Part XXIV

Appendices

A

Multinomial Theorems and Calculus Results

Given a multi-index $\alpha \in \mathbb{Z}_+^n$, let $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $\alpha! := \alpha_1! \cdots \alpha_n!$,

$$x^\alpha := \prod_{j=1}^n x_j^{\alpha_j} \text{ and } \partial_x^\alpha = \left(\frac{\partial}{\partial x}\right)^\alpha := \prod_{j=1}^n \left(\frac{\partial}{\partial x_j}\right)^{\alpha_j}.$$

We also write

$$\partial_v f(x) := \frac{d}{dt} f(x+tv)|_{t=0}.$$

A.1 Multinomial Theorems and Product Rules

For $a = (a_1, a_2, \ldots, a_n) \in \mathbb{C}^n$, $m \in \mathbb{N}$ and $(i_1, \ldots, i_m) \in \{1, 2, \ldots, n\}^m$ let $\hat{\alpha}_j (i_1, \ldots, i_m) = \# \{k : i_k = j\}$. Then

$$\left(\sum_{i=1}^n a_i\right)^m = \sum_{i_1, \ldots, i_m = 1}^n a_{i_1} \ldots a_{i_m} = \sum_{|\alpha| = m} C(\alpha) a^\alpha$$

where

$$C(\alpha) = \# \{(i_1, \ldots, i_m) : \hat{\alpha}_j (i_1, \ldots, i_m) = \alpha_j \text{ for } j = 1, 2, \ldots, n\}$$

I claim that $C(\alpha) = \frac{m!}{\alpha!}$. Indeed, one possibility for such a sequence $(a_1, \ldots, a_{i_m})$ for a given α is gotten by choosing

$$(\overbrace{a_1, \ldots, a_1}^{\alpha_1}, \overbrace{a_2, \ldots, a_2}^{\alpha_2}, \ldots, \overbrace{a_n, \ldots, a_n}^{\alpha_n}).$$

Now there are $m!$ permutations of this list. However, only those permutations leading to a distinct list are to be counted. So for each of these $m!$ permutations we must divide by the number of permutation which just rearrange the groups of a_i's among themselves for each i. There are $\alpha! := \alpha_1! \cdots \alpha_n!$ such permutations. Therefore, $C(\alpha) = m!/\alpha!$ as advertised. So we have proved

$$\left(\sum_{i=1}^{n} a_i\right)^m = \sum_{|\alpha|=m} \frac{m!}{\alpha!} a^\alpha. \tag{A.1}$$

Now suppose that $a, b \in \mathbb{R}^n$ and α is a multi-index, we have

$$(a+b)^\alpha = \sum_{\beta \le \alpha} \frac{\alpha!}{\beta!(\alpha-\beta)!} a^\beta b^{\alpha-\beta} = \sum_{\beta+\delta=\alpha} \frac{\alpha!}{\beta!\delta!} a^\beta b^\delta \tag{A.2}$$

Indeed, by the standard Binomial formula,

$$(a_i+b_i)^{\alpha_i} = \sum_{\beta_i \le \alpha_i} \frac{\alpha_i!}{\beta_i!(\alpha_i-\beta_i)!} a^{\beta_i} b^{\alpha_i-\beta_i}$$

from which Eq. (A.2) follows. Eq. (A.2) generalizes in the obvious way to

$$(a_1 + \dots + a_k)^\alpha = \sum_{\beta_1+\dots+\beta_k=\alpha} \frac{\alpha!}{\beta_1! \cdots \beta_k!} a_1^{\beta_1} \dots a_k^{\beta_k} \tag{A.3}$$

where $a_1, a_2, \dots, a_k \in \mathbb{R}^n$ and $\alpha \in \mathbb{Z}_+^n$.

Now let us consider the product rule for derivatives. Let us begin with the one variable case (write $d^n f$ for $f^{(n)} = \frac{d^n}{dx^n} f$) where we will show by induction that

$$d^n(fg) = \sum_{k=0}^{n} \binom{n}{k} d^k f \cdot d^{n-k} g. \tag{A.4}$$

Indeed assuming Eq. (A.4) we find

$$\begin{aligned}
d^{n+1}(fg) &= \sum_{k=0}^{n} \binom{n}{k} d^{k+1} f \cdot d^{n-k} g + \sum_{k=0}^{n} \binom{n}{k} d^k f \cdot d^{n-k+1} g \\
&= \sum_{k=1}^{n+1} \binom{n}{k-1} d^k f \cdot d^{n-k+1} g + \sum_{k=0}^{n} \binom{n}{k} d^k f \cdot d^{n-k+1} g \\
&= \sum_{k=1}^{n+1} \left[\binom{n}{k-1} + \binom{n}{k}\right] d^k f \cdot d^{n-k+1} g + d^{n+1} f \cdot g + f \cdot d^{n+1} g.
\end{aligned}$$

Since

$$\begin{aligned}
\binom{n}{k-1} + \binom{n}{k} &= \frac{n!}{(n-k+1)!(k-1)!} + \frac{n!}{(n-k)!k!} \\
&= \frac{n!}{(k-1)!\,(n-k)!} \left[\frac{1}{(n-k+1)} + \frac{1}{k}\right] \\
&= \frac{n!}{(k-1)!\,(n-k)!} \frac{n+1}{(n-k+1)\,k} = \binom{n+1}{k}
\end{aligned}$$

the result follows.

Now consider the multi-variable case

$$\partial^\alpha (fg) = \left(\prod_{i=1}^n \partial_i^{\alpha_i}\right)(fg) = \prod_{i=1}^n \left[\sum_{k_i=0}^{\alpha_i} \binom{\alpha_i}{k_i} \partial_i^{k_i} f \cdot \partial_i^{\alpha_i - k_i} g\right]$$

$$= \sum_{k_1=0}^{\alpha_1} \cdots \sum_{k_n=0}^{\alpha_n} \prod_{i=1}^n \binom{\alpha_i}{k_i} \partial^k f \cdot \partial^{\alpha-k} g = \sum_{k \le \alpha} \binom{\alpha}{k} \partial^k f \cdot \partial^{\alpha-k} g$$

where $k = (k_1, k_2, \ldots, k_n)$ and

$$\binom{\alpha}{k} := \prod_{i=1}^n \binom{\alpha_i}{k_i} = \frac{\alpha!}{k!(\alpha-k)!}.$$

So we have proved

$$\partial^\alpha (fg) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta f \cdot \partial^{\alpha-\beta} g. \tag{A.5}$$

A.2 Taylor's Theorem

Theorem A.1. *Suppose $X \subset \mathbb{R}^n$ is an open set, $x : [0,1] \to X$ is a C^1 – path, and $f \in C^N(X, \mathbb{C})$. Let $v_s := x(1) - x(s)$ and $v = v_1 = x(1) - x(0)$, then*

$$f(x(1)) = \sum_{m=0}^{N-1} \frac{1}{m!} \left(\partial_v^m f\right)(x(0)) + R_N \tag{A.6}$$

where

$$R_N = \frac{1}{(N-1)!} \int_0^1 \left(\partial_{\dot{x}(s)} \partial_{v_s}^{N-1} f\right)(x(s)) ds = \frac{1}{N!} \int_0^1 \left(-\frac{d}{ds} \partial_{v_s}^N f\right)(x(s)) ds. \tag{A.7}$$

and $0! := 1$.

Proof. By the fundamental theorem of calculus and the chain rule,

$$f(x(t)) = f(x(0)) + \int_0^t \frac{d}{ds} f(x(s)) ds = f(x(0)) + \int_0^t \left(\partial_{\dot{x}(s)} f\right)(x(s)) ds \tag{A.8}$$

and in particular,

$$f(x(1)) = f(x(0)) + \int_0^1 \left(\partial_{\dot{x}(s)} f\right)(x(s)) ds.$$

This proves Eq. (A.6) when $N = 1$. We will now complete the proof using induction on N.

Applying Eq. (A.8) with f replaced by $\frac{1}{(N-1)!}\left(\partial_{\dot{x}(s)}\partial_{v_s}^{N-1}f\right)$ gives

$$\begin{aligned}\frac{1}{(N-1)!}\left(\partial_{\dot{x}(s)}\partial_{v_s}^{N-1}f\right)(x(s)) &= \frac{1}{(N-1)!}\left(\partial_{\dot{x}(s)}\partial_{v_s}^{N-1}f\right)(x(0)) \\ &\quad + \frac{1}{(N-1)!}\int_0^s \left(\partial_{\dot{x}(s)}\partial_{v_s}^{N-1}\partial_{\dot{x}(t)}f\right)(x(t))dt \\ &= -\frac{1}{N!}\left(\frac{d}{ds}\partial_{v_s}^{N}f\right)(x(0)) - \frac{1}{N!}\int_0^s \left(\frac{d}{ds}\partial_{v_s}^{N}\partial_{\dot{x}(t)}f\right)(x(t))dt\end{aligned}$$

wherein we have used the fact that mixed partial derivatives commute to show $\frac{d}{ds}\partial_{v_s}^N f = N\partial_{\dot{x}(s)}\partial_{v_s}^{N-1}f$. Integrating this equation on $s \in [0,1]$ shows, using the fundamental theorem of calculus,

$$\begin{aligned}R_N &= \frac{1}{N!}\left(\partial_v^N f\right)(x(0)) - \frac{1}{N!}\int_{0\le t\le s\le 1}\left(\frac{d}{ds}\partial_{v_s}^{N}\partial_{\dot{x}(t)}f\right)(x(t))dsdt \\ &= \frac{1}{N!}\left(\partial_v^N f\right)(x(0)) + \frac{1}{(N+1)!}\int_{0\le t\le 1}\left(\partial_{w_t}^{N}\partial_{\dot{x}(t)}f\right)(x(t))dt \\ &= \frac{1}{N!}\left(\partial_v^N f\right)(x(0)) + R_{N+1}\end{aligned}$$

which completes the inductive proof. ■

Remark A.2. Using Eq. (A.1) with a_i replaced by $v_i\partial_i$ (although $\{v_i\partial_i\}_{i=1}^n$ are not complex numbers they are commuting symbols), we find

$$\partial_v^m f = \left(\sum_{i=1}^n v_i\partial_i\right)^m f = \sum_{|\alpha|=m}\frac{m!}{\alpha!}v^\alpha\partial^\alpha.$$

Using this fact we may write Eqs. (A.6) and (A.7) as

$$f(x(1)) = \sum_{|\alpha|\le N-1}\frac{1}{\alpha!}v^\alpha\partial^\alpha f(x(0)) + R_N$$

and

$$R_N = \sum_{|\alpha|=N}\frac{1}{\alpha!}\int_0^1\left(-\frac{d}{ds}v_s^\alpha\partial^\alpha f\right)(x(s))ds.$$

Corollary A.3. *Suppose $X \subset \mathbb{R}^n$ is an open set which contains $x(s) = (1-s)x_0 + sx_1$ for $0 \le s \le 1$ and $f \in C^N(X,\mathbb{C})$. Then*

$$f(x_1) = \sum_{m=0}^{N-1}\frac{1}{m!}\left(\partial_v^m f\right)(x_0) + \frac{1}{N!}\int_0^1\left(\partial_v^N f\right)(x(s))d\nu_N(s) \tag{A.9}$$

$$= \sum_{|\alpha|<N}\frac{1}{\alpha!}\partial^\alpha f(x(0))(x_1-x_0)^\alpha + \sum_{\alpha:|\alpha|=N}\frac{1}{\alpha!}\left[\int_0^1\partial^\alpha f(x(s))d\nu_N(s)\right](x_1-x_0)^\alpha \tag{A.10}$$

where $v := x_1 - x_0$ and $d\nu_N$ is the probability measure on $[0,1]$ given by

$$d\nu_N(s) := N(1-s)^{N-1}ds. \tag{A.11}$$

If we let $x = x_0$ and $y = x_1 - x_0$ (so $x + y = x_1$) Eq. (A.10) may be written as

$$f(x+y) = \sum_{|\alpha|<N} \frac{\partial_x^\alpha f(x)}{\alpha!} y^\alpha + \sum_{\alpha:|\alpha|=N} \frac{1}{\alpha!} \left(\int_0^1 \partial_x^\alpha f(x+sy) d\nu_N(s) \right) y^\alpha. \tag{A.12}$$

Proof. This is a special case of Theorem A.1. Notice that

$$v_s = x(1) - x(s) = (1-s)(x_1 - x_0) = (1-s)v$$

and hence

$$R_N = \frac{1}{N!} \int_0^1 \left(-\frac{d}{ds}(1-s)^N \partial_v^N f \right)(x(s))ds = \frac{1}{N!} \int_0^1 \left(\partial_v^N f \right)(x(s)) N(1-s)^{N-1} ds.$$

■

Example A.4. Let $X = (-1,1) \subset \mathbb{R}$, $\beta \in \mathbb{R}$ and $f(x) = (1-x)^\beta$. The reader should verify

$$f^{(m)}(x) = (-1)^m \beta(\beta-1)\dots(\beta-m+1)(1-x)^{\beta-m}$$

and therefore by Taylor's theorem (Eq. (100.75) with $x = 0$ and $y = x$)

$$(1-x)^\beta = 1 + \sum_{m=1}^{N-1} \frac{1}{m!}(-1)^m \beta(\beta-1)\dots(\beta-m+1)x^m + R_N(x) \tag{A.13}$$

where

$$\begin{aligned} R_N(x) &= \frac{x^N}{N!} \int_0^1 (-1)^N \beta(\beta-1)\dots(\beta-N+1)(1-sx)^{\beta-N} d\nu_N(s) \\ &= \frac{x^N}{N!}(-1)^N \beta(\beta-1)\dots(\beta-N+1) \int_0^1 \frac{N(1-s)^{N-1}}{(1-sx)^{N-\beta}} ds. \end{aligned}$$

Now for $x \in (-1,1)$ and $N > \beta$,

$$0 \le \int_0^1 \frac{N(1-s)^{N-1}}{(1-sx)^{N-\beta}} ds \le \int_0^1 \frac{N(1-s)^{N-1}}{(1-s)^{N-\beta}} ds = \int_0^1 N(1-s)^{\beta-1} ds = \frac{N}{\beta}$$

and therefore,

$$|R_N(x)| \le \frac{|x|^N}{(N-1)!} |(\beta-1)\dots(\beta-N+1)| =: \rho_N.$$

Since

$$\limsup_{N\to\infty} \frac{\rho_{N+1}}{\rho_N} = |x| \cdot \limsup_{N\to\infty} \frac{N-\beta}{N} = |x| < 1$$

and so by the Ratio test, $|R_N(x)| \leq \rho_N \to 0$ (exponentially fast) as $N \to \infty$. Therefore by passing to the limit in Eq. (A.13) we have proved

$$(1-x)^\beta = 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \beta(\beta-1)\dots(\beta-m+1)x^m \tag{A.14}$$

which is valid for $|x| < 1$ and $\beta \in \mathbb{R}$. An important special cases is $\beta = -1$ in which case, Eq. (A.14) becomes $\frac{1}{1-x} = \sum_{m=0}^{\infty} x^m$, the standard geometric series formula. Another another useful special case is $\beta = 1/2$ in which case Eq. (A.14) becomes

$$\begin{aligned}\sqrt{1-x} &= 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \frac{1}{2}(\frac{1}{2}-1)\dots(\frac{1}{2}-m+1)x^m \\ &= 1 - \sum_{m=1}^{\infty} \frac{(2m-3)!!}{2^m m!} x^m \text{ for all } |x| < 1.\end{aligned} \tag{A.15}$$

B

Zorn's Lemma and the Hausdorff Maximal Principle

Definition B.1. *A partial order $\leq$ on X is a relation with following properties*

(i) If $x \leq y$ and $y \leq z$ then $x \leq z$.
(ii) If $x \leq y$ and $y \leq x$ then $x = y$.
(iii) $x \leq x$ for all $x \in X$.

Example B.2. Let Y be a set and $X = 2^Y$. There are two natural partial orders on X.

1. Ordered by inclusion, $A \leq B$ is $A \subset B$ and
2. Ordered by reverse inclusion, $A \leq B$ if $B \subset A$.

Definition B.3. *Let $(X, \leq)$ be a partially ordered set we say X is **linearly** a **totally** ordered if for all $x, y \in X$ either $x \leq y$ or $y \leq x$. The real numbers $\mathbb{R}$ with the usual order $\leq$ is a typical example.*

Definition B.4. *Let $(X, \leq)$ be a partial ordered set. We say $x \in X$ is a **maximal** element if for all $y \in X$ such that $y \geq x$ implies $y = x$, i.e. there is no element larger than x. An **upper bound** for a subset E of X is an element $x \in X$ such that $x \geq y$ for all $y \in E$.*

Example B.5. Let

$$X = \left\{ a = \{1\} \; b = \{1, 2\} \; c = \{3\} \; d = \{2, 4\} \; e = \{2\} \right\}$$

ordered by set inclusion. Then b and d are maximal elements despite that fact that $b \nleq a$ and $a \nleq b$. We also have

- If $E = \{a, e, c\}$, then E has **no** upper bound.

Definition B.6. • *If $E = \{a, e\}$, then b is an upper bound.*
- *$E = \{e\}$, then b and d are upper bounds.*

Theorem B.7. *The following are equivalent.*

1. ***The axiom of choice:*** *to each collection,* $\{X_\alpha\}_{\alpha\in A}$ *, of non-empty sets there exists a "choice function,"* $x : A \to \coprod_{\alpha\in A} X_\alpha$ *such that* $x(\alpha) \in X_\alpha$ *for all* $\alpha \in A$*, i.e.* $\prod_{\alpha\in A} X_\alpha \neq \emptyset$.
2. ***The Hausdorff Maximal Principle****: Every partially ordered set has a* ***maximal*** *(relative to the inclusion order) linearly ordered subset.*
3. ***Zorn's Lemma****: If* X *is partially ordered set such that every linearly ordered subset of* X *has an upper bound, then* X *has a maximal element.*[1]

Proof. $(2 \Rightarrow 3)$ Let X be a partially ordered subset as in 3 and let $\mathcal{F} = \{E \subset X : E \text{ is linearly ordered}\}$ which we equip with the inclusion partial ordering. By 2. there exist a maximal element $E \in \mathcal{F}$. By assumption, the linearly ordered set E has an upper bound $x \in X$. The element x is maximal, for if $y \in Y$ and $y \geq x$, then $E \cup \{y\}$ is still an linearly ordered set containing E. So by maximality of E, $E = E \cup \{y\}$, i.e. $y \in E$ and therefore $y \leq x$ showing which combined with $y \geq x$ implies that $y = x$.[2]

$(3 \Rightarrow 1)$ Let $\{X_\alpha\}_{\alpha\in A}$ be a collection of non-empty sets, we must show $\prod_{\alpha\in A} X_\alpha$ is not empty. Let $\mathcal{G}$ denote the collection of functions $g : D(g) \to \coprod_{\alpha\in A} X_\alpha$ such that $D(g)$ is a subset of A, and for all $\alpha \in D(g)$, $g(\alpha) \in X_\alpha$. Notice that $\mathcal{G}$ is not empty, for we may let $\alpha_0 \in A$ and $x_0 \in X_\alpha$ and then set $D(g) = \{\alpha_0\}$ and $g(\alpha_0) = x_0$ to construct an element of $\mathcal{G}$. We now put a partial order on $\mathcal{G}$ as follows. We say that $f \leq g$ for $f, g \in \mathcal{G}$ provided that $D(f) \subset D(g)$ and $f = g|_{D(f)}$. If $\Phi \subset \mathcal{G}$ is a linearly ordered set, let $D(h) = \cup_{g\in\Phi} D(g)$ and for $\alpha \in D(g)$ let $h(\alpha) = g(\alpha)$. Then $h \in \mathcal{G}$ is an upper bound for Φ. So by Zorn's Lemma there exists a maximal element $h \in \mathcal{G}$. To finish the proof we need only show that $D(h) = A$. If this were not the case, then let $\alpha_0 \in A \setminus D(h)$ and $x_0 \in X_{\alpha_0}$. We may now define $D(\tilde{h}) = D(h) \cup \{\alpha_0\}$ and

$$\tilde{h}(\alpha) = \begin{cases} h(\alpha) & \text{if } \alpha \in D(h) \\ x_0 & \text{if } \alpha = \alpha_0. \end{cases}$$

---

[1] If X is a countable set we may prove Zorn's Lemma by induction. Let $\{x_n\}_{n=1}^\infty$ be an enumeration of X, and define $E_n \subset X$ inductively as follows. For $n = 1$ let $E_1 = \{x_1\}$, and if E_n have been chosen, let $E_{n+1} = E_n \cup \{x_{n+1}\}$ if x_{n+1} is an upper bound for E_n otherwise let $E_{n+1} = E_n$. The set $E = \cup_{n=1}^\infty E_n$ is a linearly ordered (you check) subset of X and hence by assumption E has an upper bound, $x \in X$. I claim that his element is maximal, for if there exists $y = x_m \in X$ such that $y \geq x$, then x_m would be an upper bound for E_{m-1} and therefore $y = x_m \in E_m \subset E$. That is to say if $y \geq x$, then $y \in E$ and hence $y \leq x$, so $y = x$. (Hence we may view Zorn's lemma as a " jazzed" up version of induction.)

[2] Similarly one may show that $3 \Rightarrow 2$. Let $\mathcal{F} = \{E \subset X : E \text{ is linearly ordered}\}$ and order $\mathcal{F}$ by inclusion. If $\mathcal{M} \subset \mathcal{F}$ is linearly ordered, let $E = \cup\mathcal{M} = \bigcup_{A\in\mathcal{M}} A$. If $x, y \in E$ then $x \in A$ and $y \in B$ for some $A, B \subset \mathcal{M}$. Now $\mathcal{M}$ is linearly ordered by set inclusion so $A \subset B$ or $B \subset A$ i.e. $x, y \in A$ or $x, y \in B$. Since A and B are linearly order we must have either $x \leq y$ or $y \leq x$, that is to say E is linearly ordered. Hence by 3. there exists a maximal element $E \in \mathcal{F}$ which is the assertion in 2.

Then $h \leq \tilde{h}$ while $h \neq \tilde{h}$ violating the fact that h was a maximal element.

($1 \Rightarrow 2$) Let $(X, \leq)$ be a partially ordered set. Let $\mathcal{F}$ be the collection of linearly ordered subsets of X which we order by set inclusion. Given $x_0 \in X$, $\{x_0\} \in \mathcal{F}$ is linearly ordered set so that $\mathcal{F} \neq \emptyset$.

Fix an element $P_0 \in \mathcal{F}$. If P_0 is not maximal there exists $P_1 \in \mathcal{F}$ such that $P_0 \subsetneq P_1$. In particular we may choose $x \notin P_0$ such that $P_0 \cup \{x\} \in \mathcal{F}$. The idea now is to keep repeating this process of adding points $x \in X$ until we construct a maximal element P of $\mathcal{F}$. We now have to take care of some details.

We may assume with out loss of generality that $\tilde{\mathcal{F}} = \{P \in \mathcal{F} : P \text{ is not maximal}\}$ is a non-empty set. For $P \in \tilde{\mathcal{F}}$, let $P^* = \{x \in X : P \cup \{x\} \in \mathcal{F}\}$. As the above argument shows, $P^* \neq \emptyset$ for all $P \in \tilde{\mathcal{F}}$. Using the axiom of choice, there exists $f \in \prod_{P \in \tilde{\mathcal{F}}} P^*$. We now define $g : \mathcal{F} \to \mathcal{F}$ by

$$g(P) = \begin{cases} P & \text{if } \quad P \text{ is maximal} \\ P \cup \{f(x)\} & \text{if } P \text{ is not maximal.} \end{cases} \tag{B.1}$$

The proof is completed by Lemma B.8 below which shows that g must have a fixed point $P \in \mathcal{F}$. This fixed point is maximal by construction of g. ■

Lemma B.8. *The function $g : \mathcal{F} \to \mathcal{F}$ defined in Eq. (B.1) has a fixed point.*[3]

Proof. The **idea of the proof** is as follows. Let $P_0 \in \mathcal{F}$ be chosen arbitrarily. Notice that $\Phi = \left\{g^{(n)}(P_0)\right\}_{n=0}^{\infty} \subset \mathcal{F}$ is a linearly ordered set and it is therefore easily verified that $P_1 = \bigcup_{n=0}^{\infty} g^{(n)}(P_0) \in \mathcal{F}$. Similarly we may repeat the process to construct $P_2 = \bigcup_{n=0}^{\infty} g^{(n)}(P_1) \in \mathcal{F}$ and $P_3 = \bigcup_{n=0}^{\infty} g^{(n)}(P_2) \in \mathcal{F}$, etc. etc. Then take $P_\infty = \cup_{n=0}^{\infty} P_n$ and start again with P_0 replaced by P_∞. Then keep going this way until eventually the sets stop increasing in size, in which case we have found our fixed point. The problem with this strategy is that we may never win. (This is very reminiscent of constructing measurable sets and the way out is to use measure theoretic like arguments.)

Let us now start the **formal proof.** Again let $P_0 \in \mathcal{F}$ and let $\mathcal{F}_1 = \{P \in \mathcal{F} : P_0 \subset P\}$. Notice that $\mathcal{F}_1$ has the following properties:

1. $P_0 \in \mathcal{F}_1$.
2. If $\Phi \subset \mathcal{F}_1$ is a totally ordered (by set inclusion) subset then $\cup \Phi \in \mathcal{F}_1$.
3. If $P \in \mathcal{F}_1$ then $g(P) \in \mathcal{F}_1$.

Let us call a general subset $\mathcal{F}' \subset \mathcal{F}$ satisfying these three conditions a tower and let

[3] Here is an easy proof if the elements of $\mathcal{F}$ happened to all be finite sets and there existed a set $P \in \mathcal{F}$ with a maximal number of elements. In this case the condition that $P \subset g(P)$ would imply that $P = g(P)$, otherwise $g(P)$ would have more elements than P.

$$\mathcal{F}_0 = \cap \{\mathcal{F}' : \mathcal{F}' \text{ is a tower}\}.$$

Standard arguments show that $\mathcal{F}_0$ is still a tower and clearly is the smallest tower containing P_0. (Morally speaking $\mathcal{F}_0$ consists of all of the sets we were trying to constructed in the "idea section" of the proof.)

We now claim that $\mathcal{F}_0$ is a linearly ordered subset of $\mathcal{F}$. To prove this let $\Gamma \subset \mathcal{F}_0$ be the linearly ordered set

$$\Gamma = \{C \in \mathcal{F}_0 : \text{ for all } A \in \mathcal{F}_0 \text{ either } A \subset C \text{ or } C \subset A\}.$$

Shortly we will show that $\Gamma \subset \mathcal{F}_0$ is a tower and hence that $\mathcal{F}_0 = \Gamma$. That is to say $\mathcal{F}_0$ is linearly ordered. Assuming this for the moment let us finish the proof. Let $P \equiv \cup \mathcal{F}_0$ which is in $\mathcal{F}_0$ by property 2 and is clearly the largest element in $\mathcal{F}_0$. By 3. it now follows that $P \subset g(P) \in \mathcal{F}_0$ and by maximality of P, we have $g(P) = P$, the desired fixed point. So to finish the proof, we must show that Γ is a tower.

First off it is clear that $P_0 \in \Gamma$ so in particular Γ is not empty. For each $C \in \Gamma$ let

$$\Phi_C := \{A \in \mathcal{F}_0 : \text{either } A \subset C \text{ or } g(C) \subset A\}.$$

We will begin by showing that $\Phi_C \subset \mathcal{F}_0$ is a tower and therefore that $\Phi_C = \mathcal{F}_0$.

1. $P_0 \in \Phi_C$ since $P_0 \subset C$ for all $C \in \Gamma \subset \mathcal{F}_0$. 2. If $\Phi \subset \Phi_C \subset \mathcal{F}_0$ is totally ordered by set inclusion, then $A_\Phi := \cup \Phi \in \mathcal{F}_0$. We must show $A_\Phi \in \Phi_C$, that is that $A_\Phi \subset C$ or $C \subset A_\Phi$. Now if $A \subset C$ for all $A \in \Phi$, then $A_\Phi \subset C$ and hence $A_\Phi \in \Phi_C$. On the other hand if there is some $A \in \Phi$ such that $g(C) \subset A$ then clearly $g(C) \subset A_\Phi$ and again $A_\Phi \in \Phi_C$.

3. Given $A \in \Phi_C$ we must show $g(A) \in \Phi_C$, i.e. that

$$g(A) \subset C \text{ or } g(C) \subset g(A). \tag{B.2}$$

There are three cases to consider: either $A \subsetneq C$, $A = C$, or $g(C) \subset A$. In the case $A = C$, $g(C) = g(A) \subset g(A)$ and if $g(C) \subset A$ then $g(C) \subset A \subset g(A)$ and Eq. (B.2) holds in either of these cases. So assume that $A \subsetneq C$. Since $C \in \Gamma$, either $g(A) \subset C$ (in which case we are done) or $C \subset g(A)$. Hence we may assume that

$$A \subsetneq C \subset g(A).$$

Now if C were a proper subset of $g(A)$ it would then follow that $g(A) \setminus A$ would consist of at least two points which contradicts the definition of g. Hence we must have $g(A) = C \subset C$ and again Eq. (B.2) holds, so Φ_C is a tower.

It is now easy to show Γ is a tower. It is again clear that $P_0 \in \Gamma$ and Property 2. may be checked for Γ in the same way as it was done for Φ_C above. For Property 3., if $C \in \Gamma$ we may use $\Phi_C = \mathcal{F}_0$ to conclude for all $A \in \mathcal{F}_0$, either $A \subset C \subset g(C)$ or $g(C) \subset A$, i.e. $g(C) \in \Gamma$. Thus Γ is a tower and we are done. ■

C

Nets

In this section (which may be skipped) we develop the notion of nets. Nets are generalization of sequences. Here is an example which shows that for general topological spaces, sequences are not always adequate.

Example C.1. Equip $\mathbb{C}^{\mathbb{R}}$ with the topology of pointwise convergence, i.e. the product topology and consider $C(\mathbb{R},\mathbb{C}) \subset \mathbb{C}^{\mathbb{R}}$. If $\{f_n\} \subset C(\mathbb{R},\mathbb{C})$ is a sequence which converges such that $f_n \to f \in \mathbb{C}^{\mathbb{R}}$ pointwise then f is a Borel measurable function. Hence the sequential limits of elements in $C(\mathbb{R},\mathbb{C})$ is necessarily contained in the Borel measurable functions which is properly contained in $\mathbb{C}^{\mathbb{R}}$. In short the sequential closure of $C(\mathbb{R},\mathbb{C})$ is a proper subset of $\mathbb{C}^{\mathbb{R}}$. On the other hand we have $\overline{C(\mathbb{R},\mathbb{C})} = \mathbb{C}^{\mathbb{R}}$. Indeed a typical open neighborhood of $f \in \mathbb{C}^{\mathbb{R}}$ is of the form

$$N = \{g \in \mathbb{C}^{\mathbb{R}} : |g(x) - f(x)| < \epsilon \text{ for } x \in \Lambda\},$$

where $\epsilon > 0$ and Λ is a finite subset of $\mathbb{R}$. Since $N \cap C(\mathbb{R},\mathbb{C}) \neq \emptyset$ it follows that $f \in \overline{C(\mathbb{R},\mathbb{C})}$.

Definition C.2. *A **directed set** $(A, \leq)$ is a set with a relation "$\leq$" such that*

1. *$\alpha \leq \alpha$ for all $\alpha \in A$.*
2. *If $\alpha \leq \beta$ and $\beta \leq \gamma$ then $\alpha \leq \gamma$.*
3. *A is **cofinite**, i.e. $\alpha, \beta \in A$ there exists $\gamma \in A$ such that $\alpha \leq \gamma$ and $\beta \leq \gamma$.*

*A **net** is function $x : A \to X$ where A is a directed set. We will often denote a net x by $\{x_\alpha\}_{\alpha \in A}$.*

Example C.3 (Directed sets).

1. $A = 2^X$ ordered by inclusion, i.e. $\alpha \leq \beta$ if $\alpha \subset \beta$. If $\alpha \leq \beta$ and $\beta \leq \gamma$ then $\alpha \subset \beta \subset \gamma$ and hence $\alpha \leq \gamma$. Similalry if $\alpha, \beta \in 2^X$ then $\alpha, \beta \leq \alpha \cup \beta =: \gamma$.
2. $A = 2^X$ ordered by reverse inclusion, i.e. $\alpha \leq \beta$ if $\beta \subset \alpha$. If $\alpha \leq \beta$ and $\beta \leq \gamma$ then $\alpha \supseteq \beta \supseteq \gamma$ and so $\alpha \leq \gamma$ and if $\alpha, \beta \in A$ then $\alpha, \beta \leq \alpha \cap \beta$.

3. Let $A = \mathbb{N}$ equipped with the usual ordering on $\mathbb{N}$. In this case nets are simply sequences.

Definition C.4. *Let $\{x_\alpha\}_{\alpha\in A} \subset X$ be a net then:*

1. *x_α **converges to** $x \in X$ (written $x_\alpha \to x$) iff for all $V \in \tau_x$, $x_\alpha \in V$ **eventually**, i.e. there exists $\beta = \beta_V \in A$ such that $x_\alpha \in V$ for all $\alpha \geq \beta$.*
2. *x is a **cluster point** of $\{x_\alpha\}_{\alpha\in A}$ if for all $V \in \tau_x$, $x_\alpha \in V$ **frequently**, i.e. for all $\beta \in A$ there exists $\alpha \geq \beta$ such that $x_\alpha \in V$.*

Proposition C.5. *Let X be a topological space and $E \subset X$. Then*

1. *x is an accumulation point of E (see Definition 8.28) iff there exists net $\{x_\alpha\} \subset E \setminus \{x\}$ such that $x_\alpha \to x$.*
2. *$x \in \bar{E}$ iff there exists $\{x_\alpha\} \subset E$ such that $x_\alpha \to x$.*

Proof. 1. Suppose x is an accumulation point of E and let $A = \tau_x$ be ordered by reverse set inclusion. To each $\alpha \in A = \tau_x$ choose $x_\alpha \in (\alpha \setminus \{x\}) \cap E$ which is possible sine x is an accumulation point of E. Then given $V \in \tau_x$ for all $\alpha \geq V$ (i.e. and $\alpha \subset V$), $x_\alpha \in V$ and hence $x_\alpha \to x$.
Conversely if $\{x_\alpha\}_{\alpha\in A} \subset E \setminus \{x\}$ and $x_\alpha \to x$ then for all $V \in \tau_x$ there exists $\beta \in A$ such that $x_\alpha \in V$ for all $\alpha \geq \beta$. In particular $x_\alpha \in (E \setminus \{x\}) \cap V \neq \emptyset$ and so $x \in \mathrm{acc}(E)$ – the accumulation points of E.
2. If $\{x_\alpha\} \subset E$ such that $x_\alpha \to x$ then for all $V \in \tau_x$ there exists $\beta \in A$ such that $x_\alpha \in V \cap E$ for all $\alpha \geq \beta$. In particular $V \cap E \neq \emptyset$ for all $V \in \tau_x$ and this implies $x \in \bar{E}$.
For the converse recall Proposition 8.30 implies $\overline{E} = E \cup \mathrm{acc}(E)$. If $x \in \mathrm{acc}(E)$ there exists a net $\{x_\alpha\} \subset E$ such that $x_\alpha \to x$ by item 1. If $x \in E$ we may simply take $x_n = x$ for all $n \in A := \mathbb{N}$.
■

Proposition C.6. *Let X and Y be two topological spaces and $f : X \to Y$ be a function. Then f is continuous at $x \in X$ iff $f(x_\alpha) \to f(x)$ for all nets $x_\alpha \to x$.*

Proof. If f is continuous at x and $x_\alpha \to x$ then for any $V \in \tau_{f(x)}$ there exists $W \in \tau_x$ such that $f(W) \subset V$. Since $x_\alpha \in W$ eventually, $f(x_\alpha) \in V$ eventually and we have shown $f(x_\alpha) \to f(x)$.

Conversely, if f is **not** continuous at x then there exists $W \in \tau_{f(x)}$ such that $f(V) \nsubseteq W$ for all $V \in \tau_x$. Let $A = \tau_x$ be ordered by reverse set inclusion and for $V \in \tau_x$ choose (axiom of choice) $x_V \in V$ such that $f(x_V) \notin W$. Then $x_V \to x$ since for any $U \in \tau_x$, $x_V \in U$ if $V \geq U$ (i.e. $V \subset U$). On the over hand $f(x_V) \notin W$ for all $V \in \tau_x$ showing $f(x_V) \nrightarrow f(x)$. ■

Definition C.7 (Subnet). *A net $\langle y_\beta \rangle_{\beta\in B}$ is a **subnet** of a net $\langle x_\alpha \rangle_{\alpha\in A}$ if there exists a map $\beta \in B \to \alpha_\beta \in A$ such that*

1. *$y_\beta = x_{\alpha_\beta}$ for all $\beta \in B$ and*

2. *for all $\alpha_0 \in A$ there exists $\beta_0 \in B$ such that $\alpha_\beta \geq \alpha_0$ whenever $\beta \geq \beta_0$, i.e. $\alpha_\beta \geq \alpha_0$ eventually.*

Proposition C.8. *A point $x \in X$ is a* ***cluster*** *point of a net $\langle x_\alpha \rangle_{\alpha \in A}$ iff there exists a subnet $\langle y_\beta \rangle_{\beta \in B}$ such that $y_\beta \to x$.*

Proof. Suppose $\langle y_\beta \rangle_{\beta \in B}$ is a subnet of $\langle x_\alpha \rangle_{\alpha \in A}$ such that $y_\beta = x_{\alpha_\beta} \to x$. Then for $W \in \tau_x$ and $\alpha_0 \in A$ there exists $\beta_0 \in B$ such that $y_\beta = x_{\alpha_\beta} \in W$ for all $\beta \geq \beta_0$. Choose $\beta_1 \in B$ such that $\alpha_\beta \geq \alpha_0$ for all $\beta \geq \beta_1$ then choose $\beta_3 \in B$ such that $\beta_3 \geq \beta_1$ and $\beta_3 \geq \beta_2$ then $\alpha_\beta \geq \alpha_0$ and $x_{\alpha_\beta} \in W$ for all $\beta \geq \beta_3$ which implies $x_\alpha \in W$ frequently.

Conversely assume x is a cluster point of a net $\langle x_\alpha \rangle_{\alpha \in A}$. We mak $B := \tau_x \times A$ into a directed set by defining $(U, \alpha) \leq (U', \alpha')$ iff $\alpha \leq \alpha'$ and $U \supseteq U'$. For all $(U, \gamma) \in B = \tau_x \times A$, choose $\alpha_{(U,\gamma)} \geq \gamma$ in A such that $y_{(U,\gamma)} = x_{\alpha_{(U,\gamma)}} \in U$. Then if $\alpha_0 \in A$ for all $(U', \gamma') \geq (U, \alpha_0)$, i.e. $\gamma' \geq \alpha_0$ and $U' \subset U$, $\alpha_{(U',\gamma')} \geq \gamma' \geq \alpha_0$. Now if $W \in \tau_x$ is given, then $y_{(U,\gamma)} \in U \subset W$ for all $U \subset W$. Hence fixing $\alpha \in A$ we see if $(U, \gamma) \geq (W, \alpha)$ then $y_{(U,\gamma)} = x_{\alpha_{(U,\gamma)}} \in U \subset W$ showing that $y_{(U,\gamma)} \to x$. ■

Exercise C.1. [#34, p. 121] Let $\langle x_\alpha \rangle_{\alpha \in A}$ be a net in a topological space and for each $\alpha \in A$ let $E_\alpha \equiv \{x_\beta : \beta \geq \alpha\}$. Then x is a cluster point of $\langle x_\alpha \rangle$ iff $x \in \bigcap_{\alpha \in A} \overline{E}_\alpha$.

Proof. If x is a cluster point, then given $W \in \tau_x$ we know $E_\alpha \cap W \neq \emptyset$ for all $\alpha \in E$ since $x_\beta \in W$ frequently thus $x \in \overline{E}_\alpha$ for all α, i.e. $x \in \bigcap_{\alpha \in A} \overline{E}_\alpha$. Conversely if x is **not** a cluster point of $\langle x_\alpha \rangle$ then there exists $W \in \tau_x$ and $\alpha \in A$ such that $x_\beta \notin W$ for all $\beta \geq \alpha$, i.e. $W \cap E_\alpha = \emptyset$. But this shows $x \notin \overline{E}_\alpha$ and hence $x \notin \bigcap_{\alpha \in A} \overline{E}_\alpha$. ■

Theorem C.9. *A topological space X is compact iff every net has a cluster point iff every net has a convergent subnet.*

Proof. Suppose X is compact, $\langle x_\alpha \rangle_{\alpha \in A} \subset X$ is a net and let $F_\alpha := \overline{\{x_\beta : \beta \geq \alpha\}}$. Then F_α is closed for all $\alpha \in A$, $F_\alpha \subset F_{\alpha'}$ if $\alpha \geq \alpha'$ and $F_{\alpha_1} \cap \cdots \cap F_{\alpha_n} \supseteq F_\gamma$ whenever $\gamma \geq \alpha_i$ for $i = 1, \ldots, n$. (Such a γ always exists since A is a directed set.) Therefore $F_{\alpha_1} \cap \cdots \cap F_{\alpha_n} \neq \emptyset$ i.e. $\{F_\alpha\}_{\alpha \in A}$ has the finite intersection property and since X is compact this implies there exists $x \in \bigcap_{\alpha \in a} F_\alpha$ By Eexrcise C.1, it follows that x is a cluster point of $\langle x_\alpha \rangle_{\alpha \in A}$.

Conversely, if X is not compact let $\{U_j\}_{j \in J}$ be an infinite cover with no finite subcover. Let A be the directed set $A = \{\alpha \subset J : \#(\alpha) < \infty\}$ with $\alpha \leq \beta$ iff $\alpha \subset \beta$. Define a net $\langle x_\alpha \rangle_{\alpha \in A}$ in X by choosing

$$x_\alpha \in X \setminus \left(\bigcup_{j \in \alpha} U_j \right) \neq \emptyset \text{ for all } \alpha \in A.$$

This net has no cluster point. To see this suppose $x \in X$ and $j \in J$ is chosen so that $x \in U_j$. Then for all $\alpha \geq \{j\}$ (i.e. $j \in \alpha$), $x_\alpha \notin \bigcup_{\gamma \in \alpha} U_\alpha \supseteq U_j$ and in particular $x_\alpha \notin U_j$. This shows $x_\alpha \notin U_j$ frequently and hence x is not a cluster point. ■

References

1. Vladimir I. Bogachev, *Gaussian measures*, Mathematical Surveys and Monographs, vol. 62, American Mathematical Society, Providence, RI, 1998. MR 2000a:60004
2. Isaac Chavel, *Riemannian geometry—a modern introduction*, Cambridge University Press, Cambridge, 1993. MR 95j:53001
3. ———, *The Laplacian on Riemannian manifolds*, Spectral theory and geometry (Edinburgh, 1998), Cambridge Univ. Press, Cambridge, 1999, pp. 30–75. MR 2001c:58029
4. Giuseppe Da Prato and Jerzy Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press, Cambridge, 1992. MR 95g:60073
5. E. B. Davies, *Heat kernels and spectral theory*, Cambridge Tracts in Mathematics, vol. 92, Cambridge University Press, Cambridge, 1990. MR 92a:35035
6. ———, *Heat kernel bounds for higher order elliptic operators*, Journées "Équations aux Dérivées Partielles" (Saint-Jean-de-Monts, 1995), École Polytech., Palaiseau, 1995, pp. Exp. No. III, 11. MR 96i:35020
7. Enrico Giusti, *Minimal surfaces and functions of bounded variation*, Birkhäuser Verlag, Basel, 1984. MR 87a:58041
8. Leonard Gross, *Integration and nonlinear transformations in Hilbert space*, Trans. Amer. Math. Soc. **94** (1960), 404–440. MR 22 #2883
9. ———, *Abstract Wiener spaces*, Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 1, Univ. California Press, Berkeley, Calif., 1967, pp. 31–42. MR 35 #3027
10. Ronald B. Guenther and John W. Lee, *Partial differential equations of mathematical physics and integral equations*, Dover Publications Inc., Mineola, NY, 1996, Corrected reprint of the 1988 original. MR 97e:35001
11. Hui Hsiung Kuo, *Gaussian measures in Banach spaces*, Springer-Verlag, Berlin, 1975, Lecture Notes in Mathematics, Vol. 463. MR 57 #1628
12. Serge Lang, *Real and functional analysis*, third ed., Graduate Texts in Mathematics, vol. 142, Springer-Verlag, New York, 1993. MR 94b:00005
13. Vladimir G. Maz'ja, *Sobolev spaces*, Springer Series in Soviet Mathematics, Springer-Verlag, Berlin, 1985, Translated from the Russian by T. O. Shaposhnikova. MR 87g:46056

14. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983. MR 85g:47061
15. Michael Reed and Barry Simon, *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975. MR 58 #12429b
16. ______, *Methods of modern mathematical physics. I*, second ed., Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980, Functional analysis. MR 85e:46002
17. Laurent Saloff-Coste, *Aspects of Sobolev-type inequalities*, London Mathematical Society Lecture Note Series, vol. 289, Cambridge University Press, Cambridge, 2002. MR 2003c:46048
18. Robert Schatten, *Norm ideals of completely continuous operators*, Second printing. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 27, Springer-Verlag, Berlin, 1970. MR 41 #2449
19. B. Simon, *Convergence in trace ideals*, Proc. Amer. Math. Soc. **83** (1981), no. 1, 39–43. MR 82h:47042
20. Barry Simon, *Trace ideals and their applications*, London Mathematical Society Lecture Note Series, vol. 35, Cambridge University Press, Cambridge, 1979. MR 80k:47048
21. S. L. Sobolev and E. R. Dawson, *Partial differential equations of mathematical physics*, Translated from the third Russian edition by E. R. Dawson; English translation edited by T. A. A. Broadbent, Pergamon Press, Oxford, 1964. MR 31 #2478
22. Elias M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. MR 44 #7280
23. Frank W. Warner, *Foundations of differentiable manifolds and Lie groups*, Graduate Texts in Mathematics, vol. 94, Springer-Verlag, New York, 1983, Corrected reprint of the 1971 edition. MR 84k:58001