COMPLEX NUMBERS AND GEOMETRY

by
J.B. Cooper

Inhaltsverzeichnis

1 Geometry and function theory

1.1 Normal families

We consider now the space $H(U)$ of holomorphicic functions on a domain U. We regard this as a Fréchet space (complete metrisable locally convex space) with the countable family $\{p_n\}$ of seminorms where $p_n(f)$ is the supremum of the absolute value of f on K_n and (K_n) is a (countable) basis for the compact subsets of U. We shall only require the following facts about the corresponding topology—it is metrisable and convergence means uniform convergence on compact subsets of U i.e. almost uniform convergence. (see Appendix).

Then we have the following characterisation of the relatively compact subsets of $H(U)$ which follows from the theorem of Ascoli and the Cauchy integral theorem:

Proposition 1.1 *A subset A of $C(U)$ is relatively compact if and only if it is uniformly bounded on compacta and equicontinuous on compacta. $A \subset H(U)$ is relatively compact if and only if it is uniformly bounded on compacta.*

Proof. The first statement is simply a special case of the version of Ascoli's theorem formulated in the Appendix (to Chapter 2). In order to prove the second case we begin by assuming that U is the unit disc D. We shall then show below how to deduce the general case from this one by a general localisation principle.

By Ascoli's theorem and the Weierstraß result on the almost uniform limit of analytic functions, it suffices to show that a family A which is bounded on compacta is equicontinuous on compacta. For this it suffices to show that if is equicontinuous on each set of the type

$$\overline{D}_r = \{z : |z| \leq r\}$$

for $0 < r < 1$. In fact it is uniformly Lipschitz since its the derivatives are uniformly bounded on this set (see Appendix). This follows from the Cauchy estimates

$$|f'(z)| \leq M_R \qquad (|z| \leq r)$$

where R is chosen between r and 1 and $M_R = \sup\{|f(z)| : |z| \leq R, f \in A\}$.

■

In particular we have Montel's theorem—if A is uniformly bounded, then it is relatively compact i.e. each sequence in A has a subsequence which converges almost uniformly to an analytic function.

Relatively compact subsets of $H(U)$ are traditionally called **normal families**.

We remark here that the classical definition of normality, in fact, runs as follows: A set A is normal if and only if each sequence in A has a subsequence as above, or one that converges at each point to infinity. Thus for example, the sequence $\{nz : n \in \mathbf{N}\}$ is normal in $H(G)$ where $G = \mathbf{C} \setminus \{0\}$ but not in $H(\mathbf{C})$. We shall consider this type of normality later in what we regard as its natural setting, that of meromorphic functions. The reader can verify (this requires the use of Hurwitz' theorem) that the family A is normal in this sense if and only if it is relatively compact as a subset of the space of continuous functions from U into the extended complex plane.

We now indicate in a sequence of lemmata how to deduce the general version of the above result from the case of functions on the disc.

Lemma 1.2 *A subset $\mathcal{F} \subset \mathcal{H}(\mathcal{U})$ is normal if and only if it is normal at each point i.e. for each z there is a neighbourhood of z on which the restrictions of the functions of $\mathcal{F}$ are normal).*

Proof. We sketch the proof which is a typical application of the diagonal method. It is trivial that normality on any domain implies the same property on each subdomain. Hence it suffices to show that local normality implies normality. For each $z \in U$ there is a r_z so that $\mathcal{F}$ is normal on the disc with centre z and radius r_z. The open balls $U(z, r_z)$ cover U and so by the Lindelöf property of separable metric spaces (see Appendix) we can cover U by a sequence of such balls—we denote this sequence by (U_n). We now proceed by menas of the diagonal procedure (see Appendix). Namely we choose successively subsequences (f_n^k) where

a) the sequence $(f_n^{k+1})_n$ is a subsequence of $(f_n^k)_n$. (By convention, we identify the original sequence with $(f_n^0)_n$).

b) the sequence $(f_n^k)_n$ converges uniformly onf U_k. Then the diagonal sequence $(f_k^k)_k$ can be seen to converge almost uniformly.

■

In a similar way one can show that if $\mathcal{F}$ is a subset of $H(U)$, then the following conditions are equivalent.

1. $\mathcal{F}$ is bounded on compacta i.e. for each K compact in U $\mathcal{F}$ is uniformly bounded on K;

2. $\mathcal{F}$ is locally bounded i.e. each $z \in U$ has a neighbourhood on which $\mathcal{F}$ is uniformly bounded.

Now using the Cauchy estimates it is easy to show that if $\mathcal{F}$ satisfies the second condition, then the corresponding family $\mathcal{F}' = \{\{' : \{| \in \mathcal{F}\}$ of derivatives also satisfies it. Hence we have the result

Proposition 1.3 *If $\mathcal{F}$ is a subset of $H(U)$ which is uniformly bounded on compacta then so is $\{f' : f \in \mathcal{F}\}$.*
We remark that from the above it follows that if A is a locally bounded subset of $H(U)$ then it is equicontinuous (since the derivates are locally bounded and this implies that the restrictions of A to compact subsets are uniformly Lipshitz continuous).
We close with a famous result of classical function theory which can easily proved using these ideas.

Exercise 1.4 *Prove the theorem of Vitali-Porter: Let (f_n) be a locally bounded sequence in $H(U)$, $f \in H(U)$ so that $\lim f_n(z)$ exists for each z in a subset A of U which has an accumulation point in U. Show that there is then an $f \in H(U)$ so that $f_n \to f$ almost uniformly. (Compare the (compact) topology of almost uniform convergence with the Hausdorff topology of pointwise convergence on A).*

One of the main applications of normal families is to the proof of the Riemann mapping theorem:

Proposition 1.5 *Let the domain U be a proper subset of* $\mathbf{C}$ *and simply connected (equivalently, homeomorphic to D). Then U is conformally equivalent to D.*

Proof. We begin the proof with the following two reductions.
Firstly we note that if the closure of U is a proper subset of $\mathbf{C}$ (i.e. U is not dense in $\mathbf{C}$), then we can find a disc in the exterior of U. Inversion in this disc is a conformal mapping of U onto a bounded set.
On the other hand since U is simply connected, we can define a branch of the logarithm on U. (For convenience, we assume that $0 \notin U$ and $1 \in U$—we can clearly arrange this by simple transformations. Then we define

$$\ln z = \int_\gamma \frac{1}{z}\, dz$$

where γ is a path from 1 to z in U. By the condition of simply connectedness, this is independent of the path and so is an analytic function). By the usual argument the image of U ander this function (which is then conformally equivalent to U) satisfies the conditions of the previous paragraph.
Hence combining these two arguments we can reduce to the case where U is a bounded region.
We now introduce for a fixed P in U the family

$$\mathcal{F} = \{\{ : \mathcal{U} \to \mathcal{D} : \{ \text{ is } 1-1, \text{ holomorphicic and } \{(\mathcal{P}) = \prime\}.$$

This is non-empty (since we can find such a function on any bounded disc). By Montel's theorem, $\mathcal{F}$ is normal.
Now we set $M = \sup\{|f'(P)| : f \in \mathcal{F}\}$ and find a sequence (f_n) with $|f_n{}'(P)| \geq M - \dfrac{1}{n}$.
Since $\mathcal{F}$ is normal we can, by going over to a subsequence, assume that f_n converges in $H(U)$, say to f. We claim that $f \in \mathcal{F}$ and (as is obvious) $|f'(P)| = M$ i.e. the above supremum is attained. Firstly we note that f is non-constant (why?). Secondly, f takes its values in $\overline{D}$ as the limit of the f_n. But it then follows from the maximum modulus theorem that in fact its image is a subset of D. It can easily be deduced using Hurwitz' theorem that f is also $1-1$ (Exercise).
Hence it remains to show that f is onto. This is where the concrete analysis comes in. We shall show that if f is not onto, then we can find a g in $\mathcal{F}$ with too large a value for $|g'(P)|$. Assume therefore that f is not onto. Suppose that the value β is owithted. We now compose with the Möbius transformation ϕ_β to get a function F which owis the value 0. Since this function is defined on a simply connected domain, we can again as above define a function $\ln F$ (by putting

$$\ln F(z) = \int \frac{F'(\zeta)}{F(\zeta} \, d\zeta,$$

the integral being taken along any path in U from P to z). We can then define F^α for any α by putting

$$F^\alpha(z) = \exp(\alpha \ln F(z)).$$

Putting this together we define an analytic function μ where

$$\mu(z) = (\phi_\beta \circ f(z))^{1/2}.$$

Then we put

$$\nu(z) = \frac{|\mu'(P)|}{\mu'(P)} \phi_\tau \circ \mu(z)$$

where $\tau = \mu(P)$. Then $\nu \in \mathcal{F}$ and one can compute that

$$\nu'(P)| = \frac{1 + |\beta|}{2|\beta|^{1/2}} M$$

and this is strictly greater then M. This contradiction proves the result.

■

1.2 Domains as Riemann manifolds

We shall now apply some of the results of the previous chapter to obtain two of the most remarkable results on analytic functions, the theorems of Picard. In order to do this we shall consider open subsets of $\mathbf{C} = \mathbf{R}^2$ as local Riemannian manifolds of a particularly simple type, namely where the metric tensor is a multiple of the ordinary scalar product (the factor depending on the point in the set). This means that the computations required will be particularly simple.

If $z \in \Omega$ and ξ is in $\mathbf{C} = \mathbf{R}^2$, we define

$$\|\xi\|_{\rho,z} = \rho(z)|\xi|.$$

Then the length $\ell_\rho(\gamma)$ of a path is gigen by the equation

$$\ell_\rho(\gamma) = \int_a^b \|\gamma'(s)\|_{\rho,\gamma(s)}\, ds.$$

($[a, b]$ is the parametrising interval). Note that we are not implying by the use of the symbols γ and s that γ has arc length parametrisation (as we did in chapter 2).
In the following we shall often use the notation P, Q etc. to denote points of the local manifold U when we are emphasising the geometric interpretation. Of course they are also points of the complex plane and so can be denoted by z, w etc. We shall sometimes even write z_P to denote the complex number corresponding to the geometrical point P on the manifold. If P and Q are points in U, then the **geodetic distance** is

$$\rho(P, Q) = \inf\{\ell_\rho(\gamma)\}$$

the infimum being taken over all paths in U from P to Q.
We remark that it follows from a general theorem on Riemann manifolds that the above infimum is attained (i.e. there is a path from P to Q with length $\rho(P, Q)$) for each pair P, Q if and only if U is complete under the metric ρ (it is relatively easy to see that the latter *is* a metric).
In the language of the previous chapter we are dealing with the local Riemann manifold U with metric tensor (first fundamental form)

$$G = \begin{bmatrix} \rho^2(x, y) & 0 \\ 0 & \rho^2(x, y) \end{bmatrix}$$

(i.e. "$ds^2 = \rho^2(dx^2 + dy^2)$").
Then $g = \sqrt{\det G} = \rho^2$ and

$$G^{-1} = \begin{bmatrix} 1/\rho^2 & 0 \\ 0 & 1/\rho^2 \end{bmatrix}.$$

Our main examples will be
$\Omega = D$ and $\rho = \dfrac{1}{1 - |z|^2}$
resp.
$\Omega = H_+ = \{z \in \mathbf{C} : \Im z > 0\}$ with $\rho = \dfrac{1}{y^2}$.
In fact these spaces are essentially the same and are models for non-euclidean (hyperbolic) geometry.

Example 1.6 *A routine calculation show that the length of the curve $\gamma(t) = t$ $(0 \le t \le 1-\epsilon)$ in the Poincaré metric $\dfrac{1}{1-|z|^2}$ is $\frac{1}{2} \ln \dfrac{2-\epsilon}{\epsilon}$. For the length is*

$$\int_0^{1-\epsilon} \frac{1}{1-t^2}\, dt$$

which gives the above expression. If we put $R = 1 - \epsilon$ then it takes the form

$$\frac{1}{2} \ln \frac{1+R}{1-R}.$$

This result can be reinterpreted as the fact that if $P = 0$ and $Q = (R, 0)$ where $0 < R < 1$, then $\rho(P, Q) = \frac{1}{2} \ln\left(\frac{1+R}{1-R}\right)$. (It is intuitively obvious and easy to demonstrate that the above path is the *shortest route from P and Q. In fact if we consider a curve of the form $\gamma(t) = t + ib(t)$ with $b(0) = 0$, $b(1-\epsilon) = 1-\epsilon$, then it length is*

$$\int_0^{1-\epsilon} \frac{(1+b'(t)^2)^{1/2}}{1-t^2-b(t)^2}\,dt$$

which is clearly larger than the above value).

We remark that the metrics on D and H_+ above define the usually topology on these subsets of $\mathbf{C}$. However, they are not *metrically* equivalent to the euclidean metrics there. In fact both of these metrics are complete (see below)

Definition 1.7 *Suppose now that $f : \Omega_1 \to \Omega_2$ is a non-constant holomorphic function. If ρ is a metric on Ω_2, we define the* **induced metric** *$f^*\rho$ on Ω_1 by putting*

$$f^*\rho(z) = \rho(f(z)) \left|\frac{\partial f}{\partial z}\right|.$$

If f is a bijection and ρ_1 resp. ρ_2 are metrics on the above spaces, then f is **an isometry** *if $f^*\rho_2 = \rho_1$. Then f^{-1} is also an isometry.*
A simple calculation shows that then $\ell_{\rho_1}(\gamma) = \ell_{\rho_2}(f \circ \gamma)$ for each curve in Ω_1 and so that $\rho_1(P,Q) = \rho_2(f(P), f(Q))$ for P,Q in Ω_1.

For example one can show that if h is a conformal mapping of D, then h is an isometry for the Poincaré metric. For this it suffices to consider the two cases ρ_τ and ϕ_a discussed above:

Case 1. $w = e^{i\tau}z$. Then $h'(z) = e^{i\tau}$ and so its absolute value is 1. Hence

$$h^*\rho(z) = \rho(w) = \frac{1}{1-|w|^2} = \frac{1}{1-|z|^2} = \rho(z).$$

Case 2. $w = \dfrac{z-a}{1-\bar{a}z}$. Then

$$\begin{aligned} h^*\rho(z) &= \rho(w)\left|\frac{dw}{dz}\right| \\ &= \frac{1}{1-\frac{|a-a|^2}{|1-\bar{a}z|^2}} \frac{1-|a|^2}{|1-\bar{a}z|^2} \\ &= \frac{1-|a|^2}{|1-\bar{a}z|^2 - |z-a|^2} \\ &= \frac{1-|a|^2}{(1-|a|^2)(1-|z|^2)}. \end{aligned}$$

From this we can deduce the following formula:

Proposition 1.8 *If P and Q are points of D, then*

$$\rho(P,Q) = \frac{1}{2}\ln\left(\frac{1+\left|\frac{P-Q}{1-\bar{P}Q}\right|}{1-\left|\frac{P-Q}{1-\bar{P}Q}\right|}\right).$$

PROOF. Firstly we have already seen that the formula is true for $P = 0$ and $Q = (R, 0)$. For the general case, we use the isometry ϕ_P (with the notation from above). Then

$$\rho(P, Q) = \rho(0, \phi_P(Q)) = \rho(0, |\phi_P(Q)|),$$

the last equality following from the fact that rotations about the origin are isometries. But

$$|\phi_P(Q)| = \left| \frac{P - Q}{1 - \bar{P}Q} \right|.$$

■

The same calculation shows that the curve of shortest length from P to Q is

$$\gamma_{P,Q} : t \mapsto \frac{t\frac{Q-P}{1-Q\bar{P}} + P}{1 + t\bar{P}\frac{Q-P}{1-Q\bar{P}}}.$$

This is the pre-image of the straight line from 0 to $\phi_P(Q)$ and so by the properties of Möbius transformations discussed above is (in general) an arc of a circle which cuts the unit circle at right angles.
Using the above formula one can check that $\rho(0, z) \leq r$ if and only if $\frac{1}{2} \ln \left(\frac{1+|z|}{1-|z|} \right) < r$ i.e. if and only if $|z| \leq \dfrac{e^{2r} - 1}{e^{2r} + 1}$. This allows us to show the equivalence of topologies mentioned above. For it follows that the discs with centre 0 form a basis there for both topologies. Further the above Möbius transformations are homemomorphisms for both topologies. From this it is easy to deduce that ρ induces the natural topology on the open ball. We can also see that (D, ρ) is complete. For let (z_n) be a ρ Cauchy sequence. Then it is bounded, say $\rho(z_n, 0) \leq M$ for some constant M. But then

$$|z_n| \leq \frac{e^{2M} - 1}{e^{2M} + 1}$$

and so the sequence lies in a compact subset of U. It follows easily from this that it converges (for both the natural and the ρ-topology).

Exercise 1.9 *Calculate the Christoffel symbols and the geodetic equations for D.*

We remark that it is not difficult to see that the above properties determine the Poincaré metric. In fact if $\bar{\rho}$ is a metric on D so that each conformal mapping of the disc is an isometry, then $\bar{\rho}$ is a multiple of the Poincaré metric. For suppose that $w = h(z) = \dfrac{z + z_0}{1 + \bar{z}_0 z}$. Then since

$$h^* \bar{\rho}(0) = \bar{\rho}(0)$$

we have

$$|h'(0)| \bar{\rho}(h(0)) = \bar{\rho}(0)$$

i.e.

$$\bar{\rho}(z_0) = \frac{1}{1 - |z_0|^2} \bar{\rho}(0) = \bar{\rho}(0) \rho(z_0)$$

which means that $\bar{\rho}$ is a multiple of the Poincaré metric.
On the other hand, if $f : D \to D$ is an isometry for ρ then f is automatically holomorphic.

PROOF. For by the usual reduction we can assume that $f(0) = 0$. Then as above the circle $C^R = \{z : |z| = R\}$ is mapped onto itself for each $0 < R < 1$. This means that for any P

$$\frac{|f(P) - f(0)|}{|P - 0|} = \frac{|f(P)|}{|P|}$$

and so f is conformal at 0 (in the sense that it preserves angles between curves). Once again by the homogeneity, this holds everywhere. But by the result sketched in the second chapter on the geometrical charaterisation of conformal mappings, this implies that f is either holomorphic or anti-holomorphic. The latter case is impossible (since then $\left|\frac{\partial f}{\partial \bar{z}}\right| = 0$).

■

The Lemma of Schwarz-Pick can now be interpreted as follows: Suppose that f is a holomorphic function on the disc. Then f is a contraction i.e. $f^*\rho \leq \rho$ (and so $\ell_\rho(f \circ \gamma) \leq \ell_\rho(\gamma)$ and $\rho((f(P), f(Q)) \leq \rho(P, Q))$.

Using a refinement of the Banach fixed point theorem one can deduce that if f is a holomorphic mapping on the disc with relatively compact range, then it has a (unique) fixed point. (This is known as the theorem of Farkas and Pitt).

PROOF. Under these conditions the function

$$g : z \mapsto f(z) + \epsilon(f(z) - f(0))$$

maps D into D for small enough ϵ. Then, by the above, g is a contraction in the weak sense and, since $f'(z) = \frac{1}{1+\epsilon} g'(z)$, f is a contraction in the sense of the Banach fixed point theorem.

■

It follows from the proof that the fixed point is obtained as the limit of the iterated sequence $z, f(z), f^2(z), f^3(z), \dots$ for any z in the disc.

1.3 Curvature

Since the (Gaussian) curvature is an intrinsic quantity of a Riemann surface, it can be defined in terms of the metric tensor. In fact, the formula of the Theorema Egregium simplifies in this case to

$$-\frac{\Delta \ln \rho(z)}{\rho(z)^2}$$

and we denote this quantity by $\kappa_{U,\rho}(z)$ (or simply $\kappa(z)$). (For in the case where $F = 0$, the Theorema Egregium produces the following formula for the curvature:

$$4E^2G^2\kappa = E(E_2G_2 + G_1^2) + G(E_1G_1 + E_2^2) - 2EG(E_{22} + G_{11}).$$

A simple calculation shows that this yields the above expression when $E = \rho^2 = G$). It follows immediately from these remarks that this quantity is preserved by an isometry. (This can also be deduced directly by a simple computation). It is easy to see that $\kappa = 0$ when ρ is the constant function 1 on $\mathbf{C}$, while $\kappa = -4$ for the Poincaré metric. On the other hand if $\rho = \frac{2}{1 + |z|^2}$, then $\kappa = 1$ (we shall see the geometrical reason for this shortly).

The curvature will play a crucial role in our versions of the Picard theorems. In order to see the connection note that the theorem of Liouville (in the form that each entire function with values in D is constant) can be regarded as a special case of Picard's little theorem. We shall show that the same result holds for entire functions with values in a domain U which allow a metric with certain curvature properties.

Proposition 1.10 *Let U be a domain with metric σ for which $\kappa \leq -4$. Then if $f : D \to U$ is holomorphic, $f^*\sigma \leq \rho$.*

PROOF. We consider the smaller disc $U(0,r)$ (with $r<1$) and then let r go to 1. On this set we rescale the metric to $\rho_r = \frac{r}{r^2-|z|^2}$, for which we also have constant curvature $=-4$. Define the function $v = \frac{f^*\sigma}{\rho_r}$. This is continuous and non-negative, and converges to zero at the boundary. Hence $|v|$ attains its maximum M at a point $\tau \in U(0,r)$. We show that $M \leq 1$, from which our result follows. We can suppose that $f^*\sigma(\tau) > 0$. Then the curvature of $f^*\sigma$ is defined at τ (and is ≤ -4). Now since $\ln v$ has a maximum at τ, its Laplacian there is ≤ 0 (consider the Hessean matrix). Thus

$$\begin{aligned} 0 \geq \Delta \ln v(\tau) &= \Delta \ln f^*\sigma(\tau) - \Delta \ln \rho_r \\ &= -\kappa_{f^*\sigma}(\tau)(f^*\sigma(\tau))^2 + \kappa_{\rho_r}(\rho_r(\tau))^2 \\ &\geq 4(f^*\sigma(\tau))^2 - 4(\rho_r(\tau))^2 \end{aligned}$$

and so $v(\tau) \leq 1$ which implies that $M \leq 1$.

■

We remark that this can be regarded as a form of Schwarz' lemma (apply the above to an f which vanishes at 0).
Rescaling, we get:

Proposition 1.11 *Let U be a domain with metric σ for which there is a positive constant B with $\kappa \leq -B$. Then if ρ_A^α is the metric $\frac{2\alpha}{\sqrt{A}(\alpha^2-|z|^2)}$ on $U(0,\alpha)$, we have*

$$f^*\sigma(z) \leq \frac{\sqrt{A}}{\sqrt{B}}\rho_A^\alpha(z)$$

for each holomorphic mapping f from $U(0,\alpha) \to U$.

PROOF. Exercise.

■

As an application we have

Proposition 1.12 *Let U be a domain with a metric σ so that $\kappa_\sigma \leq -B$ for some positive constant B. Then each holomorphic function f from $\mathbf{C}$ into U is constant.*

PROOF. We consider f as a mapping from $U(0,\alpha)$ with metric ρ_A^α for positive A. Then

$$f^*\sigma(z) \leq \frac{\sqrt{A}}{\sqrt{B}}\rho_A^\alpha(z)$$

for $|z|<\alpha$. Letting α go to infinity gives $f^*\sigma(z) \leq 0$ and so $f^*\sigma = 0$. But this can only happen if f' vanishes identically.

■

This result contains Liouville's theorem as a special case.
In order to motivate our proof of Picard's little theorem, consider how we can prove the following generalisation of Liouville. We show that every entire function with values in $\mathbf{C}\setminus[0,1]$ is constant. For it is a standard exercise in conformal mappings to map the above range space conformally into D and so we can deduce it from the usual version of Liouville. (Use successively the mappings $z \mapsto \frac{z}{z-1}$, $z \mapsto z^{1/2}$ and $z \mapsto \frac{z-1}{z+1}$).

Proposition 1.13 *Let U be a subset of $\mathbf{C}$ whose complement contains at least two distinct points. Then there is a metric ρ on U with $\kappa_\rho \leq -B < 0$.*

Proof.
It is no real loss of generality to assume that the omitted points are 0 and 1. Then we define

$$\rho(z) = \left[\frac{\left(1+|z|^{1/3}\right)^{1/3}}{|z|^{5/6}}\right]\left[\frac{\left(1+|z-1|^{1/3}\right)^{1/3}}{|z-1|^{5/6}}\right].$$

ρ is a smooth positive function on U and a tedious calculation shows that the curvature is

$$\kappa(z) = -\frac{1}{18}\left[\frac{\left(|z-1|^{1/3}\right)^{5/3}}{(1+|z|^{1/3})^2(1+|z-1|^{1/3})} + \frac{|z|^{5/3}}{(1+|z|)^{1/3}(1+|z-1|^{2/3}}\right].$$

Then $\kappa < 0$ and it follows from the fact that $\lim_{z\to 0}\kappa(z) = -\frac{1}{36} = \lim_{z\to 1}\kappa(z)$ and $\lim_{z\to\infty}\kappa = -\infty$ that it is bounded away from zero on the left.

■

This, combined with the above result, immediately implies Picard's little theorem.
We now turn to the great theorem. We shall be interested in functions which take their values in the extended plane $\hat{\mathbf{C}} = \mathbf{C}\cup\{\infty\}$. We identify this with the sphere S^2 in $\mathbf{R}^3$ via stereographic projection. More precisely, if we denote by (α,β,γ) the coordinates of a point P on the sphere (which is distinct from the north pole $N = (0,0,1)$) then an elementary exercise in analytic geometry shows that the point $z = x+iy$ on the plane which is the intersection of the latter with the line through N and P is given by the formula $z = \dfrac{\alpha+i\beta}{1-\gamma}$. On the other hand, if we are given a point z in the plane, then the corresponding point $P = (\alpha,\beta,\gamma)$ on the unit sphere is given by the equations

$$\alpha = \frac{2\Re z}{1+|z|^2}, \quad \beta = \frac{2\Im z}{1+|z|^2}, \quad \gamma = \frac{-1+|z|^2}{1+|z|^2}.$$

The north pole is mapped onto ∞ by convention.
If we calculate the first fundamental form of the corresponding parametrisation

$$\phi(x,y) = \left(\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2}, \frac{-1+x^2+y^2}{1+x^2+y^2}\right)$$

of the sphere, we get:

$$E = \frac{4}{(1+|z|^2)^2} = G, \quad F = 0.$$

This implies, amongst other facts, that stereographic projection is conformal and that the metric $\sigma(z) = \dfrac{2}{1+|z|^2}$ in the plane corresponds to the usual metric on S^2 as a surface in $\mathbf{R}^3$ (i.e. our correspondence is an isometry for these metrics).

Exercise 1.14 *Show that stereographic projection maps circles on the sphere onto circles on the plane. (In fact this is only true generically—the reader is invited to eludicate the exceptions).*
Using stereopgraphic projection, we can regard Möbius transformations as operators on the sphere in a natural way. The reader is invited to investigate this, in particular, to consider the action of the basic types of Möbius transformation on the sphere.

We can compute that if P_1 and P_2 are the points on the sphere which correspond to the complex numbers z_1 and z_2 in the plane, then the chordal distance from P_1 to P_2 (i.e. the distance in $\mathbf{R}^3$, not the geodetic distance on the sphere) is

$$\frac{2|z_1-z_2|}{\sqrt{1+|z_1|^2}\sqrt{1+|z_2|^2}}$$

resp. $\dfrac{2}{\sqrt{1+|z_1|^2}}$ if $z_2 = \infty$.
For if P_1 has coordinates $(\alpha_1, \beta_1, \gamma_1)$ and P_2 $(\alpha_2, \beta_2, \gamma_2)$, then the square of this distance is

$$2 - 2(\alpha_1\alpha_2 + \beta_1\beta_2 + \gamma_1\gamma_2)$$

and if we substitute the above expressions for the α's, β's and γ's, then this leads to the above formula.
We denote this quantity by $\chi(z_1, z_2)$. An easy computation shows that

$$\chi\left(\frac{1}{z_1}, \frac{1}{z_2}\right) = \chi(z_1, z_2).$$

The length of a curve in the plane (using the metric σ) is then $\int_\gamma \dfrac{2|dz|}{1+|z|^2}$. $\sigma(z_1, z_2)$, the spherical metric, is the infimum of the lengths of the paths joining z_1 and z_2.
If f is a holomorphic mapping from U into the sphere, we define the quantity $f^\#(z)$ as the limit

$$\lim_{z' \to z} \frac{\chi(f(z), f(z'))}{|z - z'|}$$

and this can be computed to be $\dfrac{|f'(z)|}{1+|f(z)|^2}$. It follows immediately from this definition that this quantity coincides for f and $\dfrac{1}{f}$.
If we compare the above formula for $f^\#$ with the definition of the induced metric then we see that the length of a curve is given by the formula $\int_\gamma 2f^\#(z)\,|dz|$. (i.e. this is the length of the curve in the plane, using the above metric or the length of its image on the sphere using the spherical metric).
We now consider the space of continuous functions from U into $\hat{\mathbf{C}}$. In particular, each meromorphic function can be so regarded—we set the value of such a function at a pole to be ∞. We are now in the situation described in the appendix to the first section. Hence we can define:

Definition 1.15 *Let (f_n) be a sequence of meromorphic functions. Then we say that (f_n) converges* **normally** *if it converges in the sense of the metric defined above to a meromorphic function or to the constant function ∞.*

Under this definition we see that both the sequence (n) and $\left(\dfrac{n}{z}\right)$ converge normally.
A family $\mathcal{F}$ of meromorphic functions on U is then said to be **normal** if each sequence in $\mathcal{F}$ contains a subsequence which converges normally.
Using the version of Ascoli's theorem quoted in the above appendix, we see that a family of meromorphic functions is normal if and only if it is spherically equicontinuous on compacta (i.e. equicontinuous as a family of functions with values in the sphere under its geodetic metric).
Normal convergence can be characterised as follows:

Proposition 1.16 *We have $f_n \to f$ normally if and only if each z_0 has a neighbourhood on which either $f_n \to f$ or $\dfrac{1}{f_n} \to \dfrac{1}{f}$ uniformly.*

Theorem 1.17 (Marty's theorem.) *Let $\mathcal{F}$ be a family of meromorphic functions on U. $\mathcal{F}$ is normal if and only if the family $\{f^\#\sigma : f \in \mathcal{F}\}$ is uniformly bounded on compact subsets of U (σ is the natural metric on the sphere).*

Remark. Using the definition of the induced metric this means that for each compact subset K of U there is a positive constant M so that for each $z \in K$ and each $f \in \mathcal{F}$, we have

$$\frac{|f'(z)|}{1+|f(z)|^2} \leq M.$$

Proof. First suppose that the above condition is satisfied. We fix z_0 and consider each z in a suitable compact disc around z_0. Then if we choose the appropriate M for this disc we have, for any path from z_0 to z within this disc and any $f \in \mathcal{F}$,

$$\chi(f(z_0), f(z)) \leq \int_\gamma f^\#(\zeta)\,|d\zeta| \leq C|z-z_0|.$$

This implies that $\mathcal{F}$ is equicontinuous on the disc.
On other hand suppose that there is a compact subset K of U, a sequence (z_n) in K and a sequence (f_n) in $\mathcal{F}$ with $f_n^\#(z_n) \to \infty$. By the normality we can suppose that f_n is convergent, say to f. Let z_0 be a limit point of (z_n). Then there is a disc around this point for which either $f_n \to f$ or $\dfrac{1}{f_n} \to \dfrac{1}{f}$ uniformly. In either case $f_n^\# \to f^\#$ uniformly on a disc around z_n (we use here the fact mentioned above that $\left(\dfrac{1}{f}\right)^\#$ coincides with $f^\#$). But this implies that $f_n^\#$ is uniformly bounded on the disc and this contradicts the assumptions.

■

Proposition 1.18 *Let U be a domain in $\mathbf{C}$ and P, Q and R three distinct points in the extended plane. Then if $\mathcal{F}$ is a family of meromorphic functions taking values in $\hat{\mathbf{C}} \setminus \{P,Q,R\}$, $\mathcal{F}$ is normal.*

Proof.
We make the customary reduction to the case where the three exceptional points are 0, 1 and ∞. Then it suffices to show that the family is normal on any disc $U(z_0,\alpha)$. It is no loss of generality to suppose that $z_0 = 0$. We use the special metric constructed above on $\mathbf{C} \setminus \{0,1\}$. By rescaling we can assume that it is ≤ -4. We denote this metric by μ. Then by the version of the Schwarz' Lemma (with $A = B = 4$) we have $f^*\mu(z) \leq \rho_\alpha^A(z)$ for z in $U(0,\alpha)$. We now compare μ with the spherical metric σ. One sees easily that $\dfrac{\sigma}{\mu}$ goes to zero near the critical points 0, 1 and ∞. Hence there is an $M > 0$ so that $\sigma \leq M.\mu$ and so

$$f^\# = f^*\sigma \leq M.f^*\mu \leq M.\rho_\alpha^A$$

on $U(0,\alpha)$. By Marty's theorem, $\mathcal{F}$ is normal.

■

In particular, we can deduce as a corollary that a family of holomorphic functions on U which omit two (finite) values is normal.
The second Corollary of the above result is Picard's great theorem.
Proof.
We suppose that we have a function on the punctured unit disc D' which omits the values 0 and 1. We prove that 0 is either a pole or a removable singularity. We consider the family of functions $f_n : z \mapsto f\left(\dfrac{z}{n}\right)$. This family also omits the values 0 and 1 and so is normal. Hence it has a subsequence which converges normally and so either to a holomorphic function on the punctured disc or to the constant function ∞. It is easy to see that f has in the first case a removable singularity and in the second case a pole at 0.

■

1.4 Appendix

Definition 1.19 *Norm, seminorm, normed space Banach space , locally convex space, Fréchet space.*

In the first paragraph, we introduce the basic definitioins (normed spaces, isomorphisms, continuous linear operators) and bring some classical examples (finite dimensional spaces, spaces of continuous and differentiable functions). We then present some basic facts on duality for normed spaces. The dual of a normed space is the space of continuous linear functionals acting on it and plays the same role in the infinite dimensional theory as does the algebraic dual for finite dimensional spaces. However, as we shall see later, this duality does not possess the symmetry of the finite dimensional case—in particular, a Banach space cannot always be identified in a natural way with its second dual. It is not even obvious that the duality is non-trivial i.e. that there are always enough continuous linear functionals on a normed space. That this is in fact the case is a consequence of one of the most famous results on normed spaces—the Hahn-Banach theorem, which is studied in detail in paragraph 2.
In order to obtain deeper results of an analytic nature, a natural completeness condition must be imposed—that of completeness with respect to the metric induced by the norm. The spaces with this properties are the Banach spaces. They are introduced in paragraph 3 and it is shown that the examples introduced in the first paragraph are Banach spaces. In paragraph 4 we consider a famous group of closely related theorems which use the completenss in an essential way (*via* an application of Baire's category theorem). They can each be regarded as precise statements of the vague principle that if a linear mapping between Banach spaces can be constructed directly, it is continuous. In the fifth paragraph we introduce one of the most important groups of Banach spaces—the L^p-spaces. For the sake of completeness we begin with a brief survey of the basic concepts of measure theory. The L^p-spaces are then defined and their properties (completenss, duality etc.) are derived. Paragraph 6 is devoted to Hilbert spaces i.e. those Banach space whose norms are defined by an inner product. Historically, these were the first Banach spaces to be studied intensively—by Hilbert in the context of integral equations. We include a proof of his spectral theorem for compact, self-adjoint operators which, together with its applications to differential equations (notably, Sturm-liouville problems), was an imporatant stage in the history of functional analysis. This leads naturally to the topic of infinite dimensional operators. Their structure can be exceedingly complicated and in the seventh paragraph we restrict attention to a class of operators whose behaviour is closest to that of the finite dimensional ones—the compact operators. It is shown that their spectrum has a particulalry simple structure.
Paragraph 8 is devoted to an introduction to one of the most important concepts in Banach space theory—that of a (Schauder) basis. An attempt has been made to show just how essential a tool this is in investigations into the structure of conrete Banach spaces and their subspaces and quotients. A final section is devoted to some more subtle constructions on Banach space—infinite sums and products, tensor products and ultraproducts.
I.1. Normed spaces: We begion with the elementary theory of normed spaces. These are vector spaces with suitable distance functions. With the help of this distance, the usual procedures involving limit operations (approximation of non-linear operators by their derivatives, approximation methods for constructing solutions of equations etc.) can be carried out. Definition 1.1 below, which was explicitly introduced by Banach and Wiener, was already implicit in earlier work on integral equations by Riesz (who employed the concrete normed spaces $C(I)$, $L^p(I)$ which will be introduced below).
The plan of the section is simple. We begin by introducing the two main concepts of the chapter—normed spaces and continuous linear operators. Their more obvious properties are discussed and some concrete examples—mainly the so-called ℓ^p-spaces—are introduced.

Definition 1.20 *Definition 1.1 A* **seminorm** *on a vector spaces E (over* $\mathbf{C}$ *or* $\mathbf{R}$*) is a mapping $x \mapsto \|x\|$ from E into* $\mathbf{R}^+$ *with the properties*

$\|x + y\| \leq \|x\| + \|y\| \quad (x, y \in E)$;

$\|\lambda x\| = |\lambda| \|x\| \quad (x \in E, \lambda \in \mathbf{C}$ *or* $\mathbf{R}$ *resp.;*

$\|\ \|$ is a **norm** *if, in addition, (3) $\|x\| = 0$ implies $x = 0$ $(x \in E)$.*
Aa **normed space** *is a pair $(E, \|\ \|)$ where E is a vector space and $\|\ \|$ is a norm on E.*
If $\|\ \|$ is a seminorm (resp. a norm) the mapping

$$d_{\|\ \|} : (x, y) \mapsto \|x - y\|$$

is a semimetric (res. a metric) on E. We call it the **semimetric (metric) induced by** *$\|\ \|$. Thus evey normed space $(E, \|\ \|)$ can be regarded in a natural way as a metric space and so as a topological space and we can use, in the context of normed spaces, such notions as continuity of mappings, convergence of sequences or nets, compactness of subsets etc.*
If $(E, \|\ \|)$ is a normed space, we write $B_{\|\ \|}$ or $B(E)$ for the closed unit ball of E i.e. the set $\|x \in E : \|x\| \leq 1\}$.

Exercise 1.21 *A. A subset A of a vector space is* **absolutely convex** *if $\lambda x + \mu y \in A$ whenever $x, y \in A$, $\lambda, \mu \in \mathbf{C}$ (respectivly $\mathbf{R}$) and $|\lambda| + |\mu| \leq 1$.*

A is **absorbing** *if for each $x \in E$ there is a $\rho > 0$ so that $\lambda x \in A$ when $\|\lambda| \leq \rho$. Show that $B(E)$ is absolutly convex and absorbing and that if A is absolutely convex and absorbing, then*

$$\|\ \|_A : x \mapsto \inf\{\rho > 0 : x \in \rho A\}$$

is a seminorm on E (it is called the **Minkowski functional** *of A). Show that it is a norm if and only if A contains no non-trivial subspace of E.*

B. Let E be a vector space, A an absolutely convex subset which does not contain a non-trivial subspace. Let $E_A = \bigcup_{n \in \mathbf{N}} nA$. Show

that E_A is a vector subspace of E;

that A absorbs E_A;

that $(E, \|\ \|_A)$ is a normed space.

The usual constructions (products, subspaces, quotients etc.) can be carried out in the context of normed spaces. For examples, if G is a vector subspace of the normed space $(E, \|\ \|)$, the restriction $\|\ \|_G$ of $\|\ \|$ to G is a norm thereon and we we can regard G in a natural way as a normed space (this norm is called **the norm induced on G by $\|\ \|$**).
Similarly, if π_G denotes the natural projection from E onto the quotient space E/G, then the mapping

$$y \mapsto \inf\{\|x\| : x \in E \text{ and } \pi_G = y\}$$

is a seminorm. The question of when it is a norm is examined in an exercise below.
There are several possibilities for defining norms on product spaces and we shall discuss these in some detail later. For our purposes, the following one on a product of two spaces will suffice: let $(E, \|\ \|_1)$ and $(F, \|\ \|_2)$ be normed spaces. The mapping

$$(x, y) \mapsto \max\{\|x\|_1, \|y\|_2\}$$

is a norm on $E \times F$ which (with this norm) is then called the **normed product** of E and F. (Note that the unit ball of $E \times F$ is then just the Cartesian product of the unit balls of E and F).

Exercise 1.22 *1. Show that the topology induced by $\| \|_G$ on G coincides with the restriction to G of the topology of E;*

2. *if $x \in E$, show that $\|pi_G(x)\|$ (the norm in E/G) is just the distance from x to G i.e.* $\inf\{\|x-y\| : y \in G\}$. *Deduce that the seminorm on E/G is a norm if and only if G is closed. Use this to give an example where it is not a norm.*

3. *Show that the topology induced by the norm on $E \times F$ is the product of the topologies on E and F.*

It follows from the very definition of the topology *via* the norm that it is very closely related to the linear structure of E. In fact, the following properties are valid:

1. the mappings $A : (x, y) \mapsto x + y$ and $M : (\lambda, x) \mapsto \lambda x$ from $E \times E \to E$ resp. $\mathbf{C} \times E$ or $\mathbf{R} \times E$ to E are continuous for the topology generated by the norms. For
$$\|(x + y) - (x_1 + y_1)\| \leq \|x - x_1\| + \|y - y_1\|$$
and
$$\|\lambda x - \lambda_1 x_1\| \leq |\lambda - \lambda_1|\|x\| + |\lambda_1|\|x - x_1\|.$$

2. Let G be a subspace of $(E, \| \|)$. Then the closure $\bar{G}$ of G is also a subspace.

Exercise 1.23 *Let $\| \|$ be a seminorm on E. Show that $E_0 = \{x \in E : \|x\| = 0\}$ is a subspace of E. If π_0 denotes the natural projection from E onto E/E_0, show that $\pi_0(x) \mapsto \|x\|$ is a well-defined mapping on E/E_0 and is, in fact, a norm. E/E_0, with this norm, is called the* **normed space associated with** *E. (This simple exercise is often useful on occasions when a natural construction "should" produce a normed space but in fact only produces a seminormed one. We simply factor our the zero subspace).*

As is customary in mathematics, we identify normed spaces which have the same structure. The appropriate concept is that of an isomorphism. It turns out that there two natural ones in this context:
Let E and F be normed space. E and F are **isomorphic** if there is a bijective linear apping $T : E \to F$ so that T is a homeomorphism for the norm topologies. T is then called an **isomorphism**. If T is, in addition, norm-preserving (i.e. $\|Tx\| = \|x\|$ for $x \in E$) T is an **isometry** and E and F are **isometrically isomorphic** (we write $E \sim F$ to indicate that E and F are isomorphic).
Two norms $\| \|$ and $\| \|_1$ on a vctor space E are **equivalent** if Id_E is an isomorphism from $(E, \| \|)$ onto $(E, \| \|_1)$ i.e. if $\| \|$ and $\| \|_1$ induce the same topology on E. Isomorphisms are characterised by the existence of estimates from above and below: let $T : E \to F$ be a bijective linear mapping. Then T is an isomorphism if and only if there exist M and m (both positive) so that
$$m\|x\| \leq \|Tx\| \leq M\|x\| \quad (x \in E).$$
(For a proof see Exercise 1.8 below).
Thus the norms $\| \|$ and $\| \|_1$ on E are equivalent if and only if there are $M, m > 0$ so that $m\|x\| \leq \|x\|_1 \leq M\|x\| \quad (x \in E)$.
We now bring a list of some simple examples of normed space. In the course of the later cahpters we shall extend it considerably.

Exercise 1.24 *A. The following mappings on* $\mathbf{C}^n$ *(resp.* $\mathbf{R}^n$*) are norms:*

$$\begin{aligned} \| \ \|_1 : (\xi_1, \dots, \xi_n) &\mapsto (|\xi_1| + \dots + |\xi_n|); \\ \| \ \|_2 : (\xi_1, \dots, \xi_n) &\mapsto (|\xi_1|^2 + \dots + |\xi_n|^2)^{1/2}; \\ \| \ \|_\infty : (\xi_1, \dots, \xi_n) &\mapsto \max(|\xi_1|, \dots, |\xi_n|). \end{aligned}$$

Each of these norms induces the usual topology on $\mathbf{C}^n$ *(res.* $\mathbf{R}^n$*). Note that the respective unit balls are (for* $n = 3$*) the octahedron, the euclidean ball and the cube (or hexahedron).*

B. Let K *be a compact space.* $C(K)$ *denotes the space of continuous, complex-valued functions on* K*. This space has a natural vector space structure and the mapping*

$$\| \ \|_\infty : x \mapsto \sup\{|x(t)| : t \in K\}$$

is a norm. $\| \ \|_\infty$ *induces the topology of uniform convergence on* K *(that is, a sequence or net of functions in* $C(K)$ *is norm-convergent if and only if it is uniformly convergent on* K*).*

C. let I *be a compact interval in* $\mathbf{R}$*,* n *a positive integer. The space*

$$C^n(I) = \{x \in C(I) : x, x', \dots, x^{(n)} \text{ exist and are continuous}\}$$

has a natural vector space structure and the mapping

$$\| \ \|_\infty^n x \mapsto \max\{\|x\|_\infty, \dots, \|x^{(n)}\|_\infty\}$$

is a norm on $C^n(I)$*. Note that* $C^n(I)$ *is a vector subspace of* $C(I)$ *but that* $(C^n(I), \| \ \|_\infty^n)$ *is a not a normed subspace of* $(C(I), \| \ \|_\infty)$*—that is* $\| \ \|_\infty^n$ *is not the norm induced on* $C^n(I)$ *from* $C(I)$*—or even equivalent to it.*

D. Let $\{(E_k, \| \ \|_k) : k = 1, \dots, n\}$ *be a family of normed spaces. On the product* $E = \prod_{k=1}^n$ *we define two norms:*

$$\begin{aligned} \| \ \|_s : (x_1, \dots, x_n) &\mapsto \sum_{k=1}^n \|x_k\|_k; \\ \| \ \|_\infty : (x_1, \dots, x_n) &\mapsto \max_{k=1}^n \{\|x_k\|_k\}. \end{aligned}$$

Then $\| \ \|_s$ *and* $\| \ \|_\infty$ *are distinct norms on* E *(if* $n > 1$*) which are, however, equivalent. In fact, we have the inequality:* $\|x\|_\infty \leq \|x\|_s \leq n\|x\|_\infty$ *(note that this means geometrically that the unit ball of* $\|\|_s$ *is containd in that of* $\| \ \|_\infty$ *resp. contains a copy of it reduced by a factor* $\frac{1}{n}$*).*

Exercise 1.25 *A. ? at on* $C([0,1])$*, the mapping* $x \mapsto \int_0^1 |x(t)|\, dt$ *is a norm which is not equivalent to* $\| \ \|_\infty$*.*

B. Show that the mapping $x \mapsto (x, x', \dots, x^{(n)})$ *is an isomorphism from* $C^n(I)$ *onto a subspace of the product space* $C(I) \times \dots \times C(I)$ *((*$n+1$*) factors).*

An important role in the thoery of infinite dimensional spaces is played by linear operators. In contrast to the finite dimensional case, we impose the following condition, which takes account of the topological resp. norm structure:

Definition 1.26 *A linear mapping* $T : ((E, \| \|_1) \to (F, \| \|_2)$ *is* **bounded** *if there is a* $C > 0$ *so that* $\|Tx\|_2 \leq C\|x\|_1 \quad (x \in E)$.

In fact, this is equivalent to continuity. Indeed we have equivalence of the following three conditions on a linear operator T between normed space E and F:

T is continuous;

T is continuous at 0;

T is bounded.

Proof. 1) implies 3) is immediate. 2) implies 3): since T is continuous at 0 and $B(F)$ is a neighbourhood of $0 = T(0)$, there is a positive δ so that $Tx \in B(F)$ if $\|x\|_1 < \delta$. Now for each $x \in E$ with x non-zero, $\left\|\frac{\delta x}{\|x\|_1}\right\|_1 \leq \delta$ and so $\left\|T\left(\frac{\delta x}{\|x\|_1}\right)\right\|_1 \leq \delta$ i.e. $\|Tx\|_2 \leq \frac{\|x\|_1}{\delta}$. 3) implies 1): we suppose that C is chosen as in 1.7. Then

$$\|Tx - Ty\|_2 = \|T(x-y)\|_2 \leq C\|x-y\|_1$$

and so T is even Lipschitz continuous.

■

We write $L(E, F)$ for the set of bounded linear mappings from E into F. This space has a natural vector space structure (*via* pointwise addition and multiplication by scalars). We define a norm on it as follows:

$$\|T\| = \inf\{C > 0 : \|Tx\|_2 \leq C\|x\|_1 \quad (x \in E)\}.$$

If E, F and G are normed spaces and $T \in L(E, F)$ and $S \in L(F, G)$, then the composed mapping ST is also bounded and we have the estimate $\|ST\| \leq \|S\|\|T\|$ for its norm.

Exercise 1.27 A. *Use 1.7 to obtain the characterisatio of isomorphisms given before 1.5.*

B . *Show that the formula given above does indeed define a norm on* $L(E, F)$ *and that*

$$\begin{aligned} \|T\| &= \sup\{\frac{\|Tx\|_2}{\|x\|_1} : x \in E, x \neq 0\} \\ &= \sup\{\|Tx\|_2 : x \in B(E)\} = \sup\{\|Tx\|_2 : \|x\|_1 = 1\}. \end{aligned}$$

C. *A subset of a normed space* E *is* **bounded** *if the norm is bounded on it. Show that this is equivalent to the fact that the set* C *is absorbed by the unit ball (i.e. there is a* $K > 0$ *so that* $C \subset KB_E$*). Show that a linear mapping* T *between normed spaces is bounded if and only if it maps bounded sets into bounded sets and that a subset of* $L(E, F)$ *is bounded if and only if it is equicontinuous.*

Example 1.28 *We bring some basic examples of operatoors:*

A. *Let* $A = [a_{ij}]$ *be an* $m \times n$ *matrix. Then the coresponding linear mapping*

$$T_A : (\xi_j) \mapsto \left(\sum_{j=1}^{n} a_{ij}\xi_j\right)_{i=1}^{m}$$

is bounded from $(\mathbf{R}^n, \| \|_p)$ *into* $\mathbf{R}^m, \| \|_p)$ *for* $p = 1, 2$ *or infty*

B. *We define a mapping $D : C^1(I) \to C(I)$ by $D : x \mapsto x'$. D is linear and bounded (in fact $\|D\| \leq 1$). More generally, one can define the operator D^k of k-times differentiation which can be regarded as a continuous linear mapping from $C^{n+k}(I) into C^n(I)$.*

C. *Let a be a function in $C(I)$. Then the mapping $M_a : x \mapsto ax$ is continuous and linear on $C(I)$ and $\|M_a\| \leq \|a\|_\infty$. D. Differential operators: let $a_0, \dots, a_n$ be elements of $C(I)$. Then we define a differential operator $L : C^n(I) \to C(I)$ by $L = \sum_{k=0}^n M_{a_k} \circ D^k$.*

E. *Let K be a bounded, continuous complex-valued function on $I \times J$ (I, J compact intervals in* **R***). We define the* **integral operator** I_K **with kernel** K *as follows:*

$$I_K : x \mapsto (s \mapsto \int_J K(s,t)x(t)\, dt.$$

I_K is a continuous linear mapping from $C(J)$ into $C(I)$. F. (Projections) An operator $T \in L(E)$ is a **projection** *if $T^2 = T$. Then $Id - T$ is also a projection (since*

$$(Id - T)^2 = Id - 2T + T^2 = Id - T).$$

Also $T(E) = \{x : Tx = x\} = Ker(Id - T)$ and so is closed.

The standard example of a projection is the mapping $(x, y) \mapsto (x, 0)$ from a product space $E_1 \times E_2$ onto the factor E_2. In a sense this is the only one since if $T \in L(E)$ is a projection, then the mapping $x \mapsto (Tx, (\mathrm{Id} - T)x)$ is an algebraic isomorphism from E onto the product space $E_1 \times E_2$ where $E_1 = T(E)$, $E_2 = (\mathrm{Id} - T)(E)$. In fact, it is also an isomorphism for the norm structure on the product. Hence a projection in this case causes a splitting up of the space into a product. A subspace of E is **complemented** if it is the range of a projection $T \in L(E)$. E is then simultaneously a subspace of E and also isomorphic to a quotient space.

Exercise 1.29 A. *Let $A = [a_{ij}]$ be as in 1.9.A and consider T_A as a mapping from $(\mathbf{R}^n, \| \|_infty)$ into $(\mathbf{R}^n, \| \|_\infty)$. Show that $\|T_A\| = \sup_i \sum_{j=1}^n |a_{ij}|$. What is its norm as an operator for the norms $\| \|_1$ and $\| \|_2$?*

B. *Show that $\|D\| = 1$ and that $\|M_a\| = \|a\|_\infty$. Give an estimate for the norms of L and I_K (notation as in 1.9.B - E).*

C. *Let E and F be normed spaces, G a closed subspace of E, T a bounded linear operator from E into F. Show that if $T(G) = \|0\|$, then it can be lifted to a continuous linear operator $\tilde{T}$ from E/G into F (i.e. $\tilde{T}$ is such that $\tilde{T} \circ \pi_G = T$).*

We shall be interested in the following properties of mappings $T \in L(E, F)$:

injectivity: this means that $\mathrm{Ker}\, T = \{0\}$;

surjectivity: i.e. that $T(E) = F$;

bijectivity: i.e. injectivity and surjectivity;

isomorphicity: c.f. definition after 1.4 above;

existence of a right inverse i.e. an $S \in L(E, F)$ so that $TS = \mathrm{Id}_F$;

existence of a left-inverse i.e. an $S \in L(E, F)$ so that $ST = \mathrm{Id}_E$.

Note that for a linear operator T on a finite dimensional space E, all of these notions coincide. In the infinite dimensional case, this is no longer true. We shall give some examples here and later.
In connection with these definitions, we can give various generalisations of the notion of an eigenvalue of an operator. We shall begin here with the most useful one. Later we shall consider refinements.
If $T \in L(E)$, the **spectrum** of T is the set of those $\lambda \in \mathbf{C}$ for which $(\lambda \mathrm{Id} - T)$ is not an isomorphism.
To illustrate these concepts, consider the following examples:

A. The identity mapping from $C([0,1])$ (with the supremum norm) into the same space with the norm $\|x\|_1 = \int_0^1 |x(t)|\, dt$ is a bijection but not an isomorphism.

B. The operator M_a on $C([0,1])$ is injective if and only if the set of zeroes of a has empty interior. It is surjective if and only if a has no zeroes. In the latter case it is an isomorphism. The spectrum of M_a is the range of a. A generalisation of the concept of a linear mapping which is often useful is that of a multi-linear mapping whereby $T : \prod_{k=1}^n E_k \to F$ (the spaces being vector spaces) is **multilinear** if for each i, the partial mapping

$$x \mapsto T(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$$

is linear for any choice of $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$.

The space of multilinear mappings is denoted by $\mathcal{L}(\mathcal{E}_\infty, \dots, \mathcal{E}_\backslash; \mathcal{F})$. It has a natural linear structure. If the E's and F are normed spaces then the following conditions on such a maapping T are easily seen to be equivalent:

T is continuous as a mapping from $\prod_{k=1}^n E_k$ with the product topology into F;

T is bounded i.e. there is a $C > 0$ so that

$$\|T(x_1, \dots, x_N)\| \leq C\|x_1\| \dots \|x_n\|$$

for each $x_1, \dots, x_n$.

This is proved exactly as for linear mappings.
The space of such multilinear mappings is denoted by $L(E_1, \dots, E_n; F)$. It is a linear subspace of $\mathcal{L}(\mathcal{E}_\infty, \dots, \mathcal{E}_\backslash; \mathcal{F})$ and the mapping

$$T \mapsto \sup\{\|T(x_1, \dots, x_n)\| : x_i \in B_{E_i}\|$$

is a norm thereon. The case $F = \mathbf{R}$ (resp. $(\mathbf{C})$ is particularly important and then we write $\mathcal{L}(\mathcal{E}_\infty\infty, \dots, \mathcal{E}_\backslash)$ and $L(E_1, \dots, E_n)$ for the corresponding spaces.
In applications we usually have the situation where all of the E_i are equal to a given space E. In this case we write $L^n(E; F)$ resp. $L^n(E)$ for the corresponding space.
Typical examples of such multilinear mappings are tensors (in the finite dimensional case) or mappings of the form

$$(x_1, \dots, x_n) \mapsto \int \dots \int x_1(s_1) \dots x_n(s_n)\, ds_1 \dots ds_n$$

on $C([0,1])$ where K is a suitable kernel e.g. a continuous function on $[0,1]^n$.
In the theory of differentiation for funtions between Banach space, we shall encounter "nested" spaces of linear operators such as $L(E, L(E, F))$, $L(E, L(E, L(E, F)))$ etc. Fortunately, these can be more conveniently represented by spaces of multilinear mappings as the following proposition shows:

Proposition 1.30 *The mapping*

$$T \mapsto (x_1 \mapsto ((x_2, \dots, x_n) \mapsto T(x_1, \dots, x_n)$$

is a linear isometric isomoprhism from $L(E_1, \dots, E_n; F)$ *onto* $L(E_1, L(E_2, \dots, E_n; F))$.

PROOF. We prove this for the case $n = 2$. First note that the mapping $S : ((x_1, x_2) \mapsto (Sx_1)(x)$ is an inverse for the one in the statement of the theorem. That both are isometries follows from the equality:

$$\begin{aligned} \|T\| &= \sup\{\|T(x_1, x_2)\| : \|x_2\| \leq 1, \|x_2\| \leq 1\} \\ &= \sup_{\|x_1\| \leq 1} \sup\{\|T(x_1, x_2)\| : \|x_2\| \leq 1\}. \end{aligned}$$

■

If we apply this result repeatedly, we see that the nested space $L(E_1, L(E_2, \dots, L(E_n, F) \dots)$ is isometrically isomorphic to $L(E_1, \dots, E_n; F)$, in particular to $L^n(E; F)$ in the case where all of the E_i coincide with E.
We continue this section with some remarks on finite dimensional spaces. These have played an increasingly important role in the theory of normed spaces in recent years as building blocks for infinite dimensional ones. They are also very helpful in providing a geometrical intuition which is useful even in the infinite dimensional case.
The first result shows that, as far as the topological structutre is concerned, all finite dimensional spaces (of the same dimension) are the same. We emphasise, however, that the isometric resp. geometric properties are very distinct.

Proposition 1.31 *Every real, finite dimensional normed space E is isomorphic to $(\mathbf{R}^n, \| \|_\infty)$ where $n = \dim E$. Similarly, every n-dimensional normed space over $\mathbf{C}$ is isomorphic to $(\mathbf{C}^n, \| \|_\infty)$.*

PROOF. Let $(x_1, \dots, x_n)$ be a basis for E and consider the continuous linear map

$$T : (\lambda_1, \dots, \lambda_n) \mapsto \sum_{i=1}^{n} \lambda_i x_i$$

from $\mathbf{R}^n \to E$.
The image of the unit sphere of $\mathbf{R}^n$ under T is a compact subset of E which does not contain 0 (since the x_i are linearly independent). Hence there is a $\delta > 0$ so that $\|T(\lambda)\| \geq \delta$ if $\lambda \in \mathbf{R}^n$, $\|\lambda\|_\infty = 1$ (since a continuous function on a compact set attains its infimum). From this it follows that $\|T^{-1}\| \leq 1/\delta$. The complex case can be proved similarly.

■

Corollary 1.32 *Any two norms on a finite dimensional normed space are equivalent.*

We shall now develop the generalisation of the concept of Banach spaces which is relevant as a framework for the topological structure of spaces of test functions and distributions. Typically the former are spaces of functions which are submitted to an infinite number of conditions, usually of growth and regularity. The appropriate concept is that of a locally convex topology i.e. one which is defined by a family of seminorms rather than by a single norm.
Recall that a **seminorm** on a vector space E is a mapping p from E into the non-negative reals so that

$$p(x + y) \leq p(x) + p(y) \quad p(\lambda x) = |\lambda| p(x)$$

for each x, y in E and λ in $\mathbf{R}$.
We shall use letters such as p, q to denote seminorms. The family of all seminorms on E is ordered in the natural way i.e. $p \leq q$ if $p(x) \leq q(x)$ for each x in E. If p is a seminorm,

$$U_p = \{x \in E : p(x) \leq 1\}$$

is the **closed unit ball** of p. This is an absolutely convex, absorbing subset of E, whereby a subset A of E is **convex** if for each $x, y \in A$ and $t \in [0,1], tx + (1-t)y \in A$ **balanced** if $\lambda A \subset A$ for λ in $\mathbf{R}$ with $|\lambda| \leq 1$ **absolutely convex** if it is convex and balanced; **absorbing** if for each x in E there is a positive ρ so that $\lambda x \in A$ for each λ with $|\lambda| < \rho$. U_p is in addition algebraically closed i.e. its intersection with each one-dimensional subspace of E is closed, whereby these subspaces carry the natural topologies as copies of the line. On the other hand, if$U$ is an absolutely convex, absorbing subset of E, then its **Minkowski functional** i.e. the mapping

$$p_U : x \mapsto \inf\{\lambda > 0 : x \in \lambda U\}$$

is a seminorm on E. In fact, the mapping $p \mapsto U_p$ is a one-one correspondence between the set of seminorms on E and the set of absolutely convex, absorbing, algebraically closed subsets of E. Also $p \leq q$ if and only if $U_p \subset U_q$.
A family S of seminorms on E is **irreducible** if the following conditions are verified:

a) if $p \in S$ and $\lambda \geq 0$, then $\lambda p \in S$;

b) if $p \in S$ and q is a seminorm on E with $q \leq p$, then $q \in S$;

c) if p_1 and p_2 are in S, then so is $\max(p_1, p_2)$;

d) S separates E i.e. if x is a non-zero element of E, then there is a $p \in S$ with $p(x) \neq 0$.

If S is a family of seminorms which satisfies only d), then there is a smallest irreducible family of seminorms which contains S. It is called the **irreducible hull** of S and denoted by $\tilde{S}$. ($\tilde{S}$ is the intersection of all irreducible families containing S – alternatively it consists of those seminorms which are majorised by one of the form

$$\max(\lambda_1 p_1, \ldots, \lambda_n p_n)$$

where the λ_i's are positive scalars and the p_i's are in S).
A **locally convex space** is a pair (E, S) where E is a vector space and S is an irreducible family of seminorms on the former. If S is a family of seminorms which separates E, then the space $(E, \tilde{S})$ is **the locally convex space generated by** S.
If (E, S) is a locally convex space, we define a topology τ_S on E as follows: a set U is said to be a **neighbourhood** of a in E if $(U - a)$ contains the unit ball of some seminorm in S. The corresponding topology is called the **topology associated with** S. This topology is Hausdorff (since S separates E) and, in fact, completely regular, since it is generated by the uniformity which is defined by the semimetrics $(d_p : p \in S)$ where $d_p(x, y) = p(x - y)$. Hence we can talk of convergence of sequences or nets, continuity and completeness in the context of locally convex spaces.
Recall that a **topological vector space** is a vector space E together with a Hausdorff topology so that the operations of addition and multiplication by scalars are continuous. It is easy to see that a locally convex space is a topological vector space. On the other hand, locally convex spaces are often defined as topological vector spaces in which the set of absolutely convex neighbourhoods of zero forms a neighbourhood basis. This is equivalent to our definition. For if (E, S) is a locally convex space as defined above, then (E, τ_S) is a topological vector space and the family $\{U_p : p \in S\}$ is a basis of convex neighbourhoods of zero. On the other hand, if (E, τ) is a topological vector space satisfying the convexity condition, then the set of all τ-continuous seminorms on E is irreducible and the corresponding topology τ_S coincides with τ.
We remark that if a family S of seminorms on the vector space E satisfies conditions a) - c) above, but not necessarily d) and we define

$$N_S = \{x \in E : p(x) = 0, p \in S\}$$

then we can define a natural locally convex structure on the quotient space E/N_S. This is a convenient method for dealing with non-Hausdorff spaces which sometimes arise.
Examples:

I. Of course, normed spaces are examples of locally convex space, where we use the single norm to generate a locally convex structure.

II. If E is a normed space and F is a separating subspace of its dual E', then the latter induces a family S of seminorms, namely those of the form $p_x : x \mapsto |f(x)|$ for $f \in F$. This induces a locally convex structure on E which we denote by $S_w(F)$. The corresponding topology $\sigma(E, F)$ is called **the weak topology induced by F.** The important cases are where $F = E'$, respectively where E is the dual G' of a normed space and F is G (regarded as a subspace of $E' = G''$).

III. (the fine locally convex structure): If E is a vector space, then the set of **all** seminorms on E defines a locally convex structure on E which we call the **fine structure** for obvious reasons.

IV. The space of continuous functions: If S is a completely regular space, we denote by $\mathcal{K}(\mathcal{S})$ or simply by $\mathcal{K}$, the family of all compact subsets of S. If $K \in \mathcal{K}$, then

$$p_K(x) = \sup\{|x(t)| : t \in K\}$$

is a seminorm on $C(S)$, the space of continuous functions from S into $\mathbf{R}$. The family of all such seminorms defines a locally convex structure $S_{\mathcal{K}}$ on $C(S)$ – the corresponding topology is that of compact convergence i.e. uniform convergence on the compacta of S.

V. Differentiable functions. If k is a positive integer, $C^k(\mathbf{R})$ denotes the family of all k-times continuously differentiable functions on $\mathbf{R}$. For each $r \leq k$ and K in $\mathcal{K}(\mathbf{R})$, the mapping

$$p_K^r : x \mapsto \sup\{|x^{(r)}(t)| : t \in K\}$$

is a seminorm on $C^k(\mathbf{R})$. The family of all such seminorms defines a locally convex structure on $C^k(\mathbf{R})$.

VI. Spaces of operators: Let H be a Hilbert space. On the operator space $L(H)$, we consider the following seminorms:

$$p_x : T \mapsto \|Tx\|$$

$$p_x^* : T \mapsto \|T^*x\|$$

$$p_{x,y} : T \mapsto |(Tx|y)|$$

for x and y in H. The family of all seminorms of the first type define the **strong locally convex structure** on $L(H)$, while those of the first two type define the **strong *-structure**. Finally, those of the third type define the **weak operator structure**.

VI. Dual pairs. We have seen that the duality between a normed space and its dual can be used to define weak topologies on E and E'. For our purposes, a more symmetrical framework for such duality is desirable. Hence we consider two vector spaces E and F, together with a bilinear form $(x, y) \mapsto \langle x, y\rangle$ from $E \times F$ into $\mathbf{R}$, which is **separating** i.e. such that

if $y \in F$ is such that $\langle x, y\rangle = 0$ for each x in E, then $y = 0$;;

if $x \in E$ is such that $\langle x, y\rangle = 0$ for each y in F, then $x = 0$.

Then we can regard F as a subspace of E^*, the algebraic dual of E, by associating to each y in F the linear functional

$$x \mapsto \langle x, y \rangle.$$

Similarly, E can be regarded as a subspace of F^*. (E, F) is then said to be a **dual pair**. The typical example is that of a normed space, together with its dual or, more generally, a subspace of its dual which separates E. For each $y \in F$, the mapping $p_y : x \mapsto |\langle x, y \rangle|$ is a seminorm on E and the family of all such seminorms generates a locally convex structure which we denote by $S_w(F)$ – the **weak structure generated by** F.

A subset B of F is said to be **bounded** for the duality if for each x in E,

$$\sup\{|\langle x, y \rangle| : y \in B\} < \infty.$$

In this case, the mapping

$$p_B : x \mapsto \sup\{p_y(x) : y \in B\}$$

is a seminorm on E. Let $\mathcal{B}$ denote a family of bounded subsets of F whose union is the whole of F. Then the family $\{p_B : B \in \mathcal{B}\}$ generates a locally convex structure $S_{\mathcal{B}}$ on E, that of **uniform convergence** on the subsets of $\mathcal{B}$.

Thus if $\mathcal{B}$ consists of the singletons of F, we rediscover the weak structure. If $\mathcal{B}$ is taken to be the family of those absolutely convex subsets of F which are compact for the topology defined by $S_w(E)$ on F, then $S_{\mathcal{B}}$ is called the **mackey structure** and the corresponding topology (which is denoted by $\tau(E, F)$) is called the **Mackey topology**. Finally, if we take for $\mathcal{B}$ the family of all bounded subsets of F, then we have the **strong structure**–the corresponding topology is called the **strong topology**.

A rich source of dual pairs is provided by the so-called **sequence spaces**. These are, by definition, subspaces of the space $\omega = \mathbf{R}^{\mathbf{N}}$ i.e. the family of all real-valued sequences which contain ϕ, the spaces of those sequences with finite support (i.e. $\phi = \{x \in \omega : \xi_n = 0$ except for finitely many $n\}$.

ϕ and ω are regarded as locally convex spaces, ω with the structure defined by the seminorms $p_n : x \mapsto |\xi_n|$ and ϕ with the fine structure.

Further examples of sequence spaces are the ℓ^p-spaces.

If E is a sequence space, we define its α**-dual** E^α as follows:

$$E^\alpha = \{y = (\eta_n) : \sum |\xi_n \eta_n| < \infty \text{ for each } x \in E\}.$$

Then (E, E^α) is a dual pair under the bilinear form

$$\langle x, y \rangle = \sum_{n=1}^{\infty} \xi_n \eta_n.$$

Of course, if $E \subset F$, then $E^\alpha \supset F^\alpha$. Also E is clearly a subspace of $(E^\alpha)^\alpha)$. A sequence space E is **perfect** if $E = E^{\alpha\alpha}$. Thus

$$\omega_\alpha = \phi \quad \phi^\alpha = \omega \quad (\ell^p)^\alpha = \ell^q$$

the latter for *all* values of p and q. Hence all of these spaces are perfect.

In particular, ℓ^2 is self-dual i.e. equal to its own α-dual. In fact, this is a characterisation of ℓ^2 as the reader can verify. More precisely, if a sequence space E is such that $E = E^\alpha$, then $E = \ell^2$.

The β**-dual** of E is the family of those sequences $y = (\eta_n)$ for which $\sum \xi_n \eta_n$ converges for each $x \in E$. This coincides with the α-dual when E is solid i.e. such that whenever $x \in E$ and z is dominated by x i.e. $|\zeta_n| \leq |\xi_n|$ for each n, then $z \in E$.

As in the case of normed spaces, we are interested in linear mappings between locally convex spaces which preserve their topological structures. Corresponding to the fact that boundedness and continuity are equivalent for linear mappings between normed spaces, we have the following equivalences: for a linear mapping $T : E \to F$ whereby (E, S) and (F, S_1) are locally convex spaces, the following are equivalent:

T is $\tau_S - \tau_{S_1}$-continuous;

T is $\tau_S - \tau_{S_1}$-continuous at zero;

for each $p \in S_1$, $p \circ T \in S$;

for each $p \in S_1$, there are finite sequences $q_1, \dots, q_n$ in S and $\lambda_1, \dots, \lambda_n$ of positive numbers so that

$$p \circ T \leq \lambda_1 q_1 + \cdots + \lambda_n q_n.$$

We remark that the last characterisation is also valid (i.e. equivalent to the continuity of T) in the case were S and S_1 are merely separating families of seminorms which generate the corresponding locally convex structures.
The most important examples of such mappings are differential operators i.e. mappings of the form

$$L : x \mapsto \sum_{i=0}^{n} a_i x^{(i)}$$

where a_0, …,a_n are smooth functions, say on $\mathbf{R}$. This operator can be regarded as a continuous linear mapping from $C^k(\mathbf{R})$ into $C^{k-n}(\mathbf{R})$.
Metrisable and Fréchet spaces: These are spaces which have representations $E = \varprojlim E_n$ of a spectrum of Banach spaces which is indexed by $\mathbf{N}$. More precisely, a space with this property is called a **Fréchet space**. A general (i.e. non-complete) space is **metrisable** if its completion is a Fréchet space. Less pedantically, they are those locally convex spaces whose structures are generated by countably many semi-norms. Of course, the name comes from the fact that this definition is equivalent to the fact that τ_S is metrisable. For suppose that this condition holds. Then 0 has a countable basis of neighbourhoods (which we can suppose to be absolutely convex) and their Minkowski functionals generate the locally convex structure. On the other hand, if E is metrisable in the above sense, then $\hat{E}$ is representable as the limit $\varprojlim E_n$ of a countable spectrum of Banach spaces and hence is homeomorphic to a subspace of the product $\prod E_N$ (as is E itself). hence it suffices to show that the latter is metrisable as a topological space. But the metric

$$d(x, y) = \sum \frac{1}{2^n} \frac{\|x_n - y_n\|}{1 + \|x_n - y_n\|}$$

where $x = (x_n)$ and $y = (y_n)$.
The closed graph theorem and its usual variants are also valid for Fréchet spaces. The reason is that the structure of a Fréchet space can be defined by a so-called **paranorm** and these are sufficiently similar to norms to allow us to carry over the proofs of these results from the case of banach space with only small changes. In fact, the results hold for an even wider class of classes which we now introduce: **Definition:** A **metric linear space** is a vector space E, provided with a metric d which is translation invariant (that is, satisfies the condition

$$d(x + x_0, y + x_0) = d(x, y) \quad (x, y, x_0 \in E)$$

and is such that the mapping $(\lambda, x) \mapsto \lambda x$ from $\mathbf{R} \times E$ into E is continuous for the topology induced by d. A **paranorm** on a linear space is a mapping p from E into $\mathbf{R}^+$ so that

$p(x) = 0$ if and only if $x = 0$;

$p(x + y) \leq p(x) + p(y)$;

$p(\lambda_n x) \to 0$ for every x in E and every null sequence of scalars.

if p is a paranorm on a space, then the mapping $d_p : (x, y) \mapsto p(x - y)$ is a translation invariant metric on E and (E, d_p) is a metric linear space. On the other hand, if (E, d) is such a space, then $pp : x \mapsto d(x, 0)$ is a paranorm. Thus the notions of a metric linear space and a space with paranorm are equivalent.

If the linear space E with paranorm is such that the metric space (E, d_p) is complete, it is called an F-space. Of course, every Banach space and indeed every Fréchet space is an F-space. An example of an F-space which is not a Fréchet space is $S(\mu)$, the set of equivalence classes of measurable functions on a measure space (Ω, μ). The mapping

$$x \mapsto \int \frac{|x|}{1 + |x|} \, d\mu$$

is a paranorm on $S(\mu)$. The reader can check that $S(\mu)$ is complete under this paranorm and that the corresponding notion of convergence is convergence in measure. The canonical example is provided by the case of lebesgue measure. In this case, if U is a neighbourhood of zero, then the absolutely convex hull of U is the whole space. This easily implies that the only continuous linear form on $S(\mu)$ is the zero form (since the set $\{|f| \leq 1\}$ is a neighbourhood of zero). Thus $S(\mu)$ cannot be locally convex (and so is not a Fréchet space).

The usual constructions shows that if (E, p) is a paranormed space, then so is each subspace and each quotient by a closed subspace. Also a countable product of paranormed spaces is paranormed (but not a non-trivial direct sum or an uncountable product). The same remark holds for F-spaces (where we only consider closed subspaces of course).

Our claim is that suitable versions of the classical theorems of Banach hold for paranormed spaces. We shall simply state these results without proof–those for Banach spaces can be carried over with only slight changes involving the substitution of norms by paranorms. We begin with the Banach-Steinhaus theorem. Here we use the term **bounded** to indicate a subset B of a paranormed space (F, p) for which $\sup\{p(x) : x \in B\} < \infty$.

Proposition 1.33 *Let E be an F-space and F an F-space or a locally convex space. Then if M is a family of continuous linear mappings from E into F which is bounded on the points of a set A of second category in E, M is equicontinuous. Hence if a sequence (T_n) of continuous linear mappings from E into F is such that the pointwise limit exists, then the latter is continuous.*

The open mapping theorem holds in the following form:

Proposition 1.34 *Let E and F be F-spaces, T a continuous linear mapping from E into F whose range $T(E)$ is of second category in F. Then T is open and surjective.*

As usual, version of the closed graph theorem and the isomorphism theorem can immediately be deduced from this result.

The following result about bounded sets resp. convergence sequences in metrisable, locally convex spaces is often useful:

Proposition 1.35 *Let (x_n) be a null-sequence resp. (B_n) a sequence of bounded sets in a metrisable locally convex space E. Then there exists*

a sequence (λ_n) of positive scalars which tends to infinity and is such that $\lambda_n x_n \to 0$; a sequence (λ_n) of positive scalars so that $\bigcup_n \lambda_n B_n$ is bounded.

Proof. We prove (1). The proof of (2) is similar. We choose an increasing sequence (p_n) of seminorms which generate the structure of E. For each k in $\mathbf{N}$ there is an n_k so that $p_k(x_n) \leq \frac{1}{k}$ if $n \geq n_k$. We can also suppose that $n_{k+1} \geq n_k$ for each k. Define the sequence (λ_n) as follows: $\lambda_n = \sqrt{k}$ where k is that positive integer for which $n_k \leq n < n_{k+1}$. Clearly this sequence increases to infinity and $\lambda_n x_n \to o$ since $p_k(\lambda_n x_n) \leq \frac{1}{\sqrt{k}}$ if $n \geq n_k$.