Miscellaneous analysis notes

Danny Nygård Hansen

26th March 2022

1 • The exponential function

THEOREM 1.1

Given $b > 0$ there exists a unique continuous function $E_b\colon \mathbb{R} \to \mathbb{R}$ such that $E_b(m/n) = (b^m)^{1/n}$ for all $m, n \in \mathbb{Z}$ with $n > 0$.

PROOF. We first assume that $b > 1$. The proof of the theorem in this case is by the following stages:

(1) Given $m, n, p, q \in \mathbb{Z}$ with $n, q > 0$ and $r = m/n = p/q$, then

$$(b^m)^{1/n} = (b^p)^{1/q}.$$

Thus defining $E_b(r) = (b^m)^{1/n}$ makes sense.

(2) If $r, s \in \mathbb{Q}$, then $E_b(r + s) = E_b(r)E_b(s)$. In particular, E_b is increasing on $\mathbb{Q}$.

(3) The map E_b is continuous on $\mathbb{Q}$.

(4) For $x \in \mathbb{R}$, let

$$L_x = \{E_b(r) \mid t \in \mathbb{Q}, t \leq x\}.$$

Then $E_b(r) = \sup L_r$ for $r \in \mathbb{Q}$, so it makes sense to define $E_b(x) = \sup L_x$ for all $x \in \mathbb{R}$. Hence E_b is increasing and therefore continuous on $\mathbb{R}$.

(5) For $x, y \in \mathbb{R}$ we have $E_b(x + y) = E_b(x)E_b(y)$.

(1) Notice that $mq = pn$, so

$$\left[(b^m)^{1/n}\right]^{pn} = (b^m)^p = (b^p)^m = \left[(b^p)^{1/q}\right]^{mq} = \left[(b^p)^{1/q}\right]^{pn}.$$

The (positive) pnth root of this number is unique, so $(b^m)^{1/n} = (b^p)^{1/q}$.

(2) Write $r = m/n$ and $s = p/q$ for appropriate $m,n,p,q \in \mathbb{Z}$ with $n,q > 0$. Then

$$[(b^m)^{1/n}(b^p)^{1/q}]^{nq} = b^{mq}b^{pn} = b^{mq+pn}.$$

Taking the nqth root implies that

$$E_b(r)E_b(s) = (b^m)^{1/n}(b^p)^{1/q} = (b^{mq+pn})^{1/nq} = E_b(r+s),$$

where we in the last equality use that

$$r + s = \frac{m}{n} + \frac{p}{q} = \frac{mq+pn}{nq}.$$

To see that E_b is increasing on $\mathbb{Q}$, notice that the assumption that $b > 1$ implies that $E_b(s) > 1$ for $s > 0$. If also $r \in \mathbb{Q}$, then

$$E_b(r+s) = E_b(r)E_b(s) \geq E_b(r),$$

so E_b is increasing.

(3) We begin by showing that E_b is continuous from above at 0. Since E_b is monotonic and $E_b(0) = 1$, it suffices to show that $\lim_{n\to\infty} E_b(r_n) = 1$ for some sequence $(r_n)_{n\in\mathbb{N}}$ in $\mathbb{Q}$ that decreases to 0. We claim that the sequence given by $r_n = 1/n$ has this property. Clearly $1/n \downarrow 0$, so assume that $E_b(1/n)$ did not converge to 1. Then there would be some $\varepsilon > 0$ such that $E_b(1/n) \geq 1+\varepsilon$, i.e. $b \geq (1+\varepsilon)^n$, for all $n \in \mathbb{N}$. But by [Bernoulli's inequality] this is impossible, so we must have $E_b(1/n) \to 1$. A similar argument shows that E_b is continuous from below at 0, using the fact that $E_b(-1/n) = (1/b)^n$.

Finally let $r,h \in \mathbb{Q}$, and notice that

$$E_b(r+h) = E_b(r)E_b(h) \to E_b(r)E_b(0) = E_b(r)$$

as $n \to \infty$. Thus E_b is also continuous at r.

(4) We clearly have $E_b(r) \leq \sup L_r$. For the opposite inequality, notice that $E_b(r)$ is an upper bound for L_r since E_b is increasing on $\mathbb{Q}$. Hence also $\sup L_r \leq E_b(r)$.

If $x \leq y$ then $L_x \subseteq L_y$, and hence $E_b(x) \leq E_b(y)$. Thus E_b is monotonic on $\mathbb{R}$. But then since E_b continuous on a dense subset of $\mathbb{R}$, it is clearly also continuous on $\mathbb{R}$.

(5) Let (r_n) and (s_n) be sequences in $\mathbb{Q}$ with limits x and y respectively. Then $r_n + s_n$ converges to $x+y$, so

$$E_b(x+y) = \lim_{n\to\infty} E_b(r_n+s_n) = \lim_{n\to\infty} E_b(r_n)E_b(s_n) = E_b(x)E_b(y). \qquad \square$$

2 • Introduction

DEFINITION 2.1

Let X be a set. A *sequence* in X is a map $a\colon \mathbb{N} \to X$. For $n \in \mathbb{N}$, we usually write a_n for $a(n)$ and denote a by $(a_n)_{n\in\mathbb{N}}$ or simply (a_n).

DEFINITION 2.2

Let (S,ρ) be a metric space, and let $(a_n)_{n\in\mathbb{N}}$ be a sequence in S. We say that (a_n) *converges* to a point $a \in S$ if for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that $n \geq N$ implies that $\rho(a_n, a) < \varepsilon$. In this case we call a the *limit* of (a_n) and write $a_n \to a$ as $n \to \infty$, and we say that (a_n) is *convergent*.

Furthermore, (a_n) is called a *Cauchy sequence* if for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that $m,n \geq N$ implies that $\rho(a_m, a_n) < \varepsilon$. If every Cauchy sequence in S is convergent, then S is said to be *complete*.

Notice that limits of sequences in metric spaces are unique. It is also clear that convergent sequences are Cauchy, and that Cauchy sequences are bounded: We say that a sequence $(a_n)_{n\in\mathbb{N}}$ in a metric space is *bounded* if the set $\{a_n \mid n \in \mathbb{N}\}$ is bounded.

If $(X,\leq)$ is a poset, a sequence $(a_n)_{n\in\mathbb{N}}$ in X is *increasing* (*decreasing*) if $m \leq n$ implies $a_m \leq a_n$ ($a_m \geq a_n$) for all $m,n \in \mathbb{N}$. If $m < n$ implies $a_m < a_n$ ($a_m > a_n$), then (a_n) is *strictly* increasing (decreasing). A sequence that is either (strictly) increasing or (strictly) decreasing is called *(strictly) monotonic*.

Given a sequence $(a_n)_{n\in\mathbb{N}}$ in a set X and a strictly increasing sequence $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$, the sequence $(a_{n_k})_{k\in\mathbb{N}}$ is called a *subsequence* of (a_n). In particular, every sequence is a subsequence of itself.

LEMMA 2.3

Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in a metric space (S,ρ). If (a_n) is both Cauchy and has a convergent subsequence, then (a_n) itself is convergent.

PROOF. Let $(a_{n_k})_{k\in\mathbb{N}}$ be a convergent subsequence of (a_n), and let $\varepsilon > 0$. Choose $N_1, N_2 \in \mathbb{N}$ such that

$$m,n \geq N_1 \quad \Rightarrow \quad \rho(a_m, a_n) < \frac{\varepsilon}{2}$$

and

$$k \geq N_2 \quad \Rightarrow \quad \rho(a_{n_k}, a) < \frac{\varepsilon}{2},$$

where $a \in S$ is the limit of (a_{n_k}). For $n \geq N_1 \vee N_2$ we thus have

$$\rho(a_n, a) \leq \rho(a_n, a_m) + \rho(a_m, a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

showing that $a_n \to a$ as $n \to \infty$. □

2.1. Sequences of real numbers

Proposition 2.4

Let $(a_n)_{n\in\mathbb{N}}$ be a monotonic sequence in $\mathbb{R}$. Then (a_n) is convergent if and only if it is bounded, in which case it converges to $\sup_{n\in\mathbb{N}} a_n$ if it is increasing and $\inf_{n\in\mathbb{N}} a_n$ if it is decreasing.

Proof. If (a_n) is convergent then it is bounded, so assume that it is bounded and let $\varepsilon > 0$. For definiteness we assume that it is increasing and let $s = \sup_{n\in\mathbb{N}} a_n$. By definition of s there exists an $N \in \mathbb{N}$ such that $s - a_N < \varepsilon$. Since (a_n) is increasing and s is an upper bound of the sequence, we thus have

$$0 \le s - a_n < \varepsilon$$

for all $n \ge N$, proving that $a_n \to s$. □

Lemma 2.5

Every sequence in $\mathbb{R}$ has a monotonic subsequence.

Proof. Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. We say that $n \in \mathbb{N}$ is a *peak* if $a_n \ge a_m$ for all $m \ge n$. If (a_n) has infinitely many peaks, the subsequence consisting of these constitute a decreasing subsequence.

Hence we assume that (a_n) only has finitely many peaks. We construct an increasing sequence $(n_k)_{n\in\mathbb{N}}$ in $\mathbb{N}$ as follows: Let $n_1 \in \mathbb{N}$ be such that all peaks are strictly less than n_1, and assume that $n_1, \dots, n_{k-1}$ have been chosen such that $a_1 \le \cdots \le a_{n_{k-1}}$. Since $a_{n_{k-1}}$ is not a peak there is an $n' > n_{k-1}$ such that $a_{n_{k-1}} < a_{n'}$. Letting $n_k = n'$ we obtain an increasing subsequence (a_{n_k}) of (a_n), proving the claim. □

Theorem 2.6: *The Bolzano–Weierstrass theorem*

Every subset of $\mathbb{R}^d$ is sequentially compact if and only if it is closed and bounded.

We recall that a topological space X is *sequentially compact* if every sequence in X has a convergent subsequence.

Proof. We begin with the case $d = 1$. Let $A \subseteq \mathbb{R}$ be closed and bounded, and let $(a_n)_{n\in\mathbb{N}}$ be a sequence in A. Let (a_{n_k}) be a monotonic subsequence of (a_n), and notice that (a_{n_k}) is convergent since it is bounded.

The case for general d follows by induction in d, by noticing that a sequence in $\mathbb{R}^d$ converges if and only if each coordinate sequence converges.

For the converse, let $A \subseteq \mathbb{R}^d$ be sequentially compact. If A were not bounded we could choose $a_n \in A \cap B(0,n)$ for all $n \in \mathbb{N}$, yielding a sequence

(a_n) with no convergent subsequence. Furthermore, if (a_n) is a sequence in A converging to a point $a \in \mathbb{R}^d$, it is in particular a Cauchy sequence. Since it has a subsequence converging to a point $a' \in A$, and by [lemma] we must have $a = a'$. Thus A is also closed. □

Theorem 2.7: *Completeness of $\mathbb{R}^d$*

The Euclidean space $\mathbb{R}^d$ is complete.

Proof. Let $(a_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathbb{R}^d$. Hence it is bounded, and so it has a convergent subsequence by the Bolzano–Weierstrass theorem. But then (a_n) itself converges by [lemma], so $\mathbb{R}^d$ is complete. □

Theorem 2.8: *The Heine–Borel theorem*

Every subset of $\mathbb{R}^d$ is compact if and only if it is closed and bounded.

Proof. Of course every compact set is closed in any Hausdorff space and bounded in any metric space, so we only consider the other implication.

We first show that closed and bounded intervals are compact. Consider the interval $[a,b]$, and let $\mathcal{U}$ be an open cover of $[a,b]$. Define the set

$$A = \left\{ x \in [a,b] \;\middle|\; [a,x] \text{ has a finite subcover in } \mathcal{U} \right\}.$$

We clearly have $a \in A$ since a point is covered by a single set in $\mathcal{U}$. If $s = \sup A$ then $a \leq s \leq b$. Suppose that $s < b$ and choose a set $U \in \mathcal{U}$ with $s \in U$. There exist $r,t \in U$ such that $r < s < t$, and so $r \in A$. Let $\mathcal{U}'$ denote a finite subcover of $[a,r]$ in $\mathcal{U}$. Then $\mathcal{U}' \cup \{U\}$ is a finite subcover of $[a,t]$, contradicting the assumption that $s < b$. Hence $s = b$.

Next, choose $V \in U$ with $b \in V$, and let $c \in V$ with $c < b$. Then $c \in A$, and adjoining V to a finite subcover of $[a,c]$ yields a finite subcover of $[a,b]$, so $b \in A$. Thus $[a,b]$ is compact.

Finally, let $K \subseteq \mathbb{R}^d$ be closed and bounded. Since it is bounded it is contained in some cube $[-a,a]^d$. But this cube is a product of compact sets and hence compact, so K is a closed subset of a compact set. The claim follows. □

If A is a subset of a metric space S, recall that A is *totally bounded* if, for every $\varepsilon > 0$, A can be covered by finitely many open balls of radius ε.

Theorem 2.9

If A is a subset of a metric space (S,ρ), then the following are equivalent:

(i) *A is complete and totally bounded.*

(ii) *A is sequentially compact.*

(iii) *A is compact.*

PROOF. *(i) $\Rightarrow$ (ii):* Assume that A is complete and totally bounded, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in A. Now A can be covered by finitely many balls of radius 1, at least one of which, say B_1, contains x_n for infinitely many n, say for $n \in N_1 \subseteq \mathbb{N}$. Similarly, $A \cap B_1$ may be covered by finitely many balls of radius 1/2, and again there is a ball B_2 containing x_n for infinitely many $n \in N_1$, say for $n \in N_2$. Continuing recursively we obtain a sequence of balls B_i of radius $1/i$ and a decreasing sequence $(N_i)_{i\in\mathbb{N}}$ of infinite subsets of $\mathbb{N}$ such that $x_n \in B_i$ for $n \in N_i$.

Next, choose a strictly increasing sequence $(n_i)_{i\in\mathbb{N}}$ of naturals numbers such that $n_i \in N_i$. Then $\rho(x_{n_i}, x_{n_j}) < 2/i$ for $i \leq j$, so $(x_{n_i})_{i\in\mathbb{N}}$ is a Cauchy sequence., and since A is complete it has a limit in A.

(ii) $\Rightarrow$ (i): Assume that A is sequentially compact. We first show that A is complete, so let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in A. This has a subsequence that converges to a point x in A, so [lemma] implies that (x_n) also converges to x.

Now suppose that A is not totally bounded, and let $\varepsilon > 0$ be such that A cannot be covered by finitely many ε-balls. We construct a sequence $(x_n)_{n\in\mathbb{N}}$ in A as follows: Choose any $x_1 \in A$, and given $x_1, \ldots, x_n$ choose $x_{n+1} \in A \setminus \bigcup_{i=1}^{n} B(x_i, \varepsilon)$. Then $\rho(x_m, x_n) \geq \varepsilon$ for all $m, n \in \mathbb{N}$ with $m \neq n$, so (x_n) has no convergent subsequence.

(i) & (ii) $\Rightarrow$ (iii): Suppose that A is complete, totally bounded and sequentially compact, and let $\mathcal{U}$ be an open cover of A. It suffices to show that there some $\varepsilon > 0$ such that any ε-ball intersecting A is contained in some $U \in \mathcal{U}$, since A can be covered by finitely many such balls.

Assume towards a contradiction that for every $n \in \mathbb{N}$ there is a ball B_n of radius $1/n$ intersecting A such that B_n is contained in no $U \in \mathcal{U}$. Picking $x_n \in B_n$ for $n \in \mathbb{N}$, we may assume that the sequence $(x_n)_{n\in\mathbb{N}}$ converges to some $x \in A$ by passing to an appropriate subsequence. Then $x \in U$ for some $U \in \mathcal{U}$, and since U is open there is an $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U$. Choosing $n \in \mathbb{N}$ large enough that $\rho(x_n, x) < \varepsilon/2$ and $1/n < \varepsilon/2$, we have $B_n \subseteq B(x, \varepsilon) \subseteq U$, which is a contradiction.

(iii) $\Rightarrow$ (ii): We prove the contrapositive, so assume that A is not sequentially compact, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in A with no convergent subsequence. Every $x \in A$ is then contained in an open ball B_x containing x_n for only finitely many n. Thus $\{B_x\}_{x\in A}$ is an open cover of A with no finite subcover, and A is not compact. $\square$

3 • Integration

3.1. Functions of bounded variation

A *partition* of an interval $[a,b]$ is a collection $P = \{x_0,\dots,x_n\}$ of real numbers such that

$$a = x_0 < \cdots < x_n = b.$$

In turn, a *tagged partition* of $[a,b]$ is a pair (P,T) where P is a partition of $[a,b]$ and $T = \{t_1,\dots,t_n\}$ is a multiset of numbers such that $t_i \in [x_{i-1},x_i]$ for all $i = 1,\dots,n$. Let $\mathcal{P}'[a,b]$ denote the set of tagged partitions of $[a,b]$. We define a direction on $\mathcal{P}'[a,b]$ by $(P,T) \preceq (P',T')$ if $P \subseteq P'$. Notice that T and T' do not appear in this definition. This also induces a direction on the set $\mathcal{P}[a,b]$ of all (non-tagged) partitions of $[a,b]$.

Given a partition $P = \{x_0,\dots,x_n\}$ of $[a,b]$ and a function $f\colon [a,b] \to \mathbb{R}$ we write $\Delta f_i = f(x_i) - f(x_{i-1})$ for $i = 1,\dots,n$. If f is the identity function we simply write $\Delta x_i = x_i - x_{i-1}$. Furthermore, we write

$$\Sigma_f(P) = \sum_{i=1}^{n} |\Delta f_i|.$$

Definition 3.1: *Total variation*

Consider a function $f\colon [a,b] \to \mathbb{R}$. The *total variation* of f on $[a,b]$ is the number

$$V_f(a,b) = \sup_{P\in\mathcal{P}[a,b]} \Sigma_f(P).$$

If $V_f(a,b) < \infty$, then we say that f is of *bounded variation* on $[a,b]$.

If f is of bounded variation on $[a,b]$, then it is clear that f is also of bounded variation on any subinterval of $[a,b]$. If $c \in (a,b)$ it is also easy to show that

$$V_f(a,b) = V_f(a,c) + V_f(c,b). \tag{3.1}$$

If $g\colon [a,b] \to \mathbb{R}$ is another function of bounded variation on $[a,b]$, then it is clear from the definition that $f + g$ is also of bounded variation.

Also note that monotonic functions are of bounded variation on any compact interval.

Lemma 3.2

Let $f\colon [a,b] \to \mathbb{R}$ be of bounded variation, and let $V(x) = V_f(a,x)$ for $x \in (a,b]$ and $V(a) = 0$. Then the functions V and $V - f$ are increasing on $[a,b]$.

PROOF. The function V is clearly increasing, so consider the function $D = V - f$. Let $x, y \in [a,b]$ with $x < y$, and notice that $f(y) - f(x) \leq V_f(x,y)$. Recalling (3.1) it follows that

$$D(y) - D(x) = V(y) - V(x) - (f(y) - f(x)) = V_f(y,x) - (f(y) - f(x)) \geq 0. \qquad \square$$

PROPOSITION 3.3

A function $f : [a,b] \to \mathbb{R}$ is of bounded variation if and only if it is the difference of two (strictly) increasing functions.

PROOF. By the lemma we can write f as the difference of two increasing functions as $f = V - (V - f)$. Adding a strictly increasing function to both V and $V - f$ yields the claim. $\square$

3.2. Integration

Next consider bounded functions $f, \alpha : [a,b] \to \mathbb{R}$. For each tagged partition (P,T) of $[a,b]$ we define the *Riemann–Stieltjes sum*

$$S_{f,\alpha}(P,T) = \sum_{i=1}^{n} f(t_i)\Delta\alpha_i.$$

This induces a net $S_{f,\alpha} : \mathcal{P}'[a,b] \to \mathbb{R}$.

DEFINITION 3.4: *Riemann–Stieltjes integral*

Let $f, \alpha : [a,b] \to \mathbb{R}$ be bounded functions. We say that f is *Riemann-integrable* with respect to α (or simply α*-integrable*) on $[a,b]$ if the net $S_{f,\alpha}$ has a limit A. In this case A is called the *Riemann–Stieltjes integral* of f with respect to α on $[a,b]$ and is denoted

$$\int_a^b f \,\mathrm{d}\alpha \quad \text{or} \quad \int_a^b f(x)\,\mathrm{d}\alpha(x).$$

We denote the set of α-integrable functions on $[a,b]$ by $\mathcal{R}_\alpha[a,b]$.

We call f the *integrand* and α the *integrator*. In the case where $\alpha(x) = x$, we use the notations

$$S_f, \quad \int_a^b f \quad \text{and} \quad \int_a^b f(x)\,\mathrm{d}x.$$

The sums S_f are then simply called *Riemann sums* and the integral the *Riemann integral* of f on $[a,b]$. With this choice of α, an α-integrable function is called *Riemann integrable* on $[a,b]$, and the set of such functions is denoted $\mathcal{R}[a,b]$.

Below we fix an interval $[a,b]$ and (bounded) integrators α and β on it.

Proposition 3.5: *Linearity of the integral*

Let $f,g \in \mathcal{R}_\alpha[a,b]$ and $c_1, c_2 \in \mathbb{R}$. Then:

(i) *$c_1 f + c_2 g$ is α-integrable on $[a,b]$ and*

$$\int_a^b (c_1 f + c_2 g)\,\mathrm{d}\alpha = c_1 \int_a^b f\,\mathrm{d}\alpha + c_2 \int_a^b g\,\mathrm{d}\alpha.$$

In particular, $\mathcal{R}_\alpha[a,b]$ is a vector space.

(ii) *f is $(c_1\alpha + c_2\beta)$-integrable on $[a,b]$ and*

$$\int_a^b f\,\mathrm{d}(c_1\alpha + c_2\beta) = c_1 \int_a^b f\,\mathrm{d}\alpha + c_2 \int_a^b f\,\mathrm{d}\beta.$$

Proof. This follows immediately from the bilinearity of the map $(f,\alpha) \mapsto S_{f,\alpha}$ along with basic properties of nets. □

Proposition 3.6: *Integration by parts*

Given functions $f,\alpha\colon [a,b] \to \mathbb{R}$, assume that f is α-integrable on $[a,b]$. Then α is f-integrable on $[a,b]$ and

$$\int_a^b f\,\mathrm{d}\alpha + \int_a^b \alpha\,\mathrm{d}f = f(b)\alpha(b) - f(a)\alpha(a).$$

Proof. Let (P,T) be a tagged partition of $[a,b]$. Then an easy calculation shows that

$$f(b)\alpha(b) - f(a)\alpha(a) - S_{\alpha,f}(P,T) = S_{f,\alpha}(P \cup T, P'),$$

where P' is obtained from P by duplicating appropriate elements such that each subinterval of $P \cup T$ contains the corresponding element from P'. Since $P \cup T$ is finer than P, the claim follows by taking the limit of $S_{\alpha,f}$. □

Proposition 3.7

Let $\alpha \in C^1[a,b]$ and $f \in \mathcal{R}_\alpha[a,b]$. Then $f\alpha' \in \mathcal{R}[a,b]$, and

$$\int_a^b f\,\mathrm{d}\alpha = \int_a^b f\alpha'.$$

Proof. Consider the Riemann(–Stieltjes) sums

$$S_{f\alpha'}(P,T) = \sum_{i=1}^n f(t_i)\alpha'(t_i)\Delta x_i \quad \text{and} \quad S_{f,\alpha}(P,T) = \sum_{i=1}^n f(t_i)\Delta\alpha_i.$$

By the mean value theorem we can write $\Delta\alpha_i = \alpha'(s_i)\Delta x_i$ for appropriate $s_i \in (x_{i-1}, x_i)$. It follows that

$$S_{f,\alpha}(P,T) - S_{f\alpha'}(P,T) = \sum_{i=1}^{n} f(t_i)(\alpha'(s_i) - \alpha'(t_i))\Delta x_i.$$

By uniform continuity of α', given $\varepsilon > 0$ there exists a $\delta > 0$ such that $|x-y| < \delta$ implies $|\alpha'(x) - \alpha'(y)| < \varepsilon$ for all $x, y \in [a,b]$. If $\|P\| < \delta$ we thus have

$$|S_{f,\alpha}(P,T) - S_{f\alpha'}(P,T)| \leq \|f\|_\infty \varepsilon(b-a).$$

But $S_{f,\alpha}(P,T)$ approaches the α-integral of f for finer and finer partitions, which proves the claim. $\square$

3.3. *Increasing integrators*

DEFINITION 3.8

Let $f, \alpha \colon [a,b] \to \mathbb{R}$ be bounded functions, and assume that α is increasing. Let $P = \{x_0, \dots, x_n\}$ be a partition of $[a,b]$, and let

$$M_i(f) = \sup\Big\{ f(x) \,\Big|\, x \in [x_{i-1}, x_i] \Big\},$$
$$m_i(f) = \inf\Big\{ f(x) \,\Big|\, x \in [x_{i-1}, x_i] \Big\}.$$

The numbers

$$U_{f,\alpha}(P) = \sum_{i=1}^{n} M_i(f)\Delta\alpha_i \quad \text{and} \quad L_{f,\alpha}(P) = \sum_{i=1}^{n} m_i(f)\Delta\alpha_i$$

are called the *upper and lower Stieltjes sums* of f with respect to α for the partition P.

It is immediate that

$$L_{f,\alpha}(P) \leq S_{f,\alpha}(P,T) \leq U_{f,\alpha}(P)$$

for any tagged partition (P,T) of $[a,b]$. It is also obvious that, if $P \subseteq P'$, then

$$U_{f,\alpha}(P) \geq U_{f,\alpha}(P') \quad \text{and} \quad L_{f,\alpha}(P) \leq L_{f,\alpha}(P'),$$

and that for any pair of partitions P_1 and P_2 we have

$$L_{f,\alpha}(P_1) \leq U_{f,\alpha}(P_2). \tag{3.2}$$

DEFINITION 3.9

Let $f,\alpha\colon [a,b] \to \mathbb{R}$ be bounded functions with α increasing. Then the numbers

$$\overline{\int_a^b} f \,\mathrm{d}\alpha = \inf\Big\{ U_{f,\alpha}(P) \,\Big|\, P \in \mathcal{P}[a,b] \Big\}$$

and

$$\underline{\int_a^b} f \,\mathrm{d}\alpha = \sup\Big\{ L_{f,\alpha}(P) \,\Big|\, P \in \mathcal{P}[a,b] \Big\}$$

are called the *upper and lower Stieltjes integrals* of f with respect to α on $[a,b]$.

It follows immediately from the definition and (3.2) that the upper integral is always greater than the lower integral. We also use the notations $\overline{I}(f,\alpha)$ and $\underline{I}(f,\alpha)$ for the upper and lower integrals, respectively, when the interval $[a,b]$ is understood.

THEOREM 3.10: *Riemann's condition*

Let $f,\alpha\colon [a,b] \to \mathbb{R}$ be bounded functions with α increasing. Then the following conditions are equivalent:

(i) $f \in \mathcal{R}_\alpha[a,b]$.

(ii) *f satisfies* Riemann's condition *with respect to α on $[a,b]$: For every $\varepsilon > 0$ there exists a partition P of $[a,b]$ such that*

$$U_{f,\alpha}(P) - L_{f,\alpha}(P) < \varepsilon. \tag{3.3}$$

(iii) $\underline{I}(f,\alpha) = \overline{I}(f,\alpha)$.

In this case we have

$$\underline{\int_a^b} f \,\mathrm{d}\alpha = \int_a^b f \,\mathrm{d}\alpha = \overline{\int_a^b} f \,\mathrm{d}\alpha.$$

PROOF. *(i) $\Rightarrow$ (ii):* Let $\varepsilon > 0$, and choose a partition $P = \{x_0,\dots,x_n\}$ of $[a,b]$ such that

$$\left| \sum_{i=1}^n f(t_i)\Delta\alpha_i - \int_a^b f \,\mathrm{d}\alpha \right| < \varepsilon$$

for all $t_i \in [x_{i-1},x_i]$. It follows that

$$\left| \sum_{i=1}^n (f(t_i) - f(t_i'))\Delta\alpha_i \right| < 2\varepsilon$$

for all $t_i, t_i' \in [x_{i-1}, x_i]$. For any $\delta > 0$ there exist t_i, t_i' such that

$$f(t_i) - f(t_i') > M_i(f) - m_i(f) - \delta.$$

From this it follows that

$$\begin{aligned} U_{f,\alpha}(P) - L_{f,\alpha}(P) &= \sum_{i=1}^{n} (M_i(f) - m_i(f))\Delta\alpha_i \\ &< \sum_{i=1}^{n} (f(t_i) - f(t_i'))\Delta\alpha_i + \delta(\alpha(b) - \alpha(a)) \\ &< 3\varepsilon \end{aligned}$$

for an appropriate choice of δ. Since ε was arbitrary, this proves (ii).

(ii) $\Rightarrow$ (iii): If P is any partition of $[a,b]$ we have

$$L_{f,\alpha}(P) \leq \underline{\int_a^b} f \,\mathrm{d}\alpha \leq \overline{\int_a^b} f \,\mathrm{d}\alpha \leq U_{f,\alpha}(P).$$

Thus (3.3) implies that $0 \leq \overline{I}(f,\alpha) - \underline{I}(f,\alpha) < \varepsilon$ for every $\varepsilon > 0$, proving (iii).

(iii) $\Rightarrow$ (i): Let $\varepsilon > 0$. There exists a partition P of $[a,b]$ such that

$$\underline{I}(f,\alpha) - \varepsilon < L_{f,\alpha}(P) \leq S_{f,\alpha}(P,T) \leq U_{f,\alpha}(P) < \overline{I}(f,\alpha) + \varepsilon$$

for any choice of points T such that (P,T) is a tagged partition. Denoting the common value of $\underline{I}(f,\alpha)$ and $\overline{I}(f,\alpha)$ by A, this shows that $|S_{f,\alpha}(P',T') - A| < \varepsilon$ for all tagged partitions (P',T') with $P \subseteq P'$. Hence $f \in \mathcal{R}_\alpha[a,b]$, and the integral of f with respect to α equals A. □

3.4. *Integrators of bounded variation*

THEOREM 3.11

Let $\alpha\colon [a,b] \to \mathbb{R}$ be of bounded variation, and let $V(x) = V_\alpha(a,x)$ for $x \in (a,b]$ and $V(a) = 0$. Then $\mathcal{R}_\alpha[a,b] \subseteq \mathcal{R}_V[a,b]$.

PROOF. Let $f \in \mathcal{R}_\alpha[a,b]$, and choose $M > 0$ such that $|f| \leq M$. Choose a partition $P = \{x_0, \ldots, x_n\}$ of $[a,b]$ such that $V(b) < \sum_{i=1}^{n} |\Delta\alpha_i| + \varepsilon$. Then

$$\begin{aligned} \sum_{i=1}^{n} (M_i(f) - m_i(f))(\Delta V_i - |\Delta\alpha_i|) &\leq 2M \sum_{i=1}^{n} (\Delta V_i - |\Delta\alpha_i|) \\ &= 2M\left(V(b) - \sum_{i=1}^{n} |\Delta\alpha_i|\right) \\ &< 2M\varepsilon. \end{aligned}$$

Also choose P such that $|\sum_{i=1}^{n}(f(t_i)-f(t_i'))\Delta\alpha_i| < \varepsilon$ for all $t_i, t_i' \in [x_{i-1}, x_i]$. Next let $\delta > 0$. For $i = 1, \dots, n$, if $\Delta\alpha_i \geq 0$ choose t_i, t_i' such that

$$f(t_i) - f(t_i') > M_i(f) - m_i(f) - \delta.$$

If instead $\Delta\alpha_i < 0$, choose t_i, t_i' such that

$$f(t_i') - f(t_i) > M_i(f) - m_i(f) - \delta.$$

It follows that

$$\sum_{i=1}^{n}(M_i(f) - m_i(f))|\Delta\alpha_i| < \sum_{i=1}^{n}(f(t_i) - f(t_i'))\Delta\alpha_i + \delta V(b) < 2\varepsilon$$

for an appropriate choice of δ. Combining these inequalities yields

$$U_{f,V}(P) - L_{f,V}(P) = \sum_{i=1}^{n}(M_i(f) - m_i(f))\Delta V_i < 2(M+1)\varepsilon,$$

and since ε was arbitrary, this shows that $f \in \mathcal{R}_V[a,b]$. □

Since $\alpha = V - (V - \alpha)$ and both V and $V - \alpha$ are increasing, this allows us to reduce questions about integrators of bounded variation to questions about monotonic integrators. In particular it lets us use Riemann's condition to prove integrability with respect to integrators of bounded variation.

Proposition 3.12

Let $\alpha\colon [a,b] \to \mathbb{R}$ be of bounded variation, and let $f \in \mathcal{R}_\alpha[a,b]$. Choose $m, M \in \mathbb{R}$ such that $m \leq f \leq M$. If $\varphi\colon [m,M] \to \mathbb{R}$ is continuous, then $\varphi \circ f \in \mathcal{R}_\alpha[a,b]$.

Proof. We may assume that α is increasing. Put $g = \varphi \circ f$ and let $\varepsilon > 0$. Uniform continuity of φ yields a $\delta > 0$ such that $|s-t| < \delta$ implies $|\varphi(s)-\varphi(t)| < \varepsilon$ for $s, t \in [m, M]$. Also choose δ such that $\delta < \varepsilon$. Let $P = \{x_0, \dots, x_n\}$ be a partition of $[a,b]$ such that

$$U_{f,\alpha}(P) - L_{f,\alpha}(P) < \delta^2.$$

Let A consist of those numbers $i \in \{1, \dots, n\}$ such that $M_i(f) - m_i(f) < \delta$, and let B consist of the remaining i. For $i \in A$ we then have $M_i(g) - m_i(g) \leq \varepsilon$.

Let $K > 0$ be such that $|\varphi| \leq K$. For $i \in B$ we then have $M_i(g) - m_i(g) \leq 2K$. Furthermore, we have

$$\sum_{i \in B}\Delta\alpha_i \leq \frac{1}{\delta}\sum_{i \in B}(M_i(f) - m_i(f))\Delta\alpha_i < \delta.$$

It thus follows that

$$
\begin{aligned}
U_{g,\alpha}(P) - L_{g,\alpha}(P) &= \sum_{i\in A}(M_i(g) - m_i(g))\Delta\alpha_i + \sum_{i\in B}(M_i(g) - m_i(g))\Delta\alpha_i \\
&\leq \varepsilon(\alpha(b) - \alpha(a)) + 2K\delta \\
&\leq (\alpha(b) - \alpha(a) + 2K)\varepsilon.
\end{aligned}
$$

Since ε was arbitrary, it follows that $g \in \mathcal{R}_\alpha[a,b]$. □

Corollary 3.13

Let $\alpha\colon [a,b] \to \mathbb{R}$ be of bounded variation, and let $f, g \in \mathcal{R}_\alpha[a,b]$. Then the functions $|f|$ and fg are also α-integrable. If α is increasing we also have

$$
\left| \int_a^b f \,\mathrm{d}\alpha \right| \leq \int_a^b |f| \,\mathrm{d}\alpha. \tag{3.4}
$$

Proof. Integrability of $|f|$ follows since $x \mapsto |x|$ is continuous. The inequality (3.4) follows since $f \leq |f|$, and since the α-integral is increasing when α is.

For the product fg, notice that

$$
2fg = (f+g)^2 - f^2 + g^2,
$$

and that the function $x \mapsto x^2$ is continuous. □

Proposition 3.14

Let $f, \alpha\colon [a,b] \to \mathbb{R}$ be functions with f continuous and α of bounded variation. Then f is α-integrable.

Proof. We may assume that α is increasing. Let $\varepsilon > 0$. Uniform continuity of f furnishes a $\delta < 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$ for $x, y \in [a,b]$. Let $P = \{x_0, \ldots, x_n\}$ be a partition with $\|P\| < \delta$. Then $M_i(f) - m_i(f) < \varepsilon$, implying that

$$
U_{f,\alpha}(P) - L_{f,\alpha}(P) = \sum_{i=1}^n (M_i(f) - m_i(f))\Delta\alpha_i \leq \varepsilon(\alpha(b) - \alpha(a)),
$$

and since ε was arbitrary, it follows from Riemann's condition that $f \in \mathcal{R}_\alpha[a,b]$. □

3.5. The fundamental theorems of calculus

THEOREM 3.15: *The first fundamental theorem of calculus*

Let $\alpha\colon [a,b] \to \mathbb{R}$ be of bounded variation, and let $f \in \mathcal{R}_\alpha[a,b]$. Define a function $F\colon [a,b] \to \mathbb{R}$ by

$$F(x) = \int_a^x f \,\mathrm{d}\alpha.$$

Then the following hold:

(i) *F is of bounded variation.*

(ii) *Every point of continuity of α is also a point of continuity of F.*

(iii) *Assume that α is increasing. If f is continuous and α differentiable at $x \in (a,b)$, then F is differentiable at x with $F'(x) = f(x)\alpha'(x)$.*

[We need that f is also integrable on the subintervals!]

PROOF. We may assume that α is increasing. Let $x, y \in [a,b]$ with $x \neq y$, and let I denote the closed interval between x and y. If $m = \inf_{t\in I} f(t)$ and $M = \sup_{t\in I} f(t)$, we claim that there exists a $c \in \mathbb{R}$ with $m \leq c \leq M$ such that

$$F(y) - F(x) = \int_x^y f \,\mathrm{d}\alpha = c(\alpha(y) - \alpha(x)). \tag{3.5}$$

If $\alpha(x) = \alpha(y)$ this is trivial, so assume otherwise. If $x < y$ we clearly have

$$m(\alpha(y) - \alpha(x)) \leq \int_x^y f \,\mathrm{d}\alpha \leq M(\alpha(y) - \alpha(x)),$$

and dividing by $\alpha(y) - \alpha(x)$ proves the above claim for $x < y$. If $y < x$, then exchanging x and y yields a sign change which cancels on each side of (3.5). From this (i) and (ii) follow immediately.

Now assume that f is continuous at $x \in (a,b)$ and that α is differentiable at x. We claim that $c \to f(x)$ as $y \to x$. Let $\varepsilon > 0$. By continuity of f at x there is a $\delta > 0$ such that $|x - x'| < \delta$ implies $|f(x) - f(x')| < \varepsilon$. Thus if $|x - y| < \delta$ we must have $f(x) - m \leq \varepsilon$ and $M - f(x) \leq \varepsilon$. Since $m \leq c \leq M$ it follows that $|c - f(x)| \leq \varepsilon$. Dividing by $y - x$ and letting $y \to x$ in (3.5) proves (iii). □

REMARK 3.16. In the case $\alpha(x) = x$, (iii) has an easier proof: Simply note that

$$\frac{F(y) - F(x)}{y - x} - f(x) = \frac{1}{y - x} \int_x^y (f(t) - f(x)) \,\mathrm{d}t,$$

and notice that the integrand can be made less than any $\varepsilon > 0$ if $|t - x| < \delta$ for an appropriate $\delta > 0$. I am not sure that this proof can be generalised. ⌟

Theorem 3.17: *The second fundamental theorem of calculus*

Let $f \in \mathcal{R}[a,b]$. If there exists a continuous function $F\colon [a,b] \to \mathbb{R}$ that is differentiable on (a,b) with $F' = f$, then

$$\int_a^b f = F(b) - F(a).$$

Proof. Let $P = \{x_0, \dots, x_n\}$ be a partition of $[a,b]$. The mean value theorem furnishes points $t_i \in (x_{i-1}, x_i)$ such that $\Delta F_i = F'(t_i)\Delta x_i = f(t_i)\Delta x_i$. It follows that

$$\left| F(b) - F(a) - \int_a^b f \right| = \left| \sum_{i=1}^n f(t_i)\Delta x_i - \int_a^b f \right| < \varepsilon$$

if P is fine enough. Since ε was arbitrary, this proves the theorem. □

3.6. Limit and continuity theorems

Proposition 3.18

Let $f\colon [a,b] \times [c,d] \to \mathbb{R}$ be continuous, and let $\alpha\colon [a,b] \to \mathbb{R}$ be of bounded variation. Then the function $F\colon [c,d] \to \mathbb{R}$ given by

$$F(y) = \int_a^b f(x,y)\,\mathrm{d}\alpha(x)$$

is continuous.

Proof. We may assume that α is increasing. By uniform continuity of f, given $\varepsilon > 0$ there is a $\delta > 0$ such that $\|z - z'\| < \delta$ implies $|f(z) - f(z')| < \varepsilon$ for $z, z' \in [a,b] \times [c,d]$. Given $y, y' \in [c,d]$ with $|y - y'| < \delta$ we thus have

$$|F(y) - F(y')| \leq \int_a^b |f(x,y) - f(x,y')|\,\mathrm{d}\alpha(x) \leq \varepsilon(\alpha(b) - \alpha(a)).$$

Since ε was arbitrary, this shows that F is continuous. □

Proposition 3.19

Let $f\colon [a,b] \times [c,d] \to \mathbb{R}$ be bounded, and let $\alpha\colon [a,b] \to \mathbb{R}$ be of bounded variation. Assume that $f(\cdot, y) \in \mathcal{R}_\alpha[a,b]$ for all $y \in [c,d]$, that $f(x,\cdot)$ is continuous on $[c,d]$ and differentiable on (c,d) for all $x \in [a,b]$, and that $D_2 f$ is continuous on $[a,b] \times (c,d)$. Then the function $F\colon [c,d] \to \mathbb{R}$ given by

$$F(y) = \int_a^b f(x,y)\,\mathrm{d}\alpha(x)$$

is differentiable on (c,d) *and*

$$F'(y) = \int_a^b D_2 f(x,y)\,\mathrm{d}\alpha(x).$$

PROOF. We may assume that α is increasing. Let $y, y_0 \in (c,d)$ with $y \neq y_0$. By the mean value theorem we have

$$\frac{F(y) - F(y_0)}{y - y_0} = \int_a^b \frac{f(x,y) - f(x,y_0)}{y - y_0}\,\mathrm{d}\alpha(x) = \int_a^b D_2 f(x,y_x)\,\mathrm{d}\alpha(x)$$

for some $y_x \in (c,d)$ lying between y and y_0, depending on x. Let $I \subseteq (c,d)$ be a non-trivial compact interval containing y. Then $D_2 f$ is uniformly continuous on $[a,b] \times I$, so given $\varepsilon > 0$ there is a $\delta > 0$ such that $\|z - z'\| < \delta$ implies $|D_2 f(z) - D_2 f(z')| < \varepsilon$ for $z, z' \in [a,b] \times I$. For $y, y_0 \in I$ with $|y - y_0| < \delta$ we also have $|y_x - y_0| < \delta$ for all $x \in [a,b]$, and so

$$\begin{aligned}
\left| \int_a^b D_2 f(x,y_x)\,\mathrm{d}\alpha(x) - \int_a^b D_2 f(x,y_0)\,\mathrm{d}\alpha(x) \right| &\leq \int_a^b |D_2 f(x,y_x) - D_2 f(x,y_0)|\,\mathrm{d}\alpha(x) \\
&\leq \varepsilon(\alpha(b) - \alpha(a)).
\end{aligned}$$

Since ε was arbitrary, this shows that F is differentiable at y_0 with derivative as claimed. □

PROPOSITION 3.20

Let α *be of bounded variation on* $[a,b]$*, and let* $(f_n)_{n\in\mathbb{N}}$ *be a sequence of* α*-integrable functions on* $[a,b]$ *that converge uniformly to a function* f*. Then* f *is also* α*-integrable on* $[a,b]$*, and*

$$\int_a^b f\,\mathrm{d}\alpha = \lim_{n\to\infty} \int_a^b f_n\,\mathrm{d}\alpha.$$

In particular, $\mathcal{R}_\alpha[a,b]$ *is a closed subspace of* $C[a,b]$ *equipped with the uniform norm.*

PROOF. We may assume that α is increasing. Let $\varepsilon_n = \|f_n - f\|_\infty$ such that

$$f_n - \varepsilon_n \leq f \leq f_n + \varepsilon_n$$

for $n \in \mathbb{N}$. It follows that

$$\int_a^b (f_n - \varepsilon_n)\,\mathrm{d}\alpha \leq \underline{\int_a^b} f\,\mathrm{d}\alpha \leq \overline{\int_a^b} f\,\mathrm{d}\alpha \leq \int_a^b (f_n + \varepsilon_n)\,\mathrm{d}\alpha,$$

and hence,

$$0 \leq \overline{\int_a^b} f \,\mathrm{d}\alpha - \underline{\int_a^b} f \,\mathrm{d}\alpha \leq 2\varepsilon_n(\alpha(b) - \alpha(a)).$$

Thus the upper and lower integrals of f are equal, so f is α-integrable. Finally we have

$$\left| \int_a^b f_n \,\mathrm{d}\alpha - \int_a^b f \,\mathrm{d}\alpha \right| \leq \int_a^b |f_n - f| \,\mathrm{d}\alpha \leq \varepsilon_n(\alpha(b) - \alpha(a)),$$

proving the claim. □

3.7. Line integrals

Recall that a *path* in a topological space X is a continuous map $\gamma\colon [a,b] \to X$. A subset $\Gamma \subseteq X$ is called a *curve* in X if there is a path α in X whose image is Γ. The image of a path γ is called its *trace* and is denoted γ^*.

Definition 3.21: *Equivalence of paths*

Let $\alpha\colon [a,b] \to X$ and $\beta\colon [c,d] \to X$ be paths in a topological space X. If there is an increasing homeomorphism $\varphi\colon [c,d] \to [a,b]$ such that $\beta = \alpha \circ \varphi$, then α and β are said to be *properly equivalent*.

If α and β are closed paths with $\alpha(a) \neq \beta(c)$, then we also say that they are properly equivalent if there is a point $e \in (c,d)$ such that α and γ are properly equivalent in the above sense, where $\gamma\colon [e, d-c+e] \to X$ is given by

$$\gamma(t) = \begin{cases} \beta(t), & t \in [e,d], \\ \beta(t-d+c), & t \in [d, d-c+e]. \end{cases}$$

If the map φ above is decreasing, then we say that α and β are *improperly equivalent*. The paths α and β are *equivalent* if they are either properly or improperly equivalent.

Note that the condition that φ be an increasing (decreasing) homeomorphism is equivalent to it being continuous, strictly increasing (decreasing) and surjective. Also note that equivalent paths trace out the same curve in X.

Definition 3.22: *Line integrals*

Let $\gamma\colon [a,b] \to \mathbb{R}^d$ be a path, and let $f\colon \gamma^* \to \mathbb{R}^d$ be a vector field. Given a tagged partition (P,T) of $[a,b]$ then, with notation as above, we form the sums

$$S_{f,\gamma}(P,T) = \sum_{i=1}^n f(\gamma(t_i)) \cdot (\gamma(x_i) - \gamma(x_{i-1})).$$

Define the *line integral* of f with respect to γ as the limit of the net S_f, if the limit exists. We denote this integral by $\int f \cdot \mathrm{d}\gamma$.

Notice that if α and β are properly equivalent paths, then

$$\int f \cdot \mathrm{d}\alpha = \int f \cdot \mathrm{d}\beta.$$

If α and β are instead improperly equivalent, then the two integrals are equal but with opposite signs.

PROPOSITION 3.23

Let $\gamma\colon [a,b] \to \mathbb{R}^d$ be a path, and let $f\colon \gamma^* \to \mathbb{R}^d$ be a bounded function. Then

$$\int f \cdot \mathrm{d}\gamma = \sum_{k=1}^{d} \int_a^b f_k \circ \gamma \,\mathrm{d}\gamma_k$$

whenever each Riemann–Stieltjes integral on the right exists. If in addition γ is piecewise C^1, then

$$\int f \cdot \mathrm{d}\gamma = \int_a^b f(\gamma(t)) \cdot \gamma'(t) \,\mathrm{d}t.$$

PROOF. Notice that

$$S_{f,\gamma}(P,T) = \sum_{k=1}^{d} \sum_{i=1}^{n} f_k(\gamma(t_i))(\gamma_k(t_i) - \gamma_k(t_{i-1})).$$

Since the inner sums on the right-hand side approximate the Riemann–Stieltjes integrals $\int_a^b f_k \circ \gamma \,\mathrm{d}\gamma_k$, the first claim follows by taking limits. The second claim follows by [reference]. □

THEOREM 3.24: *Integral of a gradient*

Let $U \subseteq \mathbb{R}^d$ be open, and let $\varphi \in C^1(U)$. For every pair of points $x, y \in U$ and every piecewise C^1 path $\gamma\colon [a,b] \to U$ with $\gamma(a) = x$ and $\gamma(b) = y$ we have

$$\int \nabla\varphi \cdot \mathrm{d}\gamma = \varphi(y) - \varphi(x).$$

If $f = \nabla\varphi$, then φ is called a *potential function* for f.

PROOF. Let $a = t_0 < \cdots < t_n = b$ be a partition of $[a,b]$ such that γ' is continuous on each subinterval. By the chain rule,

$$(\varphi \circ \gamma)'(t) = \nabla\varphi(\gamma(t)) \cdot \gamma'(t)$$

on each open subinterval (t_{i-1}, t_i). By [reference],

$$\begin{aligned}\int \nabla\varphi \cdot \mathrm{d}\gamma &= \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \nabla\varphi(\gamma(t)) \cdot \gamma'(t)\,\mathrm{d}t = \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} (\varphi \circ \gamma)'(t)\,\mathrm{d}t \\ &= \varphi(\gamma(b)) - \varphi(\gamma(a)) = \varphi(y) - \varphi(x),\end{aligned}$$

as desired. □

Theorem 3.25

Let $U \subseteq \mathbb{R}^d$ be an open, connected set, and let $f\colon U \to \mathbb{R}^d$ be a continuous function. Fix a point $x_0 \in U$. For each $x \in U$ and each pair of polygonal paths $\alpha, \beta\colon [a,b] \to U$ joining x_0 and x, assume that

$$\int f \cdot \mathrm{d}\alpha = \int f \cdot \mathrm{d}\beta.$$

Then there exists a function $\varphi \in C^1(U)$ such that $f = \nabla\varphi$.

Notice that since U is connected, such polygonal paths exist between any pair of points.

Proof. Let $x \in U$, and let $\alpha\colon [a,b] \to U$ be a polygonal curve joining x_0 and x. Define

$$\varphi(x) = \int f \cdot \mathrm{d}\alpha.$$

By hypothesis, the number $\varphi(x)$ does not depend on the particular choice of α. We show that each partial derivative $D_k\varphi(x)$ exists and equals $f_k(x)$.

Let $B(x,\delta) \subseteq U$ for some $\delta > 0$, and let $\lambda \in [-\delta/2, \delta/2]$. Define a path $\gamma\colon [0,1] \to B(x,\delta)$ by $\gamma(t) = (1-t)x + t(x + \lambda e_k)$, where e_k is the kth standard basis vector. Then

$$\varphi(x + \lambda e_k) - \varphi(x) = \int f \cdot \mathrm{d}\gamma.$$

Furthermore, $\gamma_k'(t) = \lambda$ and $\gamma_i'(t) = 0$ for $i \neq k$. Thus γ is C^1, and so

$$\begin{aligned}\varphi(x + \lambda e_k) - \varphi(x) &= \sum_{i=1}^{d} \int_0^1 f_i(\gamma(t))\gamma_i'(t)\,\mathrm{d}t \\ &= \lambda \int_0^1 f_k(\gamma(t))\,\mathrm{d}t = \lambda \int_0^1 g(t,\lambda)\,\mathrm{d}t,\end{aligned}$$

where $g(t,\lambda) = f_k((1-t)x + t(x+\lambda e_k))$. Since g is continuous on $[0,1]\times[-\delta/2, \delta/2]$, Proposition 3.18 implies that

$$\lim_{\lambda\to 0} \int_0^1 g(t,\lambda)\,\mathrm{d}t = \int_0^1 g(t,0)\,\mathrm{d}t = \int_0^1 f_k(x)\,\mathrm{d}t = f_k(x),$$

proving that $D_k\varphi(x) = f_k(x)$. Thus $\nabla\varphi(x) = f(x)$ for all $x \in U$, and $\varphi \in C^1(U)$ since f is continuous. ☐

4 • Convergence

It is well-known that a Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space S converges to some $x \in S$ if and only if (x_n) has a *subsequence* that converges to x.

In this section we highlight a similar feature of convergence in measure and convergence in mean: If $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in measure, and if there is a subsequence that converges *pointwise a.e.* to some function f, then (f_n) also converges to f in measure. Similarly for convergence in mean.

Furthermore, Markov's inequality implies that convergence in mean is stronger than convergence in measure. In particular, a sequence that is Cauchy in mean is also Cauchy in measure. Hence when we show that convergence in measure and in mean are complete, it suffices to show that being Cauchy in measure implies the existence of a pointwise a.e. convergent subsequence.

Definition 4.1: *Convergence in measure*

Let $(X,\mathcal{E},\mu)$ be a measure space. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}(\mathcal{E})$, and let further $f \in \mathcal{M}(\mathcal{E})$. We say that (f_n) *converges to f in μ-measure* if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu\Big(\{|f_n - f| > \varepsilon\}\Big) = 0.$$

Furthermore, (f_n) is called a *Cauchy sequence in μ-measure* if, for every $\varepsilon > 0$,

$$\lim_{m,n\to\infty} \mu\Big(\{|f_m - f_n| > \varepsilon\}\Big) = 0.$$

We prove that convergence in μ-measure is complete.

Lemma 4.2

Let $(X,\mathcal{E},\mu)$ be a measure space, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}(\mathcal{E})$. If there exists a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ of strictly positive numbers such that

$$\sum_{n=1}^{\infty} \varepsilon_n < \infty \quad \textit{and} \quad \sum_{n=1}^{\infty} \mu\Big(\{|f_{n+1} - f_n| > \varepsilon_n\}\Big) < \infty, \tag{4.1}$$

then there exists a function $f \in \mathcal{M}(\mathcal{E})$ such that (f_n) converges to f both μ-a.e. and in μ-measure.

Proof. For $n \in \mathbb{N}$, denote the set $\{|f_{n+1} - f_n| > \varepsilon_n\}$ by E_n, and define sets $F_k = \bigcup_{n=k}^{\infty} E_n$ and $F = \bigcap_{k\in\mathbb{N}} F_k$. Notice that $F = \limsup_{n\to\infty} E_n$, so the first Borel–Cantelli lemma implies that $\mu(F) = 0$.

For $m \geq n$ we find that

$$|f_m - f_n| \leq \sum_{i=n}^{m-1} |f_{i+1} - f_i| \leq \sum_{i=n}^{\infty} |f_{i+1} - f_i|. \tag{4.2}$$

If furthermore $x \in F_k^c$ and $m \geq n \geq k$, then

$$|f_m(x) - f_n(x)| \leq \sum_{i=n}^{\infty} \varepsilon_i.$$

The right-hand side converges to zero as $n \to \infty$, which shows that $(f_n(x))$ is a Cauchy sequence in $\mathbb{R}$ for $x \in F_k^c$, hence for $x \in F^c$. Letting $f = \lim_{n\to\infty} f_n \mathbf{1}_{F^c}$ we thus find that (f_n) converges to $f \in \mathcal{M}(\mathcal{E})$ μ-a.e.

Next we show that $f_n \to f$ in μ-measure as $n \to \infty$. Letting $m \to \infty$ in (4.2) we find that

$$|f_n - f| \leq \sum_{i=n}^{\infty} |f_{i+1} - f_i|$$

μ-a.e. Now let $\varepsilon > 0$, and choose an $N \in \mathbb{N}$ such that $\sum_{i=N}^{\infty} \varepsilon_i \leq \varepsilon$. For $n \geq N$, $|f_n - f| > \varepsilon$ then implies that

$$\sum_{i=n}^{\infty} \varepsilon_i \leq \varepsilon < |f_n - f| \leq \sum_{i=n}^{\infty} |f_{i+1} - f_i|$$

μ-a.e., which in turn implies that $\varepsilon_i < |f_{i+1} - f_i|$ μ-a.e. for some $i \geq n$. Hence it follows that

$$\mu\Bigl(\{|f_n - f| > \varepsilon\}\Bigr) \leq \mu\Bigl(\bigcup_{i=n}^{\infty} E_i\Bigr) \leq \sum_{i=n}^{\infty} \mu(E_i),$$

which converges to zero by (4.1). ◻

Theorem 4.3: *Completeness of convergence in measure*

Let $(X, \mathcal{E}, \mu)$ be a measure space, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}(\mathcal{E})$ that is Cauchy in μ-measure. Then there exists a function $f \in \mathcal{M}(\mathcal{E})$ such that $f_n \to f$ in μ-measure. Furthermore, (f_n) has a subsequence that converges to f μ-a.e.

Proof. We prove the following lemma:

> *Let $(X, \mathcal{E}, \mu)$ be a measure space, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}(\mathcal{E})$ that is Cauchy in μ-measure. If (f_n) has a subsequence that converges μ-a.e. to function $f \in \mathcal{M}(\mathcal{E})$, then (f_n) also converges to f in μ-measure.*

Let (f_{n_k}) be such a subsequence. For $\varepsilon > 0$ we then have

$$\{|f_n - f| > \varepsilon\} \subseteq \{|f_n - f_{n_k}| > \varepsilon/2\} \cup \{|f_{n_k} - f| > \varepsilon/2\},$$

and the measures of the sets on the right-hand side go to zero as $n \to \infty$ (since $n_k \geq n$). This proves the lemma.

To prove the theorem, choose a subsequence $(g_k) = (f_{n_k})$ such that

$$\mu\Big(\{|g_{k+1} - g_k| > 2^{-k}\}\Big) \leq 2^{-k}.$$

Lemma 4.2 then implies the existence of a function $f \in \mathcal{M}(\mathcal{E})$ such that $g_k \to f$ both μ-a.e. and in μ-measure. The claim then follows from the above lemma. □

DEFINITION 4.4: *Convergence in mean*

Let $(X, \mathcal{E}, \mu)$ be a measure space. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}(\mathcal{E})$, and let further $f \in \mathcal{M}(\mathcal{E})$. If $p \in (0,\infty)$ we say that (f_n) *converges to f in the μ-p-th mean* if

$$\lim_{n\to\infty} \int_X |f_n - f|^p \,\mathrm{d}\mu = 0.$$

Furthermore, (f_n) is called a *Cauchy sequence in the μ-p-th mean* if

$$\lim_{m,n\to\infty} \int_X |f_m - f_n|^p \,\mathrm{d}\mu = 0.$$

REMARK 4.5. If (f_n) converges to f in the μ-p-th mean, then $f_n - f \in \mathcal{L}^p(\mu)$ for $n \geq N$ for some $N \in \mathbb{N}$. Furthermore, if $f \in \mathcal{L}^p(\mu)$, then $f_n = (f_n - f) + f \in \mathcal{L}^p(\mu)$ for $n \geq N$.

On the other hand, if $f_n \in \mathcal{L}^p(\mu)$ for large enough n, then $f = (f - f_n) + f_n \in \mathcal{L}^p(\mu)$. In particular, $\mathcal{L}^p(\mu)$ is a closed subspace of $\mathcal{M}(\mathcal{E})$. ⌟

THEOREM 4.6: *Completeness of convergence in mean*

Let $(X, \mathcal{E}, \mu)$ be a measure space, let $p \in (0,\infty)$, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}(\mathcal{E})$ that is Cauchy in the μ-p-th mean. Then there exists a function $f \in \mathcal{M}(\mathcal{E})$ such that $f_n \to f$ in the μ-p-th mean. Furthermore, (f_n) has a subsequence that converges to f μ-a.e.

In particular, $\mathcal{L}^p(\mu)$ is complete.

PROOF. We prove the following lemma:

> *Let $(X, \mathcal{E}, \mu)$ be a measure space, let $p \in (0,1)$, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}(\mathcal{E})$ that is Cauchy in the μ-p-th mean. If (f_n) has a subsequence that converges μ-a.e. to function $f \in \mathcal{M}(\mathcal{E})$, then (f_n) also converges to f in the μ-p-th mean.*

Let $(f_{n_k})_{k\in\mathbb{N}}$ be a subsequence that converges μ-a.e. to f. For $n \in \mathbb{N}$, Fatou's lemma implies that

$$\begin{aligned}\int_X |f_n - f|^p \,\mathrm{d}\mu &= \int_X \liminf_{k\to\infty} |f_n - f_{n_k}|^p \,\mathrm{d}\mu \leq \liminf_{k\to\infty} \int_X |f_n - f_{n_k}|^p \,\mathrm{d}\mu \\ &\leq \sup_{m\geq n} \int_X |f_n - f_m|^p \,\mathrm{d}\mu,\end{aligned}$$

which converges to zero as $n \to \infty$ as desired.

We now prove the theorem. Markov's inequality implies that (f_n) is also a Cauchy sequence in μ-measure, so [reference] yields a function $f \in \mathcal{M}(\mathcal{E})$ such that $f_n \to f$ in μ-measure, and such that (f_n) has a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ that converges to f μ-a.e. The lemma then implies that (f_n) converges to f in the μ-p-th mean.

For the last claim, if $f_n \in \mathcal{L}^p(\mu)$ for all $n \in \mathbb{N}$, then since $\mathcal{L}^p(\mu)$ is closed we also have $f \in \mathcal{L}^p(\mu)$. □

5 • Portmanteau theorems

Recall that a map $f : (S,\rho) \to (T,\delta)$ between metric spaces is called *Lipschitz* if there is a real number $K > 0$ such that

$$\delta(f(x), f(y)) \leq K\rho(x,y)$$

for all $x,y \in S$. The space of Lipschitz functions $(S,\rho) \to \mathbb{R}$ is denoted $\mathrm{Lip}(S,\rho)$, and the subspace of bounded functions $\mathrm{Lip}_b(S,\rho)$. We also denote the space of continuous functions that are uniformly bounded with respect to ρ by $C_u(S,\rho)$.

If P is a Borel probability measure on a topological space X, then a subset $A \subseteq X$ is called a *P-continuity set* if $P(\partial A) = 0$.

If $(X,\mathcal{E},\mu)$ is a measure space, then μ induces a linear functional φ_μ on $\mathcal{L}^1(\mu)$ by

$$\varphi_\mu(f) = \int_X f \,\mathrm{d}\mu.$$

If μ is finite and X is a topological space, then μ in particular induces a functional φ_μ on $C_b(X)$. We also use the notation $\mu(f)$ or just μf for $\varphi_\mu(f)$.

THEOREM 5.1

Let $(P_n)_{n\in\mathbb{N}}$ and P be probability measures on a metrisable space S. Then the following conditions are equivalent:

(i) $P_n \Rightarrow P$.

(ii) *For all metrics ρ on S, $P_n f \to Pf$ as $n \to \infty$ for all $f \in C_u(S,\rho)$.*

(iii) *There exists a metric ρ on S such that $P_n f \to Pf$ as $n \to \infty$ for all $f \in C_u(S,\rho)$.*

(iv) *For all metrics ρ on S, $P_n f \to Pf$ as $n \to \infty$ for all $f \in \mathrm{Lip}_b(S,\rho)$.*

(v) *There is a metric ρ on S such that $P_n f \to Pf$ as $n \to \infty$ for all $f \in \mathrm{Lip}_b(S,\rho)$.*

(vi) $\limsup_{n\to\infty} P_n(F) \leq P(F)$ *for all closed $F \subseteq S$.*

(vii) $\liminf_{n\to\infty} P_n(G) \geq P(G)$ *for all open $G \subseteq S$.*

(viii) $P_n(A) \to P(A)$ *for all P-continuity sets $A \subseteq S$.*

Proof. We prove the following implications:

$$
\begin{tikzcd}
 & \textit{(i)} \arrow[d, Rightarrow] & \\
 & \textit{(ii)} \arrow[dl, Rightarrow] \arrow[dr, Rightarrow] & \\
\textit{(iii)} \arrow[dr, Rightarrow] & & \textit{(iv)} \arrow[dl, Rightarrow] \\
 & \textit{(v)} \arrow[d, Rightarrow] & \\
 & \textit{(vi)} \arrow[d, Leftrightarrow] & \\
 & \textit{(vii)} \arrow[d, Rightarrow] & \\
 & \textit{(viii)} \arrow[uuuuuu, Rightarrow, bend right] &
\end{tikzcd}
$$

The implications marked in red are trivial. Notice that the two existential claims follow from the the two universal claims since S is assumed metrisable.

(v) $\Rightarrow$ (vi): Let ρ be a metric on S in accordance with the hypothesis, and let $g_k(x) = (1 - k\rho(x,F)) \vee 0$ and $F_k = \{x \in S \mid \rho(x,F) < 1/k\}$ for $k \in \mathbb{N}$. First notice that g_k is Lipschitz, since

$$
\begin{aligned}
|g_k(x) - g_k(y)| &= \left|(1 - k\rho(x,F)) \vee 0 - (1 - k\rho(y,F)) \vee 0\right| \\
&\leq |k\rho(x,F) - k\rho(y,F)| \leq k\rho(x,y).
\end{aligned}
$$

Furthermore, $\mathbf{1}_F \le g_k \le \mathbf{1}_{F_k}$ so

$$\limsup_{n\to\infty} P_n(F) \le \limsup_{n\to\infty} \int_S g_k \,\mathrm{d}P_n = \int_S g_k \,\mathrm{d}P \le P(F_k).$$

Since $F_k \downarrow F$ as $k \to \infty$, continuity of P implies the claim.

(vi) ⇔ (vii): This follows easily by taking complements.

(vi) & (vii) ⇒ (viii): For $A \subseteq S$ we have

$$\begin{aligned} P(A^\circ) &\le \liminf_{n\to\infty} P_n(A^\circ) \le \liminf_{n\to\infty} P_n(A) \\ &\le \limsup_{n\to\infty} P_n(A) \le \limsup_{n\to\infty} P_n(\overline{A}) \le P(\overline{A}). \end{aligned}$$

If A is a P-continuity set then $P(A^\circ) = P(\overline{A})$, which implies (viii).

(viii) ⇒ (i): Given $f \in C_b(S)$, by linearity we may assume that $0 \le f \le 1$. Then

$$\int_S f \,\mathrm{d}P = \int_0^\infty P(f \ge t)\,\mathrm{d}t = \int_0^1 P(f \ge t)\,\mathrm{d}t,$$

and similarly for P_n. Since f is continuous, we have $\partial\{f \ge t\} \subseteq \{f = t\}$. Now notice that $\{f = t\}$ is a P-null set except for countably many $t \in (0,1)$, since P is a finite measure.[1] Hence $\{f \ge t\}$ is a P-continuity set for all but countably many t. It then follows from (viii) and the dominated convergence theorem that

$$\int_S f \,\mathrm{d}P_n = \int_0^1 P_n(f \ge t)\,\mathrm{d}t \xrightarrow[n\to\infty]{} \int_0^1 P(f \ge t)\,\mathrm{d}t = \int_S f \,\mathrm{d}P$$

as claimed. □

6 • Dynkin systems and monotone classes

DEFINITION 6.1: *Dynkin systems, π-systems*

Let X be a set. A collection $\mathcal{D}$ of subsets of X is a *Dynkin system* in X if it has the following properties:

(i) $X \in \mathcal{D}$,

(ii) $B \setminus A \in \mathcal{D}$ whenever $A, B \in \mathcal{D}$ and $A \subseteq B$, and

(iii) $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{D}$ for any increasing sequence $(A_n)_{n\in\mathbb{N}}$ of sets in $\mathcal{D}$.

[1] Indeed, f is a random variable whose discrete support is precisely this set of ts. But the discrete support of any random variable is countable.

Furthermore, a collection $\mathcal{S}$ of subsets of X is called a *π-system* if it is closed under intersections.

It is easy to show that if $\mathcal{D}$ is both a Dynkin system in X and a π-system, then it is in fact a σ-algebra in X.

DEFINITION 6.2: *Monotone classes, set algebras*

Let X be a set. A collection $\mathcal{M}$ of subsets of X is a *monotone class* if it is closed under countable increasing unions and countable decreasing intersections.

Furthermore, a collection $\mathcal{A}$ of subsets of X is called a *set algebra* in X if it is closed under finite unions and complements.

We note that a set algebra $\mathcal{A}$ in X is automatically closed under finite intersections, and that it also contains both $\emptyset$ and X. It is easy to show that if $\mathcal{M}$ is both a monotone class and a set algebra in X, then it is in fact a σ-algebra in X.

Notice that we have two pairs of properties that ensure that a collection of sets is a σ-algebra. On the one hand we should think of Dynkin systems and monotone classes as being analogous, and similarly for π-systems and set algebras on the other. The latter pair of properties are algebraic, while the first pair are somehow analytic (or continuous), in that they involve infinitely many operations.

It turns out that if $\mathcal{S}$ is a π-system, then the Dynkin system $\delta(\mathcal{S})$ generated by $\mathcal{S}$ is also a π-system. Similarly, if $\mathcal{A}$ is a set algebra, then the monotone class $M(\mathcal{A})$ generated by $\mathcal{A}$ is also a set algebra. We have the following two results whose proofs are basically identical:

THEOREM 6.3: *Dynkin's lemma*

Let $\mathcal{S}$ be a π-system in a set X. Then $\delta(\mathcal{S})$ is also a π-system, and in particular

$$\delta(\mathcal{S}) = \sigma(\mathcal{S}). \tag{6.1}$$

PROOF. The identity (6.1) follows if $\delta(\mathcal{S})$ is a π-system: For then it is also a σ-algebra, and then $\sigma(\mathcal{S}) \subseteq \delta(\mathcal{S})$.

For $A \in \delta(\mathcal{S})$ define

$$\mathcal{D}_A = \{B \in \delta(\mathcal{S}) \mid A \cap B \in \delta(\mathcal{S})\}.$$

This is easily seen to be a Dynkin system. Also notice that $B \in \mathcal{D}_A$ if and only if $A \in \mathcal{D}_B$ for all $A, B \in \delta(\mathcal{S})$. Furthermore, if $A \in \mathcal{S}$ then $\mathcal{S} \subseteq \mathcal{D}_A$, and so $\delta(\mathcal{S}) \subseteq \mathcal{D}_A$. In other words,

$$B \in \mathcal{D}_A \quad \text{for } A \in \mathcal{S} \text{ and } B \in \delta(\mathcal{S}).$$

By symmetry we then have

$$A \in \mathcal{D}_B \quad \text{for } A \in \mathcal{S} \text{ and } B \in \delta(\mathcal{S}),$$

and since $\mathcal{D}_B$ is a Dynkin system it follows that $\delta(\mathcal{S}) \subseteq \mathcal{D}_B$. Hence $\delta(\mathcal{S})$ is a π-system as desired. □

Theorem 6.4: *The monotone class lemma*

Let $\mathcal{A}$ be a set algebra in a set X. Then $M(\mathcal{A})$ is also a set algebra, and in particular

$$M(\mathcal{A}) = \sigma(\mathcal{A}). \tag{6.2}$$

Proof. The identity (6.2) follows if $M(\mathcal{A})$ is a set algebra: For then it is also a σ-algebra, and then $\sigma(\mathcal{A}) \subseteq M(\mathcal{A})$.

For $A \in M(\mathcal{A})$ define

$$\mathcal{M}_A = \{B \in M(\mathcal{A}) \mid A \setminus B,\, B \setminus A, \text{ and } A \cap B \text{ are in } M(\mathcal{A})\}.$$

This is easily seen to be a monotone class. Also notice that $B \in \mathcal{M}_A$ if and only if $A \in \mathcal{M}_B$ for all $A, B \in M(\mathcal{A})$. Furthermore, if $A \in \mathcal{A}$ then $\mathcal{A} \subseteq \mathcal{M}_A$, and so $M(\mathcal{A}) \subseteq \mathcal{M}_A$. In other words,

$$B \in \mathcal{M}_A \quad \text{for } A \in \mathcal{A} \text{ and } B \in M(\mathcal{A}).$$

By symmetry we then have

$$A \in \mathcal{M}_B \quad \text{for } A \in \mathcal{A} \text{ and } B \in M(\mathcal{A}),$$

and since $\mathcal{M}_B$ is a monotone class it follows that $M(\mathcal{A}) \subseteq \mathcal{M}_B$. Hence $M(\mathcal{A})$ is a set algebra as desired. □

7 • Complex analysis

[Should it be here?]

If $V \subseteq \mathbb{C}$ is open and $f\colon V \to \mathbb{C}$ is differentiable at every point in V, then we say that f is *holomorphic* on V. The set of functions holomorphic on V is denoted $\mathcal{H}(V)$.

Theorem 7.1: *The Cauchy–Goursat Lemma*

If $f \in \mathcal{H}(V)$, then

$$\int_{\partial T} f(x)\,\mathrm{d}z = 0$$

for every triangle $T \subseteq V$.

PROOF. Notice that any triangle T can be subdivided into four smaller triangles $T^1,\dots,T^4$ whose corners are the corners and midpoints of the sides of T. We then clearly have

$$\int_{\partial T} g(z)\,\mathrm{d}z = \sum_{i=1}^{4} \int_{\partial T^i} g(z)\,\mathrm{d}z$$

for all $g \in C(T)$.

Let $T_0 \subseteq V$ be a triangle, and consider the integral

$$I = \int_{\partial T_0} f(z)\,\mathrm{d}z.$$

By the above considerations we have

$$|I| \leq 4\left|\int_{\partial T_0^i} f(z)\,\mathrm{d}z\right|$$

for at least one value of i. For this value of i let $T_1 = T_0^i$. Continuing this process yields a sequence $(T_n)_{n\in\mathbb{N}}$ of triangles such that

$$|I| \leq 4^n\left|\int_{\partial T_n} f(z)\,\mathrm{d}z\right|$$

for $n \in \mathbb{N}_0$.

Furthermore, each of the four triangles in a subdivision of a triangle T have side lengths half of those of T, so

$$\operatorname{diam} T_n = 2^{-n} \operatorname{diam} T_0$$

for $n \in \mathbb{N}_0$. Thus there exists a point $z_0 \in \mathbb{C}$ such that $\bigcap_{n\in\mathbb{N}_0} T_n = \{z_0\}$ since (T_n) is a sequence of closed sets whose diameters tend to zero. It follows that

$$\sup_{z\in\partial T_n} |z - z_0| \leq 2^{-n} \operatorname{diam} T_0.$$

Given $\varepsilon > 0$ there exists an $r > 0$ such that

$$|f(z) - f(z_0) - f'(z_0)(z - z_0)| \leq \varepsilon |z - z_0|$$

for $z \in B(z_0, r)$, and there further exists an $N \in \mathbb{N}$ such that $n \geq N$ implies $T_n \subseteq B(z_0, r)$. Now notice that the function $z \mapsto f(z_0) + f'(z_0)(z - z_0)$ has an antiderivative, so its integral along ∂T_n is zero. Denoting the length of the curve ∂T_n by L_n we have $L_n \le 2\operatorname{diam} T_n$. Hence,

$$\begin{aligned}\left|\int_{\partial T_n} f(z)\,\mathrm{d}z\right| &\le \int_{\partial T_n} |f(z) - f(z_0) - f'(z_0)(z - z_0)|\,\mathrm{d}z \\ &\le L_n \varepsilon \sup_{z \in \partial T_n} |z - z_0| \\ &\le 2^{-2n+1} \varepsilon (\operatorname{diam} T_0)^2\end{aligned}$$

for $n \ge N$, and so

$$|I| \le 2\varepsilon(\operatorname{diam} T_0)^2.$$

Since ε was arbitrary, it follows that $I = 0$ as desired. □