INCREASING INTEGER SEQUENCES
AND GOLDBACH'S CONJECTURE *

Mauro Torelli[1]

Abstract. Increasing integer sequences include many instances of interesting sequences and combinatorial structures, ranging from tournaments to addition chains, from permutations to sequences having the *Goldbach property* that any integer greater than 1 can be obtained as the sum of two elements in the sequence. The paper introduces and compares several of these classes of sequences, discussing recurrence relations, enumerative problems and questions concerning shortest sequences.

Mathematics Subject Classification. 11Y55, 11P32, 05A15, 11B99.

1. Introduction

Long-standing conjectures are obviously difficult to settle, but we can derive some – hopefully interesting – *sequences* from them. Let us see three examples.

The well-known *Goldbach's conjecture*, open since 1742, states that *any even number greater than* 2 *can be expressed as the sum of two primes*. A relatively recent paper on the subject, with pointers to the huge number of references, is [5]. The conjecture has been verified up to 4×10^{14} [19], and several theorems have been proved, such as the following: *every positive even integer is the sum of at most* 6 *primes*, or *every sufficiently large even integer may be written as the sum of a prime and a number that has at most* 2 *prime factors*.

Let $\mathbb{Z}^+$ be the *positive integers*, $\mathbb{P}$ the *primes* and $\mathbb{E}$ the *even* positive integers. Let $\mathbf{A}$ be a subset of $\mathbb{Z}^+$: we will write $2\mathbf{A}$ to indicate the *sumset* $\mathbf{A} + \mathbf{A}$ obtained by summing any two (not necessarily distinct) elements of $\mathbf{A}$ (*cf. e.g.* [15], p. 192). The aim of this notation is to express Goldbach's conjecture in the terse form

© EDP Sciences 2006

```
2 3 5 7 11 13 17          1 2 3 5 6 8          1 2 3 5 10   22   49
 1 2 2 4   2   4           1 1 2 1 2           1 1 2 5   12   27
  1 0 2 2   2               0 1 1 1            0 1 3   7   15
   1 2 0 0                   1 0 0              1 2   4   8
    1 2 0                     1 0               1   2   4
     1 2                       1                 1   2
      1                                           1
     (a)                      (b)                 (c)
```
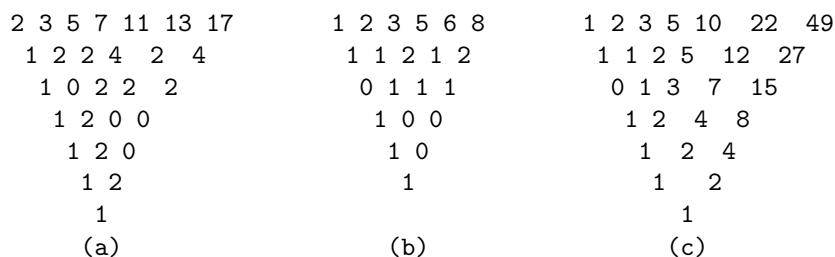
FIGURE 1. Gilbreath sequences and absolute difference triangles.

$\mathbb{E} - \{2\} \subseteq 2\mathbb{P}$. But we would like also to eliminate the irrelevant fact of 2 being the only even prime number, and are thus led to define *Goldbach sets* as those subsets **S** of $\mathbb{Z}^+$ such that $\mathbb{Z}^+ - \{1\} \subseteq 2\mathbf{S}$ (in fact, $\mathbb{Z}^+ - \{1\} = 2\mathbf{S}$). Furthermore, we will only consider *finite* subsets of any such set **S**, up to a given maximum value, and will represent them as increasing sequences, to be called *Goldbach sequences*. Some further terminology and notation about sequences will now be introduced, and Goldbach sequences will be defined formally.

Definition 1. An *increasing integer sequence (iis)* of *length* r is a finite sequence $a_0, a_1, ..., a_r$ of positive integers such that $a_0 = 1$ and $a_{i-1} < a_i, 1 \le i \le r$.

Let **s** be an iis, then the length r of **s** will be denoted by $l(\mathbf{s})$, while $\max(\mathbf{s})$ will denote the *maximum element* a_r. Observe that $l(\mathbf{s}) < \max(\mathbf{s})$.

Definition 2. An iis is a *Goldbach sequence* if and only if for any integer x such that $2 \le x \le a_r$, $x = a_j + a_k$ for some j, k with $0 \le j \le k < r$.

Example 1. $\mathbf{s} = 1\ 2\ 3\ 5\ 6\ 8\ 9\ 11\ 14\ 15\ 18$ is a Goldbach sequence, as can be easily checked. Its length is 10 and, for $0 \le i \le 10$, $2a_i + 1$ is the $(i+2)$-nd prime. $1\ 2\ 4\ 8$ is *not* a Goldbach sequence, since 7 cannot be obtained as specified in the definition.

Sequence **s** in Example 1 shows that it would be convenient to have a name for numbers q such that $2q + 1$ is prime: perhaps *primoids* would be adequate, so that we could refer to the infinite sequence starting with **s** as the *sequence of primoids*. Then, Goldbach's conjecture could be stated in the form: *any integer greater than* 1 *can be obtained as the sum of two primoids.*

Goldbach sequences clearly are a (very) particular case of *addition chains*, to be introduced in Definition 9, where only each element *in the chain* is required to be the sum of two preceding elements. Before examining more general classes of sequences, let us consider another, slightly less-well-known, number-theoretic conjecture: *Gilbreath's conjecture* (*cf. e.g.* [13, 17]). This conjecture is best explained by an example: let us take a short prefix of the sequence of primes, say 2 3 5 7 11 13 17, and construct a triangle in which each element is the *absolute value of the difference* of the two elements above it, like in Figure 1a.

Gilbreath's conjecture simply states that, for the sequence of primes, *the first element in each subsequent row of the triangle is* 1. For an iis we can adopt the following definition.

Definition 3. An iis is a *Gilbreath sequence* if and only if for any i with $0 \leq i < r$, $a_{i,0} \leq 1$, where $a_{0,k} := a_k$, $0 \leq k \leq r$; $a_{i+1,j} := |a_{i,j} - a_{i,j+1}|$, $0 \leq i < r$, $0 \leq j < r - i$.

Cases (b) and (c) of Figure 1 show, in the first row $a_{0,k}$, two examples of Gilbreath sequences, together with the subsequent rows $a_{i,j}$. Definition 3 requires the condition $a_{i,0} \leq 1$, instead of the conjecture statement $a_{i,0} = 1$, because, as before, it eliminates the peculiarities due to 2 being the only even prime number; in fact, in the examples, $a_{2,0} = 0$, and example (b) is nothing but example (a) after removing $a_{i,0}$ and dividing by 2, passing from primes to primoids. Example (c) will be considered later on.

Let us now introduce the last conjecture. Sequence 3 10 5 16 8 4 2 1 is not an iis, but should be easily recognized as an example of *Collatz* $3x+1$ *problem*: start with any integer number $x > 1$, divide it by 2 if it is even, otherwise take $3x + 1$. The *conjecture* is that we always get to 1 (*cf.*, *e.g.*, [12]). Now, to get to an iis, *reverse* the original sequence: 1 2 4 8 16 5 10 3, start from 1 and compare each element with its successor: if there is an increase, then *increase by* 1, otherwise *increase by* 2: 1 2 3 4 5 7 8 10. In the following, sequences obtained in this way, although not defined formally, will be called *Collatz sequences*.

Collatz sequences increase very slowly, since $a_i \leq a_{i-1} + 2$, $2 \leq i \leq r$ and therefore $a_i \leq 2i$, but, of course, not every iis with $a_i \leq 2i$ is Collatz. However, we shall show in next section that *every iis with $a_i \leq 2i$ is Goldbach*, therefore *every Collatz sequence is Goldbach*. On the other hand, it is easy to find a Gilbreath sequence which is not Goldbach, and vice versa, thus showing that these two conjectures for primes are – at least in this respect – independent. Here are the counterexamples: the Goldbach sequence 1 2 3 6 is not Gilbreath ($a_{3,0} = 2$), while the sequence in Figure 1c cannot be Goldbach, since $22 > 2 \times 10$, and therefore at least 21 and 22 cannot be obtained by adding two preceding elements.

The condition $a_{i+1} \leq 2a_i$ is obviously *necessary* for an iis to be Goldbach, and we are thus led to examine sequences verifying this condition. However, we have already seen that such a condition is violated by Gilbreath sequences. Therefore, looking for a set including all of the already mentioned sequences, we are forced to relax this condition. What happens if we construct the difference triangle bottom-up by putting 1 at the beginning of each row and then increasing each row as much as possible? The upper row will be made by the successive powers of 2: therefore, the appropriate condition is $a_i \leq 2^i$. Fortunately, this condition also yields an upper bound for all sequences (including Goldbach) satisfying the former necessary condition.

We will examine these classes of sequences (and a few more) in the next section. Section 3 will be devoted to enumeration, while Section 4 will discuss some questions about shortest sequences.

2. Subexponential sequences

In the present paper only sequences satisfying the condition $a_i \leq 2^i$ are considered, and this choice was motivated in the Introduction. It seems, therefore, more or less appropriate to call these sequences *subexponential sequences*, since they are superiorly bound by the exponential sequence 1 2 4 8 16 ... Actually, sequences satisfying this condition are in bijection with *full* (or *complete*) *sets* (sets whose elements are also subsets: *cf.* [3], p. 122). Sequences in this class can, of course, increase exponentially with length, so in this respect the name is misleading: we use it in the same sense in which sequences superiorly bound by Fibonacci numbers have been expressively called *sub-Fibonacci* [9].

Now we want to progressively restrict the range of possible sequences. It turns out that one of the weakest restrictions is obtained by relaxing the condition defining Goldbach sequences: instead of requiring that value x be obtained by summing *two* elements, we can require x to be obtained by summing *any number* of elements. The following lemma is general enough to allow for repeated elements.

Lemma 1. *Let $a_0 \leq a_1 \leq \ldots \leq a_r$ be positive integers and set $\sigma(k) := \sum_{j=0}^{k} a_j$ for $k \leq r$. For any x, $1 \leq x \leq \sigma(r)$, one can determine $\chi_i(x) \in \{0,1\}$, $0 \leq i \leq r$, such that $x = \sum_{i=0}^{r} a_i \cdot \chi_i(x)$, if and only if $a_0 = 1$ and $a_{i+1} \leq \sigma(i) + 1$ for $0 \leq i < r$.*

Proof. Case $x = 1$ requires $a_0 = 1$, and then the lemma is true for $r = 0$. Let it be true for $r = k \geq 0$: then any $x \leq \sigma(k)$ can be obtained, but if $a_{k+1} > \sigma(k) + 1$, then there is no way to obtain $x = \sigma(k) + 1$. On the contrary, if $a_{k+1} \leq \sigma(k) + 1$ and $a_{k+1} \leq x \leq \sigma(k+1) = \sigma(k) + a_{k+1}$, then $x = a_{k+1} + \sum_{i=0}^{k} a_i \cdot \chi_i(x - a_{k+1})$, since $x - a_{k+1} \leq \sigma(k)$.   □

We are now ready to define increasing integer sequences permitting to obtain any possible value as a sum of elements in the sequence.

Definition 4. An iis is a *complete* sequence if and only if $a_{i+1} \leq \sigma(i) + 1$, $0 \leq i < r$, where $\sigma(i) := \sum_{j=0}^{i} a_j$.

The name *complete* has been chosen after [8], where this name is used with the same meaning but for infinite sequences. Other encountered names were at least *practical* and *regular*. The term *practical* was suggested by *practical numbers*, which, after Definition 4, can be defined simply as those numbers whose sequence of divisors, when written as an iis, is complete. The infinite sequence of practical numbers has recently been proved to possess Goldbach property [14]! *Regular* sequences are *nondecreasing* sequences introduced by Fishburn *et al.* [9]: they satisfy conditions similar to those in Definition 4 and are shown to ensure uniqueness of solutions for finite measurement structures.

Let us now define sequences satisfying the necessary condition $a_{i+1} \leq 2a_i$ for a sequence to be Goldbach, and then examples and comparisons will be made.

Definition 5. An iis is a *tournament* if and only if $a_{i+1} \leq 2a_i$, $0 \leq i < r$.

Tournaments [1] have been called so because a_i can be considered as the *number of players in a knock-out tournament* after round i, starting from the *final* round as round 0, if in each round at least one and at most half of the players are knocked out. It is immediate to show that *tournaments*, according to Definition 5, *are complete sequences.*

Example 2. There are 9 subexponential sequences or full sets of length three: 1 2 3 4, 1 2 3 5, 1 2 3 6, 1 2 3 7, 1 2 3 8, 1 2 4 5, 1 2 4 6, 1 2 4 7, 1 2 4 8 (remember that the number of elements in the sequence is one more than the length, since the sequence starts with a_0). Of the 9 sequences above, 1 2 3 8 is the only one which is not complete: it is also not a tournament, the only other non-tournament being 1 2 3 7.

Let us now consider the Gilbreath sequence in the first row of Figure 1c, 1 2 3 5 10 22 49: since its last element, 49, is bigger than the sum of all the preceding elements plus 1, then that sequence is not complete. We shall not consider Gilbreath sequences any more, but would like instead to briefly introduce a subclass of complete sequences which neither contains nor is contained in the class of tournaments.

Consider a *simple graph* with m edges having weights 1 2 3... m: which increasing integer sequences may represent the sequence of weights of a *Minimum Spanning Forest* (*MSF*) of the graph? Edges with weights 1 and 2 can always be chosen, but the edge of weight 3 (let us say "edge 3" for short) might close a circuit with edges 1 and 2: in any case, either edge 3 or edge 4 (or both) will be chosen. Which is the worst case, the one with maximum weights? Suppose that i edges have been chosen (perhaps by a greedy algorithm), then they can make a tree with $i + 1$ nodes and the worst case is that of a complete graph with $\binom{i+1}{2}$ edges. Then we are led to the following definition for MSF sequences.

Definition 6. An iis is a *MSF* sequence if and only if $a_i \leq \binom{i+1}{2} + 1$, $1 \leq i \leq r$.

An interesting consequence of the definition is that we have a polynomially increasing limit sequence for this class: in fact, the sequence with maximal elements is 1 2 4 7 11 16 22 29... $(r^2 + r + 2)/2$, which is the sequence of *central polygonal numbers* (sequence A000124 in [20]). Since 1 plus the sum of elements of the smallest sequence 1 2 3... i exactly equals the maximum value for a_i, then any MSF sequence is complete; however, MSF sequence 1 2 3 7 is surely not a tournament, while 1 2 4 8 is a tournament and not an MSF. However, MSF sequences (properly) include all Goldbach sequences: we shall prove in Section 4 that, in a Goldbach sequence, a_r cannot exceed the corresponding central polygonal number $(r^2 + r + 2)/2$. We shall now abandon MSF sequences to explore only subsets of tournaments, starting with sub-Fibonacci sequences.

Definition 7. An iis is a *sub-Fibonacci* sequence if and only if $a_i \leq a_{i-1} + a_{i-2} + 1$, $2 \leq i \leq r$; $a_1 = 2$.

The original definition in [9] referred to nondecreasing sequences and did not add 1 on the right hand side. It turns out that there is an immediate bijection
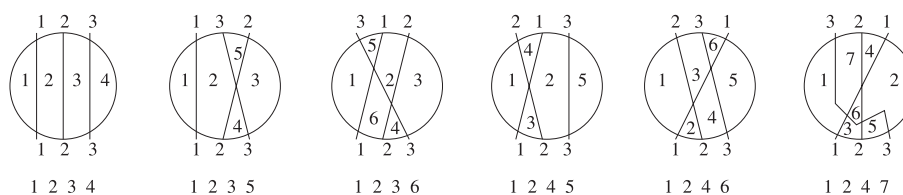
FIGURE 2. Integer sequences corresponding to 3-permutations.

between sequences "with the addition of 1" and those without, but with one more element. Although less elegant, Definition 7 permits to include sequences with $a_2 = 4$, which would otherwise be excluded, therefore designing a much richer class of sequences. In particular, as will be seen later on, this class properly includes permutations. The limit sequence 1 2 4 7 12 20... is given by *Fibonacci numbers* decreased by 1. The condition for tournaments is obviously satisfied, but 1 2 3 6 11, for instance, is both a tournament and an MSF, but it is not sub-Fibonacci. More interestingly, it may not be obvious that *a Goldbach sequence must be sub-Fibonacci*: suppose $a_i > a_{i-1} + a_{i-2} + 1$, then, which elements will generate $x = a_{i-1} + a_{i-2} + 1$? The only possibility is $x = a_{i-1} + a_{i-2} + 1 = 2a_{i-1}$, but now $a_i > 2a_{i-1}$.

Another surprise may be to find permutations among these classes: I cannot quote any paper where permutations are represented as increasing sequences. However, one of the referees pointed out the connection between arrangements of lines and permutations, as described *e.g.* in Chapter 2 of [7]. My approach was to start from the already mentioned sequence of central polygonal numbers: they count the maximum number of pieces that can be obtained by slicing a pancake with n cuts (see Fig. M1041 in [21]). Orderly numbering the cuts on one half of the pancake, then, on the other half, a permutation will be obtained, the sequence being given by the record of the number of pieces (not necessarily maximum) obtained after each cut. See Figure 2.

In the iis of a permutation, a_i is the number of pieces obtained after the i-th cut, and this is given by $a_{i-1} + 1 + $ (the number of *inversions* generated by element i in the permutation = the number of crossing points on line i). The number of inversions is the number of elements *smaller* than i which *follow* i in the permutation: since element i cannot be followed by more than $i - 1$ smaller elements, we are led to the following definition.

Definition 8. An iis is a *permutation* if and only if $a_i \leq a_{i-1} + i$, $1 \leq i \leq r$.

The table enumerating permutations according to the number of inversions (see [11] Vol. 3, Tab. 1 in Sect. 5.1.1), quite similar to that enumerating tournaments (see next section), has been illuminating in discovering how to represent permutations as a subclass of tournaments. In fact, *permutations are sub-Fibonacci*, since $i \leq a_{i-2} + 1$; on the other hand, 1 2 4 5 10 is sub-Fibonacci and Goldbach, but is not a permutation, while 1 2 4 7 is a non-Goldbach permutation.

All the definitions in this section have involved inequalities: it is time to define some more sequences through *equalities*, as we did for Goldbach sequences, where we asked for any x to be the sum of two elements in the sequence, were x in the sequence or not. By asking x to be the sum of two elements only if x itself is in the sequence, we obtain the well-studied class of (ascending) *addition chains*. Our notation for increasing integer sequences was chosen to conform to that for addition chains in [11] (Vol. 2), which is, besides, almost universally adopted.

Definition 9. An iis is an *addition chain* if and only if $a_i = a_k + a_j$ for some j, k with $0 \leq j \leq k < i$ and each $1 \leq i \leq r$. A *star chain* is an addition chain with $k = i - 1$ for each i.

1 2 4 5 8 is a non-star chain, while 1 2 4 8 is a star chain, showing that we are outside the class of sub-Fibonacci and permutation sequences, although Goldbach sequences are obviously contained in the class of addition chains, but not in that of star chains: 1 2 4 5 8 is Goldbach!

The last class we are going to introduce will be contained in all the classes we have defined (with the exception of Collatz sequences, which are, on the contrary, contained in it). In particular, we want these sequences to be Goldbach. To get to this point it is better to introduce a characterization of Goldbach sequences, having maximum n, through binary sequences $b_1, b_2, ..., b_n$ with $b_i = 1$ if there is a j such that $i = a_j$, 0 otherwise. Consider now a prefix $b_1, b_2, ..., b_k$ of the binary sequence: if the *convolution* $b_1 b_k + b_2 b_{k-1} + ... + b_k b_1$ is greater than 0, then there is a $b_i b_{k-i+1} = 1$, meaning that $i = a_j$, $k - i + 1 = a_m$, and therefore $a_j + a_m = k + 1$. If the convolution is positive for each k between 1 and $n - 1$, then the sequence is Goldbach! Suppose that, for each k, $b_1 + b_2 + ... + b_k > k/2$, *i.e.* there are more 1's than 0's: then the convolution must be positive, and the sequence is Goldbach. Now, consider the iis $a_0, a_1, ..., a_i$ and the corresponding binary sequence of length $a_i - 1$: this sequence has exactly i ones, and we want the number of zeros to be strictly less, at most $i - 1$, therefore $a_i - 1 \leq i + i - 1$ or $a_i \leq 2i$. We need a name for these sequences: since they are enumerated (according to the maximum value in the sequence) by central binomial coefficients (see next section) the name of *binomial sequences* seems to be apt.

Definition 10. An iis is a *binomial sequence* if and only if $a_i \leq 2i$, $1 \leq i \leq r$.

Having defined, at last, all the classes of increasing integer sequences we need, we can afford a digression on prime numbers and primoids. First of all, it is easy to see that *the sequence of primoids* (defined after Ex. 1) *is not a star chain*: since 4 is not in the sequence, any gap of 4 in the sequence will yield a counterexample, the first such gap occurring between 44 and 48. Element 48 is a_{23} in the sequence, thus showing also that the sequence of primoids is (obviously!) not binomial. We can state at least one simple theorem concerning primoids, Theorem 2 below.

Let us call *sub-permutation* a permutation such that $a_i \leq a_{i-1} + \lceil i/2 \rceil$, $1 \leq i \leq r$. There are only 4 sub-permutations of length 4: 1 2 3 4 5, 1 2 3 4 6, 1 2 3 5 6 and 1 2 3 5 7, and in general a_i has $\lceil i/2 \rceil$ possible values for $1 \leq i \leq r$, so that the number of sub-permutations of length r is given by sequence A010551 in [20].
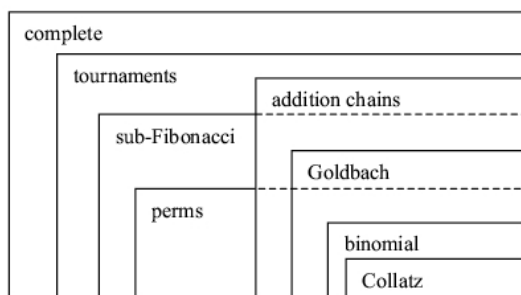
FIGURE 3. A Venn diagram showing the inclusions of the families of sequences.

Theorem 2. *Any prefix of the sequence of primoids is a sub-permutation.*

Proof. Let p_n be the nth prime. We first have to prove that $p_{n+1} \leq p_n + n$, a result not difficult to obtain, although not mentioned in any of the books on number theory I consulted. P. Dusart [6] proved that $p_n \geq n(\ln n + \ln \ln n - 1)$ for $n \geq 2$ and $p_n \leq n(\ln n + \ln \ln n - 0.9484)$ for $n \geq 39017$. Observing that $\ln(n+1) = \ln(n(1+1/n)) < \ln n + 1/n$, it is easy to write the upper bound for p_{n+1}, isolate in it the term $n(\ln n + \ln \ln n - 1)$ and substitute back p_n, obtaining $p_{n+1} \leq p_n + n$, or even a slightly better inequality, valid for $n \geq 39017$. A program check shows that $p_{n+1} \leq p_n + n$ is actually true for any $n \geq 1$.

Now we simply have to apply the inequality to primoids: since $3 = p_2 = 2a_0 + 1$, we have $p_n = 2a_{n-2} + 1$, from which $a_n \leq a_{n-1} + \lceil n/2 \rceil$ follows immediately. □

The largest gap between consecutive primes or primoids increases, as a matter of fact, much more slowly than predicted by the theorem, and, for example, $a_{94906079599} - a_{94906079598} = 326$ is the maximal gap thus far! Moreover, it seems that 6 is the only even number requiring doubling of an odd prime to obtain. If we call *elementary*, following Fishburn's terminology, the addition chains or Goldbach sequences in which doubling is forbidden (except trivially to obtain 2), then we could perhaps modify Goldbach's conjecture stating that *the sequence of primoids is an elementary Goldbach sequence.*

3. ENUMERATION

There are at least two ways of enumerating increasing integer sequences **s**: according to the value n of $\max(\mathbf{s})$ or according to the length $l(\mathbf{s}) = r$. When talking about "sequences" there is the possibility of some confusion between increasing integer sequences, which are the objects to be counted, and their *counting sequence*, which is the sequence giving their numbers as functions of r or n. We shall use the term "counting sequence" to avoid this confusion.

A first, very simple, result is that *the number of increasing integer sequences* **s** *such that* $\max(\mathbf{s}) = n$ *is* 2^{n-2}: in fact, $n - 2$ integers may appear (or not) between

$a_0 = 1$ and $a_r = n$. *Full sets* are counted according to number of elements (or corresponding *subexponential sequences* according to length) by sequence A001192 in [20]: 1 2 9 88 1802... Counting by maximum is not there, but we can adapt a formula in [2] to count the number of sequences satisfying any inequality whose right hand side increases only as a function of index i. While the formula is valid only for these classes, the recurrence relation is valid, with slight modifications, even for some sequences satisfying an inequality in which the bound is a function of a_i as well as of index i. Let us derive the recurrence equation first.

Let a class of increasing integer sequences be defined by $a_i \leq f(i)$, where f is an integer-valued function such that $f(i) > f(i-1) \geq i \geq 1$. Denoting with $S(r, n)$ the number of sequences of length r and maximum n, we can easily deduce recurrence relations, the boundary conditions being $S(r, n) = 0$ if $r \geq n$ or $n > f(r)$, $S(n-1, n) = 1$. Since from $a_0, a_1, ..., a_{r-1}$ we can obtain a sequence of length r by adding $a_r = n$ if and only if $a_{r-1} < n \leq f(r)$, the recurrence relation is the following:

$$S(r, n) = \sum_{j=r}^{n-1} S(r-1, j) \text{ for } 0 < r < n \leq f(r). \tag{1}$$

We are interested in the counting sequence according to length r: $S(r, *) := \sum_{n=r+1}^{f(r)} S(r, n)$ and the counting sequence according to maximum element n: $S(*, n) := \sum_{r=0}^{n-1} S(r, n)$.

Theorem 3. $S(*, n+1) = 2S(*, n) - \delta_{n,f(r)} S(r-1, *)$.

Proof. Kronecker function $\delta_{n,f(r)}$ is 1 if $f(r) = n$, 0 otherwise. Let $a_0, a_1, ..., a_{r-1}, n$ be an iis with maximum n and length r. From it, we can obtain two sequences with maximum $n + 1$: $a_0, a_1, ..., a_{r-1}, n, n+1$ and $a_0, a_1, ..., a_{r-1}, n+1$. While the former can always be obtained, the latter is only possible if $n + 1 \leq f(r)$. In case $n = f(r)$, all sequences $a_0, a_1, ..., a_{r-1}, n$ must be subtracted, and there are $S(r-1, *)$ of them. $\square$

Theorem 3 shows that $S(*, n+1) \leq 2S(*, n)$ and permits to derive $S(*, n)$ once $S(r, *)$ is known. For instance, there are 3 full sets with maximum 5: $S(*, 5) = 3$, $f(i) = 2^i$, therefore we can double up to $n = 8$: $S(*, 8) = 24$, then $S(*, 9) = 48 - S(2, *) = 46$, then keep on doubling up to 16... But Theorem 3 also permits to deduce, *vice versa*, $S(r, *)$ from $S(*, n)$, since, by inversion when $f(r+1) = n$, we have $S(r, *) = 2 \ S(*, f(r+1)) - S(*, f(r+1)+1))$. However, what can be done if we do not have either sequence? Then, to find $S(r, *)$, we can use the following theorem.

Theorem 4.

$$S(r, *) = \sum_{k=0}^{r-1} (-1)^k \binom{f(r-k) - (r-k)}{k+1} S(r-k-1, *).$$

TABLE 1. Values of $T(r,n)$ and $T(r) = T(r,*)$, $1 \le r \le 5$, $n = k + r$, $1 \le k \le 27$

r \ k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	T(r)
1	1																											1
2	1	1																										2
3	1	2	2	1	1																							7
4	1	3	5	6	6	6	4	4	2	2	1	1																41
5	1	4	9	15	21	26	30	31	33	30	31	26	26	20	20	14	14	10	10	6	6	4	4	2	2	1	1	397

Proof. By adapting Theorem 1 in [2] to increasing integer sequences.   □

Unfortunately, no similar theorems seem to be available when a_i satisfies inequalities involving elements of the sequence. However, these formulae solve the enumeration problem for full sets, MSF and binomial sequences. We still have to justify the name of binomial sequences: remember that binomial sequences are represented by binary vectors whose prefixes always have more ones than zeros. These sequences can be interpreted as paths in constrained regions of the plane, and a little experimenting will convince that the number of sequences with maximum n is the *central binomial coefficient* $\binom{n-2}{\lfloor (n-2)/2 \rfloor}$, proofs being similar to classical proofs for Catalan numbers. The corresponding enumeration by number of ones (which is the length of the iis) yields in fact *Catalan numbers* $\frac{1}{r+1}\binom{2r}{r}$.

Tournaments are enumerated according to maximum element as sequence A002083 in [20]: 1 2 3 6 11 22 42 84 165 ... and a simple recurrence formula is given there (*cf.* also [4] and [16]). To enumerate tournaments according to length, let $T(r,n)$ be the number of tournaments $\mathbf{t}$ with $l(\mathbf{t}) = r$ and $\max(\mathbf{t}) = n$. Recurrence (1) is still valid, provided the start value r in the summation is replaced by $\lceil n/2 \rceil$, since tournaments with one more element, of value n, can be obtained by sequences ending with an element at least $n/2$. A few values of $T(r,n)$ are tabulated in Table 1; columns are labeled by values of k to obtain $n = k + r$, so that the table can be contained in the page and resemble Knuth's table quoted after Definition 8. The last column reports values of $T(r)$, the number of tournaments of length r, irrespective of n. The sequence of values of $T(r)$ is now sequence A008934 in [20].

The enumeration of *sub-Fibonacci* sequences also requires a two-variable recurrence, which is as simple as that for tournaments. Consider a sequence ending with the two elements j and m, and let $F(j,m)$ denote the number of sub-Fibonacci sequences of this sort. From each such sequence we can generate $j + 1$ new sequences by appending $m + 1$, or $m + 2$, ..., or $m + j + 1$. Therefore, we obtain the recurrence relation $F(m,n) = \sum_{j=n-m-1}^{m-1} F(j,m)$ since $n \le m + j + 1$ yields $j \ge n - m - 1$. The boundary conditions can be expressed as $F(m,n) = 0$ if $m \ge n$ or $m < n/2$, $F(1,2) = F(2,3) = F(2,4) = 1$. The counting sequence for $F(*,n)$, the number of sequences ending with n, begins 1 1 1 2 3 6 11 21 41 81 159 316. The counting sequence according to length is sequence A005270 in [20].

Permutations are obviously very easy to count according to length, since there are $r!$ of them. Just a bit more surprising is the fact that also *star chains* are $r!$: from a chain of length $r-1$ it is possible to obtain r star chains of length r by successively adding a_{r-1} to $a_0, a_1, ..., a_{r-1}$, and there is just 1 star chain of length 1. It is also possible to exhibit an explicit bijection between the two classes. Consider the class of integer functions f such that $0 \leq f(i) < i$ for each i with $2 \leq i \leq r$. We do not require f to be increasing nor even monotone. There are $r!$ such functions, since there are i independent choices for each $f(i)$. Now we can use f in at least two ways to obtain increasing sequences: elements $a_0 = 1$ and $a_1 = 2$ of the sequence are fixed, while for $2 \leq i \leq r$

1) $a_i = a_{i-1} + a_{f(i)}$: we obtain a star chain, by Definition 9, one for each f;
2) $a_i = a_{i-1} + f(i) + 1$: we obtain a permutation, by Definition 8, since $f(i) + 1 \leq i$.

The recurrence equation for $P(r, n)$, the number of permutations of length r and maximum n, is given by (1), starting the summation from the maximum between $n - r$ and r, since $n - r$ is the minimum value that the element preceding n in the sequence can have, provided it is greater than $r - 1$. The counting sequence starts 1 1 1 2 3 6 11 21 41 80 157 310 (sequence A008930 in [20]).

Things are still not too difficult for *complete* sequences, despite the apparently complicate inequality in the definition. We only need a recurrence in *three* variables: denoting by $C(r, n, s)$ the number of complete sequences of length r, maximum n and sum $\sigma(r) = s$, any such sequence can be obtained by a sequence of length $r-1$, maximum x and sum $s-n$ by simply appending n, provided x is smaller than n and also not greater than $s - n$. Therefore, setting $y := min(n-1, s-n)$, the required recurrence is

$$C(r, n, s) = \sum_{x=r}^{y} C(r - 1, x, s - n).$$

The counting sequences are

$$C(r, *, *) = 1\ 2\ 8\ 60\ 814\ 19682\ 845368\ ...$$

$$C(*, n, *) = 1\ 1\ 1\ 2\ 3\ 6\ 12\ 23\ 45\ 90\ 180\ 359\ ...$$

Well, what about *addition chains* and *Goldbach sequences*? I wrote some programs to compute them, by explicitly generating and then counting, but recurrence formulae have eluded me up to now. We will see some motivations for this unsatisfactory situation in the next section. However, having found that binomial sequences are common to all classes helps to speed up testing, while, on the other hand, upper bounds for Goldbach sequences (see next section) allow to eliminate elements too big. We can display two comparative tables: Table 2 lists increasing integer sequences of different types, enumerated according to *maximum* element, while Table 3 lists them according to *length*.

Table 3 is quite misleading if used to compare different sequences, since the maximum element of a binomial sequence of length 8 can be at most 16, while

TABLE 2. Increasing integer sequences of different types enumerated according to maximum element. The first six values are common to all sequences: 1 1 1 2 3 6.

Max	*Chains	Binomial	Goldbach	Chains	Perms	Tournaments
7	10	10	10	10	11	11
8	20	20	20	21	21	22
9	36	35	37	38	41	42
10	70	70	73	77	80	84
11	130	126	139	144	157	165
12	252	252	275	293	310	330
13	475	462	533	563	614	654
14	916	924	1059	1131	1218	1308
15	1745	1716	2075	2205	2421	2605
16	3362	3432	4126	4434	4819	5210
17	6438	6435	8134	8711	9602	10398
18	12410	12870	16194	17466	19147	20796
19	23852	24310	32058	34506	38204	41550
20	46020	48620	63910	69169	76266	83100

TABLE 3. Increasing integer sequences of different types enumerated according to length. The first two values are common to all sequences: 1 2.

Length	Binomial	Goldbach	Perms &*	Chains	Tournam.	Full Sets
3	5	5	6	6	7	9
4	14	17	24	25	41	88
5	42	65	120	135	397	1802
6	132	292	720	913	6377	75598
7	429	1434	5040	7499	171886	6421599
8	1430	7875	40320	73191	7892642	1097780312

in a chain, tournament or full set it can be as big as 256. On the other hand, it is rather surprising to discover, in Table 2, that star chains are asymptotically less than binomial sequences, when enumerated according to maximum element, whereas permutations are not much less than tournaments.

How much less than tournaments are Goldbach sequences? We only know at present that Goldbach sequences are asymptotically more than binomial sequences, and central binomial coefficients grow slower than $2^n/\sqrt{n}$. Since the number of tournaments grows faster than 2^{n-4}, the ratio Goldbach/tournaments might approach 0 as n tends to infinity, meaning in a sense that an infinite tournament cannot be Goldbach by chance. We can, however, exhibit Goldbach sequences much shorter than the primoids, and establish a lower bound on the length of Goldbach sequences with maximum n, which will be done in the next section.

4. Looking for shortest sequences

Let us consider a Goldbach sequence $a_0, a_1, ..., a_r$: we can take (unordered) pairs of elements in $\binom{r+1}{2}$ ways, but, to obtain an x between 2 and a_r as sum of one or more pairs, we do not need a_r: however, we can use a_r merely to represent *doubling* of the other element in the pair. Since we have to obtain all the $a_r - 1$ values x between 2 and a_r, even if all the pairs yield distinct sums, a_r cannot exceed $\binom{r+1}{2} + 1 = \frac{r^2+r+2}{2}$, the *central polygonal numbers*. After inverting the inequality we can state the following theorem.

Theorem 5. *In a Goldbach sequence* $\mathbf{s}$ *with* $l(\mathbf{s}) = r$ *and* $\max(\mathbf{s}) = n$ *the following inequalities must hold:*

$$r \geq \frac{\sqrt{8n-7}-1}{2}, \ n \leq \frac{r^2+r+2}{2}.$$

Unfortunately, these inequalities only yield the order of magnitude, but are not strict and in fact grossly overestimate the strict bounds: for instance, the maximum element for Goldbach sequences of length 9 is 34 and not 46 as given by the theorem. Looking at computer-generated tables for the sequences reaching maxima, we were struck by the symmetries of the sequences up to length 9. Highly symmetric sequences attain a bound for n of

$$n = \frac{r^2 + 6r + m(2-m)}{4}, \ \text{where } m = r \bmod 4. \tag{2}$$

In fact, for each $r = 4d + m$, where $d = \lfloor r/4 \rfloor$, we can set $k = (2d+m)(d+1)$ and construct the following sequence: 1, 2, 3, ..., $d+1$, $2(d+1)$, $3(d+1)$, ..., k, $k+1$, $k+2$, ..., $k+d$, $2(k+d)$. Then $n = a_r = 2(k+d)$ is given exactly by Equation (2), while the number of elements in the sequence is $d+(2d+m)+(d+1) = 4d+m+1 = r+1$, therefore the length is r. An *example* will both make things clear and convince that such a sequence is Goldbach. For $r = 9$ we have $d = 2$, $m = 1$ and $k = 15$: the sequence is 1 2 3 6 9 12 15 16 17 34. There is no doubt that values up to 17 can be obtained, since gaps do not exceed 3, and 1 2 3 *are* in the sequence: but the sequence, excluding a_r, is *symmetric* ($a_i+a_{r-i-1} = a_{r-1}+a_0$), and so if $a_i+a_j = x$, then $a_{r-j-1} + a_{r-i-1} = a_{r-j-1} + a_{r-i-1} + a_i + a_j - (a_i+a_j) = 2a_{r-1}+2-x$: all values up to $2a_{r-1}$ are obtained as well.

The sequence of maximum values for n, according to Equation (2), is 1 2 4 6 10 14 18 22 28 34 40 ... Unfortunately, (2) is not an upper bound for lengths greater than 9: for $r = 10$ the sequence 1 2 4 5 10 12 17 18 20 21 42 attains the maximum, but $42 > 40$.

However, needing a sequence for 30, say, we know that a sequence of length 8 will not suffice, while we can use the sequence for 34 substituting 30 to 34. The biggest the number, the longest the sequence... Actually, the requirement of obtaining all values x, specific to Goldbach sequences, has been the key to Theorem 5 and Equation (2). For *addition chains*, however, this requirement is absent and "no

conjecture about addition chains is safe!", as Knuth writes. In particular, the length of the shortest chain, not being monotone, strongly depends on the value of n. Were it possible to find the counting function $S(r, n)$ for addition chains, or even some recurrence relation, then we would know (in principle) the shortest length a chain for n can have, thus perhaps solving both the *Scholz-Brauer conjecture* and several other problems of practical as well as theoretical interest.

5. Conclusions

We have rather few conclusions and a lot of open questions, or work to do. We hope to have succeeded in showing that the subject is worth some more study, so to have a commencement, not a conclusion. At least, we would like to obtain some more recurrences or closed-form formulae for the sequences just defined, in particular Goldbach sequences and addition chains, so to have information on the asymptotic behavior of their number. This is almost all we can say at present about a hypothetical contribution of this approach to the conjectures themselves which originated it.

Other conjectures in number theory might be explored by examining the set of sequences having the conjectured property; moreover, all the work concerning Gilbreath's conjecture and the $3x + 1$ problem is still to be done.

Different structures, such as *Davenport-Schinzel sequences* (E20 in[10]), might also be connected some way with increasing integer sequences. Moreover, a few known counting sequences are also interestingly "similar" to some counting sequences of ours, suggesting they might count "similar" objects: *e.g.* sequence A001258 in [20], concerning labeled trees with unlabeled end-points, is the same as the sequence of chains up to length 5, and then smaller; sequence A005130, counting alternating sign matrices [18], is 1 2 7 42 429 ..., just over tournaments... All these may well be worth exploring.

References

[1] P. Capell and T.V. Narayana, On Knock-out Tournaments. *Canad. Math. Bull.* **13** (1970) 105–109.
[2] L. Carlitz, D.P. Roselle and R.A. Scoville, Some Remarks on Ballot-Type Sequences of Positive Integers. *J. Combinatorial Theory, Ser.* A **11** (1971) 258–271.
[3] L. Comtet, *Advanced Combinatorics*. Reidel, Dordrecht (1974).
[4] M. Cook and M. Kleber, Tournament sequences and Meeussen sequences. *Electron. J. Combin.* **7** (2000), Research Paper 44, p. 16 (electronic).

[5] J-M. Deshouillers, H.J.J. te Riele and Y. Saouter, New experimental results concerning the Goldbach conjecture, in *Proc. 3rd Int. Symp. on Algorithmic Number Theory. Lect. Notes Comput. Sci.* **1423** (1998) 204–215.

[6] P. Dusart, The kth prime is greater than $k(\ln k + \ln \ln k - 1)$ for $k \geq 2$. *Math. Comp.* **68** (1999) 411–415.

[7] H. Edelsbrunner, *Algorithms in Combinatorial Geometry*. Springer, Berlin (1987).

[8] P. Erdös and M. Lewin, d-Complete Sequences of Integers. *Math. Comp.* **65** (1996) 837–840.

[9] P.C. Fishburn and F.S. Roberts, Uniqueness in finite measurement, in *Applications of Combinatorics and Graph Theory to the Biological and Social Sciences*, edited by F.S. Roberts. Springer, New York (1989) 103–137.

[10] R.K. Guy, *Unsolved Problems in Number Theory*. Springer, New York, 2nd ed. (1994).

[11] D.E. Knuth, *The Art of Computer Programming*, Vols. 2 and 3. Addison-Wesley, Reading (1997 and 1998).

[12] J.C. Lagarias, The $3x + 1$ problem and its generalizations. *American Math. Monthly* **92** (1985) 3–23.

[13] W.J. LeVeque, *Fundamentals of Number Theory*. Addison-Wesley, Reading (1977).

[14] G. Melfi, On Two Conjectures about Practical Numbers. *J. Number Theory* **56** (1996) 205–210.

[15] M.B. Nathanson, *Additive Number Theory*. Springer, New York (1996).

[16] E. Neuwirth, Computing Tournament Sequence Numbers efficiently with Matrix Techniques. *Sém. Lothar. Combin.* **47** (2002), Article B47h, p. 12 (electronic).

[17] A.M. Odlyzko, Iterated Absolute Values of Differences of Consecutive Primes. *Math. Comp.* **61** (1993) 373–380.

[18] D.P. Robbins, The Story of 1, 2, 7, 42, 429, 7436, ... *Math. Intellig.* **13** (1991) 12–19.

[19] J. Richstein, Verifying the Goldbach Conjecture up to 4×10^{14}. *Math. Comp.* **70** (2001) 1745–1749.

[20] N.J.A. Sloane, *The On-Line Encyclopedia of Integer Sequences*. World-Wide Web URL `http://www.research.att.com/~njas/sequences/`

[21] N.J.A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences*. Academic Press, New York (1995).