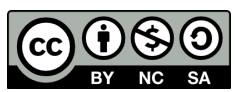
CDM

Semigroups and Groups

Klaus Sutner

Carnegie Mellon University

Spring 2025



Digression: Bourbaki

Recall Bourbaki's central idea: organize math into structures and study their properties and relationships.

As we now know, this idea works particularly well in algebra.

The axiomatic/structural style was originally pioneered by D. Hilbert; specifically for algebra the major breakthrough publication is van der Waerden's (1903-1996) classical texts that cemented the notion of algebraic structure (a first-order structure) as the fundamental concept in algebra:

B. L. van der Waerden
Moderne Algebra, Teil I
Springer Verlag, Berlin, 1930

B. L. van der Waerden
Moderne Algebra, Teil II
Springer Verlag, Berlin, 1931

> It is completely clear to me which conditions caused the gradual decadence of mathematics, from its high level some 100 years ago, down to the present hopeless nadir. Degeneration of mathematics begins with the ideas of Riemann, Dedekind and Cantor which progressively repressed the reliable genius of Euler, Lagrange and Gauss. Through the influence of textbooks like those of Hasse, Schreier and van der Waerden, the new generation was seriously harmed, and the work of Bourbaki finally dealt the fatal blow.
>
> C. L. Siegel, letter to A. Weil, 1959.

This is kind of funny, since Bourbaki was in part an attempt to create a counterweight to Hilbert's school.

More Opposition

Meanwhile, I was in a mathematics department, and this style of mathematics was not at all in fashion. Bourbaki was king: The more abstract you could be, expressing everything in terms of morphisms and categories, the better. Highly abstract methods were in favor in all the best mathematical schools. In more and more of the lectures that I was hearing at Caltech, I would find myself sitting in the audience saying to myself, "So what? So what?"

Eventually I switched fields and became a professor of computer science.

D. E. Knuth, 2014

Both Siegel and Knuth make excellent points. One needs to be careful to augment the axiomatic and semi-formal approach with lots of **intuition**, **gut-feeling**, **plausibility arguments**, **handwaving**, **pictures**, **guesswork** and so on.

Still, modern developments in symbolic computation (computer algebra, theorem provers and proof assistants) are quite closely connected to the logical presentation of mathematics.

It does take time getting used to, but the results are well worth it.

To be clear: Bourbaki won, hands down.

In a sense, that's just too bad: as far as Bourbaki is concerned, there is no computational universe, just some weird, set-theory based, logic-deprived, picture-less, non-applicable, entirely un-algorithmic wasteland.

I think this will change in the foreseeable future, mostly thanks to those pesky CS dudes and all the computing hardware/software everywhere, but right now we all still live in Bourbaki's paradise (actually, hell).

So, we will follow the axiomatic approach very closely, with a little computation added in.

An algebraic structure[†] is a logical structure

$$\mathcal{A} = \langle A; f_1, f_2, \ldots, f_k \rangle$$

where A is a non-empty set, the underlying set or carrier set.

Each of the function f_i has a fixed, finite arity n_i, so

$$f_i : A^{n_i} \to A$$

and is often rererred to as an operation in this context.

The list of arities $(n_1, n_2, \ldots, n_k)$ is the signature or type of the structure.

[†]We avoid the term *algebra* since that has a strict technical meaning.

We do allow $n_i = 0$, in which case one usually refers to a constant.

For the suspicious, the only reasonable way to think of A^0 is to construe it as a singleton set, often written $\{*\}$. Then a map $f : A^0 \to A$ essentially just picks out one element in A.

By far the most important arities are 0 through 2; the latter case is referred to as binary or dyadic.

So far, this is really just logic, all we have is a set with a bunch of functions. The algebraic character arises from the number of special properties of these functions, properties that are familiar from arithmetic operations such as addition, subtraction, multiplication, division, reciprocals and so on.

As an example, many binary operations in algebra are required to be associative. In infix notation:

$$x * (y * z) = (x * y) * z$$

Another important property is commutativity:

$$x * y = y * x$$

On occasion, it is convenient to allow additional relations. In particular order relations are very useful. For example, the rantional numbers or the reals are naturally ordered.

Similarly it can be convenient to have a carrier set that is actually composed of several disjoint parts, a multi-sorted carrier set. The standard example are vector spaces where we have a collection of scalars together with a collection of vectors.

It is straightforward to adjust our definition to handle these cases. For the time being, though, we will stick with just functions.

What are the most trivial algebraic structures? Well . . .

$$\mathcal{A} = \langle A; f \rangle$$

where $f : A \to A$ is a single unary function on A.

In this case, the only terms we can form are $f^n(x)$. An equation then looks like

$$f^n(x) = f^m(y)$$

There really is not much algebra here, we are essentially studying iteration.

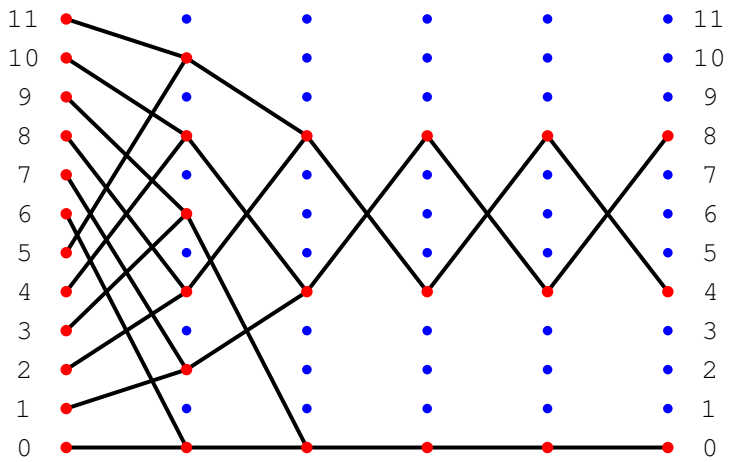
Even so, these structures are not trivial, even when A is finite and f has a simple description.

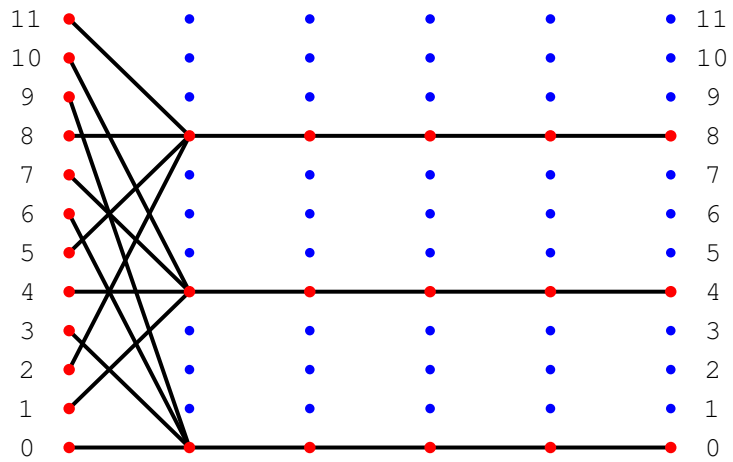
For example, consider

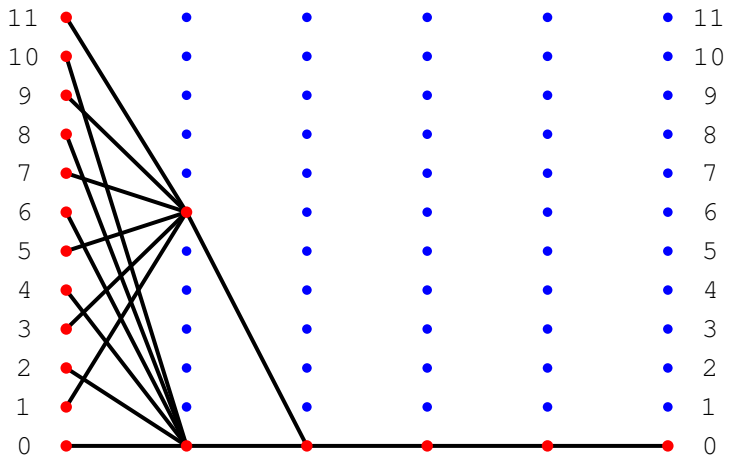
$$\mathcal{A}_{n,k} = \langle \mathbb{Z}_n;\ x \mapsto k \cdot x \bmod n \rangle$$

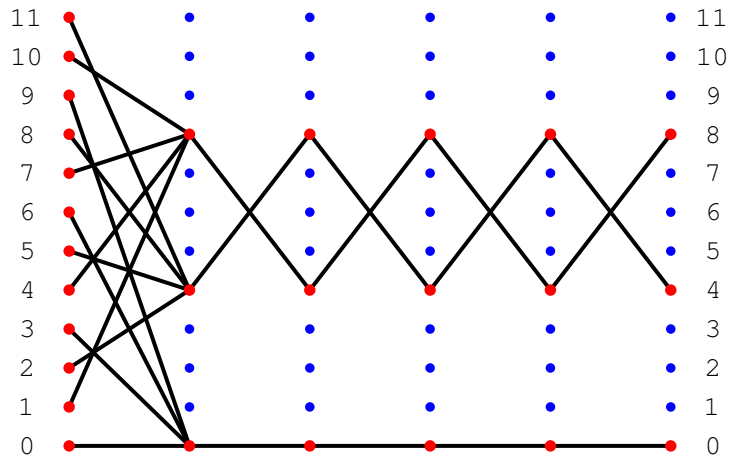
Exercise

Analyze these structures.


11
10
9
8
7
6
5
4
3
2
1
0


11
10
9
8
7
6
5
4
3
2
1
0


11
10
9
8
7
6
5
4
3
2
1
0


11
10
9
8
7
6
5
4
3
2
1
0

Slightly more complicated are structures with multiple unary operations:

$$\mathcal{A} = \langle A; f_1, \ldots, f_k \rangle$$

where all the $f_i : A \to A$ are unary operations.

The terms here are much more interesting: we can combine the given operations in an arbitrary fashion to produce expressions of the form

$$f_{e_1}\big(f_{e_2}(\ldots(f_{e_r}(x))\ldots)\big)$$

where $1 \leq e_i \leq k$.

The orbit of an element $a \in A$ is the result of evaluating all these expressions in some structure $\mathcal{A}$.

Alternatively, we have to compute the least set $B \subseteq A$ such that

- $a \in B$ and
- $x \in B$ implies $f_i(x) \in B$ for all $i = 1, \ldots, k$.

Another way of thinking about these structures is in terms of machines: in a sense, we are dealing with a DFA over a k symbol alphabet and state set A (the acceptance condition, initial and final states, are not relevant here).

Thinking of DFAs as algebras produces one of the many equivalent characterizations of regular languages.

Example: Reverse and Rotate

As a concrete example, consider the structure

$$\mathcal{A} = \langle [k]^n; L, R \rangle$$

of all binary lists of length n with operations L for rotate left and R for reverse (reverse, not rotate right) on lists of length n with elements in $[k]$ (think k colors). In combinatorics one speaks of bracelets.

Because of rotation one should think of the lists as being circular. Note that both operations are reversible, in fact

$$L^n = 1 \qquad \text{and} \qquad R^2 = 1.$$

How many different operations can one obtain by combining L and R?

The key observation is

$$R\,L = L^{n-1}\,R$$

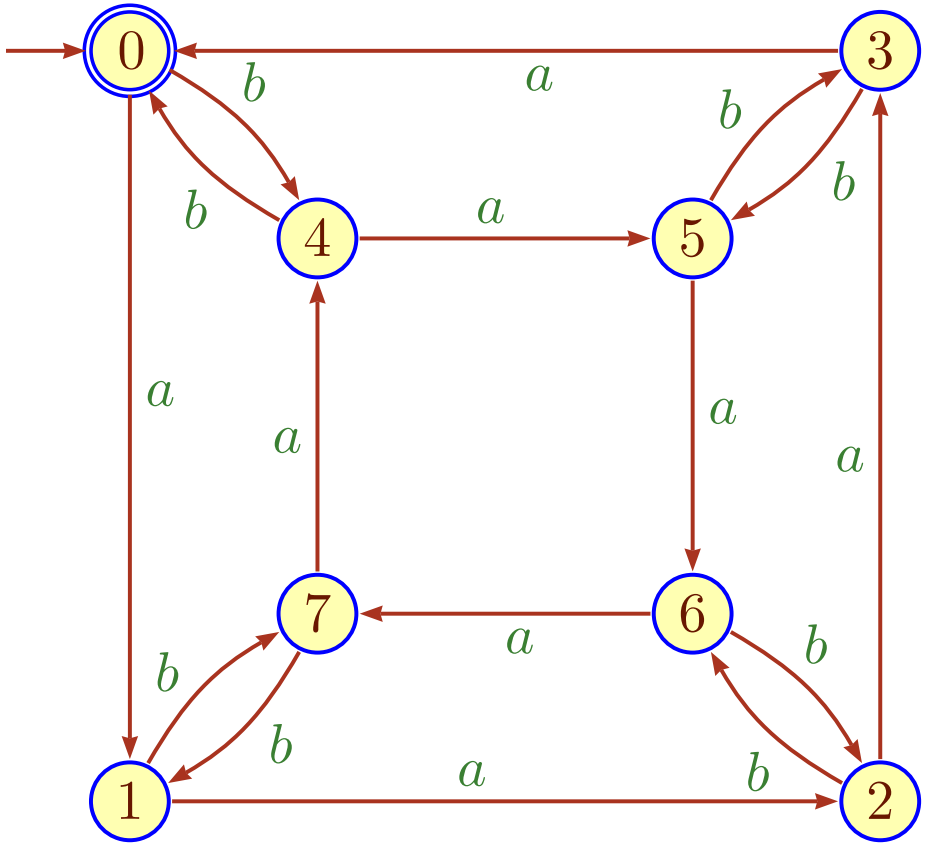
Together with the last two identities, all composite operations can be written as

$$L^l R^r(x)$$

where $0 \leq l < n$ and $0 \leq r < 2$.

So the total number is $2n$.

A DFA



Think of a as R, and b as L. Then this DFA accepts all sequences of actions that do not change any list: bb, $aaaa$, $abab$, $baba$, ...

Given a specific list $u \in \mathbf{2}^n$, what is the size of the orbit

$$\mathsf{orb}(u) = \{\, L^l R^r(u) \mid 0 \leq l < n, 0 \leq r < 2 \,\}.$$

$2n$ is an obvious, but crude upper bound. For example, $u = \mathbf{0}$ has an orbit of size 1.

We will solve this problem in detail in a while, here is just one example.

Example

For $n = 10$ the possible orbit sizes s are 1, 2, 5, 10 and 20, and their frequencies are

s	1	2	5	10	20
F_s	2	1	6	39	30

The following general ideas are crucial when dealing with algebraic structures:

Substructure A structure that is obtained by shrinking the carrier set and the algebraic operations.

Morphism A map from one structure to another that preserves the relevant operations and relations. Aka homomorphism.

Quotient A structure obtained by identifying some of the elements of a given structure (via a congruence, an equivalence relation that is compatible with the algebraic operations).

Product A structure that is defined over the Cartesian product of other structures, with appropriately defined operations.

For our very basic structures $\langle A; f \rangle$ these ideas are not hugely interesting, but we can still figure out how they would work.
First off, define two structures

$$\mathcal{A} = \langle \mathbb{Z}; S \rangle \qquad \mathcal{B} = \langle \mathbb{N}; T \rangle$$

where the operation in both cases is $x \mapsto x{+}1$.

- $\mathcal{B}$ is a substructure of $\mathcal{A}$.
- The map $f : \mathbb{N} \to \mathbb{Z}$, $f(x) = x{+}42$, is a morphism.
- The equivalence relation on $\mathbb{Z}$, $x \equiv y$ iff $x \bmod 42 = y \bmod 42$ is a congruence and produces the quotient structure $\mathbb{Z}/(42)$, the modular number modulo 42.
- The product $\mathcal{A} \times \mathcal{A}$ has carrier set the infinite grid $\mathbb{Z} \times \mathbb{Z}$ and operation $(x, y) \mapsto (x{+}1, y{+}1)$.

Time to deal with real algebraic structures.

We will start with three basic types.

- magmas
- semigroups
- groups

Definition

A magma is a structure with a single binary operation $*$:

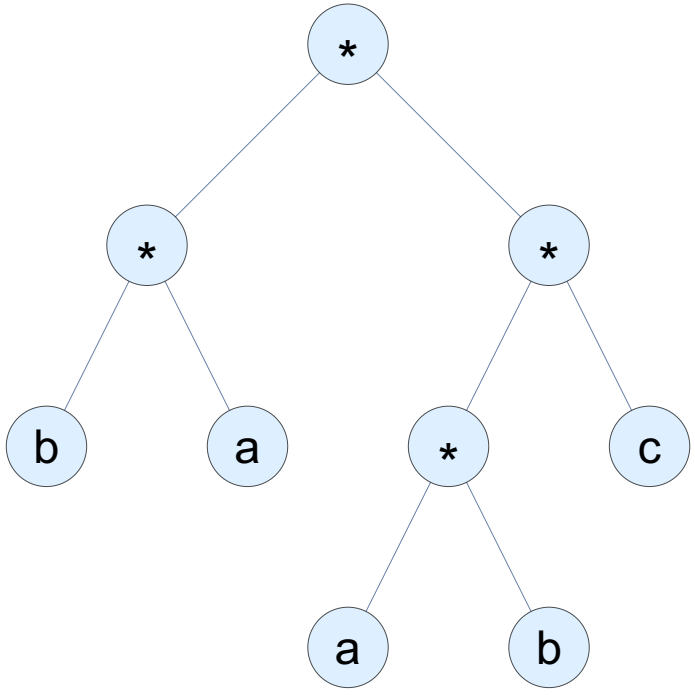
$$\mathcal{G} = \langle G; * \rangle$$

where $* : G \times G \to G$.

The operation is often referred to as multiplication.

In a magma, there are no further restrictions on the operation. Algebraists are typically not much interested in this level of generality, they want additional constraints on the operation (such as associativity).

The terms over a magma can be construed as full binary trees: the interior nodes correspond to "multiplications" and the leaves are elements of the magma.



We can evaluate these trees bottom-up eval : $\mathrm{trees} \to M$

Notation Warning

Unfortunately, magmas are sometimes referred to as "groupoids."

This is a really bad idea, since the term groupoids is defined in category theory. Roughly speaking, we can think a groupoid as a structure $\langle A; *, {}^{-1} \rangle$, a sort of groups with a partial multiplication:

- $*$ is a partial binary operation,
- $^{-1}$ is a total unary operation,
- subject to the following laws:
  - Partial associativity: if all terms are defined, then $a * (b * c) = (a * b) * c$.
  - Inverse: for all a, $a * a^{-1}$ and $a^{-1} * a$ are defined.
  - Identity: if $a * b$ is defined, so is $a^{-1} * a * b = b$ and $a * b * b^{-1} = a$.

We won't discuss groupoids here.

Some Magmas

Example

The natural numbers with exponentiation form a magma.

Example

The integers with subtraction form a magma.

Example

The positive rationals with division form a magma.

Example

Rooted binary trees can be considered as magma

$$\langle \mathcal{T}; * \rangle$$

where $\mathcal{T}$ is the collection of all rooted binary trees and $*$ denotes the operation of attaching two trees to a new root. Note that this operation is highly non-associative:

$$r * (s * t) \neq (r * s) * t$$

no matter what r, s and t are (at least in the finite case).

Example

Likewise we could consider lists over some groundset A as a magma

$$\langle \mathsf{List}(A); * \rangle$$

where $*$ is interpreted as join (concatenation). This operation is associative.

It is often helpful to translate combinatorial structures into algebraic ones. For example, suppose $\mathcal{G} = \langle V; E \rangle$ is a digraph. We can translate the graph into a magma

$$\mathcal{A}(\mathcal{G}) = \langle V_\perp; * \rangle$$

by setting $V_\perp = V \cup \{\perp\}$ where $\perp \notin V$ is a new point and

$$u * v = \begin{cases} u & \text{if } (u, v) \in E, \\ \perp & \text{otherwise.} \end{cases}$$

This operation is not associative in general.

Exercise

Figure out what left (or right) parenthesized products mean in $\mathcal{A}(\mathcal{G})$.
Is such a graph algebra commutative?

Magmas are also (mildly) helpful when dealing with more complicated structures: it is always a good idea to try to understand if a result (or even a definition) also works over magmas or whether it really requires the additional assumptions.

All the fundamental notion of sub-structure, congruence, homomorphism and so on already make sense for magmas and are perhaps a bit easier to understand there since there are no other properties lying around that can obscure the view.

Exercise

Rewrite all the definitions below in the context of magmas.

In algebra, it is interesting to understand a structure that satisfies certain specifications (i.e., a list of axioms), but has no other, special properties. These structures are called free.

So suppose we have a carrier set A. What would the free magma over A look like?

Just like the term trees from above, except that there are no free variables this time, all the leaves are labeled by elements of the ground set A.

Consider a magma $\mathcal{A} = \langle A; * \rangle$.

- A substructure of $\mathcal{A}$ consists of a set $B \subseteq A$ that is closed under $*$.
- A homomorphism from $\langle A; * \rangle$ to $\langle B; \cdot \rangle$ is a map $\varphi : A \to B$ such that $\varphi(x * y) = \varphi(x) \cdot \varphi(y)$.
- A quotient of $\langle A; * \rangle$ is given by an equivalence relation ρ such that $x \ \rho \ x'$ and $y \ \rho \ y'$ implies $x * y \ \rho \ x' * y'$ (a congruence).
- The product of $\langle A; * \rangle$ and $\langle B; \cdot \rangle$ is the structure $\langle A \times B; \otimes \rangle$ where $(x, y) \otimes (x', y') = (x * x', y \cdot y')$.

Definition

A semigroup is a magma with an associative operation.

Thus, for all x, y, z in a semigroup $\mathcal{G} = \langle G, * \rangle$ we have

$$x * (y * z) = (x * y) * z.$$

Many natural algebraic operations have this property, but not all:

- Exponentiation is not associative.
- Subtraction is not associative (but the underlying addition is).
- Graph algebras are generally not associative.

Definition

An idempotent in a semigroup is an element e such that $e * e = e$.

So an idempotent is a bit weaker than an identity.

Lemma

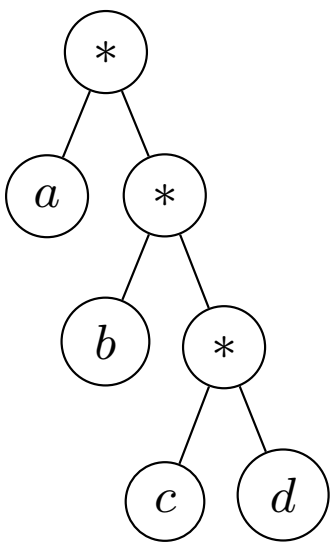
Let S be a finite semigroup. Then S contains an idempotent.

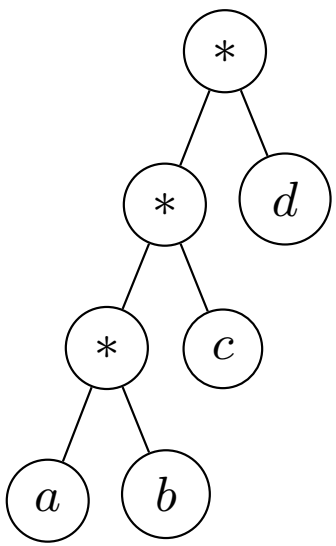
Exercise

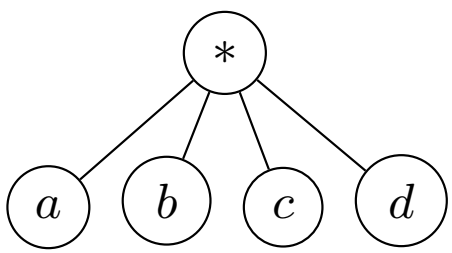
Prove the idempotent element lemma. Think Floyd.

We have seen that the free magma over A is the collection of ground terms, essentially binary trees with leaves labeled in A.

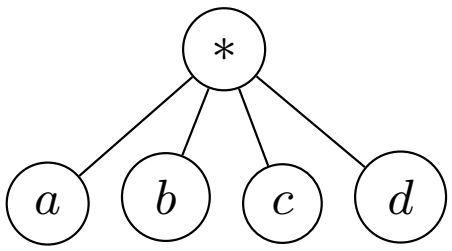
In a semigroup we have one additional specification: associativity. Hence we can identify all trees with same frontier: they correspond to the same semigroup element. Hence, we might as well think of them as a list.





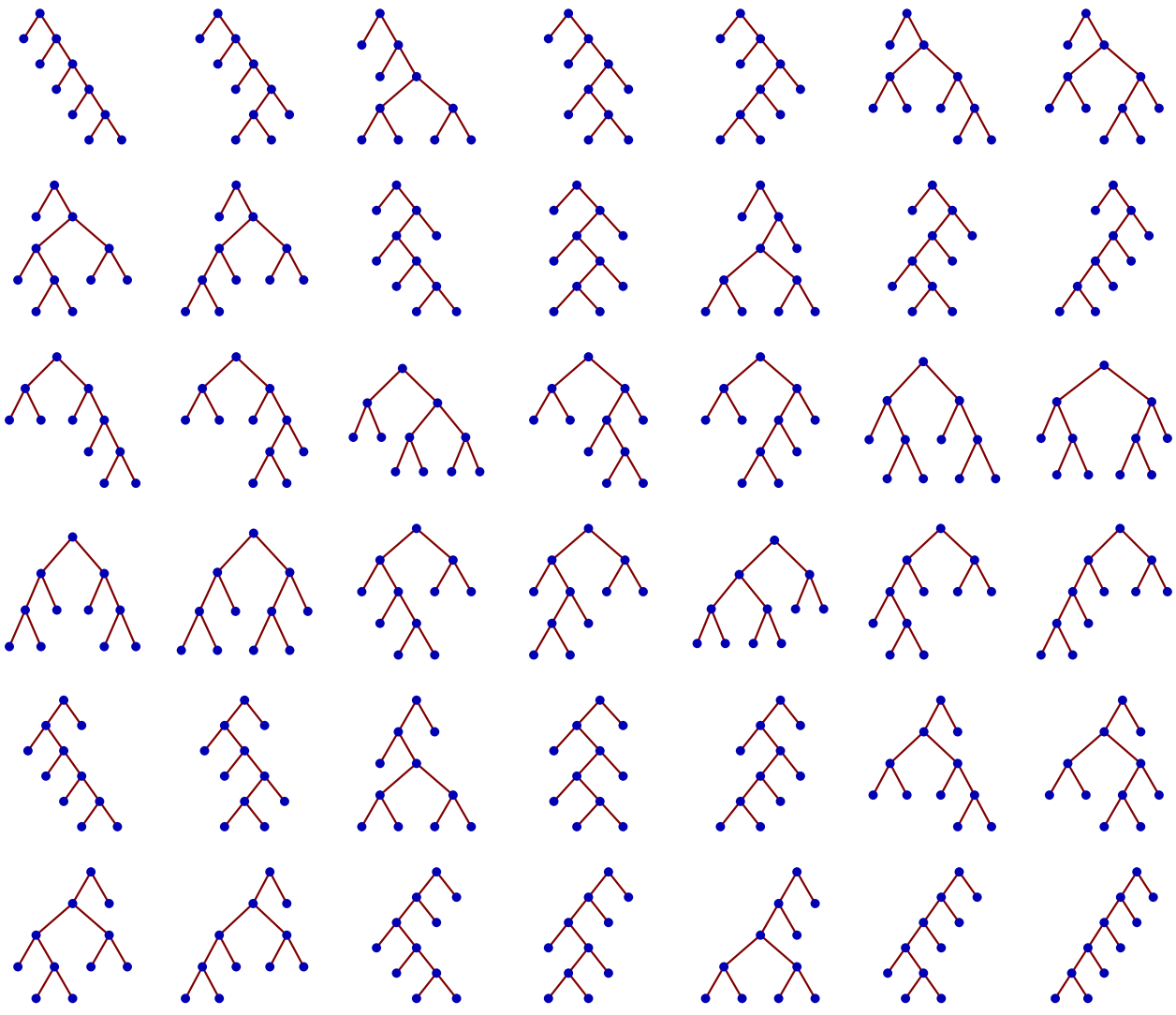


But once we think of the elements as flat lists



we might as well just use the sequence $abcd$, a word over the alphabet A.

So the set of non-empty words over A can be thought of as the free semigroup generated by A. This perspective turns out to be very useful in automata theory.



Definition

A monoid is a semigroup with an identity element e: $e * x = x * e = x$.

One usually writes monoids with signature $(2, 0)$ as

$$\mathcal{A} = \langle A; *, e \rangle$$

to indicate the neutral element. Of course, the neutral element is idempotent.

Proposition

The neutral element in a monoid is unique.

Proof. $e = e * e' = e'$. □

Example

The set of all words over a fixed alphabet forms a monoid with concatenation as operation. The neutral element is the empty word.

Example

The set of all lists over some fixed ground set forms a monoid with join as operation. The neutral element is the empty list.

Example

The set of all functions $f : A \to A$ for some arbitrary set A forms a monoid with functional composition as operation. The neutral element is the identity function.

Example

The set of all binary relations on A, for some arbitrary ground set A, forms a monoid with relational composition as operation. The neutral element is the identity relation.

Example

The set of natural numbers with addition forms a monoid; the neutral element is 0.
Ditto for integers, rationals, algebraic numbers, reals, complex numbers.

Example

The set of positive natural numbers with multiplication forms a monoid; the neutral element is 1.

Example

The set of all n by n matrices of, say, integers, with matrix multiplications forms a monoid; the neutral element is the identity matrix.

When one tries to axiomatize some sort of algebraic structure, one always has a number of civilized examples in mind. The first task is to make sure that the axioms capture all those examples.

It may very well happen, though, that the axioms have strange, unintended models that one does not really anticipate. It is well worthwhile to look around for weird examples (and possibly adjust the axioms if need be).

Example

A band is a semigroup defined on the Cartesian product $A \times B$ where A and B are two arbitrary sets (non-empty). The operation is

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

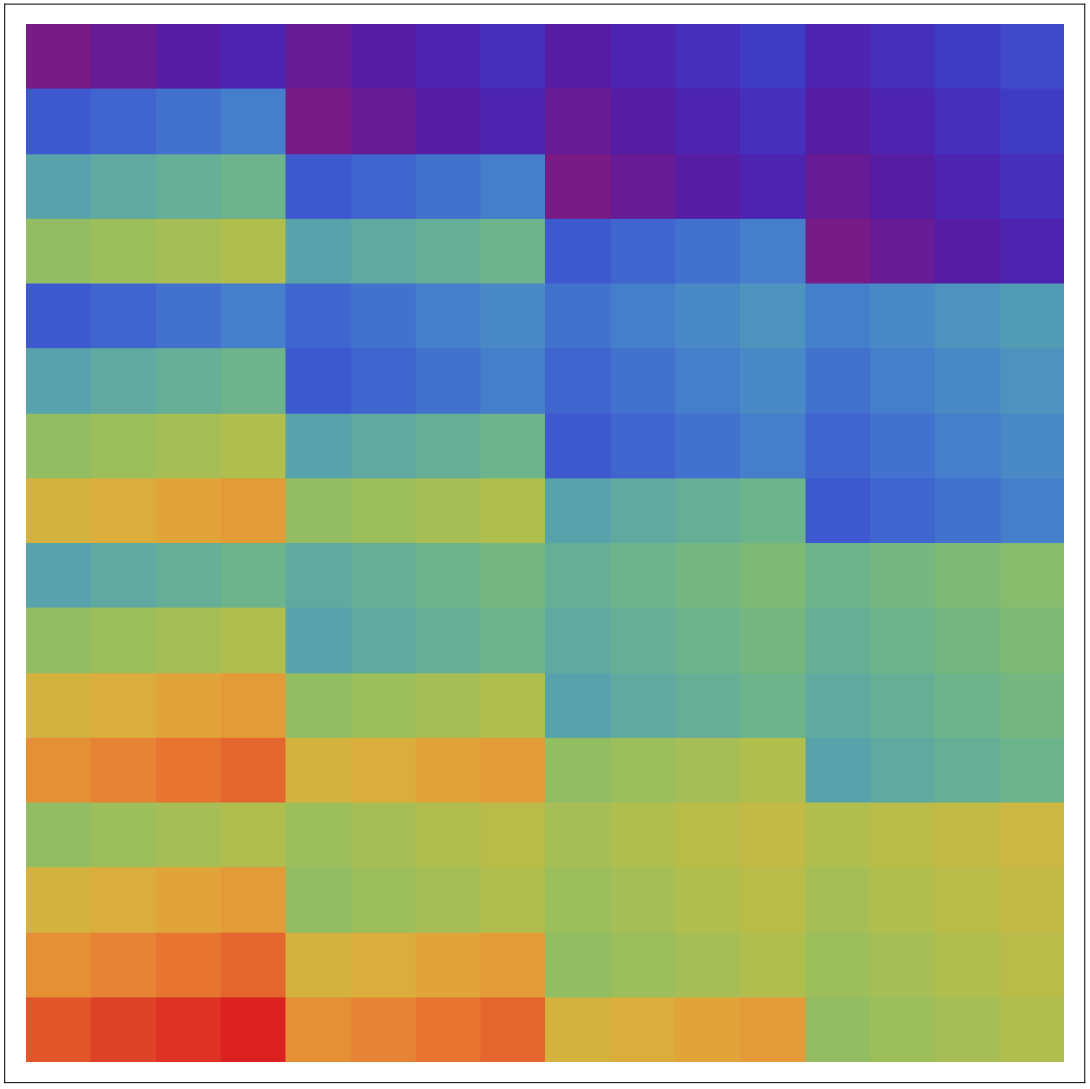
It is obvious that this operation is associative. Note that a band is idempotent: $x * x = x$ for all x.

Example

The bicyclic semigroup is defined on $\mathbb{N} \times \mathbb{N}$ by the operation

$$(a, b) * (c, d) = (a - b + \max(b, c), d - c + \max(b, c))$$

Associativity requires a little argument here. This may look strange, but it is just the free semigroup on two generators r and s subject to $sr = 1$.
The idempotents of this semigroup are exactly the elements (a, a).



Consider a semigroup $\mathcal{A} = \langle A; * \rangle$.

- A subsemigroup of $\mathcal{A}$ consists of a set $\emptyset \neq B \subseteq A$ that is closed under $*$.
- A semigroup morphism from $\langle A; * \rangle$ to $\langle B; \cdot \rangle$ is a map $\varphi : A \to B$ such that $\varphi(x * y) = \varphi(x) \cdot \varphi(y)$.
- A semigroup congruence of $\langle A; *, e \rangle$ is an equivalence relation ρ such that $x \ \rho \ u$ and $y \ \rho \ v$ implies $x * y \ \rho \ u * v$.
- The semigroup product of $\langle A; * \rangle$ and $\langle B; \cdot \rangle$ is the structure $\langle A \times B; \otimes \rangle$ where $(x, y) \otimes (u, v) = (x * u, y \cdot v)$.

The reason congruences are important is that they make it possible to form quotients.

Suppose ρ is a congruence and define the semigroup $\mathcal{A}/\rho$ by

- carrier set: $A/\rho = \{\, [x]_\rho \mid x \in A \,\}$
- operation $[x] * [y] = [x * y]$

One can check that this really produces a semigroup.

Warning: One really needs a congruence here, this does not work for an arbitrary equivalence relation (the operation is not well-defined in general).

Proposition

Suppose $f : \mathcal{A} \to \mathcal{B}$ is a monoid morphism.
Then the kernel relation $x \ \rho \ y \Leftrightarrow f(x) = f(y)$ is a congruence.

Proof.

Any kernel relation is trivially an equivalence relation. Suppose $x \ \rho \ y$ and $x' \ \rho \ y'$. Then

$$f(x * y) = f(x)f(y) = f(x')f(y') = f(x'y')$$

so that $xy \ \rho \ x'y'$. □

If you still find the congruence condition strange, here is another way to think about it: we have an equivalence relation $\rho \subseteq A \times A$ that is also a submonoid of $A \times A$ (construed as a monoid product).

Consider a monoid $\mathcal{A} = \langle A; *, e \rangle$.

- A submonoid of $\mathcal{A}$ consists of a set $\emptyset \neq B \subseteq A$ that is closed under $*$ and contains e.
- A monoid morphism from $\langle A; *, e \rangle$ to $\langle B; \cdot, e' \rangle$ is a map $\varphi : A \to B$ such that $\varphi(x * y) = \varphi(x) \cdot \varphi(y)$ and $\varphi(e) = e'$.

The only difference is that we have additional conditions regarding the neutral element. Warning: a monoid may have a subsemigroup that is not a submonoid: $\{0\}$ in $\langle \mathbb{Z}; \cdot, 1 \rangle$.

For congruences and products, nothing changes. The neutral element in $A \times B$ is (e_A, e_B).

Monoids appear quite frequently, but have one crucial flaw from the point of view of solving equations: in general we cannot solve the equation

$$a * x = b$$

As an example, consider the monoid $\langle \mathbb{N}; +, 0 \rangle$.

Then $a + x = b$ has exactly one solution whenever $a \leq b$, no solutions otherwise.

Now switch to the multiplicative monoid $\langle \mathbb{N}; \cdot, 1 \rangle$.

Then $0 \cdot x = 0$ has infinitely many solution, $2 \cdot x = 42$ has exactly one, and $2 \cdot x = 41$ has none.

Uniqueness can be ensured via the following (left) cancellation property:

$$a * x = a * y \qquad \text{implies} \qquad x = y$$

Note that this property is not equational, we need an implication between equations to express cancellation.

Exercise

Check which of the monoids from above have the cancellation property. What restrictions on the left multiplier a are necessary to guarantee cancellation?

At last, here is the kind of structure that guarantees existence and uniqueness of solutions of linear equations.

Definition

A group is a monoid $\mathcal{G} = \langle G; \cdot, e \rangle$ where for any $x \in G$:

$$\exists\, y\ (x \cdot y = y \cdot x = e)$$

The y is uniquely determined by x and called the inverse of x.

Since x uniquely determines the inverse one usually switches the signature to (2,1,0) and writes

$$\mathcal{G} = \langle G; \cdot, {}^{-1}, e \rangle$$

This has the major advantage that the group axioms are then equational, we don't need any pesky quantifiers:

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z$$
$$x \cdot e = e \cdot x = x$$
$$x \cdot x^{-1} = x^{-1} \cdot x = e$$

Actually, it turns out that we can get aways with one-sided versions of the axioms as in

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z$$
$$x \cdot e = x$$
$$x \cdot x^{-1} = e$$

Proof

Assume only the right identity and right inverses as in the one-sided axioms.

A right identity is also a left identity:

$$e \cdot x = (x \cdot x^{-1}) \cdot x = x \cdot 1 = x$$

A right inverse is also a left inverse:

$$x^{-1} \cdot x = (x^{-1} \cdot x) \cdot e = \left(x^{-1} \cdot x\right)\left(x^{-1} \cdot (x^{-1})^{-1}\right) = 1$$

In a situation like this, one often chooses the apparently weaker, if slightly cryptic, axioms.

At any rate, in a group, the equation

$$a \cdot x = b$$

always has the unique solution

$$x = a^{-1} \cdot b$$

Note that this is easier than standard arithmetic: there is no need to worry about the case $a = 0$.

The systematic study of abstract groups was one of the central accomplishments of 19th century mathematics.

They appear in many, many places and some understanding of the their basic properties is crucial.

Definition

A group is commutative or Abelian if $x \cdot y = y \cdot x$ for all x and y.

Notation:

It is customary to write Abelian groups additively as

$$\langle G; +, -, 0 \rangle \qquad \langle G; +, 0 \rangle \qquad \langle G; + \rangle$$

and general groups multiplicatively as

$$\langle G; \cdot, ^{-1}, 1 \rangle \langle G; \cdot, 1 \rangle \langle G; \cdot \rangle$$

Often one omits the multiplication operator and uses concatenation instead. xy is easier on the eye than $x \cdot y$, but this gets tricky when dealing with multiple operations.

Consider a group $\mathcal{A} = \langle A; *, {}^{-1}, e \rangle$.

- A subgroup consists of a set $\emptyset \neq B \subseteq A$ that is closed under $*$ and ${}^{-1}$.
- A group morphism from $\langle A; *, e \rangle$ to $\langle B; \cdot, e' \rangle$ is a map $\varphi : A \to B$ such that $\varphi(x * y) = \varphi(x) \cdot \varphi(y)$.

For quotients and product nothing really changes from semigroups and monoids.

Note that we dropped the condition that morphisms preserve the neutral element: this can be derived from the group axioms.

Subgroups have a simple but critical property: they make it possible to partition the group into disjoint sets of equal size. To this end, define the cosets of a subgroup $B \subseteq A$ by

$$aB = \{\, ax \mid x \in B \,\}$$

Proposition

Let $B \subseteq A$ be a subgroup.
Then the cosets aB partition A and have the same size.

Proof. Suppose $aB \cap a'B \neq \emptyset$, so $ab = a'b'$. But then $a = a'b'b^{-1} = a'b''$ and therefore $aB \subseteq a'B$. Done by symmetry.

There is a simple bijection $\widehat{a} : B \to aB : \widehat{a}(x) = ax$.

□

Example

The set of integers (rationals, reals, complexes) with addition forms a group; the neutral element is 0.

Example

The set of modular numbers relatively prime to modulus m with multiplication forms a group; the neutral element is 1.

Example

The set of non-zero rationals (reals, complexes) with multiplication forms a group; the neutral element is 1.

Example

The set of all regular n by n matrices of reals, with matrix multiplications forms a group; the neutral element is the identity matrix.

Example

The set of all permutations $f : A \to A$ for some arbitrary set A forms a group with functional composition as operation. The neutral element is the identity function.

Many important subgroups of a group G are obtained by finding the smallest subgroup that contains a given set of group elements.

Definition

Let G be a group and $A \subseteq G$. The group generated by A is the least subgroup of G that contains A. A is a set of generators for this subgroup.

In symbols: $\langle A \rangle$.

For example, in the group $\langle \mathbb{Z}, + \rangle$, 2 generates the subgroup $2\mathbb{Z}$, but $\{2, 5\}$ generates the whole group.

Abstractly we know that $\langle A \rangle$ always exists since

$$\langle A \rangle = \bigcap \{ H \subseteq G \mid A \subseteq H \text{ subgroup} \}$$

But computationally this is pretty useless: we already have to know all subgroups (containing A) to compute the intersection. Even if G is finite, there is little hope to translate this characterization into a reasonable algorithm.

So how do we actually compute $\langle A \rangle$ from A?
In particular when G is finite?

We can construct $\langle A \rangle$ from a given set of generators A by induction.

- $H_0 = A \cup \{1\}$.
- H_{n+1} is the closure of H_n with respect to multiplication and inversion.
- Let $H = \bigcup_{n \geq 0} H_n$.

When G or at least H is finite, this is a perfectly good algorithm, at least if we don't worry about efficiency.

In the infinite case, this constuction is still logically correct, but now it needs to run for infinitely many steps. Just think about $G = \langle \mathbb{Z}, + \rangle$ and $A = \{2, 11\}$.

Let G be a group.

Definition

The order of G is the cardinality of G.

The order of an element $a \in G$ is the order of the subgroup $\langle a \rangle$.

The concept of order is most important when the cardinality is finite, though we could use the set-theoretic machinery of cardinalities. Note that the order of a group element can be at most $\aleph_0$: in general

$$\langle a \rangle = \{ a^i \mid i \in \mathbb{Z} \}$$

If the order of a is finite, we even have

$$\langle a \rangle = \{ a^i \mid 0 \leq i < k \}$$

where k is the order of a.

Laziness: any property, of, say, groups derived from only the axioms holds in all groups, automatically. You only check three simple properties, and all results apply.

Psychology: it is sometimes easier to argue abstractly than in a concrete situation (you can't see the forest because of all the trees).

This a bit hard to believe, but true. E.g., consider non-singular matrices of reals. Show that

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$$

One could try to use the properties of matrix multiplication and, say, Gaussian elimination, to prove this. Any such argument would be very hard and technically difficult.

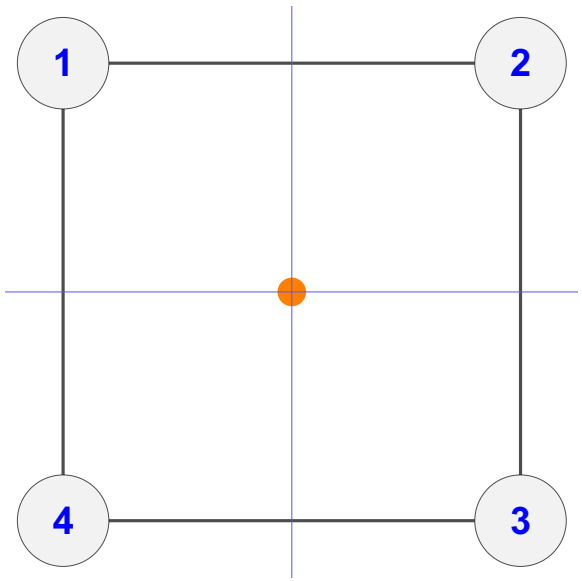
Exercise

Give a simple, abstract proof of this equation in any group whatsoever.

Felix Klein's "Erlanger Programm" of 1872 proposed to characterize geometries by studying the invariants of linear transformation groups.

Symmetries in geometry are a critical application and, in fact, the historical origin of) groups.

Consider the square with four vertices in positions $(\pm 1, \pm 1)$.



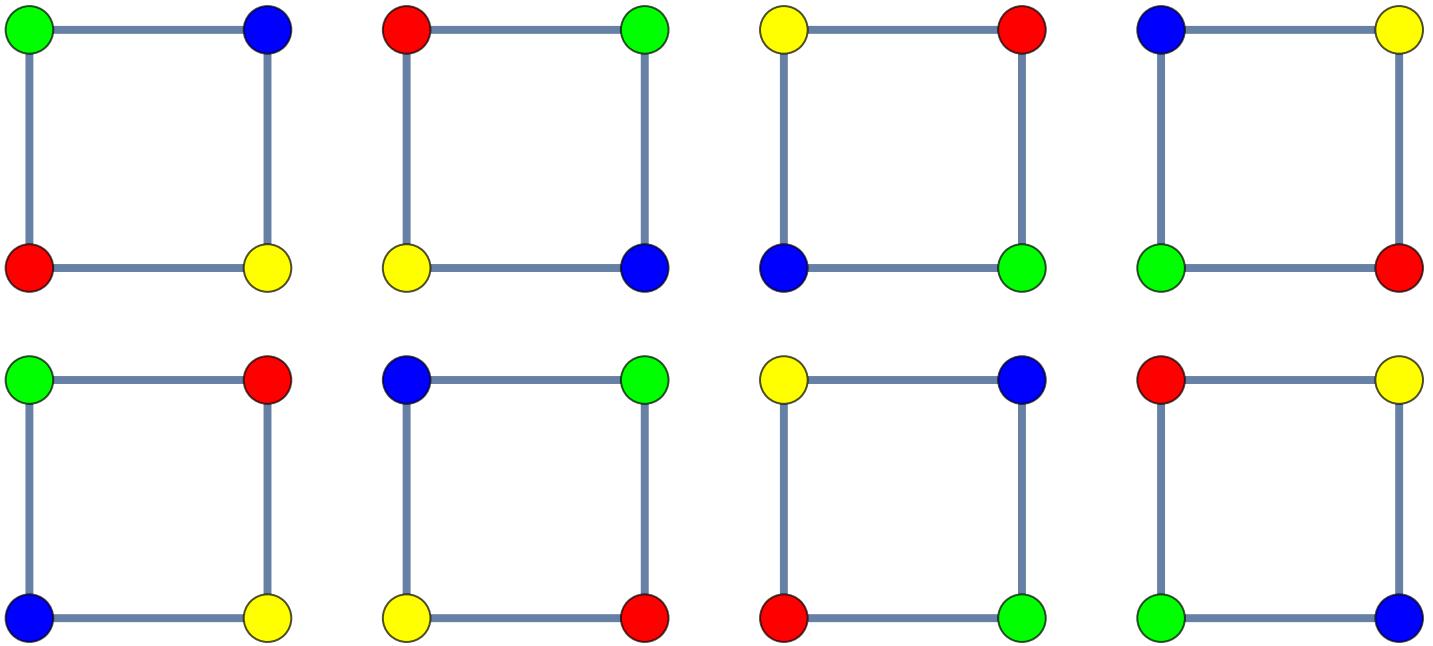
What are the rigid motions of the plane that leave the square unchanged in the sense that they place the square on top of itself?

Your geometric intuition should immediately lead to the following admissible operations.

- Rotations around the origin by multiples of $\pi/2$.
  There are essentially only 4 of these, including the trivial one.
- Reflections along the axes and diagonals.
  There are 4 of these.

There might be more, though, say a combination of a rotation and a reflection.

Let's write α for clockwise rotation by $\pi/2$ and β for reflection along the counter-diagonal.



The first row is obtained by repeated application of α.

The second row is obtained by applying β, and the α repeatedly.
Note that all other reflections appear in this row, too.

These motions naturally form a group: composition of motions is associative, the identity motion is admissible, every motion is reversible, and the composition of two admissible motions is again admissible.

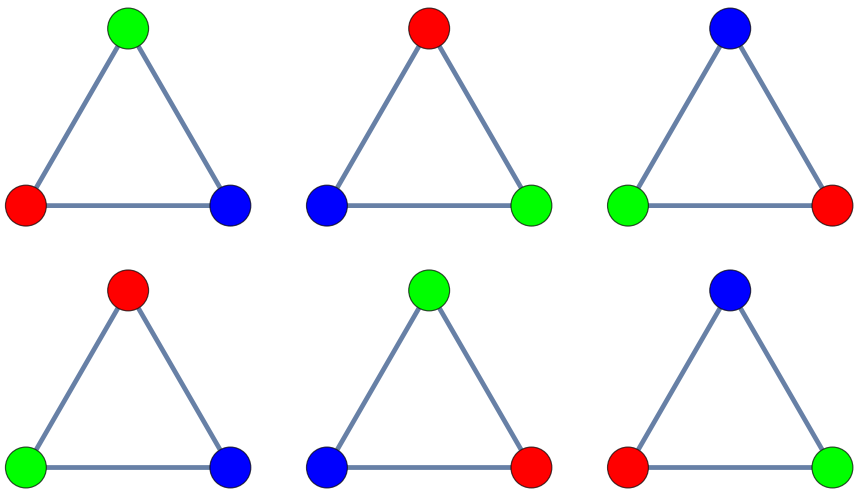
This group is called a dihedral group D_4.
It has size 8 and is generated by α and β, elements of order 4 and 2, respectively.

The important identities in this group are

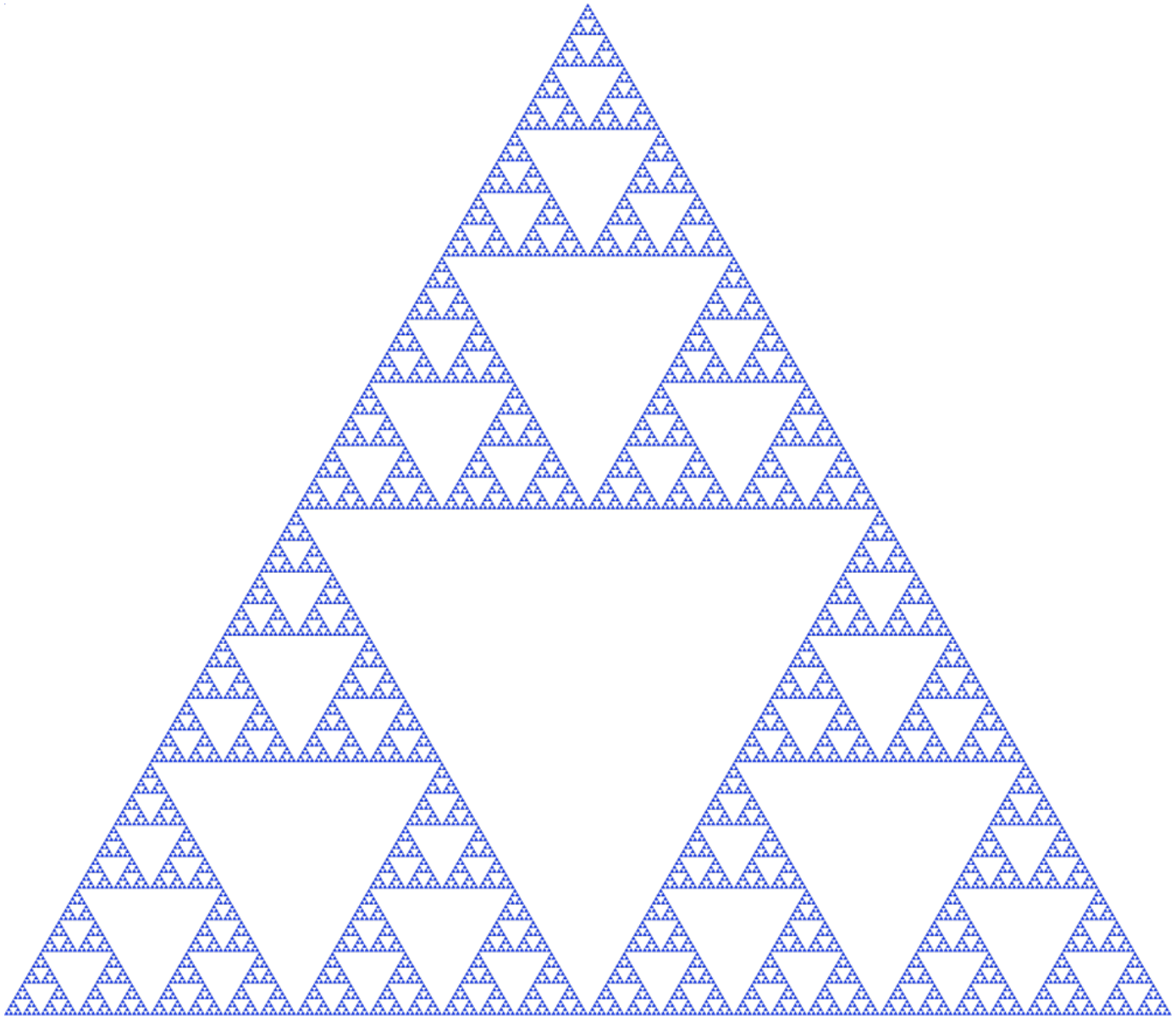
$$\alpha^4 = \beta^2 = 1 \qquad \beta\alpha = \alpha^3\beta$$

If we start with a regular n-gon instead, we get the dihedral group D_n.

For example, the dihedral group D_3 represents the symmetries of an isosceles triangle.



We can apply D_3 to more complicated sets in geometry that are somewhat similar to a plain isosceles triangle.



Alas, for the Sierpinski triangle D_3 clearly is not the whole story. There are other sorts of symmetries that are important here.

In particular, the whole figure is isomorphic to each one of the 3 sub-triangles.

And these are isomorphic to their sub-sub-triangles, and so on.

In fact, any appropriate sub-triangle in the figure is isomorphic to any other.

Obviously D_3 does not begin to capture any of this.

To handle this type of problem, we need to back off from full groups and consider weaker structures.

For any set X consider the symmetric monoid

$$\mathcal{I}(X) = \{\, f : X \nrightarrow X \mid f \text{ partial, injective} \,\}$$

Composition of these partial maps is a bit more complicated, $f \circ g$ is the largest partial map that coexists peacefully with the domain of g and the codomain of f.

Note we need to allow the empty function in $\mathcal{I}(X)$, a null element in the monoid.

Exercise

Explain why $\mathcal{I}(X)$ fails to be a group (for X non-empty).

To describe this kind of structure, one needs to relax the conditions on an inverse element in a semigroup.

Definition

Element a is said to have a generalized inverse or pseudoinverse a' if $aa'a = a$ and $a'aa' = a'$.
A semigroup is inverse if every element has exactly one generalized inverse.

$$a'' = a \qquad \text{and} \qquad (ab)' = b'a'$$

Symmetric monoids are examples of inverse semigroups.