**Section 1: (Binary) Operations**

It is assumed that you learned in Math 250 about sets and mathematical induction, the content of Section 0, and we will begin with binary operations. Technically:

**Def:** For a set $S$, any function $S \times S \to S$ is a *(binary) operation.* (Such creatures as "unary operations" exist, i.e., functions $S \to S$, as well as "ternary", $S \times S \times S \to S$, etc., but we are only interested in the binary ones, so we will usually drop the "binary" part of the term.)

So we know many, many (binary) operations on $\mathbb{R}$; for example, $(a, b) \mapsto a^3 + \sin(3b + 7)$. But of course the term is intended to capture under a single word some familiar ideas that have some things in common while still not being identical:

- Addition, multiplication and subtraction of real numbers are all operations on $\mathbb{R}$. (Division is not, because division by 0 isn't defined, so division is only a function $\mathbb{R} \times (\mathbb{R} - \{0\}) \to \mathbb{R}$.)
- Addition of $n$-vectors and cross-product of 3-vectors are operations on $\mathbb{R}^n$ and $\mathbb{R}^3$. (Dot product is not, because it is a function $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$.)
- Addition of matrices of the same dimensions, and multiplication of square matrices of the same dimension are operations on $M_{m\times n}(\mathbb{R})$ and $M_{n\times n}(\mathbb{R})$ respectively.
- "Pointwise" addition and multiplication of real-valued functions on some common set (e.g., an interval in the real line) is an operation on the set of such functions: The function $f + g$ is defined on the common domain by $(f + g)(x) = f(x) + g(x)$ for every $x$ in the domain. We might talk about these in calculus I, discussing definite integrals: For integrable functions $f, g$ on the interval $[a, b]$, $\int_a^b (f + g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx$.
- For any positive integer $n$, two operations on the set $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-2, n-1\}$ of possible remainders on long division by $n$, called "integers modulo $n$" are addition and multiplication modulo $n$: For $a, b$ in $\mathbb{Z}_n$, $a \oplus b$ and $a \odot b$ are the remainders when the integers $a + b$ and $a \cdot b$ are computed in the usual way in $\mathbb{Z}$ and then divided by $n$. (It is assumed that you learned about $\mathbb{Z}_n$ in Math 250.)
- Compositions of functions from a set $X$ to itself is an operation: If $f, g$ are functions $X \to X$, then $f \circ g$ is defined on $X$ by $(f \circ g)(x) = f(g(x))$ for every $x$ in $X$;
- On a finite set, an operation can be displayed in a table — we can even say it is defined by the table: On a set $S = \{a, b, c, d\}$,

| $*$ | $a$ | $b$ | $c$ | $d$ |
|---|---|---|---|---|
| $a$ | $a$ | $b$ | $c$ | $d$ |
| $b$ | $d$ | $c$ | $b$ | $a$ |
| $c$ | $c$ | $c$ | $d$ | $d$ |
| $d$ | $d$ | $d$ | $c$ | $c$ |

means

$$\begin{array}{llll} a*a=a & a*b=b & a*c=c & a*d=d \\ b*a=d & b*b=c & b*c=b & b*d=a \\ c*a=c & c*b=c & c*c=d & c*d=d \\ d*a=d & d*b=d & d*c=c & d*d=c \end{array}$$

In other words, the head of a given row is the first operand, the head of a column is the second operand, and the entry at that point in the table (where row and column meet) is the result.

- For a fixed set $X$, the family $S = \mathcal{P}(X)$ of all subsets of $X$ have at least three natural operations on it. (The script P is because $S$ is called the "power set" of $X$.): For subsets $A, B$ of $X$,

$$\begin{aligned} A \cup B &= \{x \in S : x \in A \text{ or } x \in B \text{ or both}\} \\ A \cap B &= \{x \in S : x \in A \text{ and } x \in B\} \\ A \triangle B &= \{x \in S : x \in A \text{ or } x \in B \text{ but not both}\} = (A \cup B) - (A \cap B) \end{aligned}$$

These are called "union," "intersection," and "symmetric difference," respectively. So if $X = \{4, 7\}$, then $\mathcal{P}(X) = \{\emptyset, \{4\}, \{7\}, X\}$, and

| $\cup$ | $\emptyset$ | $\{4\}$ | $\{7\}$ | $X$ |
|---|---|---|---|---|
| $\emptyset$ | $\emptyset$ | $\{4\}$ | $\{7\}$ | $X$ |
| $\{4\}$ | $\{4\}$ | $\{4\}$ | $X$ | $X$ |
| $\{7\}$ | $\{7\}$ | $X$ | $\{7\}$ | $X$ |
| $X$ | $X$ | $X$ | $X$ | $X$ |

| $\cap$ | $\emptyset$ | $\{4\}$ | $\{7\}$ | $X$ |
|---|---|---|---|---|
| $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ |
| $\{4\}$ | $\emptyset$ | $\{4\}$ | $\emptyset$ | $\{4\}$ |
| $\{7\}$ | $\emptyset$ | $\emptyset$ | $\{7\}$ | $\{7\}$ |
| $X$ | $\emptyset$ | $\{4\}$ | $\{7\}$ | $X$ |

| $\triangle$ | $\emptyset$ | $\{4\}$ | $\{7\}$ | $X$ |
|---|---|---|---|---|
| $\emptyset$ | $\emptyset$ | $\{4\}$ | $\{7\}$ | $X$ |
| $\{4\}$ | $\{4\}$ | $\emptyset$ | $X$ | $\{7\}$ |
| $\{7\}$ | $\{7\}$ | $X$ | $\emptyset$ | $\{4\}$ |
| $X$ | $X$ | $\{7\}$ | $\{4\}$ | $\emptyset$ |

- etc., etc.

If we really think of binary operations as functions, it is reasonable to use "Polish notation" (named for mathematicians from Poland who first used it): denoting the sum of $a$ and $b$ by $+(a, b)$. Some older calculators use "reverse Polish notation," essentially $(a, b)+$: punching "5-Enter-3-Plus" to get the sum of 5 and 3. But of course most common is "infix" notation, putting the symbol for the operation between the operands: the sum of $a$ and $b$ is denoted $a + b$. Sometimes, especially if the operation is some kind of "multiplication", we will simply denote it by juxtaposition: the "product" of $a$ and $b$ is denoted $ab$. We will do this a lot in learning about groups, because each group has only one operation, so it can't be misunderstood.

Most operations are too badly behaved to have any good properties in common, though, so we usually impose at least one or two conditions to restrict to those that act a little more like our most familiar examples. The two usual conditions are

**Def:** An operation $*$ on a set $S$ is *commutative* iff, for every two elements $a, b$ of $S$, $a * b = b * a$ (i.e., the function $*$ associates the ordered pairs $(a, b)$ and $(b, a)$ to the same element of $S$). And $*$ is *associative* iff, for all elements $a, b, c$ of $S$, $(a * b) * c = a * (b * c)$.

We know addition and multiplication of real numbers, addition of vectors and addition of matrices, pointwise addition and multiplication of functions all are both commutative and associative. Addition and multiplication modulo $n$ are commutative and associative on $\mathbb{Z}_n$ — in both cases, commutativity is easy because addition and multiplication in $\mathbb{Z}$ is commutative, and associativity is a bit more complicated to check because long division by $n$ must be done twice on both sides of the equals sign. Multiplication of matrices is associative but not commutative. Composition of functions is associative (more on this below), but it is not commutative: If $f, g : \mathbb{R} \to \mathbb{R}$ are given by $f(x) = x + 1$ and $g(x) = 2x$, then $(f \circ g)(x) = 2x + 1$ but $(g \circ f)(x) = 2x + 2$, so $f \circ g \neq g \circ f$.

Because associativity is more complicated to define, and usually also to check, it's a bit surprising that it is more basic than commutativity; but it turns out that a non-commutative operation is

somewhat inconvenient, but a non-associative operation is a mess to work with. The usual example of a non-associative operation is the cross-product of 3-vectors:

$$(\mathbf{i} \times \mathbf{i}) \times \mathbf{j} = \mathbf{0} \times \mathbf{j} = \mathbf{0} \qquad \text{but} \qquad \mathbf{i} \times (\mathbf{i} \times \mathbf{j}) = \mathbf{i} \times \mathbf{k} = -\mathbf{j} \ .$$

But a simpler non-example is subtraction of real numbers: $(3-2)-1 = 0$ but $3-(2-1) = 2$. That is why we usually think of subtraction, not as an operation in its own right, but as adding the negative. (More on that later, too.)

Commutativity is very nice when it is available, but we know that matrix multiplication is not commutative (though it *is* associative), so sometimes we make do without it. Unlike associativity, though, at least commutativity is easy to check from an operation table, just by looking for symmetry about the main diagonal (upper left to lower right) — assuming that the column heads and row heads are in the same order. For example, take two operations on the set $S = \{a, b, c\}$:

| $*$ | $a$ | $b$ | $c$ |
|---|---|---|---|
| $a$ | $a$ | $b$ | $a$ |
| $b$ | $b$ | $c$ | $b$ |
| $c$ | $a$ | $b$ | $a$ |

| $\circ$ | $a$ | $b$ | $c$ |
|---|---|---|---|
| $a$ | $a$ | $b$ | $c$ |
| $b$ | $b$ | $c$ | $b$ |
| $c$ | $a$ | $b$ | $c$ |

Of these, $\circ$ is not commutative ($a \circ c = c$ but $c \circ a = a$), while $*$ is commutative, by the symmetry of the table (though $*$ is not associative: $(b * b) * a = c * a = a$ but $b * (b * a) = b * b = c$.)

Associativity does hold "naturally" if the operation is itself, or is derived from, a function composition, because function compositions are clearly associative: $((f \circ g) \circ h)(x) = f(g(h(x))) = (f \circ (g \circ h))(x)$ — on both ends $h$ is applied to $x$, then $g$ is applied to $h(x)$, then $f$ is applied to $g(h(x))$, so the results are identical. As an example of what I mean by "derived from" a function composition, consider matrix multiplication, which is related to applying linear transformations: We can check that every linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ is given by a rule of the form $T(\begin{pmatrix} x \\ y \end{pmatrix}) = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$, so $T$ is multiplication of each vector by a fixed matrix:

$$T(\begin{pmatrix} x \\ y \end{pmatrix}) = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = B \begin{pmatrix} x \\ y \end{pmatrix} \ , \text{ say, where } B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} .$$

In linear algebra, $B$ was called the "matrix representation of $T$" (with respect to the standard basis). If $A, C$ are the matrix representations of the linear transformations $S, U : \mathbb{R}^2 \to \mathbb{R}^2$, then for every $\begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$,

$$\begin{aligned} ((AB)C) \begin{pmatrix} x \\ y \end{pmatrix} &= ((S \circ T) \circ U)(\begin{pmatrix} x \\ y \end{pmatrix}) \\ &= S(T(U(\begin{pmatrix} x \\ y \end{pmatrix}))) \\ &= (S \circ (T \circ U))(\begin{pmatrix} x \\ y \end{pmatrix}) \\ &= (A(BC)) \begin{pmatrix} x \\ y \end{pmatrix} \ , \end{aligned}$$

and because this works for every vector in $\mathbb{R}^2$, we get $(AB)C = A(BC)$. So matrix multiplication is associative *because* it reflects composition of linear transformations, which is "naturally" associative.

**Def:** If $S$ is a set and $*$ is an associative operation on $S$, then the pair $(S, *)$ (or sometimes just $S$, if there is a natural choice for $*$) is called a *semigroup.*