EQUIVALENCE RELATIONS, QUOTIENTS, AND EXAMPLES

Ivo Terek

A quick summary on equivalence relations, quotient sets, basic properties, and some examples, and constructions.

1 Equivalence relations

Definition 1

Let X be a set. An **equivalence relation** $\sim$ on X is a relation[a] which is:

(i) **reflexive**, that is, $x \sim x$ for all $x \in X$.

(ii) **symmetric**, that is, $x \sim y$ implies $y \sim x$ for all $x, y \in X$.

(iii) **transitive**, that is, $x \sim y$ and $y \sim z$ implies $x \sim z$ for all $x, y, z \in X$.

[a] A subset $\sim$ of $X \times X$, where we write $x \sim y$ to mean $(x, y) \in \sim$.

Example 1

On the set $\mathbb{Z}$, for each $m \in \mathbb{Z}$, say that $x \sim y$ if $m \mid (x - y)$. This relation is called **congruence modulo** m, and one writes $x \equiv y \pmod{m}$ or $x \equiv_m y$ instead of $\sim$.

Example 2

Let X be the set of students taking a certain math class together, and say that $x \sim y$ if x and y got the same score on the final exam.

Example 3 (Equivalence relations given by functions)

Let X and Y be sets and $f\colon X \to Y$ be a function. Say that $x \sim y$ if $f(x) = f(y)$. The above example is a particular case of the situation described here, where f is the function "score on the final exam".

Example 4 (A tragic non-example)

Let X be the set of all people on planet Earth, and say that $x \sim y$ if x loves y. The fact that $\sim$ is not symmetric is a huge source of drama and relationship problems. And the fact that $\sim$ is not reflexive can be seen as a symptom of a disease called depression.

Definition 2

Let X be a set equipped with an equivalence relation $\sim$.

(i) The **equivalence class** of an element $x \in X$ is the set $[x]_\sim \doteq \{y \in X \mid x \sim y\}$.

(ii) The **quotient of** X **by** $\sim$ is the set $X/_\sim \doteq \{[x] \mid x \in X\}$.

(iii) The map $\pi \colon X \to X/_\sim$ given by $\pi(x) = [x]_\sim$ is called **quotient projection**.

Remark. Note that, simultaneously, we have $[x]_\sim \subseteq X$ and $[x]_\sim \in X/_\sim$.

Example 5

Consider again in $\mathbb{Z}$, congruence modulo $m \in \mathbb{Z}$. We have that the congruence class of each $k \in \mathbb{Z}$ is simply $k + m\mathbb{Z} = \{k + ma \mid a \in \mathbb{Z}\}$. The quotient set, denoted by $\mathbb{Z}/m\mathbb{Z}$, is the set

$$\mathbb{Z}/m\mathbb{Z} = \{0 + m\mathbb{Z}, 1 + m\mathbb{Z}, \ldots, (m-1) + m\mathbb{Z}\}.$$

It has m elements.

Proposition 1

Let X be a set equipped with an equivalence relation $\sim$. Then:

(a) Any two equivalence classes are either equal or disjoint.

(b) The union of all equivalence classes equals X.

In other words, $X/_\sim$ is a **partition** of X.

Proof:

(a) Take $x, y \in X$ and consider $[x]_\sim, [y]_\sim \in X/_\sim$. If $[x]_\sim \cap [y]_\sim = \varnothing$, there's nothing to prove. But if there is z in such intersection, then $x \sim z$ and $y \sim z$ together imply that $x \sim y$, meaning that $[x]_\sim = [y]_\sim$.

(b) For each $x \in X$, we have $x \in [x]_\sim$.

$\square$

So, equivalence relations give rise to partitions. The converse holds:

Proposition 2

Let X be a set and $\mathscr{P} = (P_\alpha)_{\alpha\in A}$ be a partition of X. There is a unique equivalence relation $\sim$ on X for which for all $x \in X$ and $\alpha \in A$, $x \in P_\alpha$ if and only if $[x]_\sim = P_\alpha$. In other words, $X/_\sim = \mathscr{P}$.

Proof: Let $x \in y$ if there is $\alpha \in A$ such that $x, y \in P_\alpha$. This $\sim$ is reflexive because each $x \in X$ is in some P_α. It is symmetric because $x \sim y$ says that x and y are in some P_α, so y and x are in this same P_α, leading to $y \sim x$. Finally, it is transitive because if $x \sim y$ and $y \in z$, there are $\alpha, \beta \in A$ with $x, y \in P_\alpha$ and $y, z \in P_\beta$ — in particular $y \in P_\alpha \cap P_\beta \neq \varnothing$ means that $P_\alpha = P_\beta$, so that $x, z \in P_\alpha$ leads to $x \sim z$. The rest is clear. □

Hence, there is a 1-1 correspondence between equivalence relations and partitions of X. In particular, the partition corresponding to the equivalence relation given in Example 3 is just the partition of X by inverse images under f of points in Y (called **fibers** of f). We note that if $\sim$ is any equivalence relation on X, then $\sim$ arises from this construction with the quotient projection π playing the role of f. This suggests we should explore this in more detail.

Definition 3

Let X and Y be sets, and $f\colon X \to Y$ be a function. The **set-kernel** of f is the set

$$\ker_{\mathrm{s}}(f) = \{(x,y) \in X \times X \mid f(x) = f(y)\};$$

Proposition 3 (Injectiveness equals trivial kernel — set-version)

Let X and Y be sets, and $f\colon X \to Y$ be a function. Then f is injective if and only if $\ker_{\mathrm{s}}(f) = \Delta$, where $\Delta = \{(x,x) \in X \times X \mid x \in X\}$ is the diagonal of X.

Proof: Clearly $\Delta \subseteq \ker_{\mathrm{s}}(f)$ in all cases. If f is injective, then $(x,y) \in \ker_{\mathrm{s}}(f)$ implies that $f(x) = f(y)$, so $x = y$ and thus $\ker_{\mathrm{s}}(f) = \Delta$. Conversely, if such equality holds, and we take $x, y \in X$ with $f(x) = f(y)$, then $(x,y) \in \Delta$ gives that $x = y$. □

Theorem 1

Let X be a set equipped with a equivalence relation $\sim$, Y be a second set, and $f\colon X \to Y$. If f is constant along equivalence classes of $\sim$, there is a unique function $\widetilde{f}\colon X/_\sim \to Y$ such that $\widetilde{f} \circ \pi = f$, where π is the quotient projection. In particular, we have the equality $\mathrm{Im}(f) = \mathrm{Im}(\widetilde{f})$ between images.

Proof: Define $\widetilde{f}([x]_\sim) \doteq f(x)$. This is well-defined as we assume that f is constant along equivalence classes of $\sim$, and it satisfies $\widetilde{f} \circ \pi = f$ by construction. Such relation implies that $\mathrm{Im}(f) = \mathrm{Im}(\widetilde{f})$ since π is surjective. □

Remark. We say that f has **passed to the quotient**, and think of $\widetilde{f}$ as f itself, not really as a different function.

Corollary 1 (First isomorphism theorem)

Let X and Y be sets and $f\colon X \to Y$ be a function. If $\sim$ is defined via f, then there is a unique injective function $\widetilde{f}\colon X/_\sim \to Y$ such that $\widetilde{f} \circ \pi = f$, where $\pi\colon X \to X/_\sim$ is the quotient projection. In particular, we have the equality $\mathrm{Im}(f) = \mathrm{Im}(\widetilde{f})$ between images.

Remark. When f is surjective, this establishes that $X/_\sim$ is in bijection with Y.

Proof: The function $\widetilde{f}$ exists and is unique in view of the previous theorem because f is constant on the equivalence classes of $\sim$, by definition of the latter. If we start from $\widetilde{f}([x]_\sim) = \widetilde{f}([y]_\sim)$, then $f(x) = f(y)$, which means that $x \sim y$, so $[x]_\sim = [y]_\sim$. Hence $\widetilde{f}$ is injective. □

2 On vector spaces

Let $\mathbb{K}$ be a field, V be a $\mathbb{K}$-vector space, and W be a subspace of V. There is no harm in thinking that $\mathbb{K} = \mathbb{R}$ is the field of real numbers here, it makes no difference on what will happen next.

Definition 4

Let's say that two vectors $v, v' \in V$ are **congruent modulo** W, written simply as $v \equiv v' \pmod{W}$ or $v \equiv_W v'$, if $v - v' \in W$.

Lemma 1

$\equiv_W$ is an equivalence relation.

Proof:

- $\equiv_W$ is reflexive because for all $v \in V$, $v - v = 0 \in W$ says that $v \equiv_W v$.
- $\equiv_W$ is symmetric because if $v \equiv_W v'$, then $v' - v = -(v - v') \in W$ says that $v' \equiv_W v$, as W is closed under taking opposites.
- $\equiv_W$ is transitive because if $v \equiv_W v'$ and $v' \equiv_W v''$, then
$$v - v'' = (v - v') + (v' - v'') \in W$$
says that $v \equiv_W v''$, as W is closed under addition.

□

Note that the equivalence class of $v \in V$ is the translate
$$v + W = \{v + w \mid w \in W\}.$$
Since we started with a vector space V, it would make sense to ask whether the quotient set $V/_{\equiv_W}$, simply denoted by V/W, can be made into a vector space.

Proposition 4

The maps $+\colon V/W \times V/W$ and $\cdot\colon \mathbb{K} \times V/W \to V/W$ defined by

$$(v+W) + (v'+W) \doteq (v+v') + W \quad \text{and} \quad \lambda \cdot (v+W) \doteq (\lambda v) + W$$

are well-defined and turn V/W into a vector space.

Proof: If $v_1 \equiv_W v_1'$ and $v_2 \equiv_W v_2'$, let's show that $(v_1 + v_2) \equiv_W (v_1' + v_2')$. Indeed, we have that

$$(v_1 + v_2) - (v_1' + v_2') = (v_1 - v_1') + (v_2 - v_2') \in W$$

because W is closed under addition. So $+$ is well-defined on V/W. As for scalar multiplication, keeping the above notation and assumptions, let's just show that the equivalence $\lambda v_1 \equiv_W \lambda v_1'$ holds. This happens because

$$\lambda v_1 - \lambda v_1' = \lambda(v_1 - v_1') \in W,$$

as W is closed under scalar multiplication. Hence $\cdot$ is well-defined on V/W. As for the algebraic axioms that $+$ and $\cdot$ must satisfy, they're all trivial consequences of the fact that the axioms already hold for the operations on V. For example:

$$(v+W) + (v'+W) = (v+v') + W = (v'+v) + W = (v'+W) + (v+W),$$

so $+$ is commutative on V/W. The zero vector is, obviously, $0 + W$. □

Remark. $V/\{0\} \cong V$ (via $v \mapsto v + \{0\}$) and $V/V = \{0 + V\}$.

Corollary 2

The quotient projection $\pi\colon V \to V/W$ is a surjective linear map with kernel W.

Proof: By design. □

Remark. If one already knows the rank-nullity theorem, applying it to π yields the dimension relation $\dim V = \dim W + \dim(V/W)$. When the dimensions are finite, it makes sense to write $\dim(V/W) = \dim V - \dim W$. If one does not want to assume (for the sake of the presentation) that the rank-nullity theorem holds yet, we'll establish it with quotients in what follows.

As a consequence of what we have seen before, abstractly, we have the:

Theorem 2 (First isomorphism theorem)

Let $T\colon V \to W$ be a linear map. Then T passes to the quotient as an injective linear map $\widetilde{T}\colon V/\ker T \to W$, showing that $V/\ker T \cong \operatorname{Im}(T)$.

Corollary 3

Write $V = W \oplus W'$ for some complementary subspace W' to W. Then $V/W \cong W'$. In particular, $\dim V = \dim W + \dim(V/W)$.

Proof: Since $V = W \oplus W'$, we have two projection operators $\mathrm{pr}_W \colon V \to W$ and $\mathrm{pr}_{W'} \colon V \to W'$. Applying the first isomorphism theorem to $\mathrm{pr}_{W'}$ (which is surjective with kernel W) yields $V/W \cong W'$. The dimension relation follows from the direct sum decomposition, which implies that $\dim V = \dim W + \dim W'$, and we use $\dim W' = \dim(V/W)$. □

Remark. Note that $\mathrm{pr}_{W'}$ morally corresponds to $\pi|_{W'}$. The restriction of a surjective linear map to any subspace complementary to its kernel is, in fact, an isomorphism.

In practice, it is good to know how to find bases for quotient spaces.

Proposition 5 (Quotient basis algorithm)

Assume that $(e_1, \ldots, e_n)$ is a basis for V which is adapted to W, in the sense that the subcollection $(e_1, \ldots, e_k)$ is a basis for W (in other words, we complete a basis for W to a basis for V). Then

$$(e_{k+1} + W, \ldots, e_n + W)$$

is a basis for V/W.

Proof: Note that π sends $(e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)$ to

$$(0 + W, \ldots, 0 + W, e_{k+1} + W, \ldots, e_n + W).$$

Since π is surjective, the above set spans V/W (even though it is linearly dependent, as it has zeros, which must be removed). It remains to show that the surviving vectors $(e_{k+1} + W, \ldots, e_n + W)$ are linearly independent in V/W. This is done as follows: start with $a_{k+1}, \ldots, a_n \in \mathbb{K}$ such that

$$a_{k+1}(e_{k+1} + W) + \cdots + a_n(e_n + W) = 0 + W.$$

The goal is to show that $a_{k+1} = \cdots = a_n = 0$. Reorganize this linear combination, using the definition of quotient operations, as

$$(a_{k+1}e_{k+1} + \cdots + a_n e_n) + W = 0 + W,$$

so that $a_{k+1}e_{k+1} + \cdots + a_n e_n \in W$. This means that there are $b_1, \ldots, b_k \in \mathbb{K}$ such that

$$a_{k+1}e_{k+1} + \cdots + a_n e_n = b_1 e_1 + \cdots + b_k e_k,$$

simply because $(e_1, \ldots, e_k)$ is a basis for W. Now linear independence of the original basis for V together with the relation

$$-b_1 e_1 - \cdots - b_k e_k + a_{k+1}e_{k+1} + \cdots + a_n e_n = 0$$

implies that $b_1 = \cdots = b_k = a_{k+1} = \cdots = a_n = 0$, as required. □

Remark. The result still holds for infinite bases, with the same argument. Namely, the procedure for finding a basis for V/W goes as follows: start with a basis for W, complete it to a basis for V, apply π to everyone. The surviving elements in the quotient will form a basis for it. Alternatively, based on the previous result, one can just take any basis for a subspace of V complementary to W, and project it using π — the resulting collection of vectors will necessarily be a basis for V/W.

The next two results are also quick consequences of the first isomorphism theorem:

Theorem 3 (Second isomorphism theorem)

Let $W_1, W_2 \subseteq V$ be subspaces. Then

$$\frac{W_1 + W_2}{W_1} \cong \frac{W_2}{W_1 \cap W_2}.$$

Proof: The linear map $W_2 \to (W_1 + W_2)/W_1$ taking $w_2 \mapsto w_2 + W_1$ is surjective (take $v + W_1 \in (W_1 + W_2)/W_1$, write $v = w_1 + w_2$ with $w_1 \in W_1$ and $w_2 \in W_2$, and note that $w_2 \mapsto v + W_1$) and has kernel $W_1 \cap W_2$. $\square$

Theorem 4 (Third isomorphism theorem)

Let $Z \subseteq W \subseteq V$ be a chain of subspaces. Then

$$\frac{V/Z}{W/Z} \cong \frac{V}{W}.$$

Proof: The linear map $V/Z \to V/W$ taking $v + Z \mapsto v + W$ is well-defined, surjective, and has kernel W/Z. $\square$

2.1 Duals and annihilators

Let V be a vector space. Recall that

$$V^* = \{f \colon V \to \mathbb{K} \mid f \text{ is linear}\}$$

is the **dual space** to V. If $(e_1, \ldots, e_n)$ is a basis for V, then the linear functionals $e^1, \ldots, e^n \colon V \to \mathbb{K}$ defined by setting $e^i(e_j) = \delta^i_j$ for all $i, j = 1, \ldots, n$ for a basis for V^*. Now let W be a subspace of V.

Definition 5

The **annihilator (or polar space) of** W, denoted either by $\operatorname{Ann}(W)$ or W°, is defined by $W^\circ = \{f \in V^* \mid f[W] = 0\}$. In other words, $f \in W^\circ$ if and only if $f(w) = 0$ for all $w \in W$.

Clearly W° is a subspace of V^*. To understand it better, let's start with some geometric intuition. There is a natural evaluation pairing $V^* \times V \ni (f, v) \mapsto f(v) \in \mathbb{K}$. Symmetry doesn't quite make sense, but people usually think of this as an "inner product" taking elements from different spaces, and even write $f(v)$ as $\langle f, v\rangle$ (this is particularly common in quantum mechanics). The point is that W° is what the "orthogonal complement" of W is supposed to be. But talking about "orthogonal complements" doesn't really make sense, as V is not actually equipped with an inner product. So W° pays the price for our little transgression and is exiled to V^* — it cannot naturally live in V without a metric. It has properties similar to orthogonal complements.

Proposition 6

(a) $\dim W^* + \dim W^\circ = \dim V^*$ (when $\dim V < \infty$, we can drop the duals).

(b) $(W_1 + W_2)^\circ = W_1^\circ \cap W_2^\circ$.

(c) $(W_1 \cap W_2)^\circ = W_1^\circ + W_2^\circ$.

Proof:

(a) The map $V^* \to W^*$ given by $f \mapsto f|_W$ is linear, surjective (why?), and has kernel W°. By the rank-nullity theorem, we have $\dim V^* = \dim W^\circ + \dim W^*$.

(b) If f annihilates both W_1 and W_2, and hence sums of elements in W_1 and W_2, so this shows that $W_1^\circ \cap W_2^\circ \subseteq (W_1 + W_2)^\circ$. Conversely, use that taking $^\circ$ reverses inclusions (why?), so $W_1 \subseteq W_1 + W_2$ implies that $(W_1 + W_2)^\circ \subseteq W_1^\circ$, similarly for W_2, so we may take the intersection to obtain $(W_1 + W_2)^\circ \subseteq W_1^\circ \cap W_2^\circ$, as required.

(c) Exercise.

$\square$

Corollary 4

$W^* \cong V^*/W^\circ$.

With this in place, let's see how to find bases for annihilators (at least in the finite-dimensional case).

Proposition 7

Assume that $(e_1, \ldots, e_n)$ is a basis for V which is adapted to W, in the sense that the subcollection $(e_1, \ldots, e_k)$ is a basis for W (in other words, we complete a basis for W to a basis for V). If $(e^1, \ldots, e^n)$ denotes the dual basis in V^*, then $(e^{k+1}, \ldots, e^n)$ is a basis for W°.

Proof: If $i = k+1, \ldots, n$, since $e^i(e_j) = 0$ for $j = 1, \ldots k$, and those span W, it follows that e^i annihilates W. In other words, $e^{k+1}, \ldots, e^n \in W^\circ$. They are linearly independent, because they are part of a larger basis. To see that they actually span W°, one can either argue that the dimension of W° is equal to $n-k$ (so a maximal linearly independent set is a basis) or, directly take $f \in V^*$, write it as $f = \sum_{i=1}^n f_i e^i$ (with the coefficients $f_1, \ldots, f_n \in \mathbb{K}$), and use that $f \in W^\circ$ if and only if $f_1 = \cdots = f_k = 0$, so f is indeed a linear combination of the remaining functionals $e^{k+1}, \ldots, e^n$. □

3 On groups

Let G be a group and H be a subgroup of G. We write e for the identity element[1].

Definition 6

Let's say that two elements $g, g' \in G$ are **congruent modulo** H, written simply as $g \equiv g' \pmod{H}$ or $g \equiv_H g'$, if $(g')^{-1}g \in H$.

Lemma 2

$\equiv_H$ is an equivalence relation.

Proof:

- $\equiv_H$ is reflexive because for all $g \in G$, $g^{-1}g = e \in H$ says that $g \equiv_H g$.
- $\equiv_H$ is symmetric because if $g \equiv_H g'$, then $g^{-1}g' = ((g')^{-1}g)^{-1} \in H$ says that $g' \equiv_H g$, as H is closed under taking inverses.
- $\equiv_H$ is transitive because if $g \equiv_H g'$ and $g' \equiv_H g''$, then

  $$(g'')^{-1}g = (g'')^{-1}g'(g')^{-1}g \in H$$

  says that $g \equiv_H g''$, as H is closed under multiplication.

□

Note that the equivalence class of $g \in G$ is the translate (in the group setting, called a **coset**)

$$gH = \{gh \mid h \in H\}.$$

Since we started with a group G, it would make sense to ask whether the quotient set $G/_{\equiv_H}$, simply denoted by G/H, can be made into a group. Unlike what happened with vector spaces, this is not guaranteed, and we need a stronger assumption on the subgroup H.

[1]The letter e is from German, *einselement*.

Definition 7

A subgroup H of G is called **normal in G** — this is written $H \lhd G$ — if for all $g \in G$ and $h \in H$, we have $ghg^{-1} \in H$.

Remark. If G is abelian, then every subgroup is normal. In particular, this applies when we have a vector space V considered as an abelian group with addition of vectors — vector subspaces are additive subgroups, and thus normal. There are non-abelian groups whose subgroups are all normal. These are called **Hamiltonian groups** (the name is unrelated to Hamiltonian dynamics and symplectic geometry). Here's one example: $Q_8 = \{1, \pm\mathrm{i}, \pm\mathrm{j}, \pm\mathrm{k}\}$, with operations summarized by $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1$ and $\mathrm{ij} = \mathrm{k}$, $\mathrm{jk} = \mathrm{i}$ and $\mathrm{ki} = \mathrm{j}$.

Proposition 8

If $H \lhd G$, then $\cdot\colon G/H \times G/H \to G/H$ given by

$$(gH) \cdot (g'H) \doteq (gg')H$$

is well-defined and turns G/H into a group.

Proof: Exercise/maybe later. Note that the identity of G/H is eH and that inverses are given by $(gH)^{-1} = g^{-1}H$. □

Remark. Many properties for G pass to G/H. For example, if G is abelian, so will be G/H. Also note that $G/\{e\} \cong G$ (via $g \mapsto g\{e\}$) and $G/G = \{eG\}$.

Replacing linear maps with group homomorphisms, we can mimic much of what was done before.

Corollary 5

The quotient projection $\pi\colon G \to G/H$ is a surjective group homomorphism with kernel H.

Theorem 5 (First isomorphism theorem)

Let $\varphi\colon G \to H$ be a group homomorphism. Then φ passes to the quotient as an injective group homomorphism $\widetilde{\varphi}\colon G/\ker\varphi \to H$, so that $G/\ker\varphi \cong \mathrm{Im}(\varphi)$.

To proceed, recall that given two subsets $A, B \subseteq G$, we may consider the set of all products, $AB = \{ab \mid a \in A, b \in B\}$. When we take A and B to be subgroups of G, AB might still not be a subgroup! However, AB is a subgroup of G if A and B are both subgroups *and* at least one of them is normal in G.

Theorem 6 (Second isomorphism theorem)

Let $H_1, H_2 \lhd G$ be normal subgroups. Then

$$\frac{H_1H_2}{H_1} \cong \frac{H_2}{H_1 \cap H_2}.$$

Proof: The homomorphism $H_2 \to (H_1H_2)/H_1$ taking $h_2 \mapsto h_2H_1$ is surjective (take $gH_1 \in (H_1H_2)/H_1$, write $g = h_2'h_1'$ with $h_1' \in H_1$ and $h_2' \in H_2$ — we're using normality to write the product in the reverse order with possibly different elements — and note that $h_2' \mapsto gH_1$) and has kernel $H_1 \cap H_2$. □

Theorem 7 (Third isomorphism theorem)

Let $K \lhd H \lhd G$ be a chain of normal subgroups with $K \lhd G$ as well[a]. Then

$$\frac{G/K}{H/K} \cong \frac{G}{H}.$$

[a] $K \lhd H$ and $H \lhd G$ do not necessarily imply $K \lhd G$, so this has to be explictly assumed. Example?

Proof: The homomorphism $G/K \to G/H$ taking $gK \mapsto gH$ is well-defined, surjective, and has kernel H/K. □

3.1 The commutant subgroup

Let G be a group. The **commutator** of two elements $a, b \in G$ is defined to be the element $[a, b] \doteq aba^{-1}b^{-1} \in G$. The reason for the name commutator is obvious: the commutator equals e if and only if $ab = ba$. So this is measuring how far a and b are from commuting. If G is abelian, all the commutators are trivial, so this would be uninteresting. The set $\{[a, b] \mid a, b \in G\}$ of commutators is *not* a subgroup of G. But we write $[G, G]$ for the subgroup generated by such set. We call $[G, G]$ the **commutant subgroup** of G. Explictly, elements of $[G, G]$ are finite strings

$$a_1b_1a_1^{-1}b_1^{-1}a_2b_2a_2^{-1}b_2^{-1} \cdots a_kb_ka_k^{-1}b_k^{-1}$$

of commutators. To see that $[G, G] \lhd G$, it suffices to check that conjugating a single commutator yields a commutator.

Exercise 1

Show that for all $g, a, b \in G$, we have $g[a, b]g^{-1} = [gag^{-1}, gbg^{-1}]$.

So, it makes sense to consider the quotient $G/[G, G]$.

Proposition 9 (Abelianization of G)

The quotient $G/[G,G]$ is always abelian.

Proof: Let $a[G,G], b[G,G] \in G/[G,G]$. Then

$$(a[G,G])(b[G,G])(a[G,G])^{-1}(b[G,G])^{-1} = (aba^{-1}b^{-1})[G,G] = e[G,G],$$

where the very last equal sign uses $aba^{-1}b^{-1} \in [G,G]$, implies that

$$(a[G,G])(b[G,G]) = (b[G,G])(a[G,G]),$$

as required. □