

# Isoperimetric Sets of Integers

Steven J. Miller (Steven.J.Miller@williams.edu)

Frank Morgan (Frank.Morgan@williams.edu)

Edward Newkirk (Edward.S.Newkirk@williams.edu)

Lori Pedersen (lap811@aol.com)

Deividas Seferis (deividas.seferis@gmail.com)

Williams College

Williamstown, MA 01267

The celebrated isoperimetric theorem says that the circle provides the least-perimeter way to enclose a given area. In this note we discuss a generalization which arose at a departmental research seminar workshop [1] and which moves the isoperimetric problem from geometry to number theory and combinatorics. Instead of Euclidean space, let's take the set  $X$  of non-negative integers:

$$X = \{0, 1, 2, 3, \dots\}.$$

For any subset  $S$  of  $X$ , we define volume and perimeter as follows:

$\text{vol}(S) :=$  sum of elements of  $S$

$\text{per}(S) :=$  sum of elements of  $S$  whose predecessor and successor are not both in  $S$ .

For example, for  $S = \{0, 1, 4, 5, 6, 7\}$ ,  $\text{vol}(S) = 23$  and  $\text{per}(S) = 0 + 1 + 4 + 7 = 12$ . This definition is a natural generalization of the original problem; the perimeter comes from elements that are on the 'boundary' of our set, as these elements are in  $S$  but have a neighbor that is not.

We can now state the problem we want to consider:

**Isoperimetric Set Problem.** *Among all sets  $S \subset X$  whose volume is  $n$ , find the set  $S$  with smallest perimeter.*

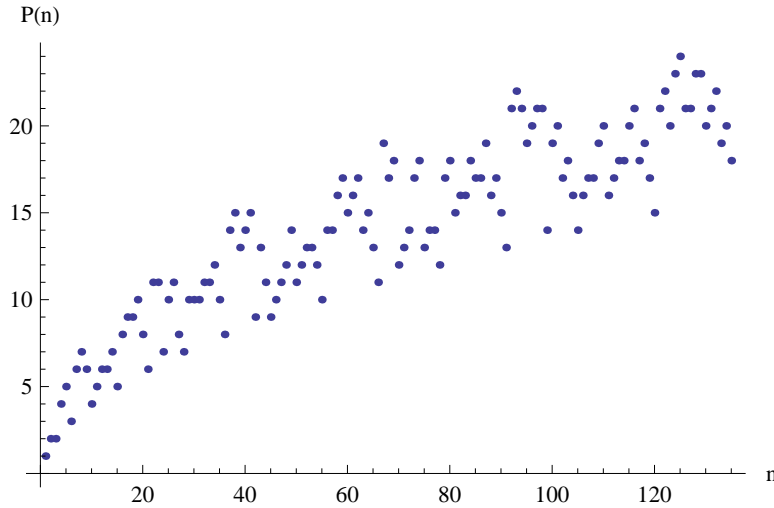


Figure 1: Plot of minimum perimeter  $P(n)$  versus prescribed volume  $n$ .

For example, among all sets of non-negative integers that sum to  $n = 9$ , the smallest possible perimeter turns out to be 6, with  $S = \{2, 3, 4\}$ . For  $n = 19$ , the smallest perimeter is 10, with  $S = \{1, 3, 4, 5, 6\}$ . For small  $n$ , we find  $P(n)$  by an exhaustive computer search; see Figure 1 for a plot.

How large is a typical  $P(n)$  relative to  $n$ ? Does  $P(n)$  grow linearly with  $n$ , or is it significantly slower? Unfortunately, our plot above has only 135 data points, and it is very easy to be misled as to the limiting behavior from such a small data set. Similar limitations of computations arise when counting prime numbers. The famous Prime Number Theorem states that the fraction of numbers from 1 to  $n$  which are prime is about  $1/\log n$ . This tends to zero, albeit very slowly. Of the numbers from 1 to 10, 40% are prime (2, 3, 5 and 7), and of the numbers from 1 to 20, still 40% are prime, without any indication that the percentage is going to 0.

As our problem is a generalization of the standard isoperimetric problem, perhaps the analysis there can suggest what the true behavior of  $P(n)$  should be. Although our problem lies on the number line, each point is weighted by its value, in some sense adding a dimension. So perhaps the two-dimensional

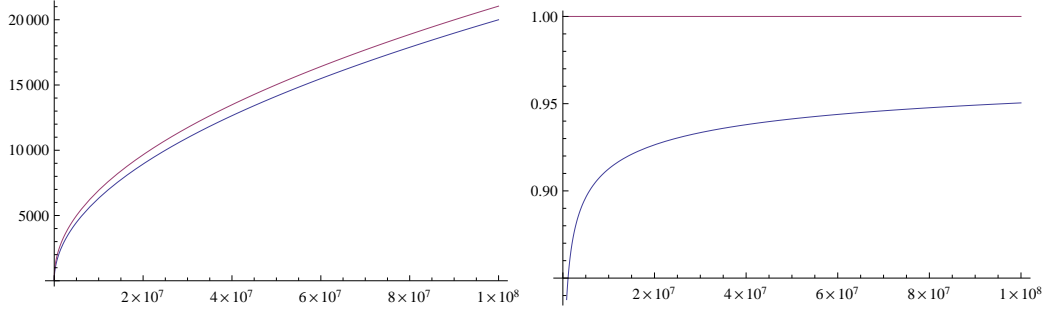


Figure 2: *Left*: plot of the upper bound  $\sqrt{2}n^{1/2} + (2n^{1/4} + 8) \log_2 \log_2 n + 58$  and the lower bound  $\sqrt{2}n^{1/2} - 1/2$ . As  $n$  approaches infinity, our two bounds asymptotically approach  $\sqrt{2}n^{1/2}$ . *Right*: plot of the ratio of the upper and lower bound. The two bounds have the same growth rate in the limit, but the convergence is very slow.

classical isoperimetric problem would be a good guide. We know there that the optimal solution is a circle. As the area of a circle of radius  $r$  is  $A = \pi r^2$  and the perimeter is  $P = 2\pi r$ , simple algebra yields that the least perimeter for a given area is  $P = 2\sqrt{\pi}A^{1/2}$ . Given this, it is not unreasonable to conjecture for large  $n$  that  $P(n)$  should approximately equal a constant times  $\sqrt{n}$ .

It turns out that the classical problem *does* provide the right intuition, as the following proposition shows. It says that  $P(n)$  is asymptotic to  $\sqrt{2}n^{1/2}$ , that is, the ratio of  $P(n)$  to  $\sqrt{2}n^{1/2}$  approaches 1 as  $n$  tends to infinity. The two bounds are graphed in Figure 2.

The proof of the lower bound is trivial. The proof of the upper bound is a tricky induction argument. The hardest part was finding exactly the right induction hypothesis, a process which involved much trial and error which we spare the kind reader.

**Proposition.**  $P(n) \sim \sqrt{2}n^{1/2}$ . Indeed,

$$\sqrt{2}n^{1/2} - 1/2 < P(n) < \sqrt{2}n^{1/2} + (2n^{1/4} + 8) \log_2 \log_2 n + 58.$$

*Proof:* Clearly the perimeter of  $S$  is at least as large as the largest element of  $S$ , which we denote by  $m$ . For the lower bound, if  $\text{vol}(S) = n$  then  $m$  must satisfy

$$n \leq 0 + 1 + \cdots + m = \frac{m(m+1)}{2}. \quad (1)$$

This means that  $m$  is bounded below by a certain function  $f(n)$ , namely (by the quadratic formula), if we set

$$f(n) := \left(2n + \frac{1}{4}\right)^{1/2} - \frac{1}{2}$$

then

$$m \geq f(n) = \left(2n + \frac{1}{4}\right)^{1/2} - \frac{1}{2} > \sqrt{2}n^{1/2} - 1/2,$$

as desired.

For the upper bound, we use the following greedy algorithm. Choose the largest term  $m_1$  in  $S$  to be as small as possible. Continue taking consecutive numbers as long as possible, through  $m_2 \leq m_1$ . Choose the next term  $m_3 < m_2$  as small as possible. Continue. For example, for  $n = 19$ , this algorithm yields  $S = \{6, 5, 4, 3, 1\}$ , which turns out to be optimal. For  $n = 11$  it yields  $\{5, 4, 2\}$  with perimeter 11, worse than the optimal  $\{5, 3, 2, 1, 0\}$  with perimeter 8.

First note that  $m_1 < f(n) + 1$ , because for any integer  $m \geq f(n)$ , (1) holds. Similarly for any odd  $k$ ,  $m_k < f(m_{k-1} - 1) + 1$  because by choice of  $m_{k-1}$ , the remaining volume is less than  $m_{k-1} - 1$ . For even  $k$ ,  $m_k - 1$  exceeded the remaining volume, so  $m = m_k - 2$  satisfies

$$\frac{m(m+1)}{2} < m_{k-1}$$

by choice of  $m_{k-1}$ . Hence  $m_k - 2 < f(m_{k-1})$ . In summary, for all  $k \geq 2$ ,

$$m_k < f(m_{k-1}) + 2. \quad (2)$$

To avoid the complications of that  $+2$ , we move to a closely related function  $g(m)$ . Note that

$$\begin{aligned} f(m) + 2 &= \left(2m + \frac{1}{4}\right)^{1/2} + \frac{3}{2} \\ &\leq \sqrt{2}m^{1/2} + \frac{1}{2} + \frac{3}{2} \\ &= \sqrt{2}m^{1/2} + 2 = g(m). \end{aligned}$$

We now consider *compositional powers*  $g^k(n)$ , where for example  $g^3(n) = g(g(g(n)))$ . By (2), for  $k \geq 2$ , the compositional power  $g^k(n)$  satisfies

$$m_k < g^k(n), \quad (3)$$

without the pesky  $+2$  of (2). Now all we need is an upper bound on  $g^k(n)$ . We'll prove by induction that for  $k \geq 1$

$$g^k(n) < 2^{1-1/2^k} n^{1/2^k} + 8. \quad (4)$$

The base case  $k = 1$  is immediate. The induction step takes just a little algebra:

$$\begin{aligned} g^{k+1}(n) &< 2^{1/2} \left(2^{1-1/2^k} n^{1/2^k} + 8\right)^{1/2} + 2 \\ &< 2^{1/2} \left(2^{1/2-1/2^{k+1}} n^{1/2^{k+1}} + 8^{1/2}\right) + 4 \\ &= 2^{1-1/2^{k+1}} n^{1/2^{k+1}} + 8. \end{aligned}$$

Now we can translate our upper bound for  $g^k(n)$  into an upper bound for  $P(n)$ . Indeed, by (3) and (4) it follows that

$$\begin{aligned} P(n) &\leq m_1 + m_2 + m_3 + \cdots \\ &< (2^{1/2} n^{1/2} + 2) + (2^{3/4} n^{1/4} + 8) + (2^{7/8} n^{1/8} + 8) + \cdots \\ &< 2^{1/2} n^{1/2} + (2n^{1/4} + 8)(\log_2 \log_2 n - 1) + 68, \end{aligned}$$

because there are at most  $\log_2 \log_2 n$  terms with  $n^{1/2^k} \geq 2$  and the rest by (4) are less than 12, with sum at most 66. Hence

$$P(n) < \sqrt{2}n^{1/2} + (2n^{1/4} + 8) \log_2 \log_2 n + 58,$$

as desired.  $\square$

It is interesting to consider other sets  $X$ . For example, for the set of harmonic numbers  $X = \{1/1, 1/2, 1/3, \dots\}$  one may attain arbitrary positive volume (via infinitely many terms). We know almost nothing about the minimum perimeter.

Roger Bolton proposed including negative numbers, say all the integers in order:

$$X = Z := \{\dots, -1, 0, 1, 2, \dots\}; \tag{5}$$

perimeter should now be defined as a sum of the absolute values. Now certain types of holes in the previous solution can be filled to reduce the perimeter. First, if the smallest term in the previous solution is  $k > 1$ , then adding the terms  $-(k-1), \dots, 0, \dots, k-1$  will leave the volume unaffected and reduce the perimeter by 1. For instance, the minimizer for volume 9 can be improved from  $\{2, 3, 4\}$  with perimeter 6 to  $\{-1, 0, 1, 2, 3, 4\}$  with perimeter 5. Second, if there is a hole in a non-negative sequence with at least two terms before and after, e.g.  $\dots, a-1, a+1, \dots$ , adding  $\{-a, a\}$  will leave the volume untouched while reducing the perimeter by  $a$ . For instance, the minimizer for volume 19 can be improved from  $\{0, 1, 3, 4, 5, 6\}$  with perimeter 10 to  $\{-2, 0, 1, 2, 3, 4, 5, 6\}$  with perimeter 8.

Often in number theory a related problem allowing differences as well as sums is significantly easier to solve than the original allowing just sums. One famous example is Waring's Problem and the Easier Waring's Problem. Waring's Problem states that for each  $k$  there is an  $s(k)$  such that all positive integers are a sum of at most  $s(k)$   $k^{\text{th}}$  powers. The Easier Waring's Problem

allows differences as well as sums. For example, 7 is not the sum of three squares, but  $7 = 3^2 - 1^2 - 1^2$ . It can be proved in half a page [2, p. 102].

Similarly, we can study the Easier Isoperimetric Sequence Problem, where we allow pluses and minuses when summing the elements of the subset  $S$  to obtain the prescribed volume  $n$ . In this case it is significantly easier to analyze the fluctuations about  $\sqrt{2}n^{1/2}$ . Let  $EP(n)$  denote the minimum perimeter (with all contributions to perimeter still positive) for our related problem, which is obviously at most  $P(n)$ . The lower bound is still  $\sqrt{2}n^{1/2} - 1/2$ , but now we can remove the  $\log_2 \log_2 n$  factor in the upper bound, as it is easy to show that  $EP(n) < \sqrt{2}n^{1/2} + 4$ . Indeed, choose the smallest  $k$  such that  $0 + 1 + 2 + \cdots + k = k(k+1)/2 \geq n$ ; then  $k < \sqrt{2}n^{1/2} + 1$ . To obtain a sum of exactly  $n$ , take a minus sign on one term, changing the sum by an even integer to  $n$  or  $n+1$ , and in the latter case drop the 1, adding 2 to the perimeter. In the exceptional case when  $k(k+1)/2 = n+3$ , drop the 1 and the 2, adding 3 to the perimeter. In any case,  $EP(n) < \sqrt{2}n^{1/2} + 4$ , and thus the fluctuations about  $\sqrt{2}n^{1/2}$  cannot be larger than 5.

We end with some questions for further research. What does the minimum perimeter function  $P(n)$  say about the number theoretical properties of a set  $X$ ? What are some interesting examples? What is the true scale of fluctuations of  $P(n)$  about  $\sqrt{2}n^{1/2}$  when  $X$  is the non-negative integers? Are the fluctuations frequently as large as  $\log_2 \log_2 n$ ? In other words, what can we say about  $P(n) - \sqrt{2}n^{1/2}$ ? Is there some  $h(n)$  so that  $(P(n) - \sqrt{2}n^{1/2})/h(n)$  has a nice limiting distribution as  $n \rightarrow \infty$ ?

**Acknowledgment** We thank the referees and our colleagues for comments. Miller and Morgan acknowledge partial support from the National Science Foundation.

## References

1. Frank Morgan, Isoperimetric sequences, <http://blogs.williams.edu/Morgan/>, 2008.
2. Melvyn B. Nathanson, Additive Number Theory: The Classical Bases, Graduate Texts in Mathematics **164**, Springer-Verlag, New York, 1996.