

# No consecutive integers

INTEGER EQUATIONS and multisets are the topics of the problem <sup>1</sup> in this note.

<sup>1</sup> Variation of Problem 1-72. on page 45 in N. Loehr. *Combinatorics*. Discrete Mathematics and Its Applications. CRC Press, 2017. ISBN 9781498780278

## Problem

Determine the number of subsets of size  $k$  from set  $\{1, 2, \dots, n\}$  that do not contain consecutive integers.

The number of subsets of size  $k$  from set  $\{1, 2, \dots, n\}$  (without any constraints) is given by the binomial coefficient  $\binom{n}{k}$ . Each subset of size  $k$  can be represented as a word of length  $n$  from alphabet  $\{\star, \square\}$  with  $k$   $\square$ 's and  $n - k$   $\star$ 's: if  $i$  is in the subset then the corresponding word has a  $\square$  at position  $i$  otherwise it has a  $\star$  at position  $i$ . This representation is clearly a bijection. The constraint of no consecutive integers in a subset implies no adjacent  $\square$ 's in the corresponding word<sup>2</sup>.

We will now associate words from  $\{\star, \square\}^n$  with other combinatorial objects: the integer equations.

**Definition 1.1.** Given fixed integers  $m > 0$  and  $t \geq 0$  a sequence  $(z_1, z_2, \dots, z_m)$  is an **integer equation** if  $\forall i : 1 \leq i \leq m : z_i \in \mathbb{N}_0$  and

$$\sum_{i=1}^m z_i = t$$

The number  $t$  is called the **target** of the integer equation.

Note that these are sequences and order matters. From an integer equation  $(z_1, z_2, \dots, z_m)$  we construct a  $\{\star, \square\}^{t+m-1}$  word in the following way: start with  $z_1$  number of  $\star$ 's, then a  $\square$ , then  $z_2$  number of  $\star$ 's, then a  $\square$  and so on finishing with the  $z_m$  number of  $\star$ 's which are **not** followed by a  $\square$ . The word will contain exactly  $t$   $\star$ 's and they will need exactly  $m - 1$   $\square$  separators to know which stars belong to which  $z_i$ . It's easy to verify that this encoding is also a bijection<sup>3</sup>.

<sup>2</sup> For example given set  $\{1, 2, 3, 4\}$  the subset  $\{2, 4\}$  corresponds to  $\star \square \star \square$ . The subset  $\{1, 2\}$  has consecutive integers and corresponds to  $\square \square \star \star$ . The reason why we chose  $\square$  to indicate inclusion into a subset will become clear soon.

<sup>3</sup> As an example let  $m = 5$  and target  $t = 10$ . The sequence  $(1, 2, 1, 3, 3)$  is an integer equation since

$$1 + 2 + 1 + 3 + 3 = 10$$

and it corresponds to the word

$\star \square \star \star \square \star \square \star \star \star \square \star \star \star$

This in turn corresponds to the subset  $\{2, 5, 7, 11\}$  of set  $\{1, \dots, 14\}$ .

**Lemma 1.2.** *If we set  $t = n - k$  (the number of  $\star$ 's) and from  $t + m - 1 = n$  we get  $m = k + 1$  (the word length) then we can associate subsets of size  $k$  from set  $\{1, 2, \dots, n\}$  with integer equations  $(z_1, z_2, \dots, z_{k+1})$  for target  $n - k$ . The constraint of not having consecutive integers in the subsets translates to integer equations  $(z_1, z_2, \dots, z_{k+1})$  where  $z_i > 0$  except for  $z_1$  and  $z_{k+1}$  (the first and the last in the sequence). This follows from the encoding not allowing adjacent  $\square$ 's so there need to be  $\star$ 's separating the  $\square$ 's.*

According to this association if we can count the number of integer equations with all but the first and last  $z_i$  strictly positive then we also have the number of subsets with no consecutive integers. To get there we will first count the number of anagrams, then the number of multisets, then the number of integer equations and finally the number of integer equations with all but the first and last  $z_i$  strictly positive. In what follows we will use  $n$  and  $k$  for other things before we bring it back in the end to our initial problem.

Let's start this journey with anagrams. Let  $\{s_1, s_2, \dots, s_k\}$  be an alphabet of distinct symbols. We can build words with these symbols, for example  $s_2 s_2 s_1 s_3 s_3 s_1$ . As a notational convenience  $s_i s_i s_i \dots s_i = s_i^j$  if  $s_i$  appears  $j$  consecutive times in a word, so the example would be  $s_2^2 s_1 s_3^2 s_1$ .

**Definition 1.3.** A word is an **anagram**<sup>4</sup> of  $s_1^{n_1} s_2^{n_2} \dots s_k^{n_k}$  with  $n_i > 0$  if it is a word containing exactly  $n_i$  number of  $s_i$  symbols for each  $1 \leq i \leq k$ . We denote with  $\mathcal{A}(s_1^{n_1} s_2^{n_2} \dots s_k^{n_k})$  the set of all anagrams of  $s_1^{n_1} s_2^{n_2} \dots s_k^{n_k}$ .

**Theorem 1.4.** *Given the set of anagrams  $\mathcal{A}(s_1^{n_1} s_2^{n_2} \dots s_k^{n_k})$  let  $n = \sum_{i=1}^k n_k$ . Then*

$$|\mathcal{A}(s_1^{n_1} s_2^{n_2} \dots s_k^{n_k})| = \binom{n}{n_1, n_2, \dots, n_k}$$

where  $\binom{n}{n_1, n_2, \dots, n_k}$  is the multinomial coefficient<sup>5</sup>.

*Proof.* We have  $n$  positions in our word that we need to fill with symbols. We are going to make the following choices: first we choose  $n_1$  positions from those  $n$  positions where we fill in the symbol  $s_1$ . Then we choose the  $n_2$  positions from the remaining unfilled positions where we fill in  $s_2$  and so on. In total we make  $k$  such choices and the number of remaining unfilled positions at each stage is independent of the previous choices, so the multiplication rule applies. For our first symbol  $s_1$  we have  $\binom{n}{n_1}$  possibilities, for our second symbol we have  $\binom{n-n_1}{n_2}$  possibilities and so on. Because of the multiplication rule the total number of choices is the product of all these binomial coefficients, so

$$|\mathcal{A}(s_1^{n_1} s_2^{n_2} \dots s_k^{n_k})| = \prod_{i=1}^k \binom{n - (\sum_{j=1}^{i-1} n_j)}{n_i}$$

<sup>4</sup> For example given word  $a^2b$  the words  $aba$  and  $baa$  are anagrams of it. The word  $abb$  is not.

<sup>5</sup> The multinomial coefficient is defined as

$$\binom{n}{n_1, n_2, \dots, n_k} = \frac{n!}{\prod_{i=1}^k n_i!}$$

Expanding<sup>6</sup> the binomial coefficients on the right-hand side into factorials according to the binomial coefficient definition and simplifying the expression gives us the desired result.  $\square$

We move on to **multisets**. Informally multisets are sets (order does not matter) where each element can appear more than once. So given a set  $A$  (the alphabet) a multiset is a tuple of  $A$  together with a function  $\mu : A \mapsto \mathbb{N}$  that determines how often an element  $a \in A$  appears in the multiset. For notational convenience we will use curly braces and list elements (with exponents if they appear more than once). For example  $\{a^2, b, c^4\}$  is a multiset where  $a$  appears twice,  $b$  once and  $c$  four times. Note that order does not matter, so  $\{a^2, b, c^4\}$  is the same multiset as  $\{b, a^2, c^4\}$ . The size of the multiset is the number of elements in it with elements appearing more than once counted accordingly, so

$$|(A, \mu)| = \sum_{a \in A} \mu(a)$$

**Theorem 1.5.** *The number of multisets of size  $k$  from an alphabet set of size  $n$  is<sup>7</sup>*

$$\binom{k+n-1}{k}$$

*Proof.* We will do an encoding of multisets to anagrams similar to what we did at the beginning of this section with  $\star$ 's and  $\mathbb{I}$ 's. To avoid confusion with that previous encoding in this proof we will use the symbols  $\circ$  and  $|$ .

Let  $A = \{a_1, a_2, \dots, a_n\}$  be our alphabet. For a multiset  $(A, \mu)$  with size  $k$  we define the following word<sup>8</sup> with symbols  $\{\circ, |\}$ :

$$\circ^{\mu(a_1)} | \circ^{\mu(a_2)} | \dots | \circ^{\mu(a_n)}$$

The first circles denote how often  $a_1$  is in the multiset. They are separated by a  $|$  from the circles that denote how often  $a_2$  is in the multiset and so on. In total there are  $k$  circles because the multiset has size  $k$  and there need to be  $n - 1$  separators because the alphabet has size  $n$  and the circles for each element need to be kept apart. It's easy to see that we have defined a bijection from the set of multisets of size  $k$  with alphabet of size  $n$  to the set of anagrams  $\mathcal{A}(\circ^k |^{n-1})$ . From theorem 1.4 we already know how to count the size of  $\mathcal{A}(\circ^k |^{n-1})$  and with the bijection it proves this theorem.  $\square$

Our next stop are the number of integer equations. Given  $m$  and  $t$  how many integer equations  $(z_1, z_2, \dots, z_m)$  for target  $t$  are there?

**Theorem 1.6.** *The number of integer equations  $(z_1, z_2, \dots, z_m)$  for target  $t$  is*

$$\binom{t+m-1}{t}$$

<sup>6</sup> After the binomial coefficients are expanded the product becomes a telescoping product that simplifies to exactly the multinomial coefficient.

<sup>7</sup> For example with alphabet set  $\{a, b\}$  the multisets of size two are  $\{a^2\}$ ,  $\{b^2\}$ ,  $\{a, b\}$ , so there are three of them.

<sup>8</sup> The multisets from the previous example would be encoded as follows:

$$\begin{aligned} \{a^2\} &\mapsto \circ \circ | \\ \{b^2\} &\mapsto | \circ \circ \\ \{a, b\} &\mapsto \circ | \circ \end{aligned}$$

*Proof.* We will associate a multiset with each integer equation<sup>9</sup>. The multiset will contain the element  $i$   $z_i$  many times, for  $1 \leq i \leq m$ . Again it can be checked that this defines a bijection. These multisets belong to the set of multisets of size  $t$  from an alphabet of size  $m$  and theorem 1.5 counts them. By the bijection rule we have proven this theorem.  $\square$

We are almost done. In the beginning of this section we encoded our subsets without consecutive integers as integer equations with all but the first and last summand strictly positive. So we need to count these types of integer equations with this constraint.

**Theorem 1.7.** *The number of integer equations  $(y_1, y_2, \dots, y_m)$  for target  $t$  with  $y_i > 0$  for all  $1 < i < m$  is*

$$\binom{t+1}{m-1}$$

*Proof.* For an integer equation  $(y_1, y_2, \dots, y_m)$  we have  $\sum_{i=1}^m y_i = t$  and  $y_i > 0$  for all  $1 < i < m$ . So we can write

$$y_1 + \sum_{i=2}^{m-1} (y_i - 1) + y_m = t - (m - 2)$$

This shows that we can transform the integer equations with the strictly positive constraints into normal integer equations without constraints but with a new target. This again is a bijection. We know how to count these from theorem 1.6. The new target is  $t - m + 2$ . Plugging it in we get

$$\binom{t - m + 2 + m - 1}{m - 1} = \binom{t + 1}{m - 1}$$

$\square$

Using 1.7 and  $t = n - k$  and  $m = k + 1$  as described by our association 1.2 of subsets of size  $k$  without consecutive integers from set  $\mathbb{N}_n$  to integer equations with all but the first and last strictly positive terms, we are finally able to solve the problem in this section. The answer is  $\binom{n-k+1}{k}$ .

<sup>9</sup>For example with  $m = 5$  and target  $t = 10$  the integer equation  $(1, 2, 1, 3, 3)$  would correspond to multiset  $\{1, 2^2, 3, 4^3, 5^3\}$ .

## *Bibliography*

N. Loehr. *Combinatorics*. Discrete Mathematics and Its Applications.  
CRC Press, 2017. ISBN 9781498780278.