Contents

Chapter 1

Point-set topology

This chapter summaries basic, point-set topology. It is roughly equivalent to Chapter 2 of Munkres, without the metric space material.

§1.1 The location trilemma

Fix a topological space X, a set A an a point x. In terms of set theory, a point has two *locations* with respect to A – either $x \in A$ or $x \notin A$. In topology, we care about a little more. Although any endpoint of a closed interval $[a, b]$ certainly is *in* the interval, we would not say it is *inside* that interval. The difference becomes even starker when looking at sets like $\mathbb{N} \subseteq \mathbb{R}$ or $\mathbb{Q} \subseteq \mathbb{R}$.

The right way – or at least *a right way* – is to say that a point x is *inside* a set A, if apart from $x \in A$, a whole neighbourhood $U \ni x$ is contained in A. Since this readily implies $x \in A$, we may dispense with that requirement in any definition. Metaphorically, this means that x *cannot escape* A or *everything* x *sees is inside of* A.

The notion that $x \notin A$ has a very similar topological analogue. We will say that x *is outside of* A if not only $x \notin A$, but a whole neighbourhood $U \ni x$ is disjoint from A. Another word for that is that x *is separated from* A. This also provides the metaphorical meaning.

With these two notions lifted up from mere set-based combinatoris to topology, we have made a bit of a problem. It could be the case that for a point x and set A neither of the above properties hold. Pictorially speaking, x is then *not far from* A, but also not really *inside of* A – a kind of in-between state. In this case we call x a *boundary point of* A.

Topological properties of operators (such as Int, Cl, Bd etc.) can often be deduced from this simple observation, which we state as a lemma. This was first taught to me by Krzysztof Omiljanowski. I am not aware of any name for the fact, so I made up my own.

Lemma 1. *(The location trilemma.) Let X be a topological space, $A \subseteq X$ and $x \in X$ (not neccesarily $x \in A$). Then exactly one of the following is true:*

1. *For some neighbourhood U, $x \in U \subseteq A$. Then x is called an* interior point*.*
2. *For some neighbourhood U, $x \in U \subseteq A^c$. Then x is called an* exterior point*. We also say that x is* separated from *A by the neighbourhood U.*
3. *For all neighbourhoods U, $U \cap A \neq \varnothing$ and $U \cap A^c \neq \varnothing$. Then x is called a* boundary point*.*

Proof. Clause (1) implies that $x \in A$ and (2) implies that $x \in A^c$, so they cannot hold at once. Moreover both of them imply that (3) is false. We will show that if both (1) and (2) do not hold, then (3) does.

Pick any neighbourhood $U \ni x$. Since (1) does not hold, $U \nsubseteq A$, so $U \cap A^c \neq \varnothing$. Analogously, $U \cap A \neq \varnothing$. □

The names we gave to the properties of points with respect to A are not random – they correspond exactly to the usual definitions of Int, Bd, Ext.

§1.2 The closure axiomatization

As one should know, a topology on X might as well be defined in terms of which sets are closed rather than which sets are open. In practice, another way of looking at closed sets might also pop up. In many scenarios, there is an operator Cl, which we might call a *closure* operator, of the signature

$$\mathrm{Cl} : \mathcal{P}(X) \to \mathcal{P}(X),$$

which adjoins to a set A some elements that are in a given sense *reachable*, *deducible*, *obtainable* etc. from A.

An example of such an operator would be, for any given topology $\mathcal{T}$ on X, the closure operator of that topology. One might wonder if from an operator one might recover a topology. If we are going to do that, there are a few questions we need to answer.

How do we recover open sets? We can get open sets as complements of closed sets. Then the question is how do we recover closed sets. The key property to use is that a set C is closed iff $\mathrm{Cl}\, C = C$.

Definition 1. *An operator*

$$\mathbf{c} : \mathcal{P}(X) \to \mathcal{P}(X)$$

is said to satisfy the Kuratowski closure axioms if it satisfies to following

TODO: add this to Anki, make a flashcard, whatever

(K1) it preserves $\varnothing$, i.e. $\mathbf{c}(\varnothing) = \varnothing$;

(K2) it is extensive*, i.e. $A \subseteq \mathbf{c}(A)$ for all A;*

(K3) it is idempotent*, i.e. $\mathbf{c}(\mathbf{c}(A)) = \mathbf{c}(A)$ for all A;*

(K4) it distributes over finite sums, i.e. $\mathbf{c}(A \cup B) = \mathbf{c}(A) \cup \mathbf{c}(B)$ for all A, B.

Now we prove that this actually defines a topology.

Lemma 2. *Let $\mathbf{c}$ be an operator*

$$\mathbf{c} : \mathcal{P}(X) \to \mathcal{P}(X)$$

satisfying the Kuratowski closure axioms. Then, the collection of its fixpoints, i.e.

$$co\mathcal{T} = \{A \,|\, \mathbf{c}(A) = A\}$$

defines a topology as its family of closed sets.

Proof. From the axiom (K1) we get that $\varnothing \in co\mathcal{T}_{\mathbf{c}}$. We also need $X \in co\mathcal{T}$, which follows from (K2), as no set in X can be larger than all of X. Suppose now that A, B are closed in the sens above. Then we have

$$\mathbf{c}(A \cup B) = \mathbf{c}(A) \cup \mathbf{c}(B) = A \cup B,$$

so $A \cup B$ is closed too. The intersection property is the tricky part.

Take $C_i \in co\mathcal{T}$. Then

$$\bigcap_i C_i \subseteq C_j$$

for all j, so

$$\forall j.\, \mathbf{c}\left(\bigcap_i C_i\right) \subseteq \mathbf{c}(C_j) = C_j,$$

so

$$\mathbf{c}\left(\bigcap_i C_i\right) \subseteq \bigcap_j C_j.$$

The other inclusion follows by extensivity of $\mathbf{c}$, so we actually have the equality

$$\mathbf{c}\left(\bigcap_i C_i\right) = \bigcap_j C_j.$$

□

You may have noticed that (K1), (K2) and (K4) were used in the proof, but not (K3). Then an operator satisying all of the Kuratowski axioms except for (K3) defines a topology via closed sets. On the other hand, the closure operator of that topology definitely has property (K3), so these two are different operators!

Luckily, we can recover the closure operator using lattice theory. First, a definition and a lemma.

Definition 2. *An operator*

$$\mathbf{c} : \mathcal{P}(X) \to \mathcal{P}(X)$$

is called a Čech closure operator *if it satisfies the Kuratowski axioms (K1), (K2) and (K4).*

Lemma 3. *Any operator satisfying (K4) is* monotonic*, i.e.*

$$A \subseteq B \Rightarrow \mathbf{c}(A) \subseteq \mathbf{c}(B).$$

for all $A, B \subseteq X$.

Proof. TODO! □

Now that we've seen these, we'll want to recover the actual A. If you think of the operator $\mathbf{c}$ as *enriching* a set, a closed set is one that *is completely enriched*. Then, how do we get $\operatorname{Cl} A$? Well, we start by enriching it

$$A \to \operatorname{Cl} A,$$

but this may not be enough, so we enrich the result and get

$$A \to \operatorname{Cl} A \to \operatorname{Cl}(\operatorname{Cl} A),$$

but this may not be enough, so we enrich the result

$$A \to \operatorname{Cl} A \to \operatorname{Cl}(\operatorname{Cl} A) \to \operatorname{Cl}(\operatorname{Cl}(\operatorname{Cl} A)),$$

and... we're stuck in an infinite loop. In such cases, it helps to take a peek past infinity and consider

$$\bigcup_{k=0}^{n} \operatorname{Cl}^k A.$$

This still does not work.

Example 1. *You can define a Cech closure operator which fails this. Let*

$$X = \mathbb{N} \cup \{\infty\}$$

and

$$\mathbf{c}(A) = \begin{cases} \varnothing & \textit{when } A = \varnothing \\ X & \textit{when } \infty \in A \\ \{0, 1, \dots, \sup A + 1\} & \textit{when } A \neq \varnothing \textit{ and } \infty \notin A \end{cases}$$

You can check that this is a Cech closure operator, but the only closed sets in $\mathbf{co}\mathcal{T}_{\mathbf{c}}$ *are* $\varnothing$ *and* X, *i.e. this is the indiscrete topology on* X.

In reality you need a transfinite number of iterations. However, for a Kuratowski operator this chain stabilizes almost immediately, so we're saved!

Lemma 4. *For a Kuratowski closure operator* $\mathbf{c}$*, the closure operator of its generated topology* $\mathcal{T}_{\mathbf{c}}$ *is precisely* $\mathbf{c}$.

Proof. Let $C \supseteq A$ be the topological closure of A. Then

$$\mathbf{c}(A) \subseteq \mathbf{c}(C) = C.$$

Since $\mathbf{c}$ is idempotent by (K3), $\mathbf{c}(A)$ is closed in $\mathcal{T}_{\mathbf{c}}$. Since a topological closure is the smallest closed set containing A, we have that

$$C \subseteq \mathbf{c}(A) \Rightarrow C = \mathbf{c}(A)$$

□

Connection with lattice theory. Note that we want to find the least (because topological closure is the smallest closed set) fixpoint (because of the definition of topology) of $\mathbf{c}$ greater than A. This is exactly the setting of the Kleene fixpoint theorem, and what we're doing here is using that.

§1.3 Literature used for this chapter

1. Munkres, chapter 2
2. Wikipedia, Kuratowski Closure Axioms
3. Wikipedia, Preclosure operator

Chapter 2

Abelian groups

We will now explore some features of abelian groups. These will later be recontextualized into features of modules (for those of you know, an abelian group is essentialy the same as a $\mathbb{Z}$-module). In this chapter, we will derive a structure theorem.

§2.1 Basic properties

§2.2 Torsion

Central to the study of abelian groups is the notion of torsion. We begin with two definitions, one for elements and one for groups.

Definition 3. *An element of an abelian group is called a* ***torsion element*** *if it is of finite order. An element of infinite order is called* ***torsion-free****.*

Definition 4. *An abelian group is called a* ***torsion group*** *if all its elements are torsion. An abelian gropu is called* ***torsion-free*** *if all its* nonzero *elements are torsion-free.*

One might ask why we define this only for abelian groups. The following crucial fact is the reason – more precisely, the fact that its proof relies on commutativity.

Theorem 1. *If α, β are two torsion elements then their sum (and difference) is also a torsion element. Thus, the torsion elements form a subgroup.*

Proof. Let α have order n and β have order m. Then we have

$$nm(\alpha + \beta) = nm\alpha + nm\beta = m(n\alpha) + n(m\beta) = m \cdot 0 + n \cdot 0 = 0.$$

The subgroup claim is proved by noting that the same computation hold as well for difference and that the identity element is of course torsion. □

Since we have a well defined kind of subgroup, we might as well give it a name.

Definition 5. *Let G be an abelian group. Its* ***torsion subgroup****, denoted*

$$\mathrm{T}(G)\mathrm{or}Tor(G)$$

is the subgroup consisting of all the torsion elements of G.

We may hope that such a nice notion of substructure is also valid for the torsion-free elements of a group. This fails, however, as the set of torsionfree elements is the complement of a subgroup, which will very often fail to be a subgroup. On the level of elements, we have the following fact.

Theorem 2. *Let $a \in G$ be torsionfree and $b \in Tor(G)$. Then $a + b$ is torsionfree.*

Proof. We give two proofs. For the first one, suppose $a + b$ is torsion. Then

$$a = (a + b) - b$$

would be torsion (since the torsion elements form a subgroup).

For the second proof, suppose $a + b$ is torsion with order n and that b has order m. Then

$$0 = mn(a + b) = mna + mnb = mna,$$

so a is torsion as well. □

All is not lost though! We cannot define a notion of *torsion-free subgroup*, but there is a simple way of killing the torsion – quotienting!

2.2.1 A weird example of groups

2.2.2 Functors

2.2.3 Tensors

§2.3 Structure theorem

§2.4 A smudge of infinite abelian group theory: the prufer groups

Prufer groups are: divisible, p-torsion, subgroups, infinitely generated

§2.5 Sources for this chapter

1. Ludomir Newelski, Algebra II
2. Ludomir Newelski, Algebra 2R
3. Wikipedia: torsion subgroup, torsion, torsion-free group, Prufer groups

Chapter 3

The Künneth formula

This chapter accomplishes a lofty goal: to calculate the homology groups of a product topological space.

§3.1 Setup

We will first try and discover how this all works for CW-homologies, and then use what we've found to generalize the statements to singular homology. The reason why CW-complexes are good for the job is that given CW-structures on X and Y, we can easily form a CW-structure on $X \times Y$. One should expect to form a simplicial structure on a product of simplicial complexes, but that would require subdividing a product of n- and m-simplices into $(n+m)$-simplices, which, if anything, is unpleasant.

We have a product structure on $X \times Y$ given by products of cells and products of characteristic maps. Let us now try to calculate the cellular boundaries.

One of the cells is larger in dimension: then we get a coefficient of zero, because the map turns out to be constant.

If the above does not hold, then we have to reduce exactly one dimension by one. However, if in the dimension we keep we use a different cell of that dimension, the same problem applies.

Generally, we have the following:

Lemma 5. *Suppose we have a continuous map*

$$f : I^k \to I^{k-1},$$

which is equal to a product map

$$f = id_I \times g$$

for some continuous

$$g : I^{k-1} \to I^{k-2}.$$

Then, the degree of

$$\widetilde{f} : \operatorname{Bd} I^k \to I^{k-1} / \operatorname{Bd} I^{k-1}$$

is equal to the degree of

$$\widetilde{g} : \operatorname{Bd} I^{k-1} \to I^{k-2} / \operatorname{Bd} I^{k-2}.$$

Remark. Note that since this is not the same sphere, the degree only makes sense *modulo orientation*.

§3.2 Orientations

Remember that when describing a CW-complex structure, there is an additional degree of freedom – the orientation of each cell. This orientation needs to be taken into account.

For the purposes of this chapter, it is very important to see how an orientation of the interior can be turned into an orientation of the boundary.

One way of doing this would be to just take the *vector to the right and up* in an Euclidean space. However, this leads to an inconsistent choice of orientation on the boundary. A correct way is to pick an inward facing normal at each point on the boundary and then pick an orientation on the boundary which

This can formally be described in homology by noting that the boundary homomorphism is an isomorphism. The determined choice of generator for I^k then projects down to $\operatorname{Bd} I^k$.

§3.3 Literature used for this chapter

1. Allen Hatcher, Algebraic Topology
2. MathOverflow, biography of Hermann Kunneth: https://mathoverflow.net/questions/114215/who-was-hermann-k
3. https://sili-math.github.io/AT2020/Lecture-22.pdf

Chapter 4

Local properties in algebra

§4.1 Localization of modules

§4.2 Local rings

This section is devoted to a class of rings very important in algebraic geometry – local rings. Although the definition is purely algebraic, it corresponds to notions of *looking around a point* on a (affine) variety.

4.2.1 Characterisations of local rings

All material in this section applies to (unital) rings. They need not be commutative for the results to work. All the proofs in this section are my own work.

Definition 6 (local ring)**.** *A ring R having a unique maximal ideal $\mathfrak{m}$ is called a* **local ring** *and denoted $(R, \mathfrak{m})$.*

It turns out there is another characterisation of such rings, which is of particular usefulness in working with the second point of interest of this chapter – Discrete Valuation Rigs, whose definition we postpone for now.

Theorem 3 (local ring characterisation)**.** *A ring R is local iff the set of noninvertible elements forms an ideal. This is the the unique maximal ideal.*

Proof. The first implication is trivial, since an ideal containing a unit is not proper.

Suppose now R is local with maximal ideal $\mathfrak{m}$. Since $\mathfrak{m}$ is proper, it does not contain any invertible element. Take a noninvertible $r \in R$. The principal ideal (r) can be extended (via Zorn's lemma) to a maximal ideal, which must be the unique maximal ideal $\mathfrak{m}$, so $r \in \mathfrak{m}$.

□

There is also an element-wise characterisation of local rings. To see it emerge, let us try a different line of attack for the previous proof.

We know that $\mathfrak{m}$ contains only noninvertible elements, but we still have to prove that it contains all such elements. By way of contradiction, suppose $r \notin \mathfrak{m}$. Now by maximality we have

$$\mathfrak{m} + (r) = R$$

and in particular

$$m + ar = 1$$

for some $a \in R, m \in \mathfrak{m}$. We would like to derive a contradiction, so we would hope that the identity above lets us conclude that r is invertible. The property below is what we need

Theorem 4 (local ring addition)**.** *A ring R is local iff it has the following property: for all a, b such that*

$$a + b \in R^{\star}$$

at least one of a, b is invertible.

Proof. The line of attack above shows how this implies R being local. Once we have this property, we can take any maximal ideal. If there was some noninvertible element it did not contain, we would be able to extend it.

On the other hand, take a local ring. We will show it is impossible to sum two noninvertibles to a unit. We have that

$$(a) + (b) \subseteq \mathfrak{m} + \mathfrak{m} = \mathfrak{m},$$

which contains no units.

□

It should be noted that the property expressed in the previous theorem might be rephrased in two ways: first, take an arbitrary finite sum instead of two elements and second, have the sum be equal to 1 instead of invertible. It should be easy to see that all such characterisations are equivalent.

4.2.2 Examples and generic constructions

4.2.3 Properties of local rings

Krull's intersection, Nakayama's, Kaplansky's theorem

§4.3 Valuations on Fields and Discrete Valuation Rings

§4.4 Literature used for this chapter

1. Piotr Kowalski, Algebraic Curves. Chapter 2.3.
2. nlab page for local rings
3. https://scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=2933&context=td (for Kaplansky's theorem)
4. Wikipedia pages

Chapter 5

Algebraic Geometry, problemset 5

Problem 4

Since v is a smooth point of C, the ring $\mathcal{O}_v$ is a DVR. A local parameter is defined as any uniformizing parameter of that DVR and having a zero of order 1 means being of valuation 1 in that DVR.

Thus is suffices to prove the following: in a DVR, an element is of valuation 1 iff it is irreducible.

Since a DVR is a local ring, let us denote its maximal ideal by $\mathfrak{m} = (r)$. Note that r must be irreducible, otherwise the ideal $\mathfrak{m}$ would not be maximal. The valuation on this DVR can be described as the usual r-adic valuation or belonging to a power of the maximal ideal (see solution of **Problem 7.**).

Let f be a local parameter, i.e. irreducible. Then, by **Remark 2.37.**

$$(f) = \mathfrak{m} = (r),$$

so f and r are associates, therefore $f = ur$ for some unit u. We have that $u \in \mathcal{O}_v \setminus \mathfrak{m}$, so $v(u) = 0$ and

$$v(f) = v(u) + v(r) = 0 + 1 = 1.$$

On the other hand, let f be of valuation 1. Then, as the valuation is the r-adic valuation we have that

$$f = ra$$

for some $r \nmid a$, so $a \in \mathcal{O}_v \setminus \mathfrak{m}$. This implies a is a unit, so f and r are associated, so f is irreducible since r is.

Problem 5

Problem 5a

Consider a tangent line $L = V(\alpha X + \beta Y + \gamma)$.

If L is tangent, the intersection number $I(0, L \cap C) > 1 > 0$, so $0 \in L$, so $\gamma = 0$. By the same reasoning, F has zero constant term.

It is easy to see that a line is a smooth curve. Therefore, the curve C is tangent to L iff F is of valuation at least two in $\mathcal{O}_0$ (which is a DVR). By **Problem 7** that is equivalent to F being a member of $\mathfrak{m}_v^2$ (where F is reinterpreted as $F + I(L)$ in $K(L)$).

Algebraically, the square of the maximal ideal corresponds to

$$I_L^2(0)\mathcal{O}_0 = \left\{ \frac{G}{H} : G \in I_L^2(0), H \notin I_L(0) \right\}.$$

Thus, F being of valuation at least 2 is equivalent to it being of the form

$$F = \frac{1}{N} \cdot \sum_i G_i H_i$$

as a rational function in $K(L)$ for some $N \in K[L] \setminus I_L(0)$ and $G_i, H_i \in I_L(0)$. This is again equivalent to the identity

$$FN = \sum_i G_i H_i$$

in $K[L]$, which in turn is equivalent to F having the form

$$FN = \sum_i G_i H_i + P(\alpha X + \beta Y)$$

for some polynomials such that $N(0) \neq 0$ and $G_i(0) = H_i(0) = 0$.

After this introduction, we will show

$$T_0 C = V\left(\partial_X F(0) X + \partial_Y F(0) Y\right)$$

via two inclusions, first from left to right.

If both partials are 0, the vanishing set is the whole space $\mathbb{A}^2$, so this inclusion is trivial. Suppose now at least one partial is nonzero. Differentiating both sides of the above identity w.r.t. X we get that

$$\partial_X F \cdot N + F \cdot \partial_X N = \left(\sum_i \partial_X G_i \cdot H_i + G_i \cdot \partial_X H_i \right) + \partial_X P \cdot (\alpha X + \beta Y) + \alpha P.$$

Evaluating both sides at 0 and remembering which polynomials vanish at 0 we obtain

$$N(0)\partial_X F(0) = P(0)\alpha.$$

We can now repeat this for differentiation w.r.t. Y. This implies that the vectors

$$[\partial_X F(0), \partial_Y F(0)]^T, [\alpha, \beta]^T \in \mathbb{K}^2$$

are linearly dependent, so they describe the same line.

Now for the other inclusion. We will lean on the fact that the first partials evaluated at zero are the coefficients of the monomials X, Y in the polynomial F. Suppose first both partials of F are zero. Then we can write F as

$$F = X^2 G(X,Y) + XYH(X,Y) + Y^2 P(X,Y) + 0 \cdot (\alpha X + \beta Y)$$

which is the form we needed. This gives that *any* line is tangent to C at 0, so the tangent space is the whole plane (as is $V(0 \cdot X + 0 \cdot Y)$).

If, on the other hand, some partial is nonzero, we can write F as

$$F = 1 \cdot (\partial_X F(0) \cdot X + \partial_Y F(0) Y) + X^2 G(X,Y) + XYH(X,Y) + Y^2 P(X,Y).$$

This proves that C is indeed tangent to $V(\partial_X F(0) \cdot X + \partial_Y F(0) Y)$.

Problem 5b

From the previous subproblem and the lecture we know that both T_0C and $I_C(0)/I_C(0)^2$ are finite-dimensional K-vector spaces. This allows us to use a theorem of linear algebra which states that a bilinear map such as is given in the problem induces an isomorphism iff for all $P \neq 0$ the function

$$x \mapsto \Phi(x, P)$$

is nonzero (i.e. is nonzero for some x) or, equivalently, that if this function is zero then so is P. Here P should be understood as the representative regular function in $K[C]$ or a polynomial which represents that function (so a representative of representatives).

Suppose then that this function is zero. First consider the case that both partials of F are zero. Then the tangent space is the whole plane and we have that for all $x, y \in K$

$$\partial_X P(0)x + \partial_Y P(0)y = 0,$$

which gives that both partials of P vanish at zero, so we have the form

$$P = X^2P_1 + XYP_2 + Y^2P_3$$

for P as a polynomial, which in particular implies

$$P \in I_C(0)^2$$

as a regular function, so

$$P = 0$$

in $I_C(0)/I_C(0)^2$.

Now consider what happens when F has a nonzero partial. This means that

$$\partial_X P(0)x + \partial_Y P(0)y = 0$$

for all points such that

$$\partial_X F(0)x + \partial_Y F(0)y = 0,$$

which gives that either the gradient of P is zero (in which case we can repeat the reasoning from the previous case) or that the gradients of P and F are linearly dependent. Then there exists a scalar α such that

$$P - \alpha F$$

has zero first partials, so

$$P = \alpha F + X^2P_1 + XYP_2 + Y^2P_3$$

as a polynomial, so

$$P = X^2P_1 + XYP_2 + Y^2P_3 + I(C)$$

as a regular function. As we have done many times, we now conclude that

$$P \in I_C(0)^2,$$

which is what we needed.

Problem 6

Problem 6a

Let

$$a = r^n \frac{\alpha}{\beta}, b = r^m \frac{\gamma}{\delta},$$

with $r \nmid \alpha, \beta, \gamma, \delta$ and wlog $n \geqslant m$. Then

$$a + b = r^m \frac{r^{n-m}\alpha\delta + \beta\gamma}{\beta\delta}.$$

Note that r does not divide the numerator (since r is irreducible and thus prime in a UFD). If $n > m$, it will also not divide the denominator, but if $n = m$ it might. In the first case the valuation is exactly m, while in the second case it might change – but only increase.

Problem 6b

We have

$$ab = r^{n+m} \frac{\alpha\gamma}{\beta\delta}.$$

By virtue of r being prime, we have $r \nmid \alpha\gamma\beta\delta$, so

$$v_r(ab) = n + m.$$

Problem 6c

For every $n \in \mathbb{Z}$ we have

$$v_r(r^n) = n,$$

so the valuation is indeed surjective.

Problem 7

We claim that

$$\mathfrak{m}^n = \{x : v_R(x) \geqslant n\},$$

from which the problem follows immediately. For $n = 0$ the claim follows from the nonnegativity of the valuation.

Let $r \in R$ be the uniformizing parameter. Then we have that

$$\mathfrak{m} = (r)$$

by **remark 2.37.**, so an element of $a \in \mathfrak{m}^n$ is of the form

$$a = \sum_i a_i r^n$$

for some $a_i \in R$, so

$$v_R(a) \geqslant \min(v_r(a_1 r^n), v_r(a_2 r^n), \ldots, v_r(a_j r^n)) \geqslant \min(n, n, \ldots n) = n.$$

This gives us one inclusion of the claim. For the other one, take an element x of valuation no less than n. The definition of valuation implies that for some $m \geqslant n$

$$x = r^m y = (r^{m-n} y) r^n \in \mathfrak{m}^n.$$

This completes the solution.

Problem 8

Take $x, y \in \mathcal{O}_v$. Then we have

$$v(xy) = v(x) + v(y) \geqslant v(y) \geqslant 0$$

and

$$v(x+y) \geqslant \min(v(x), v(y)) \geqslant 0.$$

This implies that $\mathcal{O}_v$ is a subring of L. If we take x, y such that both valuations are positive, the sum has a positive valuation, and if at least y has a positive valuation then the product does as well. This implies that $\mathfrak{m}_v$ is an ideal.

Note that for any valuation we have

$$v(1) = v(1 \cdot 1) = v(1) + v(1),$$

so

$$v(1) = 0.$$

Now take any $x \in \mathcal{O}_v \setminus \mathfrak{m}_v$. This means that $v(x) = 0$. Let y be the multiplicative inverse (in L!) of x. Then

$$0 = v(1) = v(xy) = v(x) + v(y),$$

so $v(y) = 0$ and $y \in \mathcal{O}_v$. If $v(x) > 0$, then $v(y) < 0$ and $y \notin \mathcal{O}_v$. We have just proved that the set of noninvertible elements of $\mathcal{O}_v$ is an ideal. Therefore $(\mathcal{O}_v, \mathfrak{m}_v)$ is a local ring.

Since the valuation is surjective, $\mathfrak{m}_v$ is nonempty and $\mathcal{O}_v$ has nonzero noninvertible elements, so it is not a field. To finish the proof that $\mathcal{O}_v$ is a DVR all we need to do is show that $\mathcal{O}_v$ is a PID.

To do that, take r to be any element of valuation 1. Such an element exists by surjectivity of the valuation.

Claim. Let $t \in \mathcal{O}_v$ and $v(t) = n \geqslant 0$. Then $t = ur^n$ for some unit $u \in \mathcal{O}_v$.

Proof. Let

$$\alpha = \frac{t}{r^n} \in L.$$

Then α has valuation 0, so it actually is an invertible element of $\mathcal{O}_v$.

Now take an ideal I. Note that if I has any element of valuation n, then by the claim above it contains all elements of valuation n as they all are associated with r^n. It will also contain an element (so all elements) of higher valuations by virtue of being closed under multiplication by r.

This lets us conclude that

$$I = (r^n)$$

where n is the smallest valuation achievable by an element of I.

Chapter 6

Approximation properties of smooth functions

Smoothness is a really good property that a function can have. However, a lot of the functions we need aren't smooth, i.e. ReLU, $\max(0, \min(1, x))$, $\chi_{[a,b]}$ and so on. In this chapter, we will show how to effectively approximate such functions with very good smooth lookalikes.

§6.1 Characteristic functions on intervals

Let us make our way towards approximating the interval characteristic function. In the special case of the real line, the story begins with quite a wonderful function – the magnificent

$$e^{-1/x^2}.$$

Taken literally, it's not quite defined at 0. However, as $x \to 0$, $-1/x^2 \to -\infty$, so we can define the value at 0 to be 0. Notice that the convergence of the argument to $-\infty$ is really fast – fast enough the we can retain smoothness at 0. Let us state the properties formally.

Lemma 6. *The function*

$$g(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

is smooth and satisfies the following properties:

1. $g(0) = 0$,
2. $g(x) \leq 1$,
3. $\lim_{x\to 0} g(x)/P(x) = 0$ *for any* $P \in \mathbb{R}[X]$,
4. $g^{(n)}(0) = 0$ *for all* n.

Proof. Properties (1) and (2) should be obvious. For (3) we perform the change of variables $u := 1/x^2$ to obtain

$$\lim_{x\to 0} \frac{g(x)}{P(x)} = \lim_{u\to\infty} \frac{e^{-u}}{P(\pm 1/\sqrt{u})} = \lim_{u\to\infty} \frac{u^{\deg P/2} e^{-u}}{Q(\pm\sqrt{u})} = 0,$$

where $Q \in \mathbb{R}[X]$ (actually, it's P with the coefficients written backwards). We use (3) to derive an (4). By induction, we have that at all $x \neq 0$

$$g^{(n)} = \frac{g(x)P_n(x)}{Q_n(x)}$$

for some $P, Q \in \mathbb{R}[X]$ and the only root of Q is 0. Indeed,

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{g(x)P_n(x)}{Q_n(x)} = \frac{-\frac{1}{x^2}Q_n(x)P_n(x)g(x) + P_n'(x)Q_n(x)g(x) - P_nQ_n'g(x)}{Q_n^2(x)} = g(x)\frac{P_{n+1}(x)}{Q_{n+1}(x)},$$

where we have expanded the fraction by x^2 to get the desired form. This gives that

$$Q_n = x^{a_n},$$

where

$$a_0 = 0$$
$$a_{n+1} = 2a_n + 2,$$

from which we can derive

$$a_n = 2(2^n - 1).$$

□

The magic of this function is that, since it has all the derivatives at 0 equal to 0, we might as well say it's zero on all negative numbers *and still get a smooth function*! This is very much not the case for more primitive attempts to do this – look at the case of x or more generally x^k and see that it only gives as a function of class C^{k-1}.

This maneover let us discard half of the real line a smooth fashion, and this means we're halfways towards approximating $\chi_{[0,1]}$. The idea will be to use functions like this:

$$g(x)g(1-x).$$

This has a small bump, but since g is very flat near 0, the product of two such functions will not be very large. The solution solution to this problem is to widen the interval for the moment, so that values closer to 1 get multiplied together in the formula above. Let us define

$$h_n = g(x)g(n-x).$$

This should look better, but now

$$\operatorname{supp} h_n = [0, n]$$

where we would much prefer

$$\operatorname{supp} h_n = [0, 1].$$

No worries – we can see that the *shape* is right, so squeezing the function will do.

Lemma 7. *The function sequence*

$$f_n(x) = h_n(nx) = g(nx)g(n - nx).$$

converges to $\chi_{[0,1]}$ *pointwise a.e. (outside of* $\{0, 1\}$*) and in* L^1.

Proof. Let us quantify the idea that f_n *is almost* 1 *on almost all of* $[0, 1]$. In terms of h_n, these functions should be almost 1 on almost all of $[0, n]$. We will pick an α (varying with n), for which we will be able to establish a good bound. Suppose

$$\alpha n \leq x \leq (1 - \alpha)n.$$

This gives

$$\begin{aligned}
\frac{1}{\alpha^2 n^2} &\geq \frac{1}{x^2} \\
\frac{2}{\alpha^2 n^2} &\geq \frac{1}{x^2} + \frac{1}{(n-x)^2} \\
-\frac{2}{\alpha^2 n^2} &\leq -\frac{1}{x^2} - \frac{1}{(n-x)^2} \\
e^{-\frac{2}{\alpha^2 n^2}} &\leq h_n(x)
\end{aligned}$$

We now want to pick a sequence α such that both

$$\lim_n \alpha = 0$$

and

$$\lim_n \frac{2}{\alpha^2 n^2} = 0.$$

Of course, $\alpha = 1/\sqrt{n}$ does the trick nicely. We can get better convergence estimates for a different choice. □

Another kind of useful function we might like to have are piecewise linear functions. To effectively approximate those, we need just see how to approximate ReLU and ... – all piecewise linear functions are sums of shifted/scaled combinations of these two.

§6.2 Higher dimensions

§6.3 Partitions of unity

§6.4 Approximative identities on $\mathbb{R}$

Now we will try to develop a more general way of approximating functions.

§6.5 Banach algebras

§6.6 Literature used for this chapter

Appendix A

Some values of Tor and Ext

$\mathrm{Tor}_{\mathbb{Z}}$	$\mathbb{Z}$	$\mathbb{Z}_m$	G
$\mathbb{Z}$	0	0	0
$\mathbb{Z}_n$	0	$\mathbb{Z}_{\gcd(m,n)}$	$\ker(G \xrightarrow{\cdot n} G)$