forall z \in \mathbb{C}$ such that $|z| < R$ and diverges for $|z| > R$. Note that if $\lim_{k\to\infty} \frac{c_{k+1}}{c_k}$ exists, then $\overline{\lim} |c_k|^{1/k} = \lim_{k\to\infty} \frac{c_{k+1}}{c_k}$.

Cauchy-Riemann Equations: Let $f(z) = u(z) + iv(z)$. Then the Cauchy-Riemann equations are: $u_x = v_y$ and $u_y = -v_x$ i.e. $f_y = if_x$, where $z = x + iy$.

Complex Differentiable: A function $f(z)$ is complex differentiable if f_x and f_y exist in a neighborhood of z, are continuous at z, and satisfy the Cauchy-Riemann equations there.

Analytic: A function f is analytic at z if f is differentiable in a neighborhood of z. Similarly, f is analytic on a set S if f is differentiable at all points of some open set containing S. Also f is analytic iff f has a power series expansion.

Entire: A function which is analytic in the whole complex plane is entire.

Differentiability of Power Series: Suppose that $f(z) = \sum_{k=0}^{\infty} c_k z^k$ converges for $|z| < R$. Then f is infinitely differentiable for $|z| < R$.

Uniqueness Theorem for Power Series: Suppose $f(z) = \sum_{k=0}^{\infty} c_k z^k$ is zero at all points of a nonzero sequence $\{z_n\}$ which converges to zero. Then the power series is identically zero.

Theorem: Suppose Ω is a region, $f : \Omega \to \mathbb{C}$ is analytic, and $f = u + iv$, where u, v are real valued. Then f is constant iff u is constant iff v is constant iff f^2 is constant iff $|f|$ is constant.

Properties of the Exponential: We have $|e^z| = e^x$, $e^{iy} = \cos y + i \sin y$, $e^z = \alpha$ has infinitely many solutions for $\alpha \neq 0$, and $e^z = e^x e^{iy}$.

Properties of sin and cos: We have $\sin z = \frac{1}{2i}(e^{iz} - e^{-iz})$, $\cos z = \frac{1}{2}(e^{iz} + e^{-iz})$.

Analytic Branch of the Logorithm: The function f is an analytic branch of $\log z$ in a domain D if f is analytic in D and f is an inverse of the exponential function there i.e. $e^{f(z)} = z$. Also if $g(z) = f(z) + 2\pi ki$ for some k, then g is also an analytic branch of $\log z$.

Theorem: Suppose D is simply connected and $0 \notin D$. Choose $z_0 \in D$, fix a value of $\log z_0$, and set $f(z) = \int_{z_0}^{z} \frac{dw}{w} + \log z_0$. Then f is an analytic branch of $\log z$ in D.

3.2 Line Integrals

Piecewise Differentiable and Smooth Curves: Let $z(t) = x(t) + iy(t)$ for $a \leq t \leq b$. The curve determined by $z(t)$ is called piecewise differentiable and we set $\dot{z}(t) = x'(t) + iy'(t)$ if x and y are continuous on $[a, b]$ and continuously differentiable on each subinterval $[a, x_1], [x_1, x_2], \ldots, [x_{n-1}, b]$ for some partition of $[a, b]$. The curve is said to be smooth if, in addition, $\dot{z}(t) \neq 0$ except at a finite number of points.

Line Integrals: Let C be a smooth curve given by $z(t)$ for $a \leq t \leq b$, and suppose f is continuous at all points $z(t)$. Then the integral of f along C is $\int_C f(z)\, dz = \int_a^b f(z(t))\dot{z}(t)\, dt$.

Properties of Line Integrals: Let C be a smooth curve given by $z(t)$ for $a \leq t \leq b$, and suppose f is continuous at all points $z(t)$. Let $\alpha \in \mathbb{C}$. Then $\int_{-C} f = -\int_C f$, $\int_C f + g = \int_C f + \int_C g$, and $\int_C \alpha f = \alpha \int_C f$.

M-L Formula: Suppose that C is a smooth curve of length L, f is continuous on C, and $|f| \leq M$ throughout C. Then $|\int_C f| \leq ML$.

3.3 Properties of Analytic Functions

Lemma: Suppose a is contained in the circle C_ρ, that is, C_ρ has center α, radius ρ, and $|a - \alpha| < \rho$. Then $\int_{C_\rho} \frac{dz}{z-a} = 2\pi i$.

Liouville's Theorem: A bounded entire function is constant.

Extended Liouville Theorem: If f is entire and if, for some integer $k \geq 0$, there exists positive constants A and B such that $|f(z)| \leq A + B|z|^k$, then f is a polynomial of degree at most k.

Theorem: Suppose f is analytic in $D(\alpha, r)$. If the closed curve C and the point a are both contained in $D(\alpha, r)$, then $\int_C f\, dz = \int_C \frac{f(z)-f(a)}{z-a}\, dz = 0$.

Theorem: If f is analytic in $D(\alpha, r)$ and $a \in D(\alpha, r)$, then there exists F and G analytic in $D(\alpha, r)$ such that $F'(z) = f(z)$ and $G'(z) = \frac{f(z)-f(a)}{z-a}$.

Cauchy Integral Formula: Suppose f is analytic in $D(\alpha, r)$, $0 < \rho < r$, and $|a - \alpha| < \rho$. Then $f(a) = \frac{1}{2\pi i} \int_{C_\rho} \frac{f(z)}{z-a}\, dz$, where $C_\rho = \alpha + \rho e^{i\theta}$ for $0 \leq \theta \leq 2\pi$.

Cauchy Inequalities: Let f be analytic in $D(0, r)$. Then $\exists c_k$ such that $f(z) = \sum_{k=0}^{\infty} c_k z^k$ $\forall z \in D$ and $|c_k| \leq \frac{M(r)}{r^k}$, where $M(r) = \max_{|z|=r} |f(z)|$.

Theorem: If f is analytic in $D(\alpha, r)$, then $\exists c_k$ such that $f(z) = \sum_{k=0}^{\infty} c_k(z-\alpha)^k \; \forall z \in D(\alpha, r)$ and $c_k = \frac{f^{(k)}(\alpha)}{k!} = \frac{1}{2\pi i} \int_{C_\rho} \frac{f(z)}{(z-\alpha)^{k+1}} \, dz$.

Maximum-Modulus Theorem: A non-constant analytic function in a region D does not have any interior maximum points. For each $z \in D$ and $\delta > 0$ $\exists w \in D(z, \delta) \cap D$ such that $|f(w)| > |f(z)|$. If D is compact, then f assumes its max on the boundary of D.

Minimum-Modulus Theorem: If f is a non-constant analytic function in a region D, then no point $z \in D$ can be a relative minimum of f unless $f(z) = 0$.

Anti-Calculus Proposition: Suppose that f is analytic throughout a closed disk and assumes its max modulus at a boundary point α. Then $f(\alpha) \neq 0$ unless f is constant.

Open Mapping Theorem: The image of an open set under a nonconstant analytic mapping is an open set.

Schwarz Lemma: Suppose f is analytic in $D(0, R)$ and $|f(z)| \leq M \; \forall z \in D(0, R)$ and $f(0) = 0$. Then $|f(z)| \leq \frac{M}{R}|z|$ and $|f'(0)| \leq \frac{M}{R}$ with equality only if $f(z) = e^{i\theta}\frac{M}{R}z$ for some $\theta \in \mathbb{R}$.

Morera's Theorem: Let f be a continuous function on a open set D. If $\int_\Gamma f(z)\, dz = 0$ whenever Γ is the boundary of a closed rectangle in D, then f is analytic on D.

Uniformly on Compacta: The sequence f_n converges uniformly on compacta if $f_n \to f$ uniformly $\forall K \subset D$, where K is compact and f_n, f are analytic in D.

Theorem: Suppose $\{f_n\}$ represents a sequence of functions analytic in an open domain D such that $f_n \to f$ uniformly on compacta, then f is analytic in D.

Theorem: Suppose f is continuous in an open set D and analytic there except (possibly) at the points of a line segment L. Then f is analytic throughout D.

Schwarz Reflection Principle: Suppose f is C-analytic in a region D that is contained in either the upper or lower half plane and whose boundary contains a segment L on the real axis, and suppose f is real for real z. Let $D^* = \{z : \overline{z} \in D\}$. Then we can define an analytic extension g of f to the region $D \cup L \cup D^*$ that is symmetric with respect to the real axis by setting

$$g(z) = \begin{cases} f(z), & z \in D \cup L, \\ \overline{f(\overline{z})}, & z \in D^*. \end{cases}$$

Corollary: If f is analytic in a region symmetric with respect to the real axis and if f is real-valued for all real z, then $f(z) = \overline{f(\overline{z})}$.

3.4 Singularities

Isolated Singularity: A function f is said to have an isolated singularity at z_0 if f is analytic in a deleted neighborhood D of z_0, but not analytic at z_0.

Removable Singularity: Suppose Suppose that f has an isolated singularity at z_0. If there exists g analytic at z_0 such that $f(z) = g(z)$ for all z in a deleted neighborhood of z_0, then the singularity at z_0 is removable. Also, z_0 is a removable singularity of f iff $\lim_{z\to z_0}(z-z_0)f(z) = 0$ iff f is bounded in a deleted neighborhood of z_0.

Pole: Suppose that for $z \neq z_0$ f can be written in the form $f(z) = \frac{A(z)}{B(z)}$ where A and B are analytic at z_0, $A(z_0) \neq 0$, and $B(z_0) = 0$. If B has a zero of order k at z_0, then f has a pole of order k at z_0. Also, f has a pole at z_0 iff $\lim_{z\to z_0}|f(z)| = \infty$ iff there exists k such that $\lim_{z\to z_0}(z-z_0)^k f(z) \neq 0$, but $\lim_{z\to z_0}(z-z_0)^{k+1}f(z) = 0$.

Essential Singularity: Suppose that f has an isolated singularity at z_0. If the singularity is neither a removable singularity nor a pole, then the singularity is an essential singularity. Also, f has an essential singularity at z_0 iff $c_n \neq 0$ for some sequence $n \to -\infty$ where the c_n are the coefficients of the Laurent expansion of f centered at z_0.

Casorati-Weierstrass Theorem: If f has an essential singularity at z_0 and if D is a deleted neighborhood of z_0, then the range of f for $z \in D$ is dense in $\mathbb{C}$.

Theorem: Let $f(z) = \sum_{-\infty}^{\infty} a_k z^k$ be convergent in $D = \{z : R_1 < |z| < R_2\}$, where $R_2 = 1/\overline{\lim}|a_k|^{1/k}$ and $R_1 = \overline{\lim}|a_{-k}|^{1/k}$. If $R_1 < R_2$, D is an annulus and f is analytic in D.

Theorem (Laurent): If f is analytic in the annulus $A = \{z : R_1 < |z| < R_2\}$, then f has a Laurent expansion, $f(z) = \sum_{-\infty}^{\infty} a_k z^k$ in A i.e. there exists f_1 analytic in $D(0, R_2)$ and f_2 analytic in $D(0, 1/R_1)$ with $f_2(0) = 0$ and $f(z) = f_1(z) + f_2(z^{-1})$. Note that $f_1(z)$ is called the regular part and $f_2(z^{-1})$ is called the principle part. Also, $a_k = \frac{1}{2\pi i}\int_C \frac{f(w)}{w^{k+1}}\,dw$ with $C \subset A$.

3.5 Residue Theorem

Residue: If $f(z) = \sum_{-\infty}^{\infty} c_k(z-z_0)^k$ in a deleted neighborhood of z_0, c_{-1} is called the residue of f at z_0, denoted $c_{-1} = \text{Res}(f; z_0)$. Also, $\int_\gamma f = 2\pi i c_{-1}$.

Evaluation of Reidues I: If f has a pole of order k at z_0, then $c_{-1} = \frac{1}{(k-1)!}\frac{d^{k-1}}{dz^{k-1}}[(z-z_0)^k f(z)]\Big|_{z=z_0}$. If f and g are analytic in $D(a, r)$, f has a root of multiplicity n at a and g with multiplicity m at a and $m > n$, then f/g has a pole of order $m-n$ at a and $Res(f/g; a) = \lim_{z\to a}\frac{1}{(m-n-1)!}\left(\frac{d}{dz}\right)^{m-n-1}\left[(z-a)^{m-n}\frac{f(z)}{g(z)}\right]$.

Evaluation of Reidues II: Let f have a simple pole at z_0, and let g be analytic at z_0. Then $Res_{z_0}(fg) = g(z_0)Res_{z_0}(f)$. Suppose $f(z_0) = 0$, but $f'(z_0) \neq 0$. Then $1/f$ has a pole of order 1 at z_0 and $Res_{z_0}(1/f) = 1/f'(z_0)$.

Winding Number: Let γ be a simple smooth closed curve and $a \in \mathbb{C}$ and $a \notin \gamma$. Then the winding number $\eta(\gamma, a) = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z-a}$. Then if a is outside γ, $\eta(\gamma, a) = 0$. Otherwise the winding number is the number of times the curve γ circles around a in the counter-clockwise direction.

Cauchy's Residue Theorem: Suppose f is analytic in a simply connected domain except for isolated singularities at $z_1, \ldots, z_m$. Let γ be a simple closed curve not intersecting any of the z_i. Then $\int_\gamma f = 2\pi i \sum_{k=1}^{m} \eta(\gamma, z_k) Res(f; z_k)$.

Meromorphic: A function f is meromorphic in a domain D if f is analytic there except at isolated poles.

Principle of the Argument Theorem: Let Ω be a simply connected open set and f be meromorphic in Ω. Let γ be a regular simple closed contour in Ω. Let $n_z(\gamma, f)$ = the number of zeros of f inside γ and $n_p(\gamma, f)$ = the number of poles of f inside γ. Then $\frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)}\,dz = n_z(\gamma, f) - n_p(\gamma, f)$.

Rouche's Theorem: Suppose that f and g are meromorphic inside and on a regular closed curve and $|f(z)| > |g(z)|\ \forall z \in \gamma$. Then $n_z(\gamma, f) - n_p(\gamma, f) = n_z(\gamma, f+g) - n_p(\gamma, f+g)$.

Hurwitz's Theorem: Let $\{f_n\}$ be a sequence of non-vanishing analytic functions in a region D and suppose $f_n \to f$ uniformly on compacta of D. Then either $f \equiv 0$ in D or $f(z) \neq 0\ \forall z \in D$.

Theorem: Suppose $f_n \to f$ uniformly on compacta in a region D. If f_n is 1-1 in $D\ \forall n \geq 1$, then either f is constant or f is 1-1 in D.

Residue Application to Real Integrals I: Let P and Q be polynomials such that $Q(x) \neq 0$ and $\deg(Q) - \deg(P) \geq 2$. Let z_k be poles of $P(z)/Q(z)$ such that $\mathrm{Im}(z_k) \geq 0$. Then $\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)}\,dx = 2\pi i \sum_k Res\left(\frac{P}{Q}; z_k\right)$.

Residue Application to Real Integrals II: Let $R(x) = \frac{P(x)}{Q(x)}$, where P and Q are polynomials and $Q(x) \neq 0$ (except perhaps at a zero of $\cos(x)$ or $\sin(x)$) and $\deg(Q) > \deg(P)$. Let z_k be poles of $R(z)$ such that $\mathrm{Im}(z_k) \geq 0$. Then $\int_{-\infty}^{\infty} R(x)\cos(x)\,dx = \mathrm{Re}[2\pi i \sum_k Res(R(z)e^{iz}; z_k)]$ and $\int_{-\infty}^{\infty} R(x)\sin(x)\,dx = \mathrm{Im}[2\pi i \sum_k Res(R(z)e^{iz}; z_k)]$.

Residue Application to Real Integrals III: Let P and Q be polynomials such that $Q(x) \neq 0$ for $x \geq 0$ and $\deg(Q) - \deg(P) \geq 2$. Then $\int_0^{\infty} \frac{P(x)}{Q(x)}\,dx = -\sum_k Res\left(\frac{P}{Q}\log z; z_k\right)$, where the sum is over all poles of $P(z)/Q(z)$.

3.6 Conformal Mappings

Möbius Transform: The Möbius Transform is given by $f(z) = \frac{az+b}{cz+d}$, where $ad - bc \neq 0$. These functions are also called fractional linear transformations.

Theorem: Every fractional linear transformation is a composition of dilations, translations, and inversions. Moreover, every fractional linear transformation maps circles and lines to circles and lines.

Locally 1-1: A function f is locally 1-1 at z_0 if for some $\delta > 0$ any any distinct $z_1, z_2 \in D_\delta(z_0)$, $f(z_1) \neq f(z_2)$. Also, f is locally 1-1 throughout a region D if f is locally 1-1 at every $z \in D$.

Conformal: A function is conformal if it preserves angles. More precisely, we say that a smooth complex-valued function is conformal at z_0 if whenever γ_0 and γ_1 are two curves terminating at z_0 with nonzero tangents, then the curves $f \circ \gamma_0$ and $f \circ \gamma_1$ have nonzero tangents at $f(z_0)$ and the angle from $(f \circ \gamma_0)'(z_0)$ to $(f \circ \gamma_1)'(z_0)$ is the same as the angle from $\gamma_0'(z_0)$ to $\gamma_1'(z_0)$.

Conformal Mapping: A conformal mapping of one domain D onto another V is a continuously differentiable function that is conformal at each point of D and that maps D one-to-one onto V.

Theorem: If $f(z)$ is analytic at z_0 and $f'(z_0) \neq 0$, then $f(z)$ is conformal and locally one-to-one at z_0.

Bilinear Transformations: The class of bilinear transformations that are analytic in the unit disk and bounded there by one and is given by $B_\alpha(z) = \frac{z-\alpha}{1-\overline{\alpha}z}$ for $|\alpha| < 1$.

Lemma: The only automorphisms of the unit disk with $f(0) = 0$ are given by $f(z) = e^{i\theta}z$.

Theorem: The automorphisms of the unit disk are of the form $g(z) = e^{i\theta}\left(\frac{z-\alpha}{1-\overline{\alpha}z}\right)$, $|\alpha| < 1$.

Theorem: The conformal mappings h of the upper half-plane onto the unit disk are of the form $h(z) = e^{i\theta}\left(\frac{z-\alpha}{z-\overline{\alpha}}\right)$, $\mathrm{Im}(\alpha) > 0$.

Theorem: The automorphisms of the upper half-plane are of the form $f(z) = \frac{az+b}{cz+d}$ with $a, b, c, d \in \mathbb{R}$ and $ad - bc > 0$.

Proposition: A bilinear transformation (other than the identity mapping) has at most two fixed points.

3.7 Other Useful Formulas

Calculations in $\mathbb{C}$:

$$\log z = \ln|z| + i\arg(z)$$

$$a^z = \exp(z\log a) \text{ if } a \in \mathbb{C}\backslash(-\infty, 0]$$

$$\text{if } \sqrt{a+bi} = x + iy, \text{ then } x = \pm\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}} \text{ and } y = \pm\sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}}\mathrm{sgn}(b)$$

Binomial Formula: $(a+b)^n = \sum_{k=0}^{n}\binom{n}{k} a^n b^{n-k}$

Power Series Expansions:

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty}\frac{x^n}{n!} \quad \text{for } x \in (-1,1) \\ \sin(x) &= \sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!} \quad \text{for } x \in \mathbb{R} \\ \cos(x) &= \sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!} \quad \text{for } x \in \mathbb{R} \\ \arctan(x) &= \sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{2n+1} \quad \text{for } x \in [-1,1] \end{aligned}$$

4 Topology

4.1 Set Theory

Images and Preimages: Let $f : X \to Y$, $A \subseteq X$, and $B \subseteq Y$. Then $f(A) = \{y \in Y : \exists a \in A \text{ s.t. } f(a) = y\} = \{f(a) : a \in A\}$ and $f^{-1}(B) = \{x \in X : f(x) \in B\}$.

Properties of Images: Let $f : X \to Y$ and $\{A_\lambda\} \subseteq \mathcal{P}(X)$. Then $f(\cup A_\lambda) = \cup f(A_\lambda)$. If f is injective, then $f(\cup A_\lambda) = \cup f(A_\lambda)$. If f is surjective, then $f(X - A_0) = Y - f(A_0)$.

Properties of Preimages: Let $f : X \to Y$ and $\{B_\lambda\} \subseteq \mathcal{P}(Y)$. Then $f^{-1}(\cup B_\lambda) = \cup f^{-1}(B_\lambda)$, $f^{-1}(\cap B_\lambda) = \cap f^{-1}(B_\lambda)$, and $f^{-1}(Y - B_0) = X - f^{-1}(B_0)$. Also, $f(f^{-1}(V)) \subseteq V$ with equality if f is surjective and $f^{-1}(f(U)) \supseteq U$ with equality if f is injective.

Equinumerous: Sets X and Y are equinumerous iff there exists a bijection from X to Y.

Inverse Function: When $f : X \to Y$ and $g : Y \to X$ satisfy $g \circ f = 1_X$ and $f \circ g = 1_Y$, then f and g are inverses functions. Also, f has an inverse only if f is a bijection. Note that if $V \subseteq Y$, then $f^{-1}(V) = g(V)$.

Finite Set: A set X is finite iff $\exists n \in \mathbb{N}$ and an injective function $X \to \{1, \ldots, n\}$.

Countable Set: A set X is countable iff there exists an injective function $X \to \mathbb{N}$. A set X is countable iff there exists a surjective function $\mathbb{N} \to X$.

4.2 Topological Spaces and Continuous Functions

Topological Space: A topological space is a pair $(X, \mathcal{J})$ where X is a set and $\mathcal{J}$ is a topology i.e. $\mathcal{J} \subseteq \mathcal{P}(X)$, $\emptyset, X \in \mathcal{J}$, if $A_\lambda \in \mathcal{J}$ then $\cup A_\lambda \in \mathcal{J}$, and if $A, B \in \mathcal{J}$ then $A \cap B \in \mathcal{J}$.

Interior: Let X be a space. For $A \subseteq X$, $\text{Int}_X(A) = \cup\{U \subseteq^{op} X : U \subseteq A\}$.

Properties of the Interior: (1) $\text{Int}_X(A) \subseteq^{op} X$, (2) $\text{Int}_X(A) \subseteq A$, (3) if $U \subseteq^{op} X$ and $U \subseteq A$ then $U \subseteq \text{Int}_X(A)$, (4) if $x \in X$ then $x \in \text{Int}_X(A)$ iff $\exists U \subseteq^{op} X$ s.t. $x \in U \subseteq A$, (5) $A \subseteq^{op} X$ iff $\text{Int}_X(A) = A$, (6) $\text{Int}_X(\text{Int}_X(A)) = \text{Int}_X(A)$, and (7) $\text{Int}_X(A \cap B) = \text{Int}_X(A) \cap \text{Int}_X(B)$.

Closure: Let $A \subseteq X$. Then $\text{Cl}_X(A) = \bar{A} = \cap\{F \subseteq^{cl} X : A \subseteq F\}$.

Properties of the Closure: (1) $\text{Cl}_X(A) \subseteq^{cl} X$ and $A \subseteq \text{Cl}_X(A)$, (2) if $A \subseteq F \subseteq^{cl} X$ then $\text{Cl}_X(A) \subseteq F$, (3) given $x \in X$ $x \in \text{Cl}_X(A)$ iff $U \cap A \neq \emptyset$ $\forall U \subseteq^{op} X$ s.t. $x \in U$, (4) $A \subseteq^{cl} X$ iff $\text{Cl}_X(A) = A$, (5) $\text{Cl}_X(\text{Cl}_X(A)) = \text{Cl}_X(A)$, and (6) $\text{Cl}_X(A \cup B) = \text{Cl}_X(A) \cup \text{Cl}_X(B)$.

Dense: A set D is dense in X iff $\text{Cl}_X(D) = X$ iff $D \cap U \neq \emptyset$ $\forall U \subseteq^{op} X$.

Subspace Topology: Let $A \subseteq X$. Then the subspace topology on A is $\mathcal{J}_A = \{U \cap A : U \subseteq^{op} X\}$.

Metric Topology: The metric topology is the set of all subsets of X that are unions of open balls, usually denoted $\mathcal{J}_d$.

Proposition: For $U \subseteq X$, $U \in \mathcal{J}_d$ iff $\forall x \in U\ \exists \varepsilon > 0$ s.t. $B^d_\varepsilon(x) \subseteq U$.

Proposition: Let (X, d) be a metric space. Then $\{x \in X : d(x, A) = 0\} = \mathrm{Cl}_X(A)$, where $d(x, A) = \inf\{d(x, a) : a \in A\}$.

Subspace-Metric Topology: Let (X, d) be a metric space and $A \subseteq X$. Then $d_A = d|_{A\times A}$ and thus (A, d_A) is a metric space and $B^d_\varepsilon(x) \cap A = B^{d_A}_\varepsilon(x)$, $x \in A$.

Equivalence of Metrics: Let d, d' be metrics on X, where $\mathcal{J}_d = \mathcal{J}_{d'}$, then d and d' are equivalent metrics. We then say that d and d' are topologically equivalent.

Proposition: Every metric space is topologically equivalent to a bounded metric space.

Continuity: A function $f : X \to Y$ is continuous iff $f(\mathrm{Cl}_X(A)) \subseteq \mathrm{Cl}_Y(f(A))$ for all $A \subseteq X$. The following are equivalent definitions of continuity: (1) whenever $x \in X$ and $f(x) \in V \subseteq^{op} Y$ then $\exists U \subseteq^{op} X$ s.t. $x \in U$ and $f(U) \subseteq V$, (2) whenever $V \subseteq^{op} Y$ then $f^{-1}(V) \subseteq^{op} X$, (3) whenever $k \subseteq^{cl} Y$ then $f^{-1}(K) \subseteq^{cl} X$, (4) whenever $x_n \to x$ in X then $f(x_n) \to f(x)$ in Y, and (5) if X, Y are metric spaces then $\forall x \in X, \forall \varepsilon > 0\ \exists \delta > 0$ s.t. $f(B_\delta(x) \subseteq B_\varepsilon(f(x))$.

Convergence: Then $x_n \to x$ iff $\forall U \subseteq^{op} X$ s.t. $x \in U$ and $\exists N \in \mathbb{N}$ s.t. $x_n \in U\ \forall n \geq N$. If $x \in A$, $A \subseteq X$, then $x \in \mathrm{Cl}_X(A)$ iff $\exists \{a_n\} \subseteq A$ s.t. $a_n \to x$ in X.

4.3 Topological Properties

Homeomorphism: A homeomorphism is a continuous bijection with a continuous inverse. If $f : X \to Y$ is a homeomorphism, then X is homeomorphic to Y, denoted $X \approx Y$.

Topological Property: A property P of a space is a topological property if whenever X has P and $X \approx Y$, then Y has P. Some examples: metrizability, Hausdorffness, finiteness, countability, and second countability.

Hausdorff Space: Any two distinct elements have disjoint open neighborhoods in a Hausdorff space.

Theorem: (1) Metrizable spaces are Hausdorff. (2) A finite metric space is decrete. (3) A finite space that is not Hausdorff is not metrizable.

Theorem: The following are equivalent for a continuous bijection $f : X \to Y$: (1) f is a homeomorphism, (2) f is an open function, and (3) f is a closed function.

Basis for a Topology: A collection $\mathcal{C}$ is a basis for a topology on a space X if every open set in X is the union of elements of $\mathcal{C}$.

Theorem: Let β be a collection of subsets of a set X. Then β is a basis for a topology on X iff (1) $\forall x \in X\ \exists B \in \beta$ s.t. $x \in B$ and (2) whenever $x \in B_1 \cap B_2$, where $B_1, B_2 \in \beta\ \exists B \in \beta$ s.t. $x \in B \subseteq B_1 \cap B_2$.

Second Countable: A space is second countable if it has a countable base, denoted 2°.

Separable Space: A space is separable if it has a countable, dense subset.

Theorem: Every 2° space is separable. Every separable metric space is 2°. Thus any separable space that is not 2° is not metrizable.

Theorem: Open subspaces of separable spaces and separable. The continuous image of a separable space is separable.

Theorem: Every subspace of 2° space is 2°. Second countability is invariant under continuous open surjections i.e. 2° is a topoloogical property.

Hereditary: A topological property is hereditary if it is inherited by all subspaces. Metrizability, 2°, Hausdorffness are all hereditary, but separability is not.

4.4 Quotient Spaces

Theorem: The collection $\mathcal{J} = \{V \subseteq Y : q^{-1}(V) \subseteq^{op} X\}$ is a topology on Y. If $\mathcal{J}'$ is any topology on Y, then $q : X(Y, \mathcal{J}')$ is continuous iff $\mathcal{J}' \subseteq \mathcal{J}$.

Quotient Map: A function $q : X \to Y$ is a quotient map if q is surjective and $V \subseteq^{op} Y$ iff $q^{-1}(V) \subseteq^{op} X$ or for all $V \subseteq Y$, $q^{-1}(V) \subseteq^{op} X$ implies $V \subseteq^{op} Y$.

Quotient Space: In the case above, the topology on Y is the largest (finest) topology that makes q continuous. Then space Y is a quotient space of X.

Theorem: Let $q : X \to Y$ be a continuous surjection. Then q is a quotient map if any of the following conditions hold: (1) q is an open map, (2) q is a closed map, and (3)$\forall F \subseteq Y$, $q^{-1}(F) \subseteq^{cl} X$ implies $F \subseteq clY$.

Respects the Identifications: Let $q : X \to Y$, $f : X \to Z$ be functions. The function f respects the identifications of q if for $x, x' \in X$, $q(x) = q(x')$ implies that $f(x) = f(x')$.

Universal Mapping Property (UMP) for Quotient Spaces: Let $q : X \to Y$ be a quotient map and let $f : X \to Z$ be continuous. Then there exists a unique continuous function $F : Y \to Z$ s.t. $F \circ q = f$ iff f respects the identifications of q.

Theorem: If X is a closed subset of $\mathbb{R}^n$ (euclidean topology) and X is bounded in the euclidean metric, then any continuous function from X to a Hausdorff space is a closed map.

Relation: Let X be a set. A relation on X is a subset, R, of $X \times X$. We write xRy whenever $(x, y) \in R$.

Equivalence Relation: An equivalence relation of X is a relation, $\sim$, on X that is reflexive, symmetric, and transitive.

Equivalence Class: Let $\sim$ be an equivalence relation on X. Given $x \in X$, $[x]_\sim = \{y \in X : x \sim y\}$ is the equivalence class of x under $\sim$. We set $X/\sim = \{[x]_\sim : x \in X\} \subseteq \mathcal{P}(X)$.

Partition: A partition on a set X is a set Π of subsets of X such that $\forall x \in X\ \exists A \in \Pi$ such that $x \in A$ and if $A, B \in \Pi$ and $A \cap B \neq \emptyset$ then $A = B$.

Fundamental Theorem of Equivalence Relations: If $\sim$ is an equivalence relation of X, then $X\ \sim$ is a partition of X. Conversely if Π is a partition of X, then by defining $x \sim_\Pi y$ iff $\exists A \in \Pi$ such that $x, y \in A$, we obtain an equivalence relation on X for which $X/\sim_\Pi = \Pi$.

4.5 Product Topology

Product Topology (on $X_1 \times X_2$): The product topology on $X_1 \times X_2$ is given by $\{U_1 \times U_2 : U_1 \subseteq^{op} X_1,\ U_2 \subseteq^{op} X_2\}$.

Product Set: Let $\mathcal{A}$ be a set. Then $\prod_{\alpha\in\mathcal{A}} X_\alpha = \{\sigma : \mathcal{A} \to \cup_{\alpha\in\mathcal{A}} X_\alpha : \sigma(\alpha) \in X_\alpha\ \forall\alpha \in \mathcal{A}\}$.

Projection Stuff: Let $\{X_\alpha : \alpha \in \mathcal{A}\}$ be given. Given $\beta \in \mathcal{A}$, we define $\rho_\beta : \Pi X_\alpha \to X_\beta$ by $\rho_\beta(\sigma) = \sigma(\beta)$. Then ρ_β is the β-projection on the product set. Given $\sigma \in \Pi X_\alpha$, $\rho_\beta(\sigma) = \sigma(\beta)$ is the β-coordinate of σ. The set X_β is the β-factor of ΠX_α.

Setup for UMP for Product Spaces: Let $\{X_\alpha : \alpha \in \mathcal{A}\}$ be an indexed set of nonempty sets. Let Y be a set. If $F : Y \to \Pi X_\alpha$ is a function, then for each $\beta \in \mathcal{A}$, $\rho_\beta \circ F$ is a function $Y \to X_\beta$. If we are given functions $f_\beta : Y \to X_\beta\ \forall\beta \in \mathcal{A}$, then there exists a unique function $F : Y \to \Pi X_\alpha$ such that $\rho_\beta \circ F = f_\beta\ \forall\beta \in \mathcal{A}$. The composites $\rho_\beta \circ F$ are called the coordinate functions of F.

Theorem: Let Σ be a family of subsets of a set X. Then $\mathcal{B}(\Sigma) = \{\cap_n S_i : n \geq 0,\ S_i \in \Sigma\}$ is a basis for a topology $\mathcal{J}(\Sigma)$ on X such that (1) $\Sigma \subseteq \mathcal{J}(\Sigma)$, (2) $\Sigma \subseteq \mathcal{J}' =$ topology on X, then $\mathcal{J}(\Sigma) \subseteq \mathcal{J}'$. Then we say Σ is a sub-basis for $\mathcal{J}(\Sigma)$ and $\mathcal{J}(\Sigma)$ is the topology generated by Σ.

Product Topology: Let $\{X_\alpha : \alpha \in \mathcal{A}\}$ be an indexed family of nonempty spaces and $\rho_\beta : \Pi X_\alpha \to X_\beta$, $\beta \in \mathcal{A}$. The product topology on ΠX_α is the topology generated by $\{\rho_\beta^{-1}(U_\beta) : \beta \in \mathcal{A},\ U_\beta \subseteq^{op} X_\beta\}$ with the basis $\{\cap^n \rho_{\beta_i}^{-1}(U_{\beta_i}) : n \geq 0,\ \beta_i \in \mathcal{A},\ U_{\beta_i} \subseteq^{op} X_{\beta_i}\}$.

Theorem: With the above topology on ΠX_α, $\rho_\beta : \Pi X_\alpha \to X_\beta$ are continuous $\forall\beta \in \mathcal{A}$.

Theorem (UMP for product spaces): If Y is a space and we are given functions $f_\beta : Y \to X_\beta\ \forall\beta \in \mathcal{A}$, then there exists a unique continuous map $F = \{f_\beta\} : Y \to \Pi X_\alpha$ such that $\rho_\beta \circ F = f_\beta\ \forall\beta \in \mathcal{A}$ iff each f_β is continuous.

Continuity: A function is continuous iff pre-images of sub-basis open sets are open.

Proposition: Let $\sigma \in \Pi X_\alpha$ be fixed. Let $\alpha_0 \in \mathcal{A}$. Define $\lambda^\sigma_{\alpha_0} : X_\alpha \to \Pi X_\alpha$ by $\lambda^\sigma_{\alpha_0}(x)(\alpha) = x$ if $\alpha = \alpha_0$, $\sigma(\alpha)$ otherwise. Then $\lambda^\sigma_{\alpha_0}$ is a continuous injection. Moreover, it is a homeomorphism onto its image i.e. $X_\alpha \approx \lambda^\sigma_{\alpha_0}(X_{\alpha_0})$. We say $\lambda^\sigma_{\alpha_0}$ is an embedding.

Restricted in the β-coordinate: Let $G \subseteq \Pi X_\alpha$. The subset G is restricted in the β-coordinate if $\rho_\beta(G) \subsetneq X_\beta$.

Proposition: In the product topology on ΠX_α, an open set U can be restricted in at most finitely many coordinates.

Proposition: A product space ΠX_α is Hausdorff iff each factor is Hausdorff.

Proposition: If $\mathcal{A}$ is countable and X_α is 2° (respectively separable) $\forall \alpha \in \mathcal{A}$, then ΠX_α is 2° (respectively separable).

Proposition: Assume that each X_α is metrizable and has more than one point and ΠX_α is metrizable, then $\mathcal{A}$ is countable.

Theorem: Every separable metric space can be embedded in the Hilbert Cube. The Hilbert Cube is denote H and $H = \prod_{n=1}^{\infty}[0,1] = I^\omega$ with metric $d(x,y) = \sup_n\{\frac{|x_n - y_n|}{n} : n \in \mathbb{N}\}$.

Theorem: Consider $\prod_{n=1}^{\infty} X_n$. Let d_n be the metric on X_n. Define $\overline{d}_n(x,y) = \min\{d_n(x,y), 1\}$. Then define a metric d on $\prod_{n=1}^{\infty} X_n$ by $d(\sigma, \tau) = \sup\{\overline{d}_n(\sigma_n, \tau_n)/n : n \in \mathbb{N}\}$. This is a metric that determines the product metric.

Theorem: The metric topology and the product topology agree.

Corollary: A countable product of metrizable spaces is metrizable.

Lemma: For any metric space, (X, d), and any $x_0 \in X$, the function $d(\cdot, x_0) : X \to \mathbb{R}^{euclid}$ is continuous.

Properties of the Product Topology: 1) Countable products of 2°/ separable/ metrizable spaces are 2°/ separable/ metrizable. 2) UMP applies. 3) Every separable metric space embeds in I^ω.

Box Topology: The Box Topology is $\{\prod_{\alpha \in \mathcal{A}} U_\alpha : U_\alpha \subseteq^{op} X_\alpha \ \forall \alpha \in \mathcal{A}\}$. Note that the UMP does not apply and this topology is not metrizable.

4.6 Compactness

Open Cover: An open cover of a topological space X is a family $\mathcal{U}$ of open subsets of X such that $X = \cup_{U \in \mathcal{U}} U$.

Compactness: A topological space is compact if every open cover of X has a finite subcover.

Finite Intersection Property (FIP): A family $\mathcal{F}$ of subsets of X has this property if whenever $F_1, \ldots, F_n \in \mathcal{F}$, then $\cap_{i=1}^n F_i \neq \emptyset$.

Proposition: A space X is compact iff whenever $\mathcal{F}$ is a family of closed sets of X have the FIP, then $\cap \mathcal{F} \neq \emptyset$.

Theorem: If $A \subseteq I$ is infinite, then A has an accumulation point in I.

Theorem: The continuous image of a compact space is compact.

Corollary: Compactness is topologically invariant.

Theorem: 1) A closed subspace of a compact space is compact. 2) A compact subspace of a Hausdorff space is closed.

Corollary: If $A \subseteq X$, and X is compact and Hausdorff, then A is compact iff $A \subseteq^{cl} X$.

Theorem: If $f : X \to Y$ is continuous, X is compact, and Y is Hausdorff , then f is a closed map.

Corollary: If $f : X \to Y$ is a continuous surjection from a compact space to a Hausdorff space, then f is an identification map i.e. $V \subseteq^{op} Y$ iff $f^{-1}(V) \subseteq^{op} X$.

Theorem: Suppose that X is a compact Hausdorff space. 1) If $A \subseteq^{cl} X$ and $x \in X - A$, then $\exists U, V \subseteq^{op} X$ such that $A \subseteq A$, $x \in V$, and $U \cap V = \emptyset$. 2) If $A, B \subseteq^{cl} X$ with $A \cap B = \emptyset$, then $\exists U, V \subseteq^{op} X$ such that $A \subseteq U$, $B \subseteq V$, and $U \cap V = \emptyset$.

Tychonoff's Theorem: Any product of compact spaces is compact.

Theorem: Let X be compact and Y be any space, then $\rho : X \times Y \to Y$ is a closed map.

Corollary (Tube Lemma): Let $A \subseteq Y$ and X be compact. If $X \times A \subseteq U \subseteq^{op} X \times Y$, then $\exists V \subseteq^{op} Y$ such that $X \times A \subseteq X \times V \subseteq U$.

Theorem: $X \times Y$ is compact iff X and Y are compact $(X, Y \neq \emptyset)$.

Heine-Borel Theorem: A subset A of euclidean n-space $\mathbb{R}^n$ is compact iff $A \subseteq^{cl} \mathbb{R}^N$ and A is bounded in the euclidean metric.

Proposition: If two quotient maps make the same identifications, then the resulting quotient spaces are homeomorphic.

Sequentially Compact: If every equence in X has a convergent subsequence, then X is sequentially compact.

Limit Point: Let $A \subseteq X$. We say that $x \in X$ is a limit point (or accumulation point) of A if $x \in U \subseteq^{op} X$ implies that $U \cap (A - \{x\}) \neq \emptyset$. We denote the set of limit points of A by A', A' is also called the derived set of A.

5 Probability

5.1 Basic Measure Theory

Field: $\mathcal{F}_0$ is a field if (1) $\Omega \in \mathcal{F}_0$, (2) $A \in \mathcal{F}_0$ implies that $A^c \in \mathcal{F}_0$, and (3_0) $A, B \in \mathcal{F}_0$ implies that $A \cap B \in \mathcal{F}_0$.

σ-field: $\mathcal{F}$ is a σ-field if it satisfies (1) and (2) above and (3) $\{A_k\}_1^\infty \subseteq \mathcal{F}$ implies that $\cap A_k \in \mathcal{F}$.

Probability Measure: P is a probability measure if (1) $P(\Omega) = 1$, (2) $0 \leq P(A) \leq 1$ for all $A \in \mathcal{F}$, and (3) if $\{D_k\}_1^\infty \subseteq \mathcal{F}$ is a disjoint collection, then $P(\cup D_k) = \sum P(D_k)$.

Theorem: If P_0 is a probability measure on $\mathcal{F}_0$, then
i) For $A_n \nearrow A$, $\lim P_0(A_n) = P_0(A)$ (inner continuity).
ii) For $B_n \searrow B$, $\lim P_0(B_n) = P_0(B)$ (outer continuity).
iii) For $\{C_n : n \geq 1\} \subseteq \mathcal{F}_0$ with $\cup_{n\geq 1} C_n = C \in \mathcal{F}_0$, $P_0(C) \leq \sum P_0(C_n)$ (countable subadditivity).

Lemma: If P_0 satisfies (1), (2), and (3_0) on $\mathcal{F}_0$ then inner and outer continuity is equivalent to (3), countable additivity.

Caratheodory Extension Theorem: Suppose that P_0 is a probability measure on the field $\mathcal{F}_0$ over Ω. There exists a unique probability measure P on $(\Omega, \mathcal{F})$ where $\mathcal{F} = \sigma(\mathcal{F}_0)$ with $P(A) = P_0(A)$ for all $A \in \mathcal{F}_0$.

Π-System: $\mathcal{P}$ is a π-system if $A, B \in \mathcal{P}$ imply that $A \cap B \in \mathcal{P}$.

Λ-System: $\mathcal{L}$ is a λ-system if (i) $\Omega \in \mathcal{L}$, (ii) $A \in \mathcal{L}$ imples that $A^c \in \mathcal{L}$, (iii_λ) if $\{A_n : n \geq 1\} \subseteq \mathcal{L}$ with $A_k \cap A_j = \emptyset$ for $k \neq j$, then $\cup A_n \in \mathcal{L}$.

Theorem: Suppose that P and Q are both probability measures on $\sigma(\mathcal{P})$ with $\mathcal{P}$ being a π-system. If $P(A) = Q(A)$ for all $A \in \mathcal{P}$, then $P(A) = Q(A)$ for all $A \in \sigma(\mathcal{P})$.

Π-Λ Theorem: If $\mathcal{P}$ is a π-system and $\mathcal{L}$ is a λ-system, then $\mathcal{P} \subseteq \mathcal{L}$ implies that $\sigma(\mathcal{P}) \subseteq \mathcal{L}$.

Completeness of Probability Spaces: $(\Omega, \mathcal{F}, P)$ is a complete probability space iff for $A \in \mathcal{F}$ with $P(A) = 0$, then for $B \subseteq A$ we have $B \in \mathcal{F}$.

Limit Infinum: $\liminf E_n = \cup_{n=1}^\infty \cap_{k\geq n} E_k = \{\omega \in \Omega : w \in E_k \text{ for all but finitely many } k\}$

Limit Supremum: $\limsup E_n = \cap_{n=1}^\infty \cup_{k\geq n} E_k = \{\omega \in \Omega : w \in E_k \text{ for infinitely many } k\} = [E_k \; i.o.]$

Proposition: We always have $\liminf E_n \subseteq \limsup E_n$. If $\liminf E_n = \limsup E_n$, then $\lim E_n = \liminf E_n = \limsup E_n$.

Theorem: $P(\liminf E_n) \leq \liminf P(E_n) \leq \limsup P(E_n) \leq P(\limsup E_n)$.

First Borel-Cantelli Lemma: If $\sum_{n\geq 1} P(A_n) < \infty$, then $P([A_n \; i.o.]) = 0$.

5.2 Independence

Independence of Two Sets: If $P(A \cap B) = P(A)P(B)$, then the sets A and B are independent.

Lemma: If A and B are independent, then $\sigma(\{A\})$ and $\sigma(\{B\})$ are independent.

Independence of Two Collections of Sets: Two collections of sets $\mathcal{A}_1$ and $\mathcal{A}_2$ are independent if any $A_1 \in \mathcal{A}_1$ and $A_2 \in \mathcal{A}_2$ are independent.

Independence of a Finite Collection of Sets: We have that $A_1, A_2, \ldots, A_n$ are independent if $P(\cap_{j=1}^m A_{k_j}) = \prod_{j=1}^m P(A_{k_j})$ for any $1 \leq k_1 < k_2 < \cdots < k_m \leq n$ for $m \geq 2$.

Independence of a Finite Collection of Classes of Sets: We have that $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are independent if any $A_1 \in \mathcal{A}_1, \ldots, A_n \in \mathcal{A}_n$ are independent.

Independence of an Arbitrary Collection of Classes: Then $\{\mathcal{A}_\theta : \theta \in \Theta\}$ is a collection of independent families of sets if $\mathcal{A}_{\theta_i}, \ldots, \mathcal{A}_{\theta_n}$ is an independent list of families of sets for any distinct $\{\theta_1, \ldots, \theta_n\} \subseteq \Theta$.

Theorem: If $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are independent π-systems, then $\sigma(\mathcal{A}_1), \ldots, \sigma(\mathcal{A}_n)$ are independent as well.

Corollary: If $\{\mathcal{A}_\theta : \theta \in \Theta\}$ is a collection of independent π-systems, then $\{\sigma(\mathcal{A}_\theta) : \theta \in \Theta\}$ is an independent collection as well.

Second Borel-Cantelli Lemma: If $\{A_n : n \geq 1\}$ are independent and $\sum_{n \geq 1} P(A_n) = \infty$, then $P([A_n \; i.o.]) = 1$.

Borel's Normal Number Theorem (Strong Law of Large Numbers): $\lambda(\{\omega : \lim \frac{1}{n} \sum_{k=1}^n d_k(\omega) = 1/2\}) = 1$.

5.3 Random Variables and Distributions

Random Variables: Let $(\Omega, \mathcal{F}, P)$ and $(S, \mathcal{S}, P)$ be probability measure spaces. Then $X : \Omega \to S$ is a random variable if it is measurable i.e. if $X^{-1}(B) \in \mathcal{F}$ for all $B \in \mathcal{S}$.

Theorem: We have $(\Omega, \mathcal{F}, P)$ and $(S, \mathcal{S})$. Suppose $X : \Omega \to S$ and for all $A \in \mathcal{A}$, $X^{-1}(A) \in \mathcal{F}$. If $\mathcal{S} = \sigma(\mathcal{A})$, then X is measurable $\mathcal{F}/\mathcal{S}$.

Sigma Field Generated by a Random Variable: If X is a r.v. then $\sigma(X)$ is the smallest sub-σ-field of $\mathcal{F}$ that X is measurable. Then $\sigma(X) = \{X^{-1}(A) : A \in \mathcal{S}\}$.

Distribution Function: The distribution function of a real-valued r.v. X is $F_X(x) \equiv P(\{\omega : X(\omega) \leq x\}) = P(X^{-1}((-\infty, x]))$.

Proposition: If F be a distribution, then $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$, and F is nondecreasing.

Simple Random Variable: Then X is a simple random variable if $X : \Omega \to \{a_1, \dots, a_n\}$, $k < \infty$. Let $A_j = X^{-1}(\{a_j\})$. Then $X(x) = \sum_{j=1}^{k} a_j \mathbf{1}_{A_j}(x)$, so $\sigma(X) = \sigma(\{A_1, \dots, A_k\})$.

Construction of a Convergent Sequence of Random Variables: Define a sequence of simple r.v.'s X_n as follows. For $n = 1, 2, \dots$ and $0 \le k \le 2^{2n} - 1$, let

$$E_n^k = X^{-1}((k2^{-n}, (k+1)2^{-n}]) \quad , \quad \tilde{E}_n^k = X^{-1}((-(k+1)2^{-n}, -k2^{-n}])$$

and

$$G_n = X^{-1}((2^n, \infty]) \quad \text{and} \quad \tilde{G}_n = X^{-1}((-\infty, -2^n])$$

and define

$$X_n = \sum_{k=0}^{2^{2n}-1} k2^{-n}\chi_{E_n^k} - k2^{-n}\chi_{\tilde{E}_n^k} + 2^n\chi_{F_n} - 2^n\chi_{\tilde{F}_n}.$$

Then X_n is a sequence of simple r.v.'s such that $X_n \to X$ pointwise and $0 \le |X_1| \le |X_2| \le \dots \le |X|$.

Theorem: Let Y be a $\mathbb{R}^d$ valued r.v. on $(\Omega, \mathcal{F}, P)$. A real-valued r.v. X is measurable with respect to $\sigma(Y)$ iff there exists a measurable function $g : \mathbb{R}^d \to \mathbb{R}$ with $X = g(Y)$.

Proposition: Then F is a distribution iff (i) $\lim_{x\to\infty} F(x) = 1$, $\lim_{x\to-\infty} F(x) = 0$, (ii) F is non-decreasing, and (iii) F is a cadlag function.

Theorem A: If G is a distribution function on $\mathbb{R}$, then there exists, on some probability space, a r.v. with distribution function G.

Theorem B: If G is a distribution function on $\mathbb{R}$, then there exists on $(\mathbb{R}, \mathcal{B})$ a probability measure Q such that $Q((-\infty, x]) = G(x)$. Then the r.v. $X : \mathbb{R} \to \mathbb{R}$, given by $X(x) = x$ has distribution function G.

Theorem C: There exists a r.v. X defined on $(\Omega, \mathcal{F}, P) = ((0, 1], \mathcal{B}, \lambda)$ which has distribution function G.

Levy Distance: The Levy distance $d(F, G)$ between two distribution functions is given by $d(F, G) = \inf\{\varepsilon : G(x - \varepsilon) - \varepsilon \le F(x) \le G(x + \varepsilon) + \varepsilon \text{ for all } x \in \mathbb{R}\}$.

Theorem: Suppose that $F : \mathbb{R}^n \to [0, 1]$ satisfies (i) $\lim_{\vec{x}\nearrow\infty} F(\vec{x}) = 1$, $\lim_{\vec{x}\searrow-\infty} F(\vec{x}) = 0$, (ii) $\Delta F(\vec{x}, \vec{y}) = \sum_{2^n \text{ verticies, } \vec{v}} F(\vec{v})(-1)^{\sum_{k=1}^{n} \mathbf{1}_{[v_k = x_k]}}$, and (iii) $\lim_{y_k\searrow x_k} F(\vec{y}) = F(\vec{x})$, $\lim_{y\nearrow x} F(\vec{y})$ exists. Then there exists a unique measure μ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ with $\mu((x_1, y_1] \times \dots \times (x_n, y_n]) = \Delta F(\vec{x}, \vec{y})$.

Corollary: If f is a distribution function on $\mathbb{R}^n$, then there exists a probability measure P on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ and a r.v. $\vec{X} : \mathbb{R}^n \to \mathbb{R}^n$ such that $P(\vec{x} \in (-\infty, x_1] \times \dots \times (-\infty, x_n]) = F_X(\vec{x})$.

Probability Measure Induced by a Distribution: Let $X : \Omega \to \mathbb{R}$ be a real-valued r.v. measurable $\mathcal{F}/\mathcal{B}$. This map induces a probability measure μ on $\mathbb{R}$ where $\mu((-\infty, x]) = \mu_F((-\infty, x]) = F_X(x) = P(X \le x)$. Note that $\mu(B) = P(X^{-1}(B)) = PX^{-1}(B) = P_{X^{-1}}(B)$.

Extension Theorem: Suppose that the sequence of probability spaces $\{(\mathbb{R}^n, \mathcal{B}_n, P_n) : n \geq 1\}$ satisfies $P_{n+1}(B_n \times \mathbb{R}) = P_n(B_n)$ for any $B_n \in \mathcal{B}_n$, $n \geq 1$. Then there exists a unique probability measure P on $(\mathbb{R}^\infty, \mathcal{B}_\infty)$ with $P(\{x : (x_1, \ldots, x_n) \in B_n\}) = P_n(B_n)$ for any $B_n \in \mathcal{B}_n$, $n \geq 1$.

Independent Random Variables: Random variables X and Y defined on a common probability space $(\Omega, \mathcal{F}, P)$ are independent iff $\sigma(X)$ and $\sigma(Y)$ are independent i.e. $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$.

Independence of an Arbitrary Collection of Random Variables: The collection $\{X_\theta : \theta \in \Theta\}$ (defined on $(\Omega, \mathcal{F}, P)$) is a collection of independent r.v.'s iff $\{\sigma(X_\theta) : \theta \in \Theta\}$ is a collection of independent σ-fields.

Corollary: A sequence of real-valued r.v.'s defined on a common probability space $(\Omega, \mathcal{F}, P)$ are independent iff $P(\cap_{k=1}^n X_k^{-1}(B_k)) = \prod_{k=1}^n P(X_k^{-1}(B_k))$ for $B_k \in \mathcal{B}$, $1 \leq k \leq n$, $n \geq 1$.

Theorem 20.2: Suppose that $\{X_{ij} : i, j \geq 1\}$ is an array of independent r.v.'s. Define $\mathcal{F}_i = \sigma(\{X_{ij} : j \geq 1\})$ for $i \geq 1$. Then $\{\mathcal{F}_i : i \geq 1\}$ is an independent list of σ-fields.

Theorem 20.4: For any sequence $\{\mu_n : n \geq 1\}$ of probability measures on $(\mathbb{R}, \mathcal{B})$, there exists on some probability space $(\Omega, \mathcal{F}, P)$ a sequence of independent r.v.'s $\{X_n : n \geq 1\}$ such that X_n has distribution μ_n.

5.4 Convergence in Probability and Mean

Almost Surely Convergence: $X_n \to X$ almost surely iff $\lim X_n(\omega) = X(\omega)$ for all $\omega \in \Lambda$ where $P(\Lambda) = 1$.

Proposition: Let $\{X_n : n \geq 1\}$, X be defined on $(\Omega, \mathcal{F}, P)$ and for $\varepsilon > 0$, set $B_n(\varepsilon) = [\omega : |X_k - X| \geq \varepsilon$ for some $k \geq n]$. Then $X_n \to X$ a.s. P iff $\lim P(B_n(\varepsilon)) = 0$ for all $\varepsilon > 0$.

Proposition: $X_n \to X$ a.s. iff $P([B_n(\varepsilon)\ i.o.]) = 0$ for all $\varepsilon > 0$.

Convergence in Probability: X_n converges in probability to X iff for all $\varepsilon > 0$, $\lim P(|X_n - X| \geq \varepsilon) = 0$.

Theorem 20.5: Almost surely convergence implies convergence in probability. If $X_n \to X$ in probability, then there exists a subsequence $\{X_{n_k}\}_1^\infty$ such that $X_{n_k} \to X$ a.s.

Relate to Real Analysis: Almost surely convergence is convergence a.e. and convergence in probability is convergence in measure.

Expectation of a Random Variable: The expectation of a real-valued r.v. is $EX = \int_\Omega X(\omega)\, dP(\omega)$.

Fatou's Lemma: If $P(X_n \geq 0) = 1$, then $E \liminf X_n \leq \liminf EX_n$.

Monotone Convergence Theorem: If $P(X_n \geq 0) = 1$ and $X_n \leq X_{n+1}$ a.s. P with $X_n \nearrow X$ a.s., then $\lim EX_n = EX$, provided that $X \in L^1$.

Dominated Convergence Theorem: If $X_n \to X$ a.s. and $|X_n| \leq Y$ a.s. P with $Y \in L^1$, then $\lim EX_n = EX$. Can replace $X_n \to X$ a.s. with convergence in probability.

Change of Variable: $(\Omega, \mathcal{F}, P)$, $T : \Omega \to S$ is $(S, \mathcal{S})$ measurable gives $(S, \mathcal{S}, PT^{-1})$ where $PT^{-1}(C) = P(T^{-1}(C))$. Let $X : S \to \mathbb{R}$. Then $E(X \circ T) = \int_\Omega X(T(\omega))\, dP(\omega) = \int_S X(s)\, dPT^{-1}(s)$.

Convergence in Mean: X_n converges to X in r-th mean (in L^r) if $E|X_n - X|^r \to 0$.

Proposition: If $X_n \to X$ in L^r for $r > 0$, then $X_n \to X$ in probability.

Markov's Inequality: If $P(Y \geq 0) = 1$, then for all $a > 0$, $P(Y \geq a) \leq EY/a$.

Product Measure: Measures are σ-finite, $(\Omega, \mathcal{F}, \pi)$, $\Omega = \mathbb{X} \times \mathbb{Y}$, $\omega = (x, y)$, and $\pi = \mu \times \nu$ where $(\mathbb{X}, \mathcal{X}, \mu)$ and $(\mathbb{Y}, \mathcal{Y}, \nu)$ are both measure spaces. Then π is a product measure if $\pi(A \times B) = \mu(A)\nu(B)$ for $A \in \mathcal{X}$ and $B \in \mathcal{Y}$.

Fubini-Tonelli Theorem: If $f \in L^1$, then $\int_\Omega f(x, y)\, d\pi(x, y) = \int_\mathbb{X} \int_\mathbb{Y} f(x, y)\, d\nu(y)\, d\mu(x) = \int_\mathbb{Y} \int_\mathbb{X} f(x, y)\, d\mu(x)\, d\nu(y)$ and each inner integral is finite a.e.

Proposition: If $P(Y \geq 0) = 1$, then $EY = \int_0^\infty 1 - F_Y(y)\, d\lambda(y)$.

Density Measures: Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $f : X \to [0, \infty)$ be measurable, integrable, then $\nu(A) \equiv \int_A f(\omega)\, d\mu(\omega)$ is a finite measure and f is called the density of ν with respect to μ.

Theorem: With notation as above. For all ν-integrable g, $\int_\Omega g(\omega)\, d\nu(\omega) = \int_\Omega g(\omega) f(\omega)\, d\mu(\omega)$.

Variance: The variance of X is $Var(X) = E[(X - EX)^2]$.

Covariance: The covariance of X and Y is $Cov(X, Y) = E((X - EX)(Y - EY)) = E(XY) - \mu\nu$ where $\mu = EX$ and $\nu = EY$.

Cauchy-Schwartz Estimates: $E|XY| \leq \sqrt{EX^2}\sqrt{EY^2}$ and $|Cov(X, Y)| \leq \sqrt{Var(X)}\sqrt{Var(Y)}$ provided the right sides are finite.

Proposition: If X and Y are independent L^1 r.v.'s, then $E(XY) = EX \cdot EY$ and $Cov(X, Y) = 0$.

Proposition 2: X and Y are independent iff for all $g, h : \mathbb{R} \to \mathbb{R}$ in L^1, $E(g(X)h(Y)) = E(g(X)) \cdot E(h(Y))$.

Proposition ∞: $\{X_k : k \geq 1\}$ is a sequence of independent r.v.'s iff $E(\prod_{k=1}^N g_k(X_{i_k})) = \prod_{k=1}^N E(g_k(X_{i_k}))$ for all $N < \infty$, $i_1 < i_2 < \cdots < i_N$.

Corollary 2: If X and Y are independent L^2 r.v.'s, then $Var(X+Y) = Var(X) + 2Cov(X, Y) + Var(Y) = Var(X) + Var(Y)$. This generalizes to a finite sum of r.v's.

Corollary (Weak Law of Large Numbers): Suppose $\{X_k : k \geq 1\}$ are independent L^2 r.v.'s with $EX_k = \mu$ and $Var(X_k) \leq \sigma^2 < \infty$ for all $k \geq 1$. Then $P(|\frac{1}{n}\sum^n X_k - \mu| \geq \varepsilon) \to 0$ i.e. $\frac{1}{n}\sum^n X_k \to \mu$ in probability.

Chebyshev's Inequality: If Z is such that $EZ = \mu$, $Var(Z) = \sigma^2$, then for all $b > 0$, $P(|Z - \mu| \geq b) \leq Var(Z)/b^2 = (\sigma/b)^2$.

Strong Law of Large Numbers: Suppose $\{X_n\}_1^\infty \subseteq L^1$ is a sequence of iid random variables such that $EX_n = \mu$. Then $\bar{X}_n = \frac{1}{n}\sum_{k=1}^n X_k \to \mu$ a.s.

Kolmogorov's Maximal Inequality: Let $\{X_1, \ldots, X_n\}$ be independent r.v.'s with $EX_k = 0$, $Var(X_k) < \infty$ for $1 \leq k \leq n$. Then for all $\alpha > 0$, $P(\max_{1 \leq m \leq n} |\sum_1^m X_k| \geq \alpha) \leq Var(\sum_1^n X_k)/\alpha^2$.

Convergence of Random Series (Thm 22.6): If $\{X_k\}_1^\infty$ are independent with $EX_k = 0$ and $VarX_k < \infty$ for $k \geq 1$ and $\sum_{k \geq 1} VarX_k < \infty$, then S_n converges a.s. to $S = \sum_{k \geq 1} X_k$.

Theorem: If $\{X_k\}_1^\infty$ is a sequence of independent r.v.'s and $S_n = \sum_{k=1}^n X_k$ converges in probability, then $S_n \to S$ a.s..

Maximal Inequality: Suppose that $\{X_k\}_1^n$ is a sequence of independent r.v.'s. Then for any $\alpha > 0$, $P(\max_{1 \leq k \leq n} S_n \geq 3\alpha) \leq 3\max_{1 \leq k \leq n} P(|S_k| \geq \alpha) \leq 3Var(S_n)/\alpha_2$.

Kronecker's Lemma: Let $\{x_k\}_1^\infty$ be a sequence of real numbers and $\{a_k\}_1^\infty$ be an increasing sequence of positive real numbers such that $a_k \nearrow \infty$. If $\sum_{n \geq 1} x_n/a_n$ converges, then $\frac{1}{a_n}\sum_{k=1}^n x_k \to 0$.

Corollary: Let $\{Y_k\}_1^\infty$ be iid r.v.'s with mean zero and a common variance σ^2, then $\frac{1}{n^{1/2+\varepsilon}}\sum_{k=1}^n Y_k \to 0$ a.s.

Corollary: If $\sum 1/a_k^2 < \infty$, then $1/a_n \sum_{k=1}^n Y_k \to 0$ a.s. Let $a_k = \sqrt{k}\ln(k)$, $k \geq 2$, then $\sum 1/a_k^2 < \infty$ and $1/a_n \sum_{k=1}^n Y_k \to 0$ a.s.

Law of Iterated Logarithms (LIL): We have that $P(\limsup_n 1/(\sigma\sqrt{2n\ln(\ln n)})\sum_{k=1}^n Y_k = 1) = 1$, $P(\limsup_n 1/(\sigma\sqrt{2n\ln(\ln n)})\sum_{k=1}^n Y_k = -1) = 1$, so the limit points of $1/(\sigma\sqrt{2n\ln(\ln n)})\sum_{k=1}^n Y_k \in [-1, 1]$.

5.5 Weak and Distribution Convergence

Convergence in Distribution: We say Y_n converges in distribution to Y, write $Y_n \to^d Y$, if $F_{Y_n}(y) \to F_Y(y)$ for all y such that $P(Y = y) = 0$. Note that $F_{Y_n} \to F_Y$ iff $d(F_{Y_n}, F) \to 0$, where $d(\cdot, \cdot)$ is the Levy distance.

Theorem: Probability measures μ and ν on $(S, \mathcal{S})$ are the same iff $\int_S f(s)\,d\mu(s) = \int_S f(s)\,d\nu(s)$ for all bounded and continuous $f : S \to \mathbb{R}$.

Weak Convergence of Probability Measures: Let $\{\mu_n\}_1^\infty$, μ be probability measures on $(S, \mathcal{S})$. We say that μ_n converges weakly to μ, write $\mu_n \to^w \mu$, if $\int f\,d\mu_n \to \int f\,d\mu$ for all bounded and continuous $f : S \to \mathbb{R}$.

Theorem: Let $S = \mathbb{R}$, $\mathcal{S} = \mathcal{B}(\mathbb{R})$. Then $X_n \to^d X$ iff $P_{X_n^{-1}} \to P_{X^{-1}}$, where $P_{Y^{-1}}(A) = P(Y \in A)$.

Theorem: $X_n \to^d X$ iff $E[g(X_n)] \to E[g(X)]$ for all bounded and continuous $g : \mathbb{R} \to \mathbb{R}$.

Lemma: For all $\varepsilon > 0$ there exists M_ε and N_ε such that for all $n \geq N_\varepsilon$, $P(|X_n| > M_\varepsilon) < \varepsilon$.

Theorem: $X_n \to^p X$ implies that $X_n \to^d X$.

Skorohod's Theorem: If $X_n \to^d X$, then there exists some probability space with $\tilde{X}_n$, $\tilde{X}$ defined on it with $\tilde{X}_n =^d X_n$, $\tilde{X} =^d X$ and $\tilde{X}_n \to \tilde{X}$ a.s..

Continuous Mapping Theorem: Suppose $h : \mathbb{R} \to \mathbb{R}$ is measurable. Let $D_h = \{x \in \mathbb{R} : h \text{ is not continuous at } x\}$. If $X_n \to^d X$ and $P(X \in D_h) = 0$, then $h(X_n) \to^d h(X)$.

Corollary: If $X_n \to^d a$ and h is continuous at $x = a$, then $h(X_n) \to^p h(a)$.

Portmanteau Theorem: Let $\{X_n\}_1^\infty$, X be real valued r.v.'s with induced probability measures $\{\mu_n\}_1^\infty$, μ. The following are equivalent. (1) $\mu_n \to^w \mu$, (2) $\mu_n(A) \to \mu(A)$ for all A such that $\mu(\partial A) = 0$, and (3) $X_n \to^d X$.

Theorem: Let $M_n = \max\{X_1, \ldots, X_n\}$. Suppose that for some $\alpha > 0$ $\lim_{x\to\infty} x^\alpha(1 - F(x)) = c > 0$, then $Y_n = \frac{M_n}{c n^{1/\alpha}} \to^d Y$ with distribution function H i.e. $H(y) = e^{-y^{-\alpha}}$ for $y > 0$ and 0 for $y \geq 0$.

Portmanteau Theorem II: The following are equivalent. (1) $\mu_n \to^w \mu$, (2) $\mu_n(A) \to \mu(A)$ for all A such that $\mu(\partial A) = 0$, (3) $\limsup_n \mu_n(F) \leq \mu(F)$ for all $F \in \mathcal{S}$ closed, and (4) $\liminf_n \mu_n(G) \geq \mu(G)$ for all $G \in \mathcal{S}$ open.

5.6 Tightness and Weak Convergence

Tightness: A sequence of probability measures $\{\mu_n\}_1^\infty$ on the metric space (S, ρ) is said to be tight if for each $\varepsilon > 0$ there exists a compact set M_ε such that $\mu_n(M_\varepsilon) > 1 - \varepsilon$ for all $n \geq 1$.

Lemma: If $\mu_n \to^w \mu_0$, then $\{\mu_n\}_1^\infty$ is tight.

Vague Convergence: We say that ν_n converges vaguely to ν, write $\nu_n \to^v \nu$, if $\int g\,d\nu_n \to \int g\,d\nu$ for all continuous and bounded g that vanish at $\pm\infty$.

Lemma: If $\{\mu_n\}_1^\infty$ is tight and $\lim \mu_n((-\infty, x]) = G(x)$ at all continuity points of G, where G is nondecreasing and cadlag, then there exists a probability measure μ such that $\mu_n \to^w \mu$.

Theorem: If $\{\mu_n\}_1^\infty$ is tight, then there exists a probability measure μ and a subsequence $\mu_{n_k} \to^w \mu$.

Prohorov's Theorem: $\{\mu_\alpha : \alpha \in A\}$ is tight iff each subsequence has a weakly convergent subsequence.

Helly's Selection Theorem: For any sequence of distribution functions there exists a function F that is nondecreasing and cadlag with $F_{n_j}(x) \to F(x)$ for all continuity points of F.

Corollary: If $\{\mu_n\}_1^\infty$ is tight, and each weakly convergent subsequence converges weakly to a given μ, then $\mu_n \to^w \mu$.

5.7 Characteristic Functions

Characteristic Functions: We define $\varphi_X(t) = E[e^{itX}] = \int e^{itx}\, d\mu(x)$. Note that $\varphi(0) = 1$, $|\varphi(t)| \le 1$, φ is uniformly continuous, $\varphi(-t) = \overline{\varphi(t)}$, $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$, and $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$ for X, Y independent.

Theorem: $\varphi_X(t) = \varphi_Y(t)$ for all $t \in \mathbb{R}$ iff $X =^d Y$ and $\int e^{itx}\, d\mu(x) = \int e^{itx}\, d\nu(x)$ for all $t \in \mathbb{R}$ iff $\mu = \nu$.

Theorem: There is a unique correspondence between a probability measure μ and it characteristic function.

Uniqueness Theorem: The characteristic function, $\varphi_X(t)$, of X uniquely determines the distribution of X.

Inversion Theorem: If $P(X = a) = P(X = b) = 0$ and $\varphi_X \in L^1$, then

$$F_X(b) - F_X(a) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{i}{s}(e^{-isb} - e^{-isa})\varphi_X(s)\, ds.$$

Corollary: If $\varphi \in L^1$, then F_X is differentiable and

$$F_X'(x) = f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\varphi_X(t)\, dt.$$

Lemma: For $r > 0$ and take X to be a r.v. with characteristic function φ, then $\frac{1}{r}\int_{-r}^{r} 1 - \varphi(t)\, dt \ge P(|X| > \frac{2}{r})$.
Continuity Theorem: $X_n \to^d X$ iff $\varphi_{X_n}(t) \to \varphi_X(t)$ for all $t \in \mathbb{R}$.

Binomial Central Limit Theorem: Let $\{Y_k : k \ge 1\}$ be 0,1 symmetric Bernoulli r.v.'s and $P(Y_k = 0) = P(Y_k = 1) = \frac{1}{2}$. Let $\{S_n : n \ge 1\}$ be a sequence of binomial r.v.'s defined by $S_n \sum_{k=1}^{n} Y_k$. Then $P(S_n = k) = \binom{n}{k} 2^{-n}$, $E[S_n] = \frac{n}{2}$, and $\mathrm{Var}(S_n) = n(\frac{1}{2})^2$.

Corollary: If $\{\varphi_n : n \ge 1\}$ is a sequence of characteristic functions with $\varphi_n(t) \to \theta(t)$ for all $t \in \mathbb{R}$ with θ being continuous at $t = 0$, then θ is a characteristic function and $X_n \to^d Y$ where

$\varphi_Y(t) = \theta(t)$.

Corollary: Suppose $\{\mu_n : n \geq 1\}$ is tight and $\varphi_{X_n}(t) \to \psi(t)$. Then $\psi(t) = \varphi_Y(t)$ for some Y with $X_n \to^d Y$.

5.8 Central Limit Theorem

Lemma p: If $X \in L^p$, then $\varphi_X^{(p)}(t)$ exists for all $t \in \mathbb{R}$ and $\varphi_X^{(p)}(t) = i^p E[X^p e^{itX}]$.

Corollary: If $X \in L^p$ for all $p \geq 1$, then $\varphi_X(t) = \sum_{k \geq 0} \frac{i^k t^k E[X^K]}{k!}$.

Lemma: If $X \in L^2$, then $|\varphi_X(t) - (1 + itE[X] - \frac{t^2 E[X^2]}{2}| \leq E[\min(t^2 X^2, \frac{|tX|^3}{3!})]$.

Corollary: If Z is a standard normal r.v., then $|e^{-t^2/2} - (1 - t^2/2)| \leq \frac{\sqrt{2}}{3\sqrt{\pi}}|t|^3$.

Lemma: If $\{w_k : 1 \leq k \leq m\}$ and $\{z_k : 1 \leq k \leq m\}$ are complex numbers such that $|w_k| leq 1$ and $|z_k| \leq 1$ for $k = 1, \ldots, m$, then $|\prod_{k=1}^m z_k - \prod_{k=1}^m w_k| \leq \sum_{k=1}^m |z_k - w_k|$.

Lemma: If $\{\sigma_{n,k} : 1 \leq k \leq k_n\}$ are nonnegative with $\sum_{k=1}^{k_n} \sigma_{n,k}^2 = 1$, then

$$\left| e^{-t^2/2} - \prod_{k=1}^{k_n} \left(1 - \frac{t^2 \sigma_{n,k}^2}{2} \right) \right| \leq \frac{\sqrt{2}|t|^3}{3\sqrt{\pi}} \max_{1 \leq k \leq k_n} \sigma_{n,k}.$$

Lindeberg's CLT: For each n let $\{X_{n,k} : 1 \leq k \leq k_n\}$ be a sequence of independent r.v.'s with $EX_{n,k} = 0$, $\text{Var}(X_{n,k}) = \sigma_{n,k}^2$. Let $s_n^2 = \sum_{k=1}^{k_n} \sigma_{n,k}^2$ and assume that for any $\varepsilon > 0$ Lindeberg's condition is satisfied:

$$\lim_{n \to \infty} \frac{\sum_{k=1}^{k_n} X_{n,k}^2 \mathbf{1}_{[|X_{n,k}| > \varepsilon s_n]}}{s_n^2} = 0.$$

Then $\frac{1}{s_n} \sum_{k=1}^{k_n} X_{n,k} \to^d Z$.

Feller's Theorem: Suppose that $\{X_{n,k} : 1 \leq k \leq k_n\}$ are independent with $E[X_{n,k}] = 0$, $\text{Var}(X_{n,k}) = \sigma_{n,k}^2 < \infty$. Let $s_n^2 = \sum_{k=1}^{k_n} \sigma_{n,k}^2$. If $\frac{1}{s_n} \sum_{k=1}^{k_n} X_{n,k} \to^d Z$ and $\max_{1 \leq k \leq k_n} P(|X_{n,k}| > \varepsilon s_n) \to 0$, then the Lindeberg Condition holds.

Same Type: If $Y =^d aX + b$ for $a > 0$, $b \in \mathbb{R}$ (equivalently $F_Y(y) = F_X(\frac{y-b}{a})$ or $\varphi_Y(t) = e^{itb}\varphi_X(at)$) then we say that X and Y are of the same type.

Stable Distribution: Let $\{X_k : k \geq 1\}$ be iid. If $S_n = \sum_{k=1}^n X_k$ is of the same type as X_1 for all $n \geq 1$, then we say that X_1 has a stable distribution.

Theorem: If X is a symmetric, stable, nondegenerate r.v., then $\varphi_X(t) = e^{-c|t|^\alpha}$ for some $c > 0$ and $\alpha \in (0, 2]$.

Lemma: If $\alpha > 2$, then $e^{-c|t|^\alpha}$ is not a characteristic function.

Levy's Theorem: If $\{X_k : k \geq 1\}$ is iid and there exists $a_n > 0$, $b_n \in \mathbb{R}$ such that $\frac{S_n - b_n}{a_n} \to^d Y$, then Y is a stable r.v..

Corollary 1: If X is a symmetric standard r.v., then $\frac{S_n}{n^{1/\alpha}} =^d X$.

Corollary 2: If X is stable, then $a_n = n^{1/\alpha}$.

Lemma 1: If $\{(X_n, Y_n) : n \geq 1\}$ with X_n, Y_n independent for each n, $X_n \to^d X$, $Y_n \to^d Y$, then $X_n + Y_n \to^d Z$, where $Z =^d X + Y$ and X, Y are independent.

Corollary: If for $j = 1, \ldots, k$ (k fixed), $X_{j,n} \to^d X_j$ and $\{X_{1,n}, \ldots, X_{k,n}\}$ are independent, then $\sum_{j=1}^k X_{j,n} \to^d \sum_{j=1}^k X_j$, where $\{X_1, \ldots, X_k\}$ are independent.

Lemma 2: If for $a_n > 0$, $b_n \in \mathbb{R}$, $\frac{X_n - b_n}{a_n} \to^d X$ and $\alpha_n / a_n \to 1$, $(\beta_n - b_n)/a_n \to 0$, then $\frac{X_n - \beta_n}{\alpha_n} \to^d X$.

Lemma 3: Suppose that $X_n \to^d X$ and $\frac{X_n - b_n}{a_n} \to^d Y$ for some $a_n > 0$, $b_n \in \mathbb{R}$. If neither X nor Y is degenerate, then X and Y are of the same type.

5.9 Infinitely Divisible

Infinitely Divisible: X (or F_X, $P_{X^{-1}}$, φ_X) is infinitely divisible if for each $n \geq 1$, $X =^d \sum_{k=1}^n X_{n,k}$, where $\{X_{n,k} : 1 \leq k \leq n\}$ are iid.

Lemma: Let $\{Y_k : k \geq 1\}$ be iid r.v.'s and N an independent Poisson r.v. with parameter λ. Then $S_N = \sum_{k=1}^N Y_k$ is an infinitely divisible r.v..

Compound Poisson Random Variable: The r.v. S_N above is a compound Poisson r.v..

Theorem 1: Suppose for each $n \geq 1$ $\{Y_{n,k} : 1 \leq k \leq n\}$ are iid. If $S_n = \sum_{k=1}^n Y_{n,k} \to^d X$, then X is infinitely divisible.

Theorem 2: X is infinitely divisible iff $\varphi_X(t) = \exp\left[i\beta t - \frac{\sigma^2 t^2}{2} + \int_{\mathbb{R}} \left(e^{itx} - 1 - \frac{itx}{1+x^2}\right) d\mu(x)\right]$ for some finite measure μ on $(\mathbb{R}, \mathcal{B})$ with $\mu(\{0\}) = 0$.

Theorem 3: X is infinitely divisible iff there exists a sequence of compound Poisson r.v.'s converging in distribution to X.