(19.1.1) Transfinite induction

Suppose that P is a property of ordinal numbers. Assume

- *$P(0)$ holds;*
- *if $P(n)$ holds, then $P(n+1)$ holds;*
- *if n is a limit ordinal and $P(m)$ holds for all $m < n$, then $P(n)$ holds.*

Then $P(n)$ holds for all ordinal numbers n.

Transfinite induction can be used in constructions as well as proofs, just as the more usual induction in Chapter 2.

Ordinal numbers capture the notion of succession. But they don't measure the size of a set. *Hilbert's hotel*[6] illustrates this. Consider a hotel with ω rooms (numbered 0, 1, 2, ...). One day, when all the rooms are full, a new guest arrives. To accommodate him, the manager simply moves each guest into the next room along, freeing room 0 for the newcomer. Next day, infinitely many new guests arrive. Undetered, the manager shifts the guest from room n into room $2n$ for each n, freeing the odd-numbered rooms for the new arrivals.

As we saw already, two sets *have the same cardinality* if there is a bijection between them.[7] Hilbert's hotel shows that there is a bijection between ω and $\omega+1$, and also between ω and $\omega+\omega$. So the ordinal numbers are too discriminating. There are two ways to proceed:

We may decide that, having defined what it is for sets to have the same cardinality, we have implicitly defined the cardinality of a set. Roughly speaking, cardinalities are equivalence classes for the relation 'same cardinality'; but care is required, since the equivalence classes are not sets (by the same reasoning as in Russell's paradox; singletons, for example, continue appearing at all stages).

An alternative approach depends on the fact:

Any non-empty set of ordinal numbers has a least element.

(This is proved by transfinite induction in the same way that the same assertion for the natural numbers is proved by induction — see Chapter 2.) Now, given any set X, the set of all those ordinal numbers which are bijective with X has a least element (if there are any such numbers!), and we take this least element to be the cardinality of X. In other words, a *cardinal number* is an ordinal number which is not in one-to-one correspondence with any smaller ordinal number. With this approach, all natural numbers, and ω, are cardinal numbers, but $\omega+1$ and $\omega+\omega$ are not.

[6] As described in Stanislaw Lem's story 'The Interstellar Milkman, Ion the Quiet'. See N. Ya. Vilenkin, *Stories about Sets* (1965).

[7] Thus, a set is *countable* if and only if it is bijective with $\mathbb{N}$.

In this approach, a set is finite if and only if it is bijective with some natural number. This is precisely how natural numbers are used in ordinary counting (as 'standard sets' of each possible size); our approach generalises this to the transfinite.

An alternative notation for cardinal numbers is due to Cantor, the 'aleph notation'. ($\aleph$ (aleph) is the first letter of the Hebrew alphabet.) Using transfinite induction, we define $\aleph_n$ for all *ordinal* numbers n by the rules:

- $\aleph_0 = \omega$;
- $\aleph_{n+1}$ is the next cardinal number after $\aleph_n$;
- if a is a limit ordinal and $\aleph_b$ is defined for all $b < a$, then $\aleph_a$ is the least cardinal number exceeding all these.

Life is simplified by the following theorem of Zermelo:

(19.1.2) Well-ordering Theorem. *The Axiom of Choice is equivalent to the assertion that every set admits a one-to-one function onto some ordinal number.*

Thus, if we assume that our set theory satisfies the Axiom of Choice (as is almost universally done), then every set has a unique cardinal number.

Cardinal numbers, being special ordinal numbers, are totally ordered. We have $a \le b$ if there is a one-to-one function from a set of cardinality a into one of cardinality b. It follows from (19.1.2) that, assuming the Axiom of Choice, given any two sets, there is a one-to-one function from one to the other (in some order!)

We can do arithmetic with cardinal numbers. If A and B are disjoint sets with cardinalities a and b respectively, then $a+b$, $a \cdot b$ and a^b are the cardinalities of $A \cup B$, $A \times B$ (Cartesian product), and A^B (the set of functions from B to A) respectively. (Representing subsets of B by their characteristic functions, we see that 2^b is the cardinality of the power set of B.) But the rules are a bit different. The next result assumes the Axiom of Choice.

(19.1.3) Proposition. *(a) If a and b are infinite, then $a+b = a \cdot b = \max(a,b)$.*
(b) If $a > 1$, then $a^b > b$ for all b.

In particular, $2^a > a$ for all a. It is known that 2^ω is the cardinality of the set of real numbers. Cantor's *continuum hypothesis* is the assertion that 2^ω is the smallest uncountable cardinal number; in other words, no subset of $\mathbb{R}$ has cardinality strictly between those of $\mathbb{N}$ and $\mathbb{R}$. (In aleph notation, $2^{\aleph_0} = \aleph_1$.) More generally, the *generalised continuum hypothesis* (GCH) asserts that, for any cardinal number b, 2^b is the next cardinal after b (or $2^{\aleph_n} = \aleph_{n+1}$ for all ordinals n). It is known that the GCH is undecidable; it holds in Gödel's constructible universe, but there are models in which it is false.

19.2. König's Infinity Lemma

The Axiom of Choice (which I will abbreviate to AC) is fundamental to most infinite combinatorics, and I will assume that it holds. However, some students, learning about it for the first time, overrate its influence, and worry that it is being invoked in an argument along the lines, 'The set X is non-empty, so choose an element $x \in X$ ...'. This does not require AC; one choice, or indeed finitely many choices, are permitted by the other axioms. Similarly, AC is not required if there is a rule for making the choices.[8] Only when infinitely many genuine free choices must be made is AC required.

Often, we have a situation where later choices depend on earlier ones. The so-called 'Principle of Dependent Choice' allows us to make such choices; it is a consequence of AC. (AC allows us to choose an element from any set which could conceivably arise in the process.) In this form, it is invoked in proving a result which is very useful in applying AC to combinatorics: *König's Infinity Lemma.*

A ***one-way infinite path*** in a digraph D is a sequence $v_0, v_1, v_2, \dots$ of distinct vertices such that (v_i, v_{i+1}) is an edge for all $i \geq 0$.

(19.2.1) König's Infinity Lemma. *Let v_0 be a vertex of a digraph D. Suppose that*
(a) every vertex has finite out-valency;
(b) for every positive integer n, there is a path of length n beginning at v_0.
Then there is a one-way infinite path beginning at v_0.

Remark. The result is false if condition (a) is relaxed. Take a path of length n for every finite n, all starting at the same point, but otherwise disjoint.

Proof. We call a vertex v of D *good* if, for every n, a path of length n starts at v. We claim:

If v is good, then there exists v' such that (v, v') is an edge and v' is a good vertex of $D - v$.

For let w_n be the next point after v on a path of length n starting at v. Since there are only finitely many vertices x for which (v, x) is an edge, one of them (say v') must occur infinitely often as w_n. This means that there are arbitrarily long finite paths starting at v', and hence paths of all finite lengths starting there, none of which contain v.

Now, by assumption, v_0 is good. For each $i \geq 0$, choose v_{i+1} so that (v_i, v_{i+1}) is an edge and v_{i+1} is a good vertex of $D - \{v_0, \dots, v_i\}$. Then $v_0, v_1, v_2, \dots$ is the required one-way infinite path.

Another infinite principle was invoked in the proof of the Claim above: an infinite form of the *Pigeonhole Principle* (cf. (10.1.1)).

(19.2.2) Pigeonhole Principle (infinite form)
If infinitely many objects are divided into finitely many classes, then some class contains infinitely many objects.

The infinite form of Ramsey's Theorem is a generalisation of this; see Section 19.4.

Now we give an application of König's Infinity Lemma, showing how it can be used to transfer information between the finite and the (countably) infinite.

[8] Bertrand Russell's example: If a drawer contains infinitely many pairs of shoes and we must choose one shoe from each pair, we can take all the left shoes. But, for infinitely many pairs of socks, AC is required.

(19.2.3) Proposition. *Let Γ be a countably infinite graph. Suppose that any finite induced subgraph of Γ has a vertex colouring with r colours. Then Γ has a vertex colouring with r colours.*

Proof. Let $v_1, v_2, \dots$ be the vertices of Γ. For each n, let C_n be the (non-empty) set of all vertex colourings of the induced subgraph on $\{v_1, \dots, v_n\}$ with the r colours $1, \dots, r$. Form a digraph D with $\bigcup_{n \geq 0} C_n$ as vertex set (we take C_0 to be a singleton whose only member c_0 is the empty set!) and with edges as follows: for $c_n \in C_n$ and $c_{n+1} \in C_{n+1}$, let (c_n, c_{n+1}) be an edge if and only if c_n is the restriction of the colouring c_{n+1} to the vertices $v_1, \dots, v_n$. Then each vertex has out-valency at most r (since at most r colours can be applied to v_{n+1} if $v_1, \dots, v_n$ are already coloured). Moreover, $d(c_0, c_n) = n$ for all $c_n \in C_n$.

So the hypotheses of König's Infinity Lemma are satisfied. We conclude that there is a one-way infinite path $c_0, c_1, c_2, \dots$. This gives us a rule for colouring all the vertices of Γ; for v_n is assigned a colour in c_n, and by definition it gets the same colour in all c_m for $m > n$. Moreover, it is a legitimate vertex colouring; for, if $\{v_i, v_j\}$ is an edge, then v_i and v_j are assigned different colours in c_n, where $n = \max(i, j)$.

(19.2.4) Corollary. *Any plane map, finite or infinite, can be coloured with four colours.*

Proof. For finite maps, of course, this is the Four-colour Theorem (18.6.9). A plane map has at most countably many countries, since each country contains a point with rational coordinates, and there are only a countable number of such points. So the infinite case follows from (19.2.3).

In fact, (19.2.3) holds for arbitrary infinite graphs, not just countably infinite ones. To prove this, we need a stronger principle, *Zorn's Lemma*, to be described in the next section.

19.3. Posets and Zorn's Lemma

One of the most striking differences between finite and infinite posets is that the latter need not have maximal elements, as shown by the natural numbers (for example). An important theorem giving conditions under which maximal elements exist is *Zorn's Lemma*:

(19.3.1) Zorn's Lemma. *Let P be a non-empty poset. Suppose that every chain in P has an upper bound. Then P has a maximal element.*

Proof. Recall how we showed that every finite poset has a maximal element: if not, pick an element, and repeatedly pick a larger element, yielding an infinite ascending chain. The same trick works here. Suppose that $P = (X, \leq)$ has no maximal element. By transfinite induction, define elements x_a, for all ordinal numbers a, such that $x_a < x_b$ for $a < b$. This is done as follows:

- Let x_0 be any element of x.

- If x_a is defined, let x_{a+1} be any strictly greater element (this exists since x_a is not maximal).
- If a is a limit ordinal, then the elements x_b for $b < a$ form a chain; let x_a be an upper bound for this chain.

Obviously, all the elements x_a are distinct. But this leads to a contradiction: take a to be a cardinal number greater than the cardinality of X, and there are not enough elements available in X for such a chain!

Note that we used the Axiom of Choice in this proof: we have to choose each term of the series from a set of 'admissible' elements. This is in fact inevitable: Zorn's Lemma is 'equivalent to' AC; the latter can be proved from the former and the other axioms of set theory. (See Exercise 3.)

Here is a fairly typical application of Zorn's Lemma, to an infinite version of (12.2.1):

(19.3.2) Theorem. *Any poset has a linear extension.*

Proof. Let (X, R) be a poset. We let $\mathcal{R}$ be the set of relations $R' \supseteq R$ for which (X, R') is a poset, partially ordered by inclusion. We claim:

Every chain in $(\mathcal{R}, \subseteq)$ has an upper bound.

For let C be a chain, and let R' be the union of the members of C (each member of C being a relation on X, that is, a set of ordered pairs). Then (X, R') is a partial order. (This involves checking the axioms. The arguments are all similar: here is the proof of transitivity. Suppose that $(x,y),(y,z) \in R'$. Then, say, $(x,y) \in R_1$ and $(y,z) \in R_2$ for some $R_1, R_2 \in C$. Since C is a chain, one of these relations contains the other; say $R_1 \subseteq R_2$. Then $(x,y),(y,z) \in R_2$; so $(x,z) \in R_2$ (because (X, R_2) is a poset), and $(x,z) \in R'$, as required.) Clearly $R' \supseteq R$, and R' is thus an upper bound for C in $\mathcal{R}$.

By Zorn's Lemma, there is a maximal element of R, say R'. We show that (X, R') is a total order. If it were not, then there would be some pair (a, b) of points which are incomparable in (X, R'). Now exactly the same argument as in the proof of (12.2.1) shows that we could enlarge R' to make a and b comparable, by setting $R' = R' \cup ({\downarrow}\, a \times {\uparrow}\, b)$. But this would contradict the maximality of R'.

So (X, R') is a linear extension of (X, R), as required.

Zorn's Lemma is often conveniently applied in the form of the *Propositional Compactness Theorem*, which we now develop with an application.

An *ideal* in a lattice is a non-empty down-set which is closed under taking joins. Equivalently, I is an ideal in L if

- $0 \in I$;
- $x, y \in I \Rightarrow x \vee y \in I$;
- $x \in I, a \in L \Rightarrow x \wedge a \in I$.

An ideal is *proper* if it is not the whole of L; equivalently, if it does not contain 1.

(19.3.3) Proposition. *Any lattice contains a maximal proper ideal.*

Proof. Straightforward application of Zorn's Lemma to the set of proper ideals, partially ordered by inclusion. (If no ideal in a chain contains 1, then the union doesn't contain 1 either.)

Slightly more generally, *any proper ideal I in a lattice is contained in a maximal ideal.* This is proved by modifying the argument to use only the set of proper ideals containing I.

Our application depends on the following observation: in a Boolean lattice L, if I is a maximal ideal, then for each $a \in L$, exactly one of a and a' belongs to I. (They cannot both belong, since their join is 1. If neither lies in I, then the set

$$J = \{y : y \leq a \vee x \text{ for some } x \in I\}$$

is an ideal containing I and a but not a', contradicting maximality.)

Recall the definition of propositional formulae and valuations from Section 12.4. One small piece of terminology: A set Σ of propositional formulae is *satisfiable* if there is a valuation v such that $v(\phi) = \text{TRUE}$ for all $\phi \in \Sigma$.

(19.3.4) Propositional Compactness Theorem. *Let Σ be a set of propositional formulae. Suppose that every finite subset of Σ is satisfiable. Then Σ is satisfiable.*

Proof. We work in the Boolean lattice L of equivalence classes of formulae, and identify a formula with its equivalence class. Let I be the ideal *generated by* $\Sigma' = \{(\neg\phi) : \phi \in \Sigma\}$: that is, I is the set of elements of L which lie below some finite disjunction of elements of Σ'. The hypothesis implies that $1 \notin I$. For, if $1 \in I$, then 1 would be a (finite) disjunction of elements of Σ'. By assumption, there is a valuation giving all these elements the value FALSE; but then 1 would have the value FALSE, which is impossible.

By the extension of (19.3.3), there is a maximal ideal I^* containing I. Now define a valuation v^* by

$$v^*(\phi) = \begin{cases} \text{TRUE} & \text{if } \phi \notin I^*, \\ \text{FALSE} & \text{if } \phi \in I^*. \end{cases}$$

Check that v really is a valuation; clearly $v(\Sigma) = \text{TRUE}$.

The Propositional Compactness Theorem is a more powerful tool than König's Infinity Lemma, allowing arguments to be extended to arbitrary infinite cardinality, as we'll see shortly. It is in fact less powerful than the Axiom of Choice: there are models of set theory in which AC fails but Propositional Compactness is true.

As an application, we extend (19.2.3) to arbitrary infinite graphs.

(19.3.5) Proposition. *Suppose that every finite subgraph of Γ has a vertex colouring with r colours. Then Γ has a vertex colouring with r colours.*

Proof. We take the set

$$\{p_{x,i} : x \text{ a vertex of } \Gamma,\ i = 1, \ldots, r\}$$

of propositional variables. Let Σ be the set of formulae of the following types:

- for each vertex x of Γ, a formula asserting that $p_{x,i}$ is true for exactly one value of i;
- for each edge $\{x, y\}$ of Γ, a formula asserting that $p_{x,i}$ and $p_{y,i}$ are not true for the same value of i.

For example, if $r = 3$, these formulae would be

$$(p_{x,1} \vee p_{x,2} \vee p_{x,3}) \wedge (\neg(p_{x,1} \wedge p_{x,2}) \wedge \neg(p_{x,1} \wedge p_{x,3}) \wedge \neg(p_{x,2} \wedge p_{x,3}))$$

and

$$\neg(p_{x,1} \wedge p_{y,1}) \wedge \neg(p_{x,2} \wedge p_{y,2}) \wedge \neg(p_{x,3} \wedge p_{y,3})$$

respectively.

This set of formulae is satisfiable if and only if a vertex colouring with r colours exists. For, if v is a valuation making Σ true, then give vertex x the colour i if $v(p_{x,i}) =$ TRUE; and conversely.

By assumption, any finite subset of Σ is satisfiable. For a finite subset involves the variables $p_{x,i}$ for only finitely many vertices x; these form a finite subgraph which can be coloured with r colours; use this colouring to define $v(p_{x,i})$ for vertices x in the subgraph, and define the other values arbitrarily.

So the Propositional Compactness Theorem gives the desired result.

19.4. Ramsey theory

The infinite form of Ramsey's Theorem can be stated as follows.

> **(19.4.1) Ramsey's Theorem (infinite form)**
> *Suppose that k and r are positive integers, and let X be an infinite set. Suppose that the set of k-element subsets of X are partitioned into r classes. Then there is an infinite subset Y of X, all of whose k-element subsets belong to the same class.*

For example, the case $k = 1$ is the infinite form of the Pigeonhole Principle (19.2.2). I will give a proof for $k = 2$; the general case is an exercise (with hints — Exercise 4).

We may suppose that X is countable, say $X = \{x_1, x_2, \ldots\}$. (Simply choose a sequence of distinct elements of X and use these.) Now we define a subsequence $y_1, y_2, \ldots$ of distinct elements, and a sequence of infinite subsets $Y_0, Y_1, Y_2, \ldots$ such that

(a) $Y_1 \supseteq Y_2 \supseteq \ldots$;
(b) $y_i \notin Y_i$, and all pairs $\{y_i, z\}$ for $z \in Y_i$ have the same colour;
(c) $y_j \in Y_i$ for all $j > i$.

This construction is done by induction, starting with $Y_0 = X$. In the i^{th} step, choose $y_i \in Y_{i-1}$; observe that the infinitely many pairs $\{y_i, x\}$, for $x \in Y_{i-1} \setminus \{y_i\}$, fall into r disjoint 'colour classes'; so there is an infinite subset Y_i of $Y_{i-1} \setminus \{y_i\}$ for which (b) holds, by the Pigeonhole Principle.

At the conclusion of the inductive argument, we have arranged that the colour of a pair $\{y_i, y_j\}$ for $j > i$ depends only on i, not on j. Let c_i be this colour. By the Pigeonhole Principle again, there is an infinite subset M of the natural numbers such that c_i is constant for $i \in M$. Then $\{y_i : i \in M\}$ is the required infinite monochromatic set.

The finite version of Ramsey's Theorem can be deduced from the infinite, using König's Infinity Lemma. The argument is very similar to (19.2.3). We suppose that the finite version is false, for some choice of r, k, l; that is, for every positive integer n there is a colouring of the k-subsets of $\{1, \ldots, n\}$ with no monochromatic l-set. Let C_n be the set of such colourings. Form a digraph with vertex set $\bigcup_{n \geq 0} C_n$, edges from C_n to C_{n+1} being defined by restriction just as before. König's Infinity Lemma gives us an infinite path, which tells us how to colour the k-subsets of the natural numbers without creating a monochromatic l-set, contrary to the infinite Ramsey Theorem.

The remainder of this section concerns possible infinite extensions or quantifications of Ramsey's Theorem. The proofs are sketched or omitted; you should regard it as a Project.

There are three natural ways in which we could try to extend Ramsey's Theorem:
(a) quantify the two infinities in the statement (as infinite cardinals);
(b) allow infinitely many colours;
(c) colour subsets of infinite size.
These three will be considered in turn.

(a) QUANTIFYING THE INFINITIES. For simplicity, we assume that $k = r = 2$. Here are one positive and one negative result.

(19.4.2) Theorem. *(a) Let a be an infinite cardinal. If $|X| > 2^a$, and the 2-subsets of X are coloured with two colours, there must exist a monochromatic set of cardinality greater than a.*
(b) The 2-subsets of $\mathbb{R}$ can be coloured so that no uncountable set is monochromatic.

(Since $|\mathbb{R}| = 2^{\aleph_0}$, part (b) says that the result of (a) is best possible for $a = \aleph_0$. In the notation of Chapter 10, $R(2, 2, \aleph_1)$ is the next cardinal after $2^{\aleph_0}$.)

I won't prove (a) — for the proof, which is not difficult, see for example *Ramsey Theory*, by R. L. Graham *et al.* (1990) — but the construction for (b) is quite easy. It depends on the following fact. Let a family (x_a) of real numbers indexed by ordinal numbers be given, and suppose that, if $a < b$, then $x_a < x_b$. Then the family is at most countable. For there is a 'gap' between x_a and the next number in the sequence, and this gap (an interval of $\mathbb{R}$) contains a rational number q_a. All these rationals are distinct. The result follows since there are only countably many rationals.

Now, by the Axiom of Choice, there is a bijection between $\mathbb{R}$ and an ordinal number. Let x_a be the real corresponding to the ordinal a. For $a < b$, colour $\{x_a, x_b\}$ red if $x_a < x_b$, blue if $x_a > x_b$. Now, according to the last paragraph, a monochromatic red set is at most countable; the same holds for a monochromatic blue set, by reversing the order of $\mathbb{R}$ in the argument.

(b) INFINITELY MANY COLOURS. There are two different directions possible here. The first is a simple extension, illustrated by the following negative result:

(19.4.3) Theorem. *The 2-subsets of a set of size 2^a can be coloured with a colours without creating a monochromatic triangle.*

PROOF. We take our set of size 2^a to be the set of all functions from the ordinal number a to $\{0, 1\}$. Now, for each $b \in a$, we colour the pair $\{f, g\}$ with colour b if b is the smallest point at which f and g disagree. Now there cannot be three functions pairwise disagreeing at the same point!

To motivate the other approach, we have to return to the basic philosophy of Ramsey theory, as expressed in the phrase 'complete disorder is impossible'. We expect that, if an infinite set carries an arbitrary colouring, there should be an infinite subset on which the colouring is particularly simple. With only finitely many colours, 'simple' has to mean 'monochromatic'; but in general there are other possibilities, for example, all the colours may be different! This leads to so-called 'canonical' forms of the theorems, first developed by Erdős and Rado. For example:

(19.4.4) Canonical Pigeonhole Principle. *If the elements of an infinite set are coloured with arbitrarily many colours, then there is an infinite subset in which either all the colours used are the same, or all the colours are different.*

This is clear because, if the first alternative fails, then each colour appears only finitely often, so infinitely many colours must be used; and using AC we can choose one point of each colour.

Erdős and Rado proved the canonical Ramsey theorem (sometimes called the *Erdős–Rado Canonisation Theorem*. Here is the formulation for $k = 2$.