What is the source of the commutativity of the basic arithmetical operations?

Our recognition of the commutativity of the basic arithmetical operations of addition and multiplication can be seen as arising from the mind's grasp of the invariance of the content, or size, of discrete assemblages under simple combinatorial and geometric procedures.

In the case of addition, there is in the first instance a straightforward linear account of what seems to be going on. Fundamentally, *addition is juxtaaposition*. In adding two natural numbers m and n one first represents them in as sequences of dots **m, n** and then juxtaposes the results in the given order to obtain a sequence **m** ★ **n** of $m + n$ dots. Thus, for example, in calculating 2 + 3 = 5 one starts with 2 = $\boxed{\bullet\bullet}$ objects ("dots") and then juxtaposes 3 = $\boxed{\bullet\bullet\bullet}$ dots; this yields the assemblage

$$\boxed{\bullet\bullet\quad\bullet\bullet\bullet}$$

in which the "gap" between the two components $\boxed{\bullet\bullet}$, $\boxed{\bullet\bullet\bullet}$ reflects the order in which the operation of juxtaposition has been performed. This assemblage is a perfect iconic representation of 2 + 3, in which the `gap`` is, so to speak, waiting to be closed (through addition). Closing the gap naturally yields the homogeneous assemblage of dots

$$\boxed{\bullet\bullet\bullet\bullet\bullet}$$

which can then be counted smoothly as 5 *independently* of how it was assembled to begin with: its combinatorial origins have, so to speak, been effaced.

Now in particular, one could have started with 3 dots and then juxtaposed 2 dots, so obtaining

$$\boxed{\bullet\bullet\bullet}\ \boxed{\bullet\bullet}$$

Closing the gap in this case yields what on inspection is seen to be precisely the same result as before, namely

$$\boxed{\bullet\bullet\bullet\bullet\bullet}$$

Another way of drawing the same conclusion is to note that if the assemblage

$$\boxed{\bullet\bullet}\quad\boxed{\bullet\bullet\bullet}$$

is reflected in a mirror (i.e., rotated through 180°) it is transformed into the assemblage

$$\boxed{\bullet\bullet\bullet}\quad\boxed{\bullet\bullet}$$

When the gap is closed in the "real" world, and so also in the "mirror" world, the "gapless" result

$$\boxed{\bullet\bullet\bullet\bullet\bullet}$$

is the *same* in both, since homogeneous sequences of dots are invariant under reflections. In the "real" world, the result, 5, has been obtained by adding 2 and 3, while in the "mirror" world, the same result, again 5, has been obtained by adding 3 and 2. The invariance of the result $\boxed{\bullet\bullet\bullet\bullet\bullet}$ **= 5** under reflections then leads to the conclusion that 2 + 3 = 3 + 2.

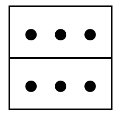
Presented in this way, by "closing the gap" the commutative law of addition *for small numbers presentable in the forms of discrete dots* becomes clear. The extension of the commutative law of addition to *arbitrary* numbers is then made on the basis of *generalization*, namely, that what holds for small numbers holds for all of them.

The analysis of the commutative law using reflections has the advantage of being *completely independent* of the size of the numbers (of dots) involved, and so is not based on generalization from special cases. If one were presented with the sum 1233495 + 23956, say, and claimed, *without calculating*, that the result was the same as 23956 + 1233495, one would simply argue that, if the sums involved are represented as "gapped" dot sequences, then the mirror image of the first sum is the second sum, which are then, because of reflection invariance, the same.

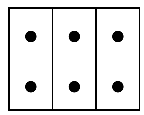
In this way we are led to accept the *commutative law of addition:* $m + n = n + m$ for any natural numbers m, n.

The *associative law of addition* $(m + n) + p = m + (n + p)$ can be seen to arise from an even more rudimentary use of the idea of " gap closing" of juxtaposed dot sequences. It can be immediately seen that , once the gaps are closed, $(\mathbf{m} \bigstar \mathbf{n}) \bigstar \mathbf{p}$ is identical with $\mathbf{m} \bigstar (\mathbf{n} \bigstar \mathbf{p})$.

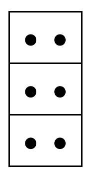
The *commutative law for multiplication* can be justified by similar considerations of invariance. Multiplication arises from repeated addition: 2×3 , for example, corresponds to the combination of 2 sets of 3 dots:



This contains altogether 6 dots, so $2 \times 3 = 6$. Regrouping the above arrangement gives



and by rotating this through 90° we obtain



that is, 3 sets of 2 dots corresponding to 3×2 dots. Since throughout the regrouping and rotation the assemblage of dots remains unchanged, the number of dots must be the same in both cases, so that $2 \times 3 = 3 \times 2 = 6$. Thus we are led to adopt the general rule known as the *commutative law of multiplication*: $m \times n = n \times m$ for any natural numbers m, n.

Similar combinatorial-geometric arguments lead to the intuitive justification of the associative law of multiplication and the distributive law for multiplication over addition.

It is remarkable that, while our intuitive grasp of the commutative laws of arithmetic (as well as the others) must surely have been the result of basic combinatorial-geometrical considerations like the above , *all of these laws can be formally derived from the simple idea of immediate succession, rules for of adding and multiplying by an immediate successor, and the law of mathematical induction.*

To see how this can be done we set up a *formal language for arithmetic* (with 0) . This will have a unary operation symbol *s*, corresponding to (*immediate*) *succession:* thus, if *n* is a sequence of dots

$$\boxed{\bullet\bullet\bullet\bullet\bullet}$$

sn can be pictured as the result of by adding one new dot (on the right) to the given sequence of dots *n*

$$\boxed{\bullet\bullet\bullet\bullet\bullet}\ \boxed{\bullet}$$

and then closing the gap to obtain

$$\boxed{\bullet\bullet\bullet\bullet\bullet\bullet}$$

The successor operation accordingly corresponds to the operation of *juxtaposing a single new dot.*

We also equip our formal language with two binary operation symbols **+** and **×** of addition and multiplication, corresponding to the operations on assemblages of dots we have analyzed above. Our language will also be equipped with a symbol *0,* which will correspond to the "dotless", or *empty* assemblage

$$\square$$

Now in the language we have set up, using standard logical symbols, we can write down the postulates of *Basic Arithmetic.* These are the following:

B1 $\forall x \forall y (sx = sy \rightarrow x = y)$

B2 $\forall x \quad 0 \neq sx$

B3 $\forall x \quad x + 0 = x$

B4 $\forall x \forall y \quad x + sy = s(x + y)$

B5 $\forall x \quad x \times 0 = 0$

B6 $\forall x \forall y \quad x \times sy = (x \times y) + x.$

The first two of these express familiar facts about the successor operation:

B1 Natural numbers with identical successors are themselves identical.

B2 Zero is the successor of no natural number.

The next two postulates tell us how to add in this notation (where we have introduced the symbol *1* for *s0*)

B3 Adding 0 has no effect.

B4 $(m + n) + 1 = m + (n + 1).$

Finally, the two remaining postulates reduce multiplication to repeated addition:

B5 Multiplying by 0 yields 0.

B6 $x \times (y + 1) = (x \times y) + x.$

It is easy to justify **B1 - B6** using dot assemblages.

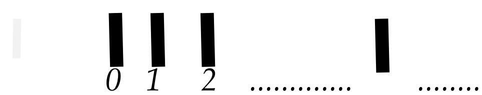
Full arithmetic is now obtained by adding to **B1 - B6** the following

Principle of Mathematical Induction. Suppose given any property P of natural numbers. Suppose also that

- 0 has P
- for any natural number n, if n has P, so does sn.

Then every natural number has P.

This principle is usually justified by the following argument. Suppose that $P(0)$ and that , for any number n, $P(sn)$ follows from $P(n)$. Then since we have $P(0)$, it follows that $P(1)$; from this follows $P(2)$, hence $P(3)$, etc. *ad infinitum*. Another way of justifying mathematical induction is the *domino argument.* Let us suppose we represent the sequence of natural numbers by a series of dominoes, initially all standing upright:



Let us suppose that, if any domino falls forward, it strikes the following one, causing it to fall. Now let us represent the assertion $P(n)$ by the nth domino falling forward. Then $P(0)$ means that domino 0 falls forward. This causes domino 1 and hence all the remaining dominoes to do the same. Since all the dominoes (eventually) fall, it follows

that $P(n)$ for all natural numbers n.

In full arithmetic, using the principle of mathematical induction, the commutative laws for addition and multiplication are *derivable.*

We derive the commutative law for addition. This is most easily done in 5 stages.

(1) Prove the associative law $(m + n) + p = m + (n + p)$.

(2) Show that, for any n, $0 + n = n$.

(3) Show that, for any n, $sn = n + 1$.

(4) Show that, for any n, $n + 1 = 1 + n$.

(5) Prove the commutaive law of addition: for all m, n, $m + n = n + m$.

Proof of (1). Fix arbitrary natural numbers m, n. Let P be the property of natural numbers p that $(m + n) + p = m + (n + p)$. Then 0 has P since

$$(m + n) + 0 = m + n = m + (n + 0)$$

Now supposing that p has P, *i.e.* $(m + n) + p = m + (n + p)$, we deduce that sp has P as follows:

$(m + n) + sp = s((m + n) + p) = s(m + (n + p)) = m + s(n + p) = m + (n + sp)$.

So $(m + n) + p = m + (n + p)$ holds for any p by the Induction Principle. Since m and n were arbitrary, the equation holds for all m, n, p.

Proof of (2). Let P be the property $0 + n = n$. Then 0 has P by **B3.** Supposing that n has P, we deduce that sn has P:

$$0 + sn = s(0 + n) = sn.$$

So (2) follows by the Induction Principle.

Proof of (3). $sn = s(n + 0) = n + s0 = n + 1.$

Proof of (4). Let P be the property $n + 1 = 1 + n$. Then 0 has P since $0 + 1 = 1$ *(by (2))*

$= 1 + 0$. Now suppose that n has P. Then we deduce that sn has P:

$$sn + 1 = sn + s0 = s(sn + 0) = ssn = s(n + 1) \text{ (by (3))} = s(1 + n) = 1 + sn.$$

(4) now follows by the Induction Principle.

Proof 0f (5). Fix m and let P be the property of n $m + n = n + m$. Then 0 has P since $m + 0 = m = 0 + m$ by (2). Now suppose that n has P. Then we deduce, using associativity, (3) and (4) that sn has P:

$$m + sn = m +(n + 1) = (m + n) + 1= (n + m) + 1 = n + (m + 1) = n +(1 + m) = (n + 1) + m = sn + m.$$

(5) now follows by the Induction Principle.

It will be observed that in this derivation of the commutative law of addition, the associative law is derived first (and much more easily), thus providing further evidence that associativity is a more fundamental property of operations than commutativity. Of course, this is already clear in the case of *composition* of operations or functions, which is associative but not usually commutative.