

1 Groups and Homomorphisms

1.1 Binary Operations

Definition. A **(binary) operation** on a **nonempty set** G is a **function** $\mu : G \times G \rightarrow G$.

An **operation** μ **assigns** to each **ordered pair** (a, b) of **elements** of G a **third element** of G , **namely**, $\mu(a, b)$. In **practice**, μ is **regarded** as a “**multiplication**” of **elements** of G , and, **instead** of $\mu(a, b)$, **more suggestive notions** are **used**, such as ab , $a + b$, $a \circ b$, or $a * b$. In this first **chapter**, we **shall use** the **star notation** $a * b$.

This is a little unclear. Rotman soon refers to ‘an operation $*$ on a set G ’, which makes it clear that $*$ is regarded as a specific binary operation. And he goes on to refer to groups $(G, *)$ and (H, \circ) for which ‘ $f(a * b) = f(a) \circ f(b)$ ’. This in turn suggests that ‘ $*$ ’ in ‘ $a * b$ ’ is not a fixed token like ‘ \times ’ in ‘ $a \times b$ ’, but rather refers to a particular binary operation, thought of as a mathematical object. So we want to write something like

If $*$ is a binary operation on a set G , and $a, b \in G$, then we denote $*(a, b)$ by $a * b$.

There is one other important change. We only want to license infix notation for objects which are *explicitly presented* as “binary operations”, not for any functions which might turn out to be binary operations. So rather than saying that binary operations are functions, we will say that binary operations ‘consist of’ functions; as discussed in §???, this means that binary operations can freely be used as functions, but not vice versa.

Finally, a minor point; Rotman speaks of a ‘(binary) operation’, in order to define synonyms ‘operation’ and ‘binary operation’. We need to explicitly separate these.

Applying these changes leads us to

Definition. A **binary operation** or **operation** on a **nonempty set** G consists of a **function** $\mu : G \times G \rightarrow G$.

If $*$ is a **binary operation** on a **set** G , and $a, b \in G$, then we denote $*(a, b)$ by $a * b$.

It is not normal to refer to an operation as ‘multiplication’ except in the context of a semigroup, so we will introduce that term when we come to define semigroups below.

1.2 Associativity

Definition. An operation $*$ on a set G is *associative* if

$$(a * b) * c = a * (b * c)$$

for every $a, b, c \in G$.

I'm not sure whether we need to change this to 'a nonempty set G ' to satisfy the presuppositions. (Should we have to? Opinions?) No other changes are necessary. So we get

Definition. A operation $*$ on a nonempty set G is *associative* if

$$(a * b) * c = a * (b * c)$$

for every $a, b, c \in G$.

The source goes on to discuss the way in which associativity removes the need for parentheses. There is a long, informal, discussion, but the key parts are as follows:

Associativity allows one to multiply every ordered triple of elements in G unambiguously; parentheses are unnecessary, and there is no confusion in writing $a * b * c$. If we are confronted by four elements of G , or, more generally, by a finite number of elements of G , must we postulate more intricate associativity axioms to avoid parentheses?

Definition. An expression $a_1 * a_2 * \dots * a_n$ *needs no parentheses* if, no matter what choice of multiplications of adjacent factors are made, the resulting elements of G are all equal.

Theorem 1.8 (Generalised Associativity). If $*$ is an associative operation on a set G , then every expression $a_1 * a_2 * \dots * a_n$ needs no parentheses.

Actually formalising the definition of 'needs no parentheses' into something which has clear truth-conditional content would be time-consuming. Further, the parser automatically understands that if $a * b * c$ is unambiguous (for a, b, c elements of the same group), then $a_1 * a_2 * \dots * a_n$ is also unambiguous (for a_i elements of the same group) — in fact, this is a very simple special case of the kinds of facts that the parser needs to understand about ambiguity and disambiguation. So there is no need to state more than

If $*$ is an associative operation on a set G , and a, b and c are elements of G , then " $a * b * c$ " is unambiguous.

A further note on expressions ‘ $a_1 * a_2 * \dots * a_n$ ’ could be added as a comment if desired.

1.3 Semigroups

Definition. A *semigroup* $(G, *)$ is a nonempty set G equipped with an associative operation $*$.

Usually, one says “Let G be a semigroup ...,” displaying the set G , but tacitly assuming that the operation $*$ is known. The reader must realise, however, that there are many possible operations on a set making it a semigroup.

The use of language in “Let G be a semigroup ...,” is subtler than Rotman realises. We discuss extensively this elsewhere (§...), but for the moment it suffices to note that the ‘tacit assumption’ is part of a general language-of-maths phenomenon which the compiler understands, so that the remark is not needed by the compiler. Since it may be useful to human readers, we will turn it into a comment. We were also introduced the notion of ‘multiplication’ discussed above

Definition. A *semigroup* $(G, *)$ is a nonempty set G equipped with an associative operation $*$.

[Usually, one says “Let G be a semigroup ...,” displaying the set G , but tacitly assuming that the operation $$ is known. The reader must realise, however, that there are many possible operations on a set making it a semigroup.]*

If $(G, *)$ is a semigroup, then we will call $*$ the *multiplication* in $(G, *)$.

1.4 Powers

The source continues:

Definition. Let G be a semigroup and let $a \in G$. Define $a^1 = a$ and, for $n \geq 1$, define $a^{n+1} = a * a^n$.

We have real problems with this definition. For a start, without “common-sense”, we have no way of knowing that ‘1’ in ‘ $a^1 = a$ ’ refers to a number rather than a fixed token. (Compare ‘ $\sin^{-1} x$ ’ — here it is arguably the case that ‘-’ and ‘1’ are fixed tokens and that ‘ $\sin^k x$ ’ is ambiguous if it happens to be the case that $k = -1$.) Similarly, there is no way of knowing that ‘+1’ in ‘ a^{n+1} ’

should be read as part of a number rather than a required sequence of tokens. The workaround we will use is:

Definition. Let G be a [semigroup](#) and let $a \in G$. Let $f : \mathbb{N}^+ \rightarrow G$ be the [function](#) defined by

1. $f(1) = a$ and
2. $f(n + 1) = a * f(n)$ for all $n \in \mathbb{N}^+$.

For $n \in \mathbb{N}^+$, define a^n to be $f(n)$.

(There may be room for improvement here. Suggestions?) Note, incidentally, that we have written ‘for $n \in \mathbb{N}^+$ ’ rather than ‘for $n \geq 1$ ’. The latter is clearly more idiomatic; there are several ways in which to introduce it (some easy), but we are not yet satisfied that we have found the cleanest.

The final result of the section on semigroups is:

Corollary 1.9. Let G be a [semigroup](#), let $a \in G$ and let m and n be [positive integers](#). Then $a^m * a^n = a^{m+n} = a^n * a^m$ and $(a^m)^n = a^{mn} = (a^n)^m$.

Proof. Both [sides](#) of the first (or [second](#)) [equations](#) arise from an [expression](#) having $m + n$ (or mn) [factors](#) all equal to a . But [these expressions](#) need no [parentheses](#), by [Theorem 1.8](#). ■

The statement of the result is entirely unproblematic. The proof is a perfect example of an entirely sound mathematical argument that is very difficult to translate into truth-conditional terms. We will simply replace it by:

Corollary 1.9. Let G be a [semigroup](#), let $a \in G$ and let m and n be [positive integers](#). Then $a^m * a^n = a^{m+n} = a^n * a^m$ and $(a^m)^n = a^{mn} = (a^n)^m$.

Proof. The result follows from the [Principle of Induction](#). □

(Note that ‘The Principle of Induction’ is the name of a result, just like ‘Cauchy’s Theorem’.) This reformulation does not really captured the heart of the argument — the original text refers to something much closer to Haskell [concat](#) than to induction — but pursuing that observation would take us very far from the present work.

1.5 Groups

The **most important** semigroups are groups.

Definition. A **group** is a semigroup G containing an element e such that:

- (i) $e * a = a * e$ for all $a \in G$;
- (ii) for every $a \in G$, there is an element $b \in G$ with

$$a * b = e = b * a.$$

Many authors misstep here by, for example, relying on the uniqueness of e before they prove it. Rotman is exquisitely careful, and so the only change we need to make is to refer to ‘a semigroup $(G, *)$ ’ to introduce the binary operation ‘*’. (If we had chosen to treat ‘*’ as a fixed token used in the multiplication operation, even this would not be necessary.)

[The most important semigroups are groups.]

Definition. A **group** is a semigroup $(G, *)$ containing an element e such that:

- (i) $e * a = a * e$ for all $a \in G$;
- (ii) for every $a \in G$, there is an element $b \in G$ with

$$a * b = e = b * a.$$

The next definition is:

Definition. A pair of elements a and b in a semigroup **commutes** if $a * b = b * a$. A group (or a semigroup) is **abelian** if every pair of its elements commute.

If ‘pair’ is taken to refer to a mathematical object, such as an ordered pair, then this definition only licences remarks like ‘the pair consisting of x and y commutes’, and does not allow us to simply say ‘ x and y commute’. This is clearly not what is intended, and so we are forced to take ‘pair’ as a purely linguistic, extramathematical term. Since this term is not in the fragment vocabulary, we need to rewrite the definition to avoid it. We will also remove the unsupported aside, ‘(or a semigroup)’, by relying on the fact that groups are semigroups.

Definition. Elements a and b in a semigroup *commute* if $a * b = b * a$. A semigroup G is *abelian* if whenever x and y are elements of G , x and y commute.

1.6 Uniqueness of Identity and Inverses

Theorem 1.10. If G is a group, there is a unique element e with $e * a = a = a * e$ for all $a \in G$. Moreover, for each $a \in G$, there is a unique $b \in G$ with $a * b = e = b * a$.

Proof. Suppose that $e' * a = a * e$ for all $a \in G$. In particular, if $a = e$, then $e' * e = e$. On the other hand, the defining property of e gives $e' * e = e'$, and so $e' = e$.

Suppose that $a * c = e = c * a$. Then $c = c * e = c * (a * b) = (c * a) * b = e * b = b$, as desired. ■

There are three points of logic to note here. First, ‘if $a = e$, then $e' * e = e$ ’ looks like a conditional statement but is in fact making an absolute assertion. Second, the compiler has no way of knowing that the second paragraph on the proof does not lie inside it is the scope of ‘Suppose that $e' * a = a * e$ for all $a \in G$.’ Third, e is unbound in the theorem statement and ‘ a ’, ‘ b ’, ‘ c ’, ‘ e ’ and ‘ e' ’ are all unbound at various points in the proof. Also, some extramathematical phrases which are not in the fragment vocabulary (‘moreover’, ‘in particular’, ‘defining property’, ‘gives’) need to be replaced. Fixing these issues gives us:

Theorem 1.10. If G is a group, there is a unique element $e \in G$ with $e * a = a = a * e$ for all $a \in G$. Further, for each $a \in G$, there is a unique $b \in G$ with $a * b = e = b * a$.

Proof. Since G is a semigroup, there is some element $e \in G$ with $e * a = a = a * e$ for all $a \in G$. Suppose that there is an element $e' \in G$ with $e' * a = a * e$ for all $a \in G$. [In particular, if $a = e$,] then $e' * e = e$. On the other hand, the definition of e shows that $e' * e = e'$, and so $e' = e$.

Now fix $a \in G$ and let $b \in G$ be such that $a * b = e = b * a$. Suppose that $a * c = e = c * a$ for some $c \in G$. Then $c = c * e = c * (a * b) = (c * a) * b = e * b = b$, as desired. □

(Note in particular the use of ‘now’ to cancel the previous supposition. ‘now’ usually appears as part of ‘now suppose...’.)

This result is immediately followed by some important definitions:

As a **result** of the **uniqueness assertions** of the theorem, we **may** now **give names** to e and to b . We call e the **identity** of G and, if $a * b = e = b * a$, then we call b the **inverse** of a and denote it by a^{-1} .

At the moment, variables defined inside a result are not automatically visible from outside it, and so none of the variables in this last remark are bound. We will bind G and e by ‘extracting’ them from the result, and bind a and b by explicit quantification. More seriously, e is being silently upgraded from a variable (like ‘ a ’) to a piece of notation (like ‘ \bullet^{-1} ’); we need to make this explicit.

*[As a result of the uniqueness assertions of the theorem, we may now give names to e and to b .] Let G and e be as in Theorem 1.10. We call e the **identity** of G and denote it by e or e_G ; and, if $a * b = e = b * a$ for some $a, b \in G$, then we call b the **inverse** of a and denote it by a^{-1} .*

The uniqueness of inverses has an important consequence:

Corollary 1.11. If G is a **group**, and $a \in G$, then

$$(a^{-1})^{-1} = a.$$

Proof. By definition, $(a^{-1})^{-1}$ is that **element** $g \in G$ with $a^{-1} * g = e = g * a^{-1}$. But a is such an **element**, and so the **uniqueness** gives $g = a$.

The fragment vocabulary does not include ‘such an’ or ‘uniqueness’, and we need to rewrite the proof accordingly.

Corollary 1.11. If G is a **group**, and $a \in G$, then

$$(a^{-1})^{-1} = a.$$

Proof. By definition, $(a^{-1})^{-1}$ is the unique **element** $g \in G$ with $a^{-1} * g = e = g * a^{-1}$. But $a^{-1} * a = e = a * a^{-1}$, and so $(a^{-1})^{-1} = a$ as required.

Once inverses are defined, it is possible to define negative powers:

Definition. If G is a **group**, and $a \in G$, define the **powers** of a as follows: if n is a **positive integer**, then a^n is defined as in any **semigroup**; define $a^0 = e$; define $a^{-n} = (a^{-1})^n$.

The language does not support “copying” earlier definitions, as in ‘ a^n is defined as in any semigroup’; since the definition is unnecessary (a group is a semigroup) we will omit it. We encounter the same problem as earlier regarding ‘ $a^0 = e$ ’: ‘0’ could be a fixed token. And finally, ‘as follows’ lies outside the fragment vocabulary. So this definition requires considerable rewriting:

Definition. Let G be a **group** and let $a \in G$. If n is a **non-negative integer**, then we define a^n to be e if $n = 0$ and $(a^{-1})^{-n}$ otherwise. Note that “ a^n ” is unambiguous for all **integers** n . The **powers** of a are the objects of the form a^n for some **integer** n .

(The final note is necessary because we have defined ‘ a^n ’ in two separate definitions; we need to explain that these are compatible parts of one concept rather than two conflicting notations. Cf. §?? on reanalysis and the extension of notations.)

1.7 Characterisation

Even though the **list** of **axioms** defining a **group** is **short**, it is **worthwhile** to **make** it **even shorter** so it will be as **easy** as **possible** to **verify** that a **particular** example is, in **fact**, a **group**.

Theorem 1.12. If G is a **semigroup** with an **element** e such that:

- (i’) $e * a = a$ for all $a \in G$; and
- (ii’) for each $a \in G$, there is an **element** $b \in G$ with $b * a = e$, then G is a **group**.

This is unproblematic apart from the curious way in which ‘, then G is a group.’ is typeset as part of the item (ii’).

[Even though the list of axioms defining a group is short, it is worthwhile to make it even shorter so it will be as easy as possible to verify that a particular example is, in fact, a group.]

Theorem 1.12. If G is a **semigroup** with an **element** e such that:

- (i’) $e * a = a$ for all $a \in G$; and
 - (ii’) for each $a \in G$, there is an **element** $b \in G$ with $b * a = e$
- then G is a **group**.

And now to the proof of the theorem:

Proof. We claim that if $x * x = x$ in G , then $x = e$. There is an element $y \in G$ with $y * x = e$, and $y * (x * x) = y * x = e$. On the other hand, $y * (x * x) = (y * x) * x = e * x = x$. Therefore, $x = e$.

If $b * a = e$, let us show that $a * b = e$. Now $(a * b) * (a * b) = a * [(b * a) * b] = a * [e * b] = a * b$, and so our claim gives that $a * b = e$. (Observe that we have used associativity for an expression having four factors.)

If $a \in G$, we must show that $a * e = a$. Choose $b \in G$ with $b * a = e = a * b$ (using our just finished calculation). Then $a * e = a * (b * a) = (a * b) * a = e * a = a$, as desired. ■

The content of this proof requires only minor modifications: square brackets must be replaced with round, some words outside the fragment vocabulary ('us', 'our', 'gives', 'must') must be removed and several variables need to be bound. But the rhetorical structure of the proof is subtle and difficult to extract without using common sense: there are no clear flags showing where the claim begins and ends, or that the second and third paragraphs are essentially independent. We will edit the proof substantially to make the rhetorical structure transparent.

Proof.

Claim. If $x * x = x$ for some x in G , then $x = e$.

Proof. There is an element $y \in G$ with $y * x = e$, and $y * (x * x) = y * x = e$. On the other hand, $y * (x * x) = (y * x) * x = e * x = x$. Therefore, $x = e$. □

Suppose $b * a = e$ for some $a, b \in G$; we will show that $a * b = e$. Now $(a * b) * (a * b) = a * ((b * a) * b) = a * (e * b) = a * b$, and so the claim shows that $a * b = e$. [Observe that we have used associativity for an expression having four factors.]

Now suppose $a \in G$; it suffices to prove that $a * e = a$. Choose $b \in G$ with $b * a = e = a * b$ [using our just finished calculation]. Then $a * e = a * (b * a) = (a * b) * a = e * a = a$, as desired. □

The degree of rewriting required here is rather unsatisfactory; I would like future versions to infer more of the rhetorical structure. (And in particular, to directly support more of the cases where 'If $a \in G$, ...' means 'Suppose $a \in G$; ...'.) Any thoughts on this issue would be welcome.

1.8 Homomorphisms

Definition. Let $(G, *)$ and (H, \circ) be groups. A function $f : G \rightarrow H$ is a *homomorphism* if, for all $a, b \in G$,

$$f(a * b) = f(a) \circ f(b).$$

An *isomorphism* is a homomorphism that is also a bijection. We say that G is *isomorphic* to H , denoted by $G \cong H$, if there exists an isomorphism $f : G \rightarrow H$.

This definition needs no changes.

Definition. Let $(G, *)$ and (H, \circ) be groups. A function $f : G \rightarrow H$ is a *homomorphism* if, for all $a, b \in G$,

$$f(a * b) = f(a) \circ f(b).$$

An *isomorphism* is a homomorphism that is also a bijection. We say that G is *isomorphic* to H , denoted by $G \cong H$, if there exists an isomorphism $f : G \rightarrow H$.

The section closes with a result on the basic properties of homomorphisms.

Theorem 1.13. Let $f : (G, *) \rightarrow (G', \circ)$ be a homomorphism.

- (i) $f(e) = e'$, where e' is the identity in G' .
- (ii) If $a \in G$, then $f(a^{-1}) = f(a)^{-1}$.
- (iii) If $a \in G$ and $n \in \mathbb{Z}$, then $f(a^n) = f(a)^n$.

This version of the language does not support pattern matching in ‘compound’ let-expressions; i.e. it can’t work back from the fact that ‘ $f : (G, *) \rightarrow (G', \circ)$ ’ is a homomorphism to deduce that $(G, *)$ and (G', \circ) are groups. (The basic reason for this is that one could define new notation ‘ $f : X \rightarrow Y$ ’ for f a homomorphism but X, Y objects of some strange type; so the required inference is not sound. I believe this problem is surmountable.) Apart from this, the text is unproblematic.

Theorem 1.13. Let $(G, *)$ and (G', \circ) be groups, and let $f : (G, *) \rightarrow (G', \circ)$ be a homomorphism.

- (i) $f(e) = e'$ where e' is the identity in G'
- (ii) If $a \in G$ then $f(a^{-1}) = f(a)^{-1}$
- (iii) If $a \in G$ and $n \in \mathbb{Z}$ then $f(a^n) = f(a)^n$

The proof of the result is as follows:

- Proof.* (i) Applying f to the equation $e = e * e$ gives $f(e) = f(e * e) = f(e) \circ f(e)$. Now multiply each side of the equation by $f(e)^{-1}$ to obtain $e' = f(e)$.
- (ii) Applying f to the equations $a * a^{-1} = e = a^{-1} * a$ gives $f(a) * f(a^{-1}) = e = f(a^{-1}) * f(a)$. It follows from Theorem 1.10, the uniqueness of the inverse, that $f(a^{-1}) = f(a)^{-1}$.
- (iii) An easy induction proves $f(a^n) = f(a)^n$ for all $n \geq 0$, and then $f(a^{-n}) = f((a^{-1})^n) = f(a^{-1})^n = f(a)^{-n}$. ■

(This excerpt contains typographical errors — it should read ‘ $f(a) \circ f(a^{-1}) = e = f(a^{-1}) \circ f(a)$ ’, with \circ in place of $*$.)

This version of the language does not allow reference to entities other than mathematical objects and results. So ‘each side of the equation’ and ‘an easy induction’ need to be removed. Also, the terms ‘applying’ (a function to an equation) and ‘multiply’ are *mathematical* terms that have never been defined, so we need to remove them. And finally, we need to bind n in (iii). Making these changes gives:

- Proof.* (i) The equation $e = e * e$ gives $f(e) = f(e * e) = f(e) \circ f(e)$. It follows that $f(e)^{-1} \circ f(e) = f(e)^{-1} \circ f(e) \circ f(e)$, so that $e' = f(e)$ as required.
- (ii) The equation $a * a^{-1} = e = a^{-1} * a$ gives $f(a) \circ f(a^{-1}) = e = f(a^{-1}) \circ f(a)$. It follows from Theorem 1.10 [the uniqueness of the inverse,] that $f(a^{-1}) = f(a)^{-1}$.
- (iii) It follows from the Principle of Induction that $f(a^n) = f(a)^n$ for all $n \geq 0$; then $f(a^{-n}) = f((a^{-1})^n) = f(a^{-1})^n = f(a)^{-n}$ for all $n \geq 0$. □

Note that if we had named Theorem 1.10 “The Uniqueness of the Inverse”, then we could have used that phrase in the text rather than relegating it to a comment.

2 The Isomorphism Theorems

We now drop the $*$ notation for the operation in a group. Henceforth, we shall write ab instead of $a * b$, and we shall denote the identity element by 1 instead of e .

This is an informal remark; to move it into the formal register we need to state that a, b are elements of the same group, etc.:

If G is a group, and $a, b \in G$, then we write ab for $a * b$, and 1 or 1_G for the identity element of G .

Now that we have not revoked the old notations ‘ $a*b$ ’ and ‘ e ’; allowing this would raise a host of complications for the language. (For example: what happens if notation is introduced in one document, used in a second, revoked in a third, and a fourth document references all the others? Does the notation remain ‘live’?)

2.1 Subgroups

Definition. A nonempty subset S of a group G is a *subgroup* of G if $s \in S$ implies $s^{-1} \in S$ and $s, t \in S$ imply $st \in S$.

Apart from the typographical errors (G for S , repeatedly), this needs no changes.

Definition. A nonempty subset S of a group G is a *subgroup* of G if $s \in S$ implies $s^{-1} \in S$ and $s, t \in S$ imply $st \in S$.

The definition is immediately followed by some notation:

If X is a subset of a group G , we write $X \subset G$; if X is a subgroup of a group G , we write $X \leq G$.

The first half of this is superfluous; because the group is a set, it has picked up all the terminology (cf. ‘subset of a group G ’) and notation associated with sets. If we repeated the definition of ‘ \subset ’ we would then have to tell the compiler that the two notions of ‘ \subset ’ are in fact the same — and while this is quite possible, we prefer to avoid making work for ourselves.

If X is a subgroup of a group G , we write $X \leq G$.

The following result is crucial because, as well as being a verifiable assertion, it causes the language to treat subgroups in a different way (as groups rather than as mere sets); this is discussed extensively in §???. Of the eight well-regarded group theory texts we studied, Rotman is the only one which makes this fact explicit.

Theorem 2.1. If $S \leq G$ (i.e. if S is a subgroup of G), then S is a group in its own right.

We need to state that G is a group. Also, the language does not support asides (bracketed or otherwise), and ‘own right’ lies outside the fragment vocabulary; moving these into comments gives:

Theorem 2.1. Let G be a group. If $S \leq G$ [i.e. if S is a subgroup of G], then S is a group [in its own right].

The proof bears careful examination:

Proof. The hypothesis “ $s, t \in S$ imply $st \in S$ ” shows that S is equipped with an operation (if $\mu : G \times G \rightarrow G$ is the given multiplication in G , then its restriction $\mu|_{S \times S}$ has its image contained in S). Since S is nonempty, it contains an element, say, s , and the definition of subgroup says that $s^{-1} \in S$; hence $1 = ss^{-1} \in S$. Finally, the operation on S is associative because $(a * b) * c = a * (b * c)$ for every $a, b, c \in G$ implies, in particular, that $(a * b) * c = a * (b * c)$ for every $a, b, c \in S$.

This text relies on the intelligence of the reader to fill in an important point of precision. The cited restriction $\mu|_{S \times S}$ of μ is not the group operation on S — $\mu|_{S \times S}$ has co-domain G (even if its image is contained in S), whereas the group operation on S has co-domain S . We need to make this explicit, and to remove some words outside the fragment vocabulary, such as ‘finally’, ‘in particular’ and (outside a definition) ‘say’:

Proof. Let $\mu : G \times G \rightarrow G$ be the multiplication in G . Since $s, t \in S$ imply $st \in S$, there is an operation μ' on S defined by

$$\mu' : S \times S \rightarrow S \quad (a, b) \mapsto \mu(a, b).$$

Since S is nonempty, it contains an element s , and the definition of a subgroup shows that $s^{-1} \in S$; hence $1 = ss^{-1} \in S$. [Finally,] μ' is associative because $(a * b) * c = a * (b * c)$ for every $a, b, c \in G$ implies [in particular,] that $(a * b) * c = a * (b * c)$ for every $a, b, c \in S$. It follows that (S, μ') is a group, as required.

This is followed by some useful tests for group-hood and subgroup-hood.

Verifying associativity is the most tedious part of showing that a given set G equipped with a multiplication is actually a group. Therefore, if G is given as a subset of a group G^* , then it is much simpler to show that G is a subgroup of G^* than to verify all the group axioms for G .

Theorem 2.2. A subset S of a group G is a subgroup if and only if $1 \in S$ and $s, t \in S$ imply $st^{-1} \in S$.

Proof. If $s \in S$, then $1s^{-1} = s^{-1} \in S$, and if $s, t \in S$, then $s(t^{-1})^{-1} = st \in S$. The converse is also easy. ■

A human reader will infer most of the rhetorical structure of the proof using common sense; our language needs this to be explicitly spelt out. While doing this we will remove certain terms outside the fragment vocabulary ('converse', 'also', 'easy').

[Verifying associativity is the most tedious part of showing that a given set G equipped with a multiplication is actually a group. Therefore, if G is given as a subset of a group G^ , then it is much simpler to show that G is a subgroup of G^* than to verify all the group axioms for G .]*

Theorem 2.2. A subset S of a group G is a subgroup if and only if $1 \in S$ and $s, t \in S$ imply $st^{-1} \in S$.

Proof. Suppose that S is a subgroup. If $s \in S$, then $1s^{-1} = s^{-1} \in S$, and if $s, t \in S$, then $s(t^{-1})^{-1} = st \in S$.

Now suppose that $1 \in S$ and $s, t \in S$ imply $st^{-1} \in S$. It remains to prove that S is a subgroup. *[This is easy.]* □

2.2 Examples of Subgroups

Definition. If G is a group and $a \in G$, then the *cyclic subgroup generated by a* , denoted by $\langle a \rangle$, is the set of all the powers of a . A group G is called *cyclic* if there is $a \in G$ with $G = \langle a \rangle$; that is, G consists of all the powers of a .

It is plain that $\langle a \rangle$ is, indeed, a subgroup of G . Notice that different elements can generate the same cyclic subgroup. For example, $\langle a \rangle = \langle a^{-1} \rangle$.

'consists of' is an undefined mathematical term and must be replaced; and the informal remark needs to be moved into a comment.

Definition. If G is a group and $a \in G$, then the *cyclic subgroup generated by a* , denoted by $\langle a \rangle$, is the set of all the powers of a . A group G is called *cyclic* if there is $a \in G$ with $G = \langle a \rangle$ [;] that is, G contains precisely the powers of a .

[It is plain that $\langle a \rangle$ is, indeed, a subgroup of G . Notice that different elements can generate the same cyclic subgroup. For example, $\langle a \rangle = \langle a^{-1} \rangle$.]

The cyclic subgroup just introduced is immediately used to define the order of an element:

Definition. If G is a group and $a \in G$, then the *order of a* is $|\langle a \rangle|$, the number of elements in $\langle a \rangle$.

This needs no changes.

Definition. If G is a group and $a \in G$, then the *order of a* is $|\langle a \rangle|$, the number of elements in $\langle a \rangle$.

Some other important subgroups are introduced in examples:

EXAMPLE 2.1. If G is a group, then G itself and $\{1\}$ are always subgroups of G ; we shall henceforth denote the subgroup $\{1\}$ by 1 . Any subgroup H of G other than G is called *proper*, and we denote this by $H < G$. The subgroup 1 of G is often called the trivial subgroup of G .

We need to remove a few words outside the fragment vocabulary ('itself', 'shall', 'henceforth', 'other', 'often') as well as the use of 'this' to refer to a fact (as opposed to a mathematical object).

EXAMPLE 2.1. If G is a group, then G [itself] and 1 are always subgroups of G ; we will [henceforth] denote the subgroup 1 by 1 . Any subgroup H of G which is not equal to G is called *proper*, and if H is a proper subgroup of G we write $H < G$. The subgroup 1 of G is [often] called the trivial subgroup of G .

The second example relates subgroups to homomorphisms:

EXAMPLE 2.2. Let $f : G \rightarrow H$ be a homomorphism, and define

$$\text{kernel } f = \{a \in G : f(a) = 1\}$$

and

$$\text{image } f = \{h \in H : h = f(a) \text{ for some } a \in G\}.$$

Then $K = \text{kernel } f$ is a subgroup of G and image f is a subgroup of H . To see that $K \leq G$ note first that $f(1) = 1$ so that $1 \in K$. Also if $s, t \in K$ then $f(s) = 1 = f(t)$ and so $f(st^{-1}) = f(s)f(t)^{-1} = 1$ hence $st^{-1} \in K$ and so K is a subgroup of G . It is equally easy to see that image f is a subgroup of H .

Notation. We usually write $\ker f$ instead of kernel f and $\text{im } f$ instead of image f .

This is somewhat eccentric. One would normally introduce textual names ('the kernel of f ', 'the image of f ') and then introduce the symbolic names which are transparently contractions of the textual names. Rotman instead introduces long symbolic names (kernel f , image f), uses these to motivate short symbolic names (again as abbreviations), and then discards the long names. Since the text later refers to 'the kernel of a group', we will normalise the definitions to use the standard approach.

The other distinctive feature of this definition is an inversion in the rhetorical structure: the sentence 'Then $K = \text{kernel } f$ is a subgroup of G and image f is a subgroup of H .' precedes the content that justifies it. The compiler is unable to infer the existence of this inversion (the argument is always assumed to run forwards, sentence by sentence), and so the sentence needs an explicit rhetorical marker.

Apart from this, the changes needed are minor. Certain terms lying outside the fragment vocabulary ('first', 'equally', 'easy', 'see') must be removed. Also, as previously, the system is unable to 'pattern match' to deduce that ' G ' and ' H ' in ' $f : G \rightarrow H$ ' are groups, and must be told this explicitly.

EXAMPLE 2.2. Let G and H be groups and let $f : G \rightarrow H$ be a homomorphism. We define the kernel of f , $\ker f$, to be

$$\{a \in G : f(a) = 1\}$$

and the image of f , $\text{im } f$, to be

$$\{h \in H : h = f(a) \text{ for some } a \in G\}.$$

We will show that $K = \ker f$ is a subgroup of G and $\text{im } f$ is a subgroup of H . To see that $K \leq G$, note first that $f(1) = 1$, so that $1 \in K$. Also, if $s, t \in K$, then $f(s) = 1 = f(t)$, and so $f(st^{-1}) = f(s)f(t)^{-1} = 1$; hence $st^{-1} \in K$, and so K is a subgroup of G . [It is equally easy to see that] image f is a subgroup of H .

2.3 Creating Subgroups

Theorem 2.5. The intersection of any family of subgroups of a group G is again a subgroup of G .

Here we only need to comment out the word ‘again’, which is not in the fragment vocabulary.

Theorem 2.5. The intersection of any family of subgroups of a group G is [again] a subgroup of G .

The proof is as follows:

Proof. Let $\{S_i : i \in I\}$ be a family of subgroups of G . Now $1 \in S_i$ for every i , and so $1 \in \bigcap S_i$. If $a, b \in \bigcap S_i$, then $a, b \in S_i$ for every i , and so $ab^{-1} \in S_i$ for every $i \in I$; hence $ab^{-1} \in \bigcap S_i$, and $\bigcap S_i \leq G$. ■

There are two problems here. First, the language does not currently support introduction of compound objects in let-statements, as in ‘Let $\{S_i : i \in I\}$ ’. Second, I and many instances of i are unbound. Fixing these gives:

Proof. Let I be a set and let S be a I -indexed family of subgroups of G . Now $1 \in S_i$ for every i , and so $1 \in \bigcap_i S_i$. If $a, b \in \bigcap_i S_i$, then $a, b \in S_i$ for every i , and so $ab^{-1} \in S_i$ for every i ; hence $ab^{-1} \in \bigcap_i S_i$, and $\bigcap_i S_i \leq G$. □

This result is immediately followed by a corollary:

Corollary 2.6. If X is a subset of a group G , then there is a *smallest* subgroup H of G containing X ; that is, if $X \subset S$ and $S \leq G$, then $H \leq S$.

It is not clear whether ‘smallest subgroup containing’ is being used or defined. Also, the language does not understand any connection between ‘smallest’ and notions of ‘size’, or between ‘contains’ and ‘containing’. We will rewrite the definition to deal with these issues, and also to bind S in ‘ $X \subset S$ ’.

Corollary 2.6. If X is a subset of a group G , then there is a subgroup H of G such that $X \subset H$ and such that if $X \subset S$ for some $S \leq G$, then $H \leq S$. H is called the *smallest subgroup* of G containing X .

The proof is as follows:

Proof. There are subgroups of G containing X ; for example, G itself contains X ; define H as the intersection of all the subgroups of G which contain X . Note that H is a subgroup, by Theorem 2.5, and $X \subset H$. If $S \leq G$ and $X \subset S$, then S is one of the subgroups of G being intersected to form H ; hence, $H \leq S$, and so H is the smallest such subgroup. ■

There are many points here. ‘for example’ lies outside the fragment vocabulary; and the sentence containing it is in a small way a rhetorical inversion (a conclusion preceding its justification), which the language does not support. We can deal with both of these issues by joining the two sentences using ‘since’. ‘define H as...’ is invalid — ‘define’ may not be used to introduce variables. We have only defined intersections of families of subgroups, not intersections of (linguistic collections of) subgroups. Also, the language does not understand the connection between ‘containing’ and ‘contains’, between ‘intersection’ and ‘being intersected’, or between ‘smallest’ and ‘size’. And finally, ‘itself’ lies outside the fragment vocabulary. Dealing with these issues gives:

Proof. There are subgroups of G containing X , since G [itself] contains X ; let H be the intersection of the family of subgroups of G which contain X . Note that H is a subgroup, by Theorem 2.5, and $X \subset H$. If $S \leq G$ and $X \subset S$, then $H \leq S$, as required. □

Finally, these results are used to introduce the notion of generation for subgroups.

Definition. If X is a subset of a group G , then the smallest subgroup of G containing X , denoted by $\langle X \rangle$, is called the *subgroup generated by X* . One also says that X *generates* $\langle X \rangle$.

In particular, if H and K are subgroups of G , then the subgroup $\langle H \cup K \rangle$ is denoted by $H \vee K$.

We need to replace or comment several words outside the fragment vocabulary (‘one (pronoun)’, ‘also’, ‘particular’).

Definition. If X is a subset of a group G , then the smallest subgroup of G containing X , denoted by $\langle X \rangle$, is called the *subgroup generated by X* . We [also] say that X *generates* $\langle X \rangle$.

[In particular,] if H and K are subgroups of G , then the subgroup $\langle H \cup K \rangle$ is denoted by $H \vee K$.

2.4 Cosets

Definition. If S is a subgroup of G and if $t \in G$, then a *right coset* of S in G is the subset of G

$$St = \{st : s \in S\}$$

(a *left coset* is $tS = \{ts : s \in S\}$). One calls t a *representative* of St (and also of tS).

This definition is tangled. It makes sense to say that a X is a Y , or to say that the X is the Y , but it makes no sense to say that *a* X is *the* Y . Also, there is nothing to indicate that ‘ St ’ is being defined rather than being used to refer to an object. The confusion arises from an attempt to simultaneously define a class of textual objects (‘right coset’) and an individual symbolic object ‘ St ’. We need to separate these in order to make sense of the definition.

Also, the language does not support **parenthetical asides** like ‘(a left coset is ...)’ and ‘(and also of tS)’.

As we do not need to refer to left cosets for the material we are covering, we will simply excise the asides. The text needs to be emended to actually state that G is a group (a recurring issue in this section). And finally, ‘one’ is outside the fragment vocabulary and needs to be replaced.

Definition. If S is a subgroup of a group G and if $t \in G$, then we denote the set

$$\{st : s \in S\}$$

by St . We call t a *representative* of St . A *right coset* of S in G is a set of the form St for some $t \in G$.

This definition is followed by a remark about representatives:

A right coset St has many representatives; every element of the form st for $s \in S$ is a representative of St . The next lemma gives a criterion for determining whether two right cosets of S are the same when a representative of each is known.

Most of this is informal and unformalisable; for example, some right cosets only have one representative. But we can formalise the remark about ‘every element of the form...’:

[A right coset St has many representatives;] if S is as in the definition, then every element of the form st for $s \in S$ is a representative of St . [The next lemma gives a criterion for determining whether two right cosets of S are the same when a representative of each is known.]

The lemma referred to is:

Lemma 2.8. If $S \leq G$, then $Sa = Sb$ if and only if $ab^{-1} \in S$ ($aS = bS$ if and only if $b^{-1}a \in S$).

We need to actually bind several variables (G , a and b) as well as removing the **parenthetical aside**:

Lemma 2.8. If G is a **group**, $S \leq G$, and $a, b \in G$, then $Sa = Sb$ if and only if $ab^{-1} \in S$.

The proof of the lemma is as follows:

Proof. If $Sa = Sb$, then $a = 1a \in Sa = Sb$, and so there is $s \in S$ with $a = sb$; hence, $ab^{-1} = s \in S$. Conversely, assume that $ab^{-1} = \sigma \in S$; hence, $a = \sigma b$. To **prove** that $Sa = Sb$, we **prove** two **inclusions**. If $x \in Sa$, then $x = sa$ for some $s \in S$, and so $x = s\sigma b \in Sb$; **similarly**, if $y \in Sb$, then $y = s'b$ for some $s' \in S$, and so $y = s'\sigma^{-1}a \in Sa$. Therefore, $Sa = Sb$. ■

There are several points to note here. First, ‘If $Sa = Sb$, ...’ is not expressing a conditional remark, but introducing an assumption in the same way as ‘assume’ and ‘suppose’ do. The language has no way of deducing this, and requires an explicit emendation. Second, σ needs to be bound. Third, ‘To prove that $Sa = Sb$, we prove two inclusions.’, which is a vague, high-level remark and refers to an undefined mathematical term (‘inclusions’) will be treated as a comment. Fourth, ‘similarly’ lies outside the fragment vocabulary and will also be put inside a comment.

Proof. Assume $Sa = Sb$. Then $a = 1a \in Sa = Sb$, and so there is $s \in S$ with $a = sb$; hence, $ab^{-1} = s \in S$. Conversely, assume that $ab^{-1} \in S$; hence, $a = \sigma b$, where $\sigma = ab^{-1}$. *[To prove that $Sa = Sb$, we prove two inclusions.]* If $x \in Sa$, then $x = sa$ for some $s \in S$, and so $x = s\sigma b \in Sb$; *[similarly,]* if $y \in Sb$, then $y = s'b$ for some $s' \in S$, and so $y = s'\sigma^{-1}a \in Sa$. Therefore, $Sa = Sb$. □

The next theorem takes a major step towards an important result, Lagrange’s Theorem:

Theorem 2.9. If $S \leq G$, then any two **right** (or any two **left**) **cosets** of S in G are either **identical** or **disjoint**.

As previously, we will remove the **parenthetical aside** and explicitly state that G is a group. We will also replace the undefined term ‘identical’ with the synonym ‘equal’.

Theorem 2.9. If G is a group and $S \leq G$, then any two right cosets of S in G are either equal or disjoint.

The proof of this result is quite short:

Proof. We show that if there exists an element $x \in Sa \cap Sb$, then $Sa = Sb$. Such an x has the form $sb = x = ta$, where $s, t \in S$. Hence, $ab^{-1} = t^{-1}s \in S$, and so the lemma gives $Sa = Sb$. ■

The first line, ‘we show that ...’ looks like a rhetorical inversion (a conclusion preceding its justification), which the language does not support; we will reword it to ‘it is sufficient to prove that ...’. We also need to bind a and b and to replace some terms (‘such’, ‘form’, ‘gives’) outside the fragment vocabulary:

Proof. Let $a, b \in G$. It is sufficient to prove that if there exists an element $x \in Sa \cap Sb$, then $Sa = Sb$. If $x \in Sa \cap Sb$, then x satisfies $sb = x = ta$ for some $s, t \in S$. Hence, $ab^{-1} = t^{-1}s \in S$, and so the lemma gives $Sa = Sb$. □

The theorem is followed by a remark which gives the ‘underlying idea’ or ‘spirit’ of Lagrange’s Theorem:

Theorem 2.9 may be paraphrased to say that the right cosets of a subgroup S comprise a partition of G (each such coset is nonempty, and G is their disjoint union). This being true, there must be an equivalence relation on G lurking somewhere in the background: it is given, for $a, b \in G$, by $a \equiv b$ if $ab^{-1} \in S$, and its equivalence classes are the right cosets of S .

The choice of language (‘paraphrased’, ‘lurking’, ‘background’) makes it clear that this is an informal remark, which should be converted into a comment:

[Theorem 2.9 may be paraphrased to say that the right cosets of a subgroup S comprise a partition of G (each such coset is nonempty, and G is their disjoint union). This being true, there must be an equivalence relation on G lurking somewhere in the background: it is given, for $a, b \in G$, by $a \equiv b$ if $ab^{-1} \in S$, and its equivalence classes are the right cosets of S .]

2.5 Lagrange’s Theorem

We needed a little more notation before we can state Lagrange’s Theorem:

Definition. If $S \leq G$, then the *index* of S in G , denoted by $[G : S]$, is the number of right cosets of S in G .

Definition. If G is a group, then the *order* of G , denoted by $|G|$, is the number of elements in G .

We need to remove the square brackets, which are reserved for comments; we could choose any token(s) we liked as replacements, but the standard convention will be to use ‘\[' and ‘\]’ (each of which is a single token). We also need to bind G in the first definition.

Definition. If G is a group and $S \leq G$, then the *index* of S in G , denoted by $\backslash[G : S\backslash]$, is the number of right cosets of S in G .

Definition. If G is a group, then the *order* of G , denoted by $|G|$, is the number of elements in G .

(Note that ‘the number of \mathbf{D} ’ is defined as ‘the size of the set of \mathbf{D} ’ in the document on set theory.) And finally we come to the theorem itself:

Theorem 2.11 (Lagrange). If G is a finite group and $S \leq G$, then $|S|$ divides $|G|$ and $\backslash[G : S\backslash] = |G|/|S|$.

No mathematical changes are needed. We will however relabel the theorem as “Lagrange’s Theorem” so that we can refer back to it by name.

Theorem 2.11 (Lagrange’s Theorem). If G is a finite group and $S \leq G$, then $|S|$ divides $|G|$ and $\backslash[G : S\backslash] = |G|/|S|$.

The given proof is very short:

Proof. By Theorem 2.9, G is partitioned into its right cosets

$$G = St_1 \cup St_2 \cup \cdots \cup St_n,$$

and so $|G| = \sum_{i=1}^n |St_i|$. But it is easy to see that $f_i : S \rightarrow St_i$, defined by $f_i(s) = st_i$, is a bijection, and so $|St_i| = |S|$ for all i . Thus $|G| = n|S|$, where $n = |G|/|S|$.

(A question: is ‘ $f_i : S \rightarrow St_i$ ’ inside the scope of ‘for all i ’? We need to read it this way in order to bind all occurrences of i , but it doesn’t feel right... .)

There is a problem with the content here. A relative clause with commas around it, like

Professor James, **who is an expert in Victorian poetry**, will be giving a lecture tonight.

is a non-restrictive relative clause; it adds parenthetical information which can be removed without affecting the main sentence. So

$f_i : S \rightarrow St_i$, defined by $f_i(s) = st_i$, is a bijection

really asserts two separate statements, which should be verified separately:

1. $f_i : S \rightarrow St_i$ is a bijection
2. $f_i : S \rightarrow St_i$ is defined by $f_i(s) = st_i$

This is of course nonsense — $f_i : S \rightarrow St_i$ does not in itself refer to a unique object. The text should contain a restrictive relative clause here.

On the logical level, we may not quantify over an unknown number of variables (t_1, t_2, \dots, t_n); we need to rewrite the proof to avoid doing so. As we do so, we will remove the words ‘partitioned’ (an undefined mathematical term) and ‘easy’ and ‘see’ (outside the fragment vocabulary).

Proof. By Theorem 2.9, there exist a natural number n and a function $t : \{1, 2, \dots, n\} \rightarrow G$ such that

$$\{St(1), St(2), \dots, St(n)\}$$

is a partition of G and so $|G| = \sum_{i=1}^n |St(i)|$. But *[it is easy to see that]* the function $f_i : S \rightarrow St(i)$ defined by $f_i(s) = st(i)$ for $s \in S$ is a bijection, and so $|St(i)| = |S|$ for all i . Thus $|G| = n|S|$, where $n = |G|/|S|$.

(The second sentence is a little unwieldy. Suggestions?)

We finish with a couple of exercises that are needed for a later proof.

EXERCISES

2.11 Let $a \in G$ have order $n = mk$, where $m, k \geq 1$. Prove that a^k has order m .

2.14 If $a \in G$ has finite order and $f : G \rightarrow H$ is a homomorphism, then the order of $f(a)$ divides the order of a .

The language does not currently have much support for exercises. We need to reformat the headings to a stand-alone result format (‘Exercise 2.11’, etc.), and remove the imperative ‘prove’, which lies outside the fragment vocabulary. We also need to bind G and H .

Exercise 2.11. Let G be a group and let $a \in G$ have order $n = mk$, where $m, k \geq 1$. [Prove that] a^k has order m .

Exercise 2.14. Let G and H be groups. If $a \in G$ has finite order and $f : G \rightarrow H$ is a homomorphism, then the order of $f(a)$ divides the order of a .

2.6 Products of Subsets

Definition. If S and T are nonempty subsets of a group G , then

$$ST = \{st : s \in S \text{ and } t \in T\}.$$

We specifically disallow definitions of symbolic material where the definiendum and definiens are only separated by an equals sign. So we emend the definition to:

Definition. If S and T are nonempty subsets of a group G , then we write ST for

$$\{st : s \in S \text{ and } t \in T\}.$$

The definition is followed by some remarks:

If $S \leq G$, $t \in G$ and $T = \{t\}$, then ST is the right coset St . Notice that the family of all the nonempty subsets of G is a semigroup under this operation: if S , T and U are nonempty subsets of G , then $(ST)U = S(TU)$, for either side consists of all the elements of G of the form $(st)u = s(tu)$ with $s \in S$, $t \in T$ and $u \in U$.

Once again, we need to specify that G is the group. We also need to rewrite the reference to ‘either side’ of an equation, as this is neither a mathematical object nor a result. The term ‘this operation’ is too informal for the compiler; fleshing it out would be time-consuming and turns out to be unnecessary, so we will comment it out. And finally, we will explicitly note that ‘ STU ’ is unambiguous, for the benefit of the parser.

If G is a group, $S \leq G$, $t \in G$ and $T = \{t\}$, then ST is the right coset St . [Notice that the family of all the nonempty subsets of G is a semigroup under this operation:] if S , T and U are nonempty subsets of G , then

$(ST)U = \{(st)u : s \in S, t \in T \text{ and } u \in U\} = \{s(tu) : s \in S, t \in T \text{ and } u \in U\} = S(TU)$. It follows that “ STU ” is unambiguous.

2.7 Normal Subgroups

Definition. A subgroup $K \leq G$ is a **normal subgroup**, denoted by $K \trianglelefteq G$, if $gKg^{-1} = K$ for every $g \in G$.

There are two important points here. First, we want not just to define a term “normal subgroup”, but an adjective “normal”. (We need to be explicit about this because (for example) a “complex number” is not a “number which is complex”.) Second, the expression ‘ gKg^{-1} ’ is not just a combination of coset notation (‘ gH ’ for H a subgroup) and subset product notation (‘ ST ’); so it has never actually been defined. And a minor point recurs: we need to state that G is a group!

If S is a subgroup of a group G and if $a, b \in G$, then we denote the set $\{asb : s \in S\}$ by aSb .

Definition. Let G be a group. A subgroup $K \leq G$ is **normal**, denoted by $K \triangleleft G$, if $gKg^{-1} = K$ for every $g \in G$.

The definition is followed by some important comments:

If $K \leq G$ and there are inclusions $gKg^{-1} \leq K$ for every $g \in G$, then $K \triangleleft G$: replacing g by g^{-1} , we have the inclusion $g^{-1}Kg \leq K$, and this gives the reverse inclusion $K \leq gKg^{-1}$.

The kernel K of a homomorphism $f : G \rightarrow H$ is a normal subgroup: if $a \in K$, then $f(a) = 1$; if $g \in G$, then $f(gag^{-1}) = f(g)f(a)f(g^{-1}) = f(g)f(g^{-1}) = 1$, and so $gag^{-1} \in K$. Hence, $gKg^{-1} \leq K$ for all $g \in G$, and so $K \triangleleft G$. Conversely, we shall see later that every normal subgroup is the kernel of some homomorphisms.

(The first paragraph contains typographical errors and should read “inclusions $gKg^{-1} \hookrightarrow K$ ”, etc.)

The rhetorical structure in this remarks needs clarification; for example, there is no textual marker showing when we have finished showing that K is a normal subgroup. The sentence ‘Replacing...’ is multiply problematic; it contains terms outside the fragment (‘replacing’, ‘reverse’), uses g without binding it, and has a strange and illegal use of ‘this’ which does not refer to a mathematical object (it refers, apparently, to the fact of our having a particular inclusion). This sentence

may be intended to be informal. The rest of the proof is unproblematic; we only need to bind G and H .

Lemma. If $K \leq G$ and there are inclusions $gKg^{-1} \hookrightarrow K$ for every $g \in G$, then $K \triangleleft G$.

Proof. [replacing g by g^{-1}], we have that there exist inclusions $g^{-1}Kg \hookrightarrow K$ for every $g \in G$, and hence that there exist inclusions $K \hookrightarrow gKg^{-1}$ for every $g \in G$. \square

Lemma. Let G and H be groups. The kernel K of a homomorphism $f : G \rightarrow H$ is a normal subgroup.

Proof. If $a \in K$, then $f(a) = 1$; if $g \in G$, then $f(gag^{-1}) = f(g)f(a)f(g^{-1}) = f(g)f(g^{-1}) = 1$, and so $gag^{-1} \in K$. Hence, $gKg^{-1} \leq K$ for all $g \in G$, and so $K \triangleleft G$. [Conversely, we shall see later that every normal subgroup is the kernel of some homomorphisms.] \square

The section closes with a final definition:

If $x \in G$, then a **conjugate** of x in G is an **element** of the form axa^{-1} for some $a \in G$.

Again, we need to bind G .

Let G be a group. If $x \in G$, then a **conjugate** of x in G is an **element** of the form axa^{-1} for some $a \in G$.

2.8 Quotient Groups

The **construction** of the **quotient group** (or **factor group**) G/N in the **next theorem** is of **fundamental importance**.

Theorem 2.21. If $N \triangleleft G$, then the **cosets** of N in G **form** a **group**, denoted by G/N , of **order** $[G : N]$.

We have never defined what it means for some objects to ‘form a group’ (although we could); similarly, we have not defined a ‘coset’, but only a right coset. We also need to extract the definition-in-passing, ‘, denoted by G/N ,’ into a separate sentence, to remove the use of square brackets and, of course, to bind G .

[The construction of the quotient group (or factor group) G/N in the next theorem is of fundamental importance.]

Theorem 2.21. Let G be a group. If $N \triangleleft G$ then we write G/N for the set of right cosets of N in G . G/N is a group of order $|[G : N]|$.

The proof of this theorem is compressed and often informal, so we will consider it in parts. It begins:

Proof. To define a group, one needs a set and an operation. The set here is the family of all cosets of N in G (notice that we need not bother with the adjectives “left” and “right” because N is a normal subgroup). As operation, we propose the multiplication of nonempty subsets of G defined earlier. We have already observed that this operation is associative.

There is a subtlety here. There is a difference between an abstract concept like ‘addition’, ‘conjugation’ or ‘multiplication’ and specific mathematical objects like ‘the operation μ on \mathbb{Z} defined by $\mu(a, b) = a + b$ for all $a, b \in \mathbb{Z}$ ’. Multiplication of cosets has only been introduced (and proved associative) as an abstract concept; we need to introduce a corresponding concrete mathematical object.

Proof. Let $*$ be the operation on G/N defined by $x * y = xy$ for all $x, y \in G/N$. $*$ is associative.

(Note that one needs to show that $*$ is closed; this proof obligation is generated as a presupposition of the phrase ‘the operation on...’.)

The proof continues:

Now

$$\begin{aligned} NaNb &= Na(a^{-1}Na)b && \text{(because } N \text{ is normal)} \\ &= N(aa^{-1})Nab = NNab = Nab && \text{(because } N \leq G). \end{aligned}$$

Thus, $NaNb = Nab$, and so the product of two cosets is a coset.

(This is the required closure proof.)

Here we need to bind a and b (taking care to avoid ‘compound let-s’) and remove the parenthetical asides. We also need to remove the undefined mathematical term ‘product’.

If $x, y \in G/N$, then there exist $a, b \in G$ such that $x = Na$ and $y = Nb$. Now

$$NaNb = Na(a^{-1}Na)b = N(aa^{-1})Nab = NNab = Nab$$

Thus, $NaNb = Nab$, and so xy is a coset.

(Note that multiplication of subsets of a group has been set up in such a way that well-definedness is not an issue here.)

The proof ends as follows:

We let the **reader prove** that the **identity** is the **coset** $N = N1$ and that the **inverse** of Na is $N(a^{-1})$. This **group** is denoted by G/N , and the definition of **index** gives $|G/N| = [G : N]$. ■

A subtle point: we can't refer to 'the identity' or 'the inverse' *until* we have shown that G/N is a group. If we had set up the terminology differently, we could have referred to 'an identity' or 'an inverse' in a semigroup — but as it is, we need to state the equations explicitly in the fashion of Theorem 1.12. We also need to replace the square brackets.

Thus, G/N equipped with $*$ is a **semigroup** such that:

- (i') $N1 * a = a$ for all $a \in G/N$; and
- (ii') for each $x \in G/N$, there is an **element** $a \in G$ with $x = Na$; and $N(a^{-1}) * Na = N1$.

It follows that G/N equipped with $*$ is a **group**, and the definition of **index** gives $|G/N| = [G : N]$. □

The section finishes with a result showing the connection between homomorphisms and normal subgroups:

Corollary 2.22. If $N \triangleleft G$, then the **natural map** (i.e., the **function** $\nu : G \rightarrow G/N$ defined by $\nu(a) = Na$) is a **surjective homomorphism** with **kernel** N .

The language does not support **parenthetical asides** or definitions 'in passing'; so we need to separate out the definition. Also, in the current version of the language, 'the (singular)' is always taken to refer to unique objects — and since there are many natural maps, we may define 'a natural map' or 'the natural map from G to G/N ', but not 'the natural map'. And, as usual, we need to bind G and a .

Corollary 2.22. Let G be a **group**. If $N \triangleleft G$, then we call the **function** $\nu : G \rightarrow G/N$ defined by $\nu(a) = Na$ for all $a \in G$ the **natural map** from G to G/N . The **natural map** from G to G/N is a **surjective homomorphism** with **kernel** N .

(Note that the natural map is defined as a function and upgraded to a homomorphism.)

The proof is as follows:

The equation $\nu(a)\nu(b) = \nu(ab)$ is just the formula $NaNb = Nab$; hence, ν is a homomorphism. If $Na \in G/N$, then $Na = \nu(a)$, and so ν is surjective. Finally, $\nu(a) = Na = N$ if and only if $a \in N$, by Lemma 2.8, so that $N = \ker \nu$. ■

The references to equations and formulas need to be removed, and various instances of a, b need to be bound. Note in particular that ‘If $Na \in G/N$, then...’ is a compound variable introduction and needs to be rewritten.

If $a, b \in G$, then $\nu(a)\nu(b) = NaNb = Nab = \nu(ab)$; hence, ν is a homomorphism. If $x \in G/N$, then $x = Na$ for some $a \in G$; $Na = \nu(a)$, and so ν is so surjective. Finally, $\nu(a) = Na = N$ if and only if $a \in N$, by Lemma 2.8, so that $N = \ker \nu$. □

2.9 The Isomorphism Theorems

Theorem 2.24 (First Isomorphism Theorem). Let $f : G \rightarrow H$ be a homomorphism with kernel K . Then K is a normal subgroup of G and $G/K \cong \text{im } f$.

All we need to do here is to bind G and H .

Theorem 2.24 (First Isomorphism Theorem). Let G and H be groups and let $f : G \rightarrow H$ be a homomorphism with kernel K . Then K is a normal subgroup of G and $G/K \cong \text{im } f$.

(I’m not sure whether we can handle the introduction of K as it is embedded in the let statement. On the other hand, I’m not sure what principle rules this out.) The proof is as follows:

We have already noted that $K \triangleleft G$. Define $\phi : G/K \rightarrow H$ by

$$\phi(Ka) = f(a).$$

To see that ϕ is well-defined, assume that $Ka = Kb$; that is, $ab^{-1} \in K$. Then $1 = f(ab^{-1}) = f(a)f(b^{-1})$, and $f(a) = f(b)$; it follows that $\phi(Ka) = \phi(Kb)$, as desired. Now ϕ is a homomorphism:

$$\phi(KaKb) = \phi(Kab) = f(ab) = f(a)f(b) = \phi(Ka)\phi(Kb).$$

It is **plain** that $\text{im } \phi = \text{im } f$. Finally, we show that ϕ is an **injection**. If $\phi(Ka) = \phi(Kb)$ then $f(a) = f(b)$, hence $f(a)f(b^{-1}) = 1$, $ab^{-1} \in K$ and $Ka = Kb$ (**note that ϕ being an injection is the converse of ϕ being well defined**). We have shown that ϕ is an **isomorphism**.

It follows that there is no **significant difference between a quotient group and a homomorphic image**.

The main difficulties with this proof lie in its rhetorical structure; for example, it requires common-sense to see where the assumption that $Ka = Kb$ terminates. (We could use ‘as desired’ as an end-of-assumption marker, but this seems brittle enough to be dangerous.) We will restructure the proof so as to sidestep this issue. We also need to avoid using a number of words that are outside the fragment vocabulary (‘already’, ‘noted’, ‘see’, ‘plain’ and ‘finally’), to avoid using ‘define’ to introduce a variable, and to explicitly bind a number of variables.

[We have already noted that] $K \triangleleft G$. Let $\phi : G/K \rightarrow H$ be defined by

$$\phi(Ka) = f(a) \text{ for all } a \in G.$$

Claim. ϕ is well-defined

Proof. Assume that $Ka = Kb$ for some $a, b \in G$; that is, $ab^{-1} \in K$. Then $1 = f(ab^{-1}) = f(a)f(b^{-1})$, and $f(a) = f(b)$; it follows that $\phi(Ka) = \phi(Kb)$, as desired. \square

ϕ is a **homomorphism**, since for every $a, b \in G$,

$$\phi(KaKb) = \phi(Kab) = f(ab) = f(a)f(b) = \phi(Ka)\phi(Kb).$$

[It is plain that] $\text{im } \phi = \text{im } f$.

Claim. ϕ is an **injection**.

Proof. If $\phi(Ka) = \phi(Kb)$ for some $a, b \in G$, then $f(a) = f(b)$; hence $f(a)f(b^{-1}) = 1$, $ab^{-1} \in K$, and $Ka = Kb$ *[note that ϕ being an injection is the converse of ϕ being well-defined]*.

We have shown that ϕ is an **isomorphism**. \square

[It follows that there is no significant difference between a quotient group and a homomorphic image.]

Before we can even state the Second Isomorphism Theorem, we need to prove

that the product of any subgroup with a normal subgroup is itself a subgroup.

Lemma 2.25. If S and T are subgroups of G and if one of them is normal, then $ST = S \vee T = TS$.

All we need to do here is to bind G .

Lemma 2.25. If S and T are subgroups of a group G and if one of them is normal, then $ST = S \vee T = TS$.

The proof is straightforward manipulation:

Proof. Recall that ST is just the set of all products of the form st , where $s \in S$ and $t \in T$; hence ST and TS are subsets of $S \vee T$ containing $S \cup T$. If ST and TS are subgroups, then the reverse inclusion will follow from Corollary 2.6. Assume that $T \triangleleft G$. If $s_1 t_1$ and $s_2 t_2 \in ST$, then

$$\begin{aligned} (s_1 t_1)(s_2 t_2)^{-1} &= s_1 t_1 t_2^{-1} s_2^{-1} \\ &= s_1 (s_2^{-1} s_2) t_1 t_2^{-1} s_2^{-1} \\ &= s_1 s_2^{-1} t_3 \\ &= (s_1 s_2^{-1}) t_3 \in ST, \end{aligned}$$

where $t_3 = s_2 (t_1 t_2^{-1}) s_2^{-1} \in T$ because $T \triangleleft G$. Therefore, $ST = S \vee T$. A similar proof shows that TS is a subgroup, and so $TS = S \vee T = ST$. ■

This proof contains references to entities other than that magical objects and results ('the reverse inclusion', 'a similar proof'), undefined mathematical terms ('product', 'containing') and terms outside the fragment vocabulary ('recall', 'just'). Dealing with these gives:

Proof. [Recall that] ST is just the set of all products of the form st , where $s \in S$ and $t \in T$; hence ST and TS are subsets of $S \vee T$ which contain $S \cup T$. If ST and TS are subgroups, then Corollary 2.6 shows that $S \vee T$ is a subset of ST and TS . Assume that $T \triangleleft G$. If $s_1 t_1$ and $s_2 t_2 \in ST$, then

$$\begin{aligned} (s_1 t_1)(s_2 t_2)^{-1} &= s_1 t_1 t_2^{-1} s_2^{-1} \\ &= s_1 (s_2^{-1} s_2) t_1 t_2^{-1} s_2^{-1} \\ &= s_1 s_2^{-1} t_3 \\ &= (s_1 s_2^{-1}) t_3 \in ST, \end{aligned}$$

where $t_3 = s_2 (t_1 t_2^{-1}) s_2^{-1} \in T$ because $T \triangleleft G$. Therefore, $ST = S \vee T$. Also, TS is a subgroup, and so $TS = S \vee T = ST$. □

We are now in a position to set up the objects need to state the Second Isomorphism Theorem:

Suppose that $S \leq H \leq G$ are subgroups with $S \triangleleft G$. Then $S \triangleleft H$ and the quotient H/S is defined; it is the subgroup of G/S consisting of all those cosets Sh with $h \in H$. In particular, if $S \triangleleft G$ and T is any subgroup of G , then $S \leq ST \leq G$ and ST/S is the subgroup of G/S consisting of all those cosets Sst , where $st \in ST$. Since $Sst = St$, it follows that ST/S consists precisely of all those cosets of S having a representative in T .

‘Suppose that $S \leq H \leq G$ are subgroups’ is very informal mathematical usage that would not normally reach a textbook; we will replace this, bind G , etc. . ‘where $st \in ST$ ’ is a compound variable introduction and needs to be broken up. And on a more shallow level, this remark uses various undefined mathematical terms (‘quotient’, ‘consisting of’, ‘consists precisely of’) and two phrases outside the fragment (‘in particular’, ‘having’); replacing these and binding assorted variables is enough to rehabilitate the remark.

Suppose that G is a group and that S and H are subgroups of G with $S \leq H$ and $S \triangleleft G$. Then $S \triangleleft H$ and H/S is defined; it is the subgroup of G/S which contains precisely those cosets Sh with $h \in H$. [In particular,] if $S \triangleleft G$ and T is any subgroup of G , then $S \leq ST \leq G$ and ST/S is the subgroup of G/S which contains precisely those cosets of the form Sst , where $s \in S$ and $t \in T$. Since $Sst = St$, it follows that ST/S contains precisely those cosets of S which have a representative in T .

And now we are finally in a position to state the Second Isomorphism Theorem itself.

Theorem 2.26 (Second Isomorphism Theorem). Let N and T be subgroups of G with N normal. Then $N \cap T$ is normal in T and $T/(N \cap T) \cong NT/N$.

We have not define what it means to be ‘normal in T ’, only what it means to be a ‘normal subgroup of T ’. I would like a future version of the language to be able to make this connection, but for now explicit enmendation is necessary. We also need to bind G (again).

Theorem 2.26 (Second Isomorphism Theorem). Let N and T be subgroups of a group G with N normal. Then $N \cap T$ is a normal subgroup of T and $T/(N \cap T) \cong NT/N$.

The proof is as follows:

Proof. Let $\nu : G \rightarrow G/N$ be the natural map, and let $\nu' = \nu|_T$, the restriction of ν to T . Since ν' is a homomorphism whose kernel is $N \cap T$, Theorem 2.24 gives $N \cap T \triangleleft T$ and $T/(N \cap T) \cong \text{im } \nu'$. Our remarks above show that ν' is just the family of all those cosets of N having a representative in T ; that is, $\text{im } \nu'$ consists of all the cosets in NT/N .

Here we need to eliminate two asides ('...', the restriction of ν to T ' and 'that is, ...'), some phrases outside the fragment vocabulary ('our remarks above', 'just' and 'having') and an undefined mathematical term ('consists of').

Proof. Let $\nu : G \rightarrow G/N$ be a natural map, and let $\nu' = \nu|_T$, the restriction of ν to T . Since ν' is a homomorphism whose kernel is $N \cap T$, Theorem 2.24 gives $N \cap T \triangleleft T$ and $T/(N \cap T) \cong \text{im } \nu'$. [Our remarks above show that] $\text{im } \nu'$ is just the family of all those cosets of N which have a representative in T ; [that is,] $\text{im } \nu'$ contains precisely the cosets in NT/N .

And we move immediately to the Third Isomorphism Theorem:

Theorem 2.27 (Third Isomorphism Theorem). Let $K \leq H \leq G$, where both K and H are normal subgroups of G . Then H/K is a normal subgroup of G/K and

$$(G/K)/(H/K) \cong G/H.$$

Proof. Again we let the first isomorphism theorem do the dirty work. Define $f : G/K \rightarrow G/H$ by $f(Ka) = Ha$ (this "enlargement of coset" map f is well-defined because $K \leq H$). The reader may easily check that f is a surjection with kernel H/K . ■

Imagine trying to prove the third isomorphism theorem directly; the elements of $(G/K)/(H/K)$ are cosets whose representatives are cosets!

The statement of the theorem is fine, apart from the need to bind G . The proof, however, is quite informal. As well as commenting out some remarks, we need to extract the bracketed aside into a separate sentence, to use 'let' instead of 'define' to introduce f and bind a .

Theorem 2.27 (Third Isomorphism Theorem). Let $K \leq H \leq G$, where both K and H are normal subgroups of a group G . Then H/K is a normal subgroup of G/K and

$$(G/K)/(H/K) \cong G/H.$$

Proof [Again we let the first isomorphism theorem do the dirty work.] Let $f : G/K \rightarrow G/H$ be defined by $f(Ka) = Ha$ for all $a \in G$. This [“enlargement of coset”] map f is well defined because $K \leq H$ [The reader may easily check that] f is a surjection with kernel H/K .

[Imagine trying to prove the third isomorphism theorem directly; the elements of $(G/K)/(H/K)$ are cosets whose representatives are cosets!]

3 G-Sets

3.1 Conjugates

Lemma 3.1. If G is a group, then the relation “ y is a *conjugate* of x in G ”, that is, $y = gxg^{-1}$ for some $g \in G$, is an *equivalence relation*.

Proof. Routine. ■

(Note that although ‘conjugate’ is highlighted, it has been defined previously rather than being defined during this statement.)

Relations are defined mathematical objects, and we have not introduced a way of referring to a relation either by a textual phrase (“ y is a conjugate of x in G ”) or by an unaided equation (“ $y = gxg^{-1}$ for some $g \in G$ ”). So we need to rewrite this to use only terms that we have defined. And as ‘routine’ is outside the fragment vocabulary, we need to replace it.

Lemma 3.1. If G is a group, then the relation \sim on G defined by
for x, y in G , $x \sim y$ if and only if $y = gxg^{-1}$ for some $g \in G$
is an *equivalence relation*.

Proof. Trivial. □

Whenever one has an equivalence relation, it is natural to look at the equivalence classes:

If G is a group, then the *equivalence class* of $a \in G$ under the relation “ y is a *conjugate* of x in G ” is called the *conjugacy class* of a ; it is denoted by a^G .

We want to avoid having to specify \sim ab initio, so we will borrow it from the previous lemma.

Let G and \sim be as in Lemma 3.1. The **equivalence class** of $a \in G$ under \sim is called the **conjugacy class** of a ; it is denoted by a^G .

(Note that we have also recalled G from the previous result. If we had instead written ‘If G is a group, ...’, then this would introduce a fresh variable called G which would have no resemblance to the G referred to by \sim .)

The concept of conjugacy class is used to motivate another definition:

If $a \in G$ is the **sole resident** of its **conjugacy class**, then $a = gag^{-1}$ for all $g \in G$; that is, a **commutes** with every **element** of G .

Definition. The **center** of a **group** G , denoted by $Z(G)$, is the **set** of all $a \in G$ that **commute** with every **element** of G .

It is **easy** to **check** that $Z(G)$ is a **normal abelian subgroup** of G .

We need to comment the words ‘sole’, ‘easy’, ‘check’, which are not in the fragment vocabulary, and the word ‘resident’, which is an undefined mathematical term. And, as ever, we need to bind various (independent) instances of G .

If G is a **group** and $a \in G$ is the only **element** in its **conjugacy class**, then $a = gag^{-1}$ for all $g \in G$; that is, a **commutes** with every **element** of G .

Definition. The **center** of a **group** G , denoted by $Z(G)$, is the **set** of all $a \in G$ that **commute** with every **element** of G .

*[It is easy to check that] if G is a **group**, then $Z(G)$ is a **normal abelian subgroup** of G .*

The next definition is motivated in the text itself:

The **following subgroup** is **introduced** to **count** the **number** of **elements** in a **conjugacy class**.

Definition. If $a \in G$, then the **centralizer** of a in G , denoted by $C_G(a)$, is a **set** of all $a \in G$ that **commute** with a .

It is **immediate** that $C_G(a)$ is a **subgroup** of G .

Again we need to separate the aside introducing symbolic notation. We also need to remove a word outside the fragment vocabulary (‘immediate’), and bind some variables (a and G).

[The following subgroup is introduced to count the number of elements in a conjugacy class.]

Definition. If G is a group and $a \in G$ then the **centralizer** of a in G is the set of all $a \in G$ that commute with a it is denoted by $C_G(a)$

If G is a group and $a \in G$ then *[It is immediate that]* $C_G(a)$ is a subgroup of G

The remaining definitions in this section lead towards Sylow's Theorem, one of the 'famous' results of group theory.

One may conjugate subgroups as well as elements.

Definition. If $H \leq G$ and $g \in G$, then the **conjugate** gHg^{-1} is $\{ghg^{-1} : h \in H\}$. The conjugate gHg^{-1} is often denoted by H^g .

Note that a subgroup H is a normal subgroup if and only if it has only one conjugate.

In order to match this definition to actual usage, we should speak of the 'conjugate of H by g in G .' The notation ' gHg^{-1} ' was defined during the introduction of normal subgroups and should not be reintroduced. We need to comment a word outside the fragment vocabulary ('often') and, once again, to bind G .

[One may conjugate subgroups as well as elements.]

Definition. If G is a group, $H \leq G$ and $g \in G$, then the **conjugate** of H by g in G is $\{ghg^{-1} : h \in H\}$. The conjugate gHg^{-1} is *[often]* denoted by H^g .

Note that a subgroup H is a normal subgroup if and only if it has only one conjugate.

(I'm beginning to think the background object mechanism should handle 'subgroup H ' rather than requiring 'subgroup H of G '. Taking that approach means I should revisit the transformations of 'coset' into e.g. 'coset of p H of G ', to remain consistent.)

Definition. If $H \leq G$, then the **normaliser** of H in G , denoted by $N_G(H)$, is

$$N_G(H) = \{a \in G : aHa^{-1} = H\}.$$

It is immediate that $N_G(H)$ is a subgroup of G . Notice that $H \triangleleft N_G(H)$; indeed, $N_G(H)$ is the largest subgroup of G in which H is normal.

Again we need to remove terms outside the fragment vocabulary ('immediate', 'notice', 'indeed'), and again, we need to bind G . Also, 'largest' is an undefined mathematical term; in this instance, to avoid breaking the flow of the text, we

will simply put the last remark into a comment.

Definition. If $H \leq G$, then the *normaliser* of H in G , denoted by $N_G(H)$, is

$$N_G(H) = \{a \in G : aHa^{-1} = H\}.$$

[It is immediate that] $N_G(H)$ is a subgroup of G . Note that $H \triangleleft N_G(H)$; [indeed, $N_G(H)$ is the largest subgroup of G in which H is normal.]

3.2 G-Sets

Definition. If X is a set and G is a group, then X is a *G-set* if there is a function $\alpha : G \times X \rightarrow X$ (called an *action*), denoted by $\alpha : (g, x) \mapsto gx$, such that:

- (i) $1x = x$ for all $x \in X$; and
- (ii) $g(hx) = (gh)x$ for all $g, h \in G$ and $x \in X$.

One also says that G *acts* on X . If $|X| = n$, then n is called the *degree* of the G -set X .

First and foremost, we need to look at an issue with the mathematical content of the above. Any group can be made to act on any set (by fixing every element); equivalently, it is possible to construct an action of any group on any set. So the following:

... if there is a function $\alpha : G \times X \rightarrow X$ (called an *action*),

cannot be taken to mean what it seems to mean, in truth-conditional terms. Instead we want to say something like:

... if we have a *particular* action of G on X in mind, and have explicitly communicated this.

The key phrase that we will use to capture this requirement is 'we are considering a D'; this is taken to require not just that a D exist, but that one has been explicitly presented to the user. (Compare the treatment of ideals above: an ideal is not just a set which happens to have certain properties, but a set which has those properties and has been explicitly presented as an ideal.) Similarly, we will say that an action 'consists of ...' to make sure that the incidental presentation of some unrelated function is not taken as explicit presentation of an action.

And now to the presentation. The original combines three definitions in one sentence; as well as the main definition of ' G -set', 'action' is defined in a *parenthetical aside* and ' α ', denoted by $\alpha : (g, x) \mapsto gx$, in a relative clause. The *parenthetical aside* is unsupported by the language, and this way of

defining things via some use of function notation (involving \mapsto) was not defined when functions were defined. So we need to separate these definitions out. We also need to amend $|X| = n$, as only the LHS of an equation can introduce a fresh variable.

Definition. If X is a **set** and G is a **group**, then an **action** of G on X consists of a **function** $\alpha : G \times X \rightarrow X$ such that

- (i) $\alpha(1, x) = x$ for all $x \in X$; and
- (ii) $\alpha(g, \alpha(h, x)) = \alpha(gh, x)$ for all $g, h \in G$ and $x \in X$.

If we are considering an **action** of G on X , then we say that X is a **G -set** or that G **acts** on X ; and, given $g \in G$ and $x \in X$, we denote $\alpha(g, x)$ by gx . If $n = |X|$, then n is called the **degree** of the G -set X .

The following examples will actually be used in the proofs of important results:

Example 3.4. Every **group** G **acts** on itself by **conjugation**.

Example 3.5. Every **group** G **acts** on the **family** of all its **subgroups** by **conjugation**.

‘conjugation’ is not a defined mathematical term — so we will need to rewrite the examples to actually define it. (Or rather, to define ‘action by conjugation’.)

Definition. The **action** of a **group** G on itself by **conjugation** is the **map** $\alpha : G \times G \rightarrow G$ defined by

$$\alpha(g, x) = gxg^{-1} \quad \text{for all } g, x \in G$$

. If we are considering the **action** of G on itself by **conjugation**, then we say that G **acts** on itself by **conjugation**.

Definition. Let G and let F be the **family** of **subgroups** of G . The **action** of G on F by **conjugation** is the **map** $\alpha : G \times F \rightarrow F$ defined by

$$\alpha(g, X) = gXg^{-1} \quad \text{for all } g \in G, X \in F$$

. If we are considering the **action** of G on F by **conjugation**, then we say that G **acts** on F by **conjugation**.

Two important definitions follow; the first is:

There are two **fundamental aspects** of a G -set.

Definition. If X is a G set and $x \in X$ then the G orbit of x is

$$\mathcal{O}(x) = \{gx : g \in G\} \subset X$$

One often denotes the orbit $\mathcal{O}(x)$ by Gx . Usually we will say orbit instead of G orbit. The orbits of X form a partition indeed the relation $x \equiv y$ defined by $y = gx$ for some $g \in G$ is an equivalence relation whose equivalence classes are the orbits.

First, there is a subtle point of content here. We have defined an ‘orbit of x ’, but later refer to ‘the orbits of X ’; we need to modify the definition to make the connection between these.

As for the presentation: the language forbids definitions where definiens and definiendum are separated only by an equals-sign (essentially as a safety mechanism). So we will need to move ‘ $\mathcal{O}(x)$ ’ into some bracketing textual clause. This in turn makes ‘is $\{gx : g \in G\} \subset X$ ’ ungrammatical (in the formal register of maths used in textbooks); we need to separate the ‘ $\subset X$ ’ into a separate assertion. Making these changes also allows us to move the definitions of ‘ Gx ’ and ‘orbit’ from the comment into the main definition.

Apart from this, we only need to make the definition of the equivalence relation more formal, by binding x, y , etc, and comment the word ‘indeed’, which is outside the fragment vocabulary.

[There are two fundamental aspects of a G -set.]

Definition. If X is a G -set and $x \in X$, then the orbit or G -orbit of x , denoted by $\mathcal{O}(x)$ or Gx , is

$$\{gx : g \in G\}.$$

The orbit of x is a subset of X . A subset of X is an orbit of X if it is the orbit of some $x \in X$.

The orbits of X form a partition; indeed, the relation \equiv on X defined by

$$\text{given } x, y \in X, x \equiv y \text{ if and only if } y = gx \text{ for some } g \in G$$

is an equivalence relation whose equivalence classes are the orbits.

The second “fundamental aspect” is:

Definition. If X is a G -set and $x \in X$, then the *stabilizer* of x , denoted by G_x , is

$$G_x = \{g \in G : gx = x\} \leq G.$$

Again we need to need to avoid defining notation using nothing but an equals-sign, and again this forces us to move ' $\leq G$ ' into a separate statement to keep the text in a 'textbook' register.

Definition. If X is a G -set and $x \in X$, then the *stabilizer* of x , denoted by G_x , is

$$\{g \in G : gx = x\}.$$

The *stabilizer* of x is a *subgroup* of G .

We now illustrate these definitions by returning to the previous examples:

Example 3.4. If G acts on itself by *conjugation* and $x \in G$, then $\mathcal{O}(x)$ is the *conjugacy class* of x and $G_x = C_G(x)$.

Example 3.5. If G acts by *conjugation* on the *family* of all its *subgroups* and if $H \leq G$, then $\mathcal{O}(H) = \{\text{alltheconjugatesof}H\}$ and $G_H = N_G(H)$.

The only problems here are that G is not bound and the notation ' $\{\text{alltheconjugatesof}H\}$ ' is not defined. The language can in fact support this notation without difficulty, but it is rather inform and so we will replace it instead.

Example 3.4. If G acts on itself by *conjugation* and $x \in G$, then $\mathcal{O}(x)$ is the *conjugacy class* of x and $G_x = C_G(x)$.

Example 3.5. If G acts by *conjugation* on the *family* of all its *subgroups* and if $H \leq G$, then $\mathcal{O}(H)$ is the *set* of *conjugates* of H and $G_H = N_G(H)$.

The next theorem is the famous 'Orbit-Stabiliser Theorem'.

Theorem 3.19. If X is a G -set and $x \in X$, then

$$|\mathcal{O}(x)| = [G : G_x].$$

This is fine, apart from the uses square brackets and usual need to bind G .

Theorem 3.19. If G is a group, X is a G -set and $x \in X$, then

$$|\mathcal{O}(x)| = \backslash[G : G_x\backslash].$$

The proof is as follows:

Proof. If $x \in X$, let G/G_x denote the family of all left cosets of G_x in G . Define $f : \mathcal{O}(x) \rightarrow G/G_x$ by $f(ax) = aG_x$. Now f is well-defined: if $ax = bx$ for some $b \in G$, then $b^{-1}ax = x$, $b^{-1}a \in G_x$ and $aG_x = bG_x$. The function f is an injection: if $aG_x = f(ax) = f(cx) = cG_x$ for some $c \in G$, then $c^{-1}a \in G_x$, $c^{-1}ax = x$, and $ax = cx$; the function f is a surjection: if $a \in G$, then $aG_x = f(ax)$. Therefore, f is a bijection and $|\mathcal{O}(x)| = |G/G_x| = [G : G_x]$. ■

The notation ' G/G_x ' is already defined, because G_x is a subgroup of G ; attempting to reintroduce it is unnecessary and will cause confusion. The variable f cannot be introduced using 'define', as this would be ambiguous between variable introduction and notation introduction. The three sentences of the form 'P:Q' are rhetorical inversions (conclusions preceding their proofs) and need to be un-inverted. Square brackets are reserved for comments and need to be replaced with some other tokens. And finally, some instances of a need to be bound.

Proof. Let $f : \mathcal{O}(x) \rightarrow G/G_x$ be defined by $f(ax) = aG_x$ for all $a \in G$.

Claim. f is well-defined.

Proof. If $ax = bx$ for some $b \in G$, then $b^{-1}ax = x$, $b^{-1}a \in G_x$ and $aG_x = bG_x$. □

Claim. The function f is an injection.

Proof. If $aG_x = f(ax) = f(cx) = cG_x$ for some $c \in G$, then $c^{-1}a \in G_x$, $c^{-1}ax = x$, and $ax = cx$. □

Claim. The function f is a surjection.

Proof. If $a \in G$, then $aG_x = f(ax)$. □

Therefore, f is a bijection and $|\mathcal{O}(x)| = |G/G_x| = \backslash[G : G_x\backslash]$. □

The section finishes with some immediate consequences of the Orbit-Stabiliser Theorem.

Corollary 3.20. If a finite group G acts on a set X , then the number of elements in any orbit is a divisor of $|G|$.

Corollary 3.21.

- (i) If G is a finite group and $x \in G$, then the number of conjugates of x in G is $[G : C_G(x)]$.
- (ii) If G is a finite group and $H \leq G$, then the number of conjugates of H in G is $[G : N_G(H)]$.

Proof. Use Examples 3.4' and 3.5'. ■

The only issues here are the uses of square brackets and of the term 'use', which is not in the fragment vocabulary.

Corollary 3.20. If a finite group G acts on a set X , then the number of elements in any orbit is a divisor of $|G|$.

Corollary 3.21.

- (i) If G is a finite group and $x \in G$, then the number of conjugates of x in G is $\lfloor [G : C_G(x)] \rfloor$.
- (ii) If G is a finite group and $H \leq G$, then the number of conjugates of H in G is $\lfloor [G : N_G(H)] \rfloor$.

Proof. The result follows from Examples 3.4' and 3.5'. □

4 The Sylow Theorems

4.1 p-Groups

Definition. If p is a prime, then a p -group is a group in which every element has order a power of p .

The language only supports subject relative clauses, and not other kinds of relatives like 'a X in which P '.

Definition. If p is a prime, then a p -group is a group G such that every element of G has order a power of p .

We start the theory with an easy lemma.

Lemma 4.1. If G is a finite abelian group whose order is divisible by a prime p , then G contains an element of order p .

This needs no changes.

Lemma 4.1. If G is a finite abelian group whose order is divisible by a prime p , then G contains an element of order p .

The proof is easy, but long.

Proof. Write $|G| = pm$, where $m \geq 1$. We proceed by induction on m after noting that the base step is clearly true. For the inductive step, choose $x \in G$ of order $t > 1$. If $p|t$, then Exercise 2.11 shows that $x^{t/p}$ has order p , and the lemma is proved. We may, therefore, assume that the order of x is not divisible by p . Since G is abelian, $\langle x \rangle$ is a normal subgroup of G , and $G/\langle x \rangle$ is an abelian group of order $|G|/t = pm/t$. Since $p|t$, we must have $m/t < m$ an integer. By induction, $G/\langle x \rangle$ contains an element y^* of order p . But the natural map $\nu : G \rightarrow G/\langle x \rangle$ is a surjection, and so there is $y \in G$ with $\nu(y) = y^*$. By Exercise 2.14, the order of y is a multiple of p , and we have returned to the first case. ■

The macroscopic structuring of the proof is approved by induction is reflected in the use of terms such as ‘induction’, ‘base step’ and ‘inductive step’. These are undefined mathematical terms — and in the language as it stands, there is no way to introduce them. (The issue is that we do not yet know the right way to introduce them *in a library*, rather than hardwiring them into the language.) As a result, the proof needs to be restructured.

There are also more localised issues. The phrase ‘write ...’ is used for definitions, so ‘write $|G| = pm$ ’ would be taken as a definition and needs to be avoided. And the terms ‘proved’, ‘may’, ‘must’ and ‘first case’, which lie outside the fragment vocabulary, need to be replaced or removed.

Proof. $|G| = pm$ for some $m \geq 1$. If $m = 1$, the result is [clearly] true.

Claim. Suppose that whenever G' is a finite abelian group of order pm' for some positive integer $m' < m$, G' contains an element of order p . The lemma follows.

Proof. Choose $x \in G$ of order $t > 1$. If $p|t$, then Exercise 2.11 shows that $x^{t/p}$ has order p , and the lemma follows. [We may, therefore,] assume that the order of x is not divisible by p . Since G is abelian, $\langle x \rangle$ is a normal subgroup of G , and $G/\langle x \rangle$ is an abelian group of order $|G|/t = pm/t$. Since $p|t$, we [must] have $m/t < m$ an integer. [By induction,] $G/\langle x \rangle$ contains an

element y^* of order p . But the natural map $\nu : G \rightarrow G/\langle x \rangle$ is a surjection, and so there is $y \in G$ with $\nu(y) = y^*$. By Exercise 2.14, the order of y is a multiple of p , and the lemma follows from Exercise 2.11 [again]. \square

The result follows from the Principle of Strong Induction. \square

The next theorem is needed to prove Sylow's Theorems, but is also famous in its own right.

Theorem 4.2 (Cauchy, 1845). If G is a finite group whose order is divisible by a prime p , then G contains an element of order p .

The label attached to the theorem ('Cauchy, 1845') is not actually the name of the theorem, so we will comment it out.

Theorem 4.2 [Cauchy, 1845]. If G is a finite group whose order is divisible by a prime p , then G contains an element of order p .

The proof begins as follows:

One needs to be careful here. It is only the fact that there are a finite number of classes that makes the choice of the x_i straightforward; in other circumstances such choice requires the Axiom of Choice.

TEMP

Sylow's Theorems

Sylow's Theorems.

Let G be a finite group whose order is divisible by the prime p . Suppose p^m is the highest power of p which is a factor of $|G|$ and set

$$k = \frac{|G|}{p^m}.$$

Then

1. the group G contains at least one subgroup of order p^m ,
2. any two subgroups of G of order p^m are conjugate, and

3. the number of subgroups of G of order p^m is congruent to 1 modulo p and is a factor of k .

Theorem 72 (“Sylow’s Theorems”).

Let G be a finite group whose order is divisible by a prime p . Let m be the integer s.t. p^m is the largest power of p which divides $|G|$ and set

$$k = |G|/p^m.$$

Then

1. the group G contains a subgroup of order p^m ,
2. any two subgroups of G of order p^m are conjugate, and
3. the number of subgroups of G which have order p^m is congruent to 1 modulo p and is a factor of k .