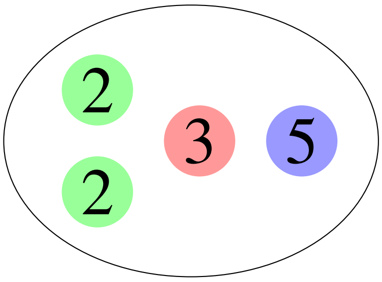
MULTISETS IN ARITHMETIC

Antonella Perucca

The fundamental theorem of arithmetic states that every positive integer is a product of prime numbers, and this in a unique way up to rearranging the prime factors (prime numbers are considered to be products with just one factor, and the number 1 is the empty product). To any positive integer we may then associate the *multiset*[1] of its prime factors, and we have a very elegant correspondence between the positive integers and the finite multisets of prime numbers. We first show that several arithmetic notions related to divisibility translate to basic notions for multisets, and then we use multisets to express the least common multiple of several numbers in terms of the greatest common divisor.



$$60 = 2 \times 2 \times 3 \times 5$$

Figure 1: The multiset of the prime factors for the number 60.

Given two or more 'numbers' – by which we mean positive integers, the *sum*[2] of their multisets of prime factors clearly corresponds to the product of the numbers. We also have a useful criterion for divisibility involving the *inclusion*[3] of multisets. Indeed, if some number a divides some number b, then we can write $b = a \cdot c$ for some number c, and then the multiset of prime factors for a is included in the multiset for b. Conversely, if the multiset of prime factors for a is included in the one for b, then a is a partial product of the prime factors of b and hence it divides b.

With the divisibility criterion one can easily prove that the distinct prime factors of a number are exactly the prime divisors (in other words, the *support*[4] for the multiset of prime factors is the set of prime divisors). Moreover, the *multiplicity*[5] of a prime divisor is the exponent of the largest power of that prime which divides the number.

We now deal with the *intersection*[6] of the multisets of prime factors for two or more numbers. By the divisibility criterion, the number corresponding to this intersection is

---

[1] Multisets are like sets, but the elements can be repeated: the number of times that an element appears in a multiset is its *multiplicity* (if an element does not appear, then its multiplicity is zero).

[2] The *sum* of two multisets is the multiset obtained by taking all their elements together, which amounts to summing the multiplicities. We may straight-forwardly generalize this definition to more than two multisets. We can also define the sum of one multiset as the multiset itself.

[3] You have an *inclusion* between two multisets if all elements of the first are also contained in the second with at least the same multiplicity.

[4] If we only keep the distinct elements of a multiset, then we get a set, which is called the *support* of the multiset.

[5] Multisets are like sets, but the elements can be repeated: the number of times that an element appears in a multiset is its *multiplicity* (if an element does not appear, then its multiplicity is zero).

[6] The *intersection* of two multisets is the largest multiset which is included in both (simply take the common elements). The multiplicity for an element in the intersection is the minimum of the two multiplicities. This generalizes straight-forwardly to several multisets, and to one multiset. Aside remark: one also finds a different definition for the intersection of multisets.

a common divisor and a multiple of every common divisor, and hence it is the greatest common divisor (in short, $\gcd$). Similarly consider the *union*[7] of the multisets of prime factors: the number corresponding to this union is a common multiple which divides every common multiple, so it is the least common multiple (in short, lcm).
It follows from what above that two or more numbers are coprime (i.e. the $\gcd$ is 1) if and only if the multisets of their prime factors are *disjoint*[8]. They are pairwise coprime (i.e. the $\gcd$ of any two of them is 1) if and only if the multisets of their prime factors are *pairwise disjoint*[9].
Finally, if a number is a multiple of another, then the *difference*[10] of their multisets of prime factors corresponds to the quotient of the two numbers. In general, the difference of the two multisets corresponds to the quotient of the first number by the $\gcd$ (this can easily be seen by considering the multiplicities for each prime factor).

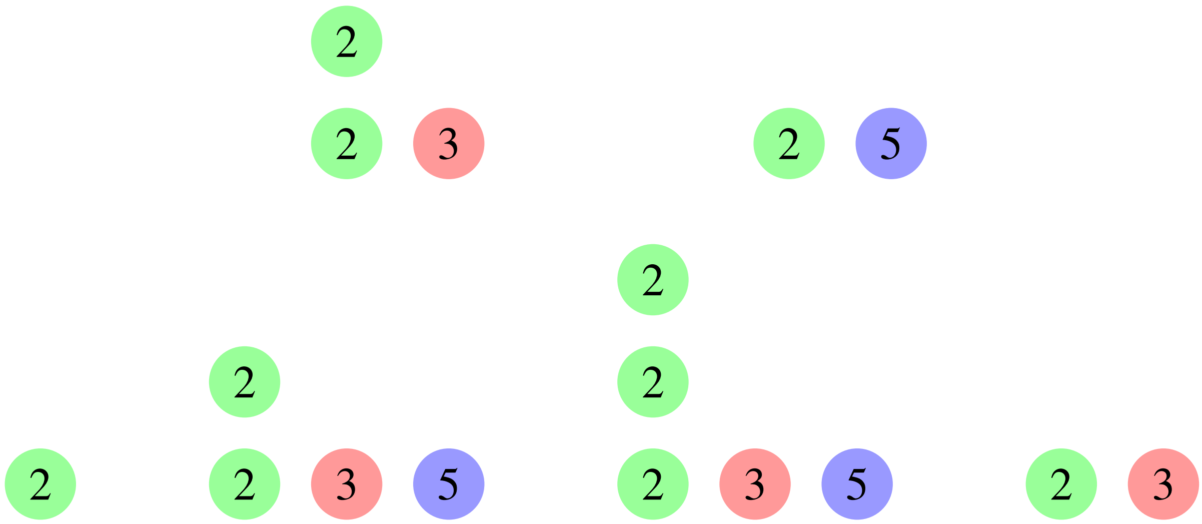


Figure 2: The four multisets below are the intersection, the union, the sum, and the difference of the two multisets above. The corresponding numbers are 12; 10; $\gcd(12, 10) = 2$; $\operatorname{lcm}(12, 10) = 60$; $12 \cdot 10 = 120$; $12/\gcd(12, 10) = 6$.

There are several arithmetic properties that can be (visualized and) proven with multisets. The basic idea is either applying some property of multisets to the multisets of prime factors, or directly work with the multiplicities of the prime factors.
As an example, we analyze with multisets how to express the least common multiple in terms of the greatest common divisor. Observe that the sum of two multisets equals the sum of their intersection and their union (can you prove this property, for example by

[7]The *union* of two multisets is the smallest multiset which includes both. The multiplicity for an element in the union is the maximum of the two multiplicities. This generalizes straight-forwardly to several multisets, and to one multiset.

[8]Two or more multisets are *disjoint* if their intersection is empty. Considering the multiplicities of an element, we have the following characterization: ‘disjoint’ means that there is at least one multiplicity which is zero.

[9]Two or more multisets are *pairwise disjoint* if any two of them are disjoint (in particular, they are disjoint multisets). Considering the multiplicities of an element, we have the following characterization: ‘pairwise disjoint’ means that all multiplicities are zero, with at most one exception.

[10]If one multiset includes another, then we obtain their *difference* by removing all elements of the latter from the former. The multiplicity for the difference is thus the difference of the multiplicites. For two general multisets we may adapt this definition by setting the multiplicity for the difference to be zero in case that the difference of the multiplicities is negative. Aside remark: one can also find another definition for the difference of multisets.

considering the multiplicities for each element?). By translating this property into the arithmetic setting, we deduce that the following formula holds for two numbers a and b:

$$\operatorname{lcm}(a,b) = \frac{a \cdot b}{\gcd(a,b)} \,.$$

For three numbers a, b, c we have

$$\operatorname{lcm}(a,b,c) = \frac{a \cdot b \cdot c \cdot \gcd(a,b,c)}{\gcd(a,b) \cdot \gcd(a,c) \cdot \gcd(b,c)}$$

because this formula is the translation of the inclusion-exclusion principle for three multisets A, B, C, which is the identity

$$A \cup B \cup C = A + B + C + (A \cap B \cap C) - (A \cap B) - (A \cap C) - (B \cap C)\,.$$

Finally, the general formula for the lcm of n numbers is quite impractical to write, but it exists as the direct translation of the inclusion-exclusion principle for multisets A_1 to A_n, which is the identity

$$\bigcup_{i=1}^{n} A_i = \sum_{\substack{\emptyset \neq J \subseteq \{1,2,\ldots,n\} \\ \#J \text{ odd}}} \bigcap_{j \in J} A_j \quad - \quad \sum_{\substack{\emptyset \neq J \subseteq \{1,2,\ldots,n\} \\ \#J \text{ even}}} \bigcap_{j \in J} A_j \,.$$

Note: To prove this identity, first rewrite it for the multiplicities of an element (in short, replace the sets by the multiplicities, $\cup$ by $\max$, and $\cap$ by $\min$), and then prove it by putting aside the maximum of the multiplicities.