FARMER TED GOES NATURAL

GREG MARTIN

1. SETTING THE STAGE

We've all been given a problem in a calculus class remarkably similar to the following one:

> Farmer Ted is building a chicken coop. He decides he can spare 190 square feet of his land for the coop, which will be built in the shape of a rectangle. Being a practical man, Farmer Ted wants to spend as little as possible on the chicken wire for the fence. What dimensions should he make the chicken coop?

By solving a simple optimization problem, we learn that Farmer Ted should make his chicken coop a square with side lengths $\sqrt{190}$ feet. And that, according to the solution manual, is that.

But the calculus books don't tell the rest of the story:

> So Farmer Ted went over to Builders Square and told the salesman, "I'd like $4\sqrt{190}$ feet of chicken wire, please." The salesman, however, replied that he could sell one foot or two feet or a hundred feet of chicken wire, but what the heck was $4\sqrt{190}$ feet of chicken wire? Farmer Ted was taken aback, explaining heatedly that his family had been buying as little chicken wire as possible for generations, and he really wanted $4\sqrt{190}$ feet of chicken wire measured off for him immediately! But the salesman, fearing more irrational behavior from Farmer Ted, told him, "I don't want to hear about your roots. We do business in a natural way here, and if you don't like it you can leave the whole store." Well, Farmer Ted didn't feel that this treatment was commensurate with his request, but he left Builders Square to rethink his coop from square one.

At first, Farmer Ted thought his best bet would be to make a $10' \times 19'$ chicken coop, necessitating the purchase of 58 feet of chicken wire—certainly this was better than 86 feet of chicken wire for a $5' \times 38'$ coop, say. But then he realized that he could be more cost-effective by not using all of the 190 square feet of land he had reserved for the coop. For instance, he could construct an $11' \times 17'$ coop (187 square feet) with only 56 feet of chicken wire; this would give him about 3.34 square feet of coop space per foot of chicken wire purchased, as opposed to only 3.28 square feet per chicken-wire-foot for the $10' \times 19'$ coop. Naturally, the parsimonious farmer wondered: could he do even better?

2. POSING THE PROBLEM

Jon Grantham posed the following problem at the 1998 SouthEast Regional Meeting On Numbers in Greensboro, North Carolina: given a positive integer N, find the dimensions of the rectangle with integer side lengths and area at most N whose area-to-perimeter ratio is largest among all such rectangles. In the story above, Farmer Ted is trying to solve this problem for $N = 190$.

Let's introduce some notation so we can formulate Grantham's problem more precisely. For a positive integer n, let $s(n)$ denote the least possible semiperimeter (length plus width) of a rectangle with integer side lengths and area n. (Since the area-to-semiperimeter ratio of a rectangle is always twice the area-to-perimeter ratio, it doesn't really change the problem

if we consider semiperimeters instead of perimeters; this will eliminate annoying factors of 2 in many of our formulas.) In other (and fewer) words,

$$s(n) = \min_{cd=n}(c+d) = \min_{d|n}(d + n/d),$$

where $d \mid n$ means that d divides n.

Let $F(n) = n/s(n)$ denote the area-to-semiperimeter ratio in which we are interested. We want to investigate the integers n such that $F(n)$ is large, and so we define the set $\mathcal{A}$ of "record-breakers" for the function F as follows:

$$\mathcal{A} = \{n \in \mathbb{N} \colon F(k) \le F(n) \text{ for all } k \le n\}. \tag{1}$$

(Well, the "record-tiers" are also included in $\mathcal{A}$.) Then it is clear after a moment's thought that to solve Grantham's problem for a given number N, we simply need to find the largest element of $\mathcal{A}$ not exceeding N.

By computing all possible factorizations of the numbers up to 200 by brute force, we can make a list of the first 59 elements of $\mathcal{A}$:

$\mathcal{A} = \{1, 2, 3, 4, 6, 8, 9, 12, 15, 16, 18, 20, 24, 25, 28, 30, 35, 36, 40, 42, 48, 49, 54, 56, 60, 63, 64, 70, 72, 77, 80, 81, 88, 90, 96, 99, 100, 108, 110, 117, 120, 121, 130, 132, 140, 143, 144, 150, 154, 156, 165, 168, 169, 176, 180, 182, 192, 195, 196, \ldots \}$

If we write, in place of the elements $n \in \mathcal{A}$, the dimensions of the rectangles with area n and least semiperimeter, we obtain

$\mathcal{A} = \{1\times1, 1\times2, 1\times3, 2\times2, 2\times3, 2\times4, 3\times3, 3\times4, 3\times5, 4\times4, 3\times6, 4\times5, 4\times6, 5\times5, 4\times7, 5\times6, 5\times7, 6\times6, 5\times8, 6\times7, 6\times8, 7\times7, 6\times9, 7\times8, 6\times10, 7\times9, 8\times8, 7\times10, 8\times9, 7\times11, 8\times10, 9\times9, 8\times11, 9\times10, 8\times12, 9\times11, 10\times10, 9\times12, 10\times11, 9\times13, 10\times12, 11\times11, 10\times13, 11\times12, 10\times14, 11\times13, 12\times12, 10\times15, 11\times14, 12\times13, 11\times15, 12\times14, 13\times13, 11\times16, 12\times15, 13\times14, 12\times16, 13\times15, 14\times14, \ldots \},$

a list that exhibits a tantalizing promise of pattern! The interested reader is invited to try to determine the precise pattern of the set $\mathcal{A}$, before reading into the next section in which the secret will be revealed. One thing we immediately notice, though, is that the dimensions of each of these rectangles are almost (or exactly) equal. For this reason, we will call the elements of $\mathcal{A}$ *almost-squares*. This supports our intuition about what the answers to Grantham's problem should be, since after all, Farmer Ted would build his rectangles with precisely equal sides if he weren't hampered by the integral policies of (the ironically-named) Builders Square.

From the list of the first 59 almost-squares, we find that 182 is the largest almost-square not exceeding 190. Therefore, Farmer Ted should build a chicken coop with area 182 square feet; and indeed, a $13' \times 14'$ coop would give him about 3.37 square feet of coop space per foot of chicken wire purchased, which is more cost-effective than the options he thought of back in Section 1. But what about next time, when Farmer Ted wants to build a supercoop on the 8,675,309 square feet of land he has to spare, or even more? Eventually, computations will need to give way to a better understanding of $\mathcal{A}$.

Our specific goals in this paper will be to answer the following questions:

1. Can we describe $\mathcal{A}$ more explicitly? That is, can we characterize when a number n is an almost-square with a description that refers only to n itself, rather than all the numbers smaller than n? Can we find a formula for the number of almost-squares not exceeding a given positive number x?

2. Can we quickly compute the largest almost-square not exceeding N, for a given number N? We will describe more specifically what we mean by "quickly" in the next section, but for now we simply say that we'll want to avoid both brute force searches and computations that involve factoring integers.

In the next section, we will find that these questions have surprisingly elegant answers.

3. REMARKABLE RESULTS

Have you uncovered the pattern of the almost-squares? One detail you might have noticed is that all numbers of the form $m \times m$ and $(m-1) \times m$, and also $(m-1) \times (m+1)$, seem to be almost-squares. (If not, maybe we should come up with a better name for the elements of $\mathcal{A}$!) This turns out to be true, as we will see in Lemma 3 below. The problem is that there are other almost-squares than these—3×6, 4×7, 5×8, 6×9, 6×10—and the "exceptions" seem to become more and more numerous .... Even so, it will be convenient to think of the particular almost-squares of the form $m \times m$ and $(m-1) \times m$ as "punctuation" of a sort for $\mathcal{A}$. To this end, we will define a *flock* to be the set of almost-squares between $(m-1)^2 + 1$ and $m(m-1)$, or between $m(m-1) + 1$ and m^2, including the endpoints in both cases.

If we group the rectangles corresponding to the almost-squares into flocks in this way, indicating the end of each flock by a semicolon, we obtain:

$\mathcal{A} = \{$1×1; 1×2; 1×3, 2×2; 2×3; 2×4, 3×3; 3×4; 3×5, 4×4; 3×6, 4×5; 4×6, 5×5; 4×7, 5×6; 5×7, 6×6; 5×8, 6×7; 6×8, 7×7; 6×9, 7×8; 6×10, 7×9, 8×8; 7×10, 8×9; 7×11, 8×10, 9×9; 8×11, 9×10; 8×12, 9×11, 10×10; 9×12, 10×11; 9×13, 10×12, 11×11; 10×13, 11×12; 10×14, 11×13, 12×12; 10×15, 11×14, 12×13; 11×15, 12×14, 13×13; 11×16, 12×15, 13×14; 12×16, 13×15, 14×14; ... $\}$

It seems that all of the rectangles in a given flock have the same semiperimeter; this also turns out to be true, as we will see in Lemma 4 below. The remaining question, then, is to determine which rectangles of the common semiperimeter a given flock contains. At first it seems that all rectangles of the "right" semiperimeter will be in the flock as long as their area exceeds that of the last rectangle in the preceding flock, but then we note a few omissions—2×5, 3×7, 4×8, 5×9, 5×10—which also become more numerous if we extend our computations of $\mathcal{A}$ ....

But as it happens, this question can be resolved, and we can actually determine exactly which numbers are almost-squares, as our main theorem indicates. Recall that $\lfloor x \rfloor$ denotes the greatest integer not exceeding x.

Main Theorem. *For any integer $m \geq 2$, the set of almost-squares between $(m-1)^2 + 1$ and m^2 (inclusive) consists of two flocks, the first of which is*

$$\{(m + a_m)(m - a_m - 1), (m + a_m - 1)(m - a_m), \ldots, (m+1)(m-2), m(m-1)\}$$

where $a_m = \lfloor (\sqrt{2m-1} - 1)/2 \rfloor$, and the second of which is

$$\{(m + b_m)(m - b_m), (m + b_m - 1)(m - b_m + 1), \ldots, (m+1)(m-1), m^2\}$$

where $b_m = \lfloor \sqrt{m/2} \rfloor$.

The Main Theorem allows us to easily enumerate the almost-squares in order, but if we simply want an explicit characterization of almost-squares without regard to their order, there turns out to be one that is extremely elegant. To describe it, we recall that the *triangular numbers* $\{0, 1, 3, 6, 10, 15, \ldots\}$ are the numbers $t_n = \binom{n}{2} = n(n-1)/2$ (Conway

4

GREG MARTIN

and Guy [1] describe many interesting properties of these and other "figurate" numbers). We let $T(x)$ denote the number of triangular numbers not exceeding x. (Notice that in our notation, $t_1 = \binom{1}{2} = 0$ is the first triangular number, so that $T(1) = 2$, for instance.) Then an alternate interpretation of the Main Theorem is the following:

Corollary 1. *The almost-squares are precisely those integers that can be written in the form $k(k+h)$, for some integers $k \geq 1$ and $0 \leq h \leq T(k)$.*

It is perhaps not so surprising that the triangular numbers are connected to the almost-squares—after all, adding t_m to itself or to t_{m+1} yields almost-squares of the form $m(m-1)$ or m^2, respectively (Figure 1 illustrates this for $m = 6$). In any case, the precision of this characterization is quite attractive and unexpected, and it is conceivable that Corollary 1 has a direct proof that doesn't use the Main Theorem. We leave this as an open problem for the reader.

$$m(m-1) = t_m + t_m$$
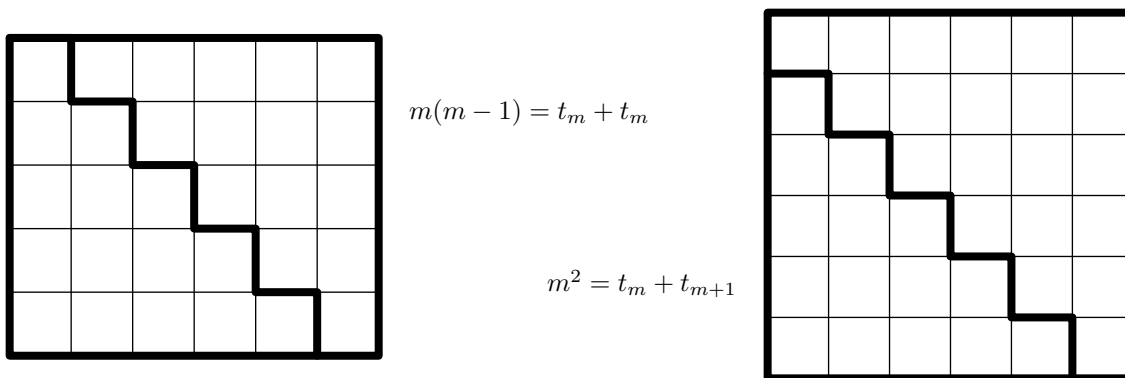
$$m^2 = t_m + t_{m+1}$$

FIGURE 1. Two triangular integers invoke an almost-square

In a different direction, we can use the Main Theorem's precise enumeration of the almost-squares in each flock to count the number of almost-squares quite accurately.

Corollary 2. *Let $A(x)$ denotes the number of almost-squares not exceeding x. Then for $x \geq 1$,*

$$A(x) = \frac{2\sqrt{2}}{3}x^{3/4} + \frac{1}{2}x^{1/2} + R(x),$$

where $R(x)$ is an oscillating term whose order of magnitude is $x^{1/4}$.

A graph of $A(x)$ (see Figure 2) exhibits a steady growth with a little bit of a wiggle. When we isolate $R(x)$ by subtracting the main term $2\sqrt{2}x^{3/4}/3 + x^{1/2}/2$ from $A(x)$, the resulting graph (Figure 2, where we have plotted a point every time x passes an almost-square) is a pyrotechnic, almost whimsical display that seems to suggest that our computer code needs to be rechecked. Yet this is the true nature of $R(x)$. When we prove Corollary 2 (in a more specific and precise form) in Section 6, we will see that there are two reasons that the "remainder term" $R(x)$ oscillates: there are oscillations on a local scale because the almost-squares flock towards the right half of each interval of the form $((m-1)^2, m(m-1)]$ or $(m(m-1), m^2]$, and oscillations on a larger scale for a less obvious reason.

These theoretical results about the structure of the almost-squares address question 1 nicely, and we turn our attention to the focus of question 2, the practicality of actually
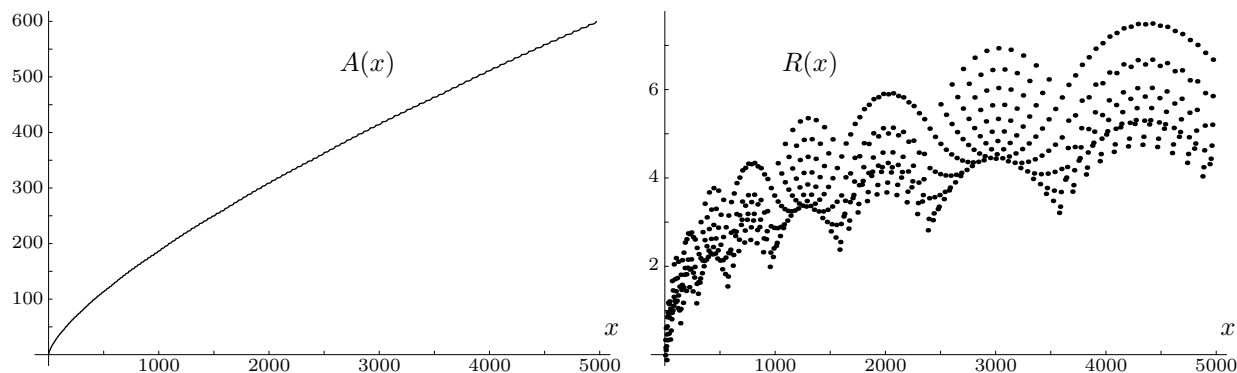
FIGURE 2. Superficial steadiness of $A(x)$, mesmerizing meanderings of $R(x)$

computing answers to questions about almost-squares. Even simple tasks like printing out a number or adding two numbers together obviously take time for a computer to perform, and they take longer for bigger numbers. To measure how the computing time needed for a particular computation increases as the size of the input grows, let $f(k)$ denote the amount of time it takes to perform the calculation on a k-digit number. Of course, the time could depend significantly on which k-digit number we choose; what we mean is the worst-case scenario, so that the processing time is at most $f(k)$ no matter which k-digit number we choose.

We say that a computation runs in *polynomial time* if this function $f(k)$ grows only as fast as a polynomial in k, i.e., if there are positive constants A and B such that $f(k) < Ak^B$. Generally speaking, the computations that we consider efficient to perform on very large inputs are those that run in polynomial time. (Because we are only concerned with this category of computations as a whole, it doesn't matter if we write our numbers in base 10 or base 2 or any other base, since this only multiplies the number of digits by a constant factor like $\log_2 10$.)

All of our familiar arithmetic operations $+$, $-$, $\times$, $\div$, $\sqrt{\cdot}$, $\lfloor \cdot \rfloor$ and so on have polynomial-time algorithms. On the other hand, performing a calculation on each of the numbers from 1 to the input n, or even from 1 to $\sqrt{n}$, etc., is definitely not polynomial-time. Thus computing almost-squares by their definition, which involves comparing $F(n)$ with all of the preceding $F(k)$, is not efficient for large n. Furthermore, the obvious method of factoring numbers—testing all possible divisors in turn—is not polynomial-time for the same reason. While there are faster ways to factor numbers, at this time there is still no known polynomial-time algorithm for factoring numbers; so even factoring a single number would make an algorithm inefficient. (Dewdney [2] writes about many facets of algorithms, including this property of running in polynomial time, while Pomerance [4] gives a more detailed discussion of factoring algorithms and their computational complexity.)

Fortunately, the Main Theorem provides a way to compute almost-squares that avoids both factorization and brute-force enumeration. In fact, we can show that all sorts of computations involving almost-squares are efficient:

Corollary 3. *There are polynomial-time algorithms to perform each of the following tasks, given a positive integer N:*

(a) *determine whether N is an almost-square, and if so determine the dimensions of the optimal rectangle;*

(b) *find the greatest almost-square not exceeding N, including the dimensions of the optimal rectangle;*

(c) *compute the number $A(N)$ of almost-squares not exceeding N;*

(d) *find the Nth almost-square, including the dimensions of the optimal rectangle.*

We reiterate that these algorithms work without ever factoring a single integer. Corollary 3, together with our lack of a polynomial-time factoring algorithm, has a rather interesting implication: for large values of N, it is much faster to compute the most cost-effective chicken coop (in terms of area-to-semiperimeter ratio) with area *at most* N than it is to compute the most cost-effective chicken coop with area *equal to* N, a somewhat paradoxical state of affairs! Nobody ever said farming was easy ... .

4. The Theorem Thought Through

Before proving the Main Theorem, we need to build up a stockpile of easy lemmas. The first of these simply confirms our expectations that the most cost-effective rectangle of a given area is the one whose side lengths are as close together as possible, and also provides some inequalities for the functions $s(n)$ and $F(n)$. Let us define $d(n)$ to be the largest divisor of n not exceeding $\sqrt{n}$ and $d'(n)$ the smallest divisor of n that is at least $\sqrt{n}$, so that $d'(n) = n/d(n)$.

Lemma 1. *The rectangle with integer side lengths and area n that has the smallest semiperimeter is the one with dimensions $d(n) \times d'(n)$. In other words,*

$$s(n) = d(n) + d'(n).$$

We also have the inequalities

$$s(n) \geq 2\sqrt{n} \quad and \quad F(n) \leq \sqrt{n}/2.$$

Proof: For a fixed positive number n, the function $f(t) = t + n/t$ has derivative $f'(t) = 1 - n/t^2$ which is negative for $1 \leq t < \sqrt{n}$, and therefore $t + n/t$ is a decreasing function of t in that range. Thus if we restrict our attention to those t such that both t and n/t are positive integers (in other words, t is an integer dividing n), we see that the expression $t + n/t$ is minimized when $t = d(n)$. We therefore have

$$s(n) = d(n) + n/d(n) \geq 2\sqrt{n},$$

where the last inequality follows from the Arithmetic Mean/Geometric Mean inequality; and the inequality for $F(n)$ then follows directly from the definition of F. $\square$

If we have a number n written as $c \times d$ where c and d are pretty close to $\sqrt{n}$, when can we say that there isn't some better factorization out there, so that $s(n)$ is really equal to $c + d$? The following lemma gives us a useful criterion.

Lemma 2. *If a number n satisfying $(m-1)^2 < n \leq m(m-1)$ has the form $n = (m-a-1)(m+a)$ for some number a, then $s(n) = 2m-1$, and $d(n) = m-a-1$ and $d'(n) = m+a$. Similarly, if a number n satisfying $m(m-1) < n \leq m^2$ has the form $n = m^2 - b^2$ for some number b, then $s(n) = 2m$, and $d(n) = m - b$ and $d'(n) = m + b$.*

Proof: First let's recall that for any positive real numbers α and β, the pair of equations $r + s = \alpha$ and $rs = \beta$ have a unique solution (r, s) with $r \leq s$, as long as the Arithmetic Mean/Geometric Mean inequality $\alpha/2 \geq \sqrt{\beta}$ holds. This is because r and s will be the roots of the quadratic polynomial $t^2 - \alpha t + \beta$, which has real roots when its discriminant $\alpha^2 - 4\beta$ is nonnegative, i.e., when $\alpha/2 \geq \sqrt{\beta}$.

Now if $n = (m - a - 1)(m + a)$, then clearly $s(n) \leq (m - 1 - a) + (m + a) = 2m - 1$ by the definition of $s(n)$. On the other hand, by Lemma 1 we know that $s(n) \geq 2\sqrt{n} > 2(m - 1)$, and so $s(n) = 2m - 1$ exactly. We now know that $d(n)d'(n) = n = (m - a - 1)(m + a)$ and

$$d(n) + d'(n) = s(n) = 2m - 1 = (m - a - 1) + (m + a),$$

and of course $d(n) \leq d'(n)$ as well; by the argument of the previous paragraph, we conclude that $d(n) = m - a - 1$ and $d'(n) = m + a$. This establishes the first assertion of the lemma, and a similar argument holds for the second assertion. $\qquad \square$

Of course, if a number n satisfies $s(n) = 2m$ for some m, then n can be written as $n = cd$ with $c \leq d$ and $c + d = 2m$; and letting $b = d - m$, we see that $n = cd = (2m - d)d = (m - b)(m + b)$. A similar statement is true if $s(n) = 2m - 1$, and so we see that the converse of Lemma 2 also holds. We also remark that in the statement of the lemma, the two expressions $m(m - 1)$ can be replaced by $(m - 1/2)^2 = m(m - 1) + 1/4$ if we wish.

Lemma 2 implies in particular that for $m \geq 2$,

$$s(m^2) = 2m, \quad s(m(m - 1)) = 2m - 1, \quad \text{and } s((m - 1)(m + 1)) = 2m,$$

and so

$$F(m^2) = \frac{m}{2}, \quad F(m^2 - m) = \frac{m(m - 1)}{2m - 1}, \quad \text{and } F(m^2 - 1) = \frac{m^2 - 1}{2m}.$$

Using these facts, we can verify our theory that these numbers are always almost-squares.

Lemma 3. *Each positive integer of the form m^2, $m(m - 1)$, or $m^2 - 1$ is an almost-square.*

It is interesting to note that these are precisely those integers n that are divisible by $\lfloor \sqrt{n} \rfloor$ (see [3]), one of the many interesting things that can be discovered by referring to Sloane and Plouffe [6].

Proof: We verify directly that such numbers satisfy the condition in the definition (1) of $\mathcal{A}$. If $k < m^2$, then by Lemma 1 we have $F(k) \leq \sqrt{k}/2 < m/2 = F(m^2)$, and so m^2 is an almost-square. Similarly, if $k < m(m - 1)$, then again

$$F(k) \leq \frac{\sqrt{k}}{2} \leq \frac{\sqrt{m^2 - m - 1}}{2} < \frac{m(m - 1)}{2m - 1} = F(m(m - 1)),$$

where the strict inequality can be verified as a "fun" algebraic exercise. Thus $m(m - 1)$ is also an almost-square. A similar argument shows that $m^2 - 1$ is also an almost-square. $\quad \square$

Now we're getting somewhere! Next we show that the semiperimeters of the rectangles corresponding to the almost-squares in a given flock are all equal, as we observed at the beginning of Section 3.

Lemma 4. *Let $m \geq 2$ be an integer. If n is an almost-square satisfying $(m - 1)^2 < n \leq m(m - 1)$, then $s(n) = 2m - 1$; similarly, if n is an almost-square satisfying $m(m - 1) < n \leq m^2$, then $s(n) = 2m$.*

Proof: If $n = m(m - 1)$, we have already shown that $s(n) = 2m - 1$. If n satisfies $(m - 1)^2 < n < m(m - 1)$, then by Lemma 1 we have $s(n) \geq 2\sqrt{n} > 2(m - 1)$. On the other hand, since n is an almost-square exceeding $(m - 1)^2$, we have

$$\frac{m - 1}{2} = F((m - 1)^2) \leq F(n) = \frac{n}{s(n)} < \frac{m(m - 1)}{s(n)},$$

and so $s(n) < 2m$. Therefore $s(n) = 2m - 1$ in this case.

Similarly, if n satisfies $m(m - 1) < n < m^2$, then $s(n) \geq 2\sqrt{n} \geq 2\sqrt{m^2 - m + 1} > 2m - 1$; on the other hand,

$$\frac{m(m - 1)}{2m - 1} = F(m(m - 1)) \leq F(n) = \frac{n}{s(n)} \leq \frac{m^2 - 1}{s(n)},$$

and so $s(n) < (m + 1)(2m - 1)/m < 2m + 1$. Therefore $s(n) = 2m$ in this case.   □

Finally, we need to exhibit some properties of the sequences a_m and b_m defined in the statement of the Main Theorem.

Lemma 5. *Define $a_m = \lfloor (\sqrt{2m - 1} - 1)/2 \rfloor$ and $b_m = \lfloor \sqrt{m/2} \rfloor$. For any integer $m \geq 2$:*
  (a) $a_m \leq b_m \leq a_m + 1$;
  (b) $b_m = \lfloor m/\sqrt{2m - 1} \rfloor$; *and*
  (c) $a_m + b_m = \lfloor \sqrt{2m} \rfloor - 1$.

We omit the proof of this lemma since it is tedious but straightforward. The idea is to show that in the sequences a_m, b_m, $\lfloor m/\sqrt{2m - 1} \rfloor$, and so on, two consecutive terms are either equal or else differ by 1, and then to determine precisely for what values of m the differences of 1 occur.

Armed with these lemmas, we are now ready to furnish a proof of the Main Theorem.

Proof of the Main Theorem: Fix an integer $m \geq 2$. By Lemma 4, every almost-square n with $(m - 1)^2 < n \leq m(m - 1)$ satisfies $s(n) = 2m - 1$; while by Lemma 2, the integers $(m - 1)^2 < n \leq m(m - 1)$ satisfying $s(n) = 2m - 1$ are precisely the elements of the form $n_a = (m - a - 1)(m + a)$ that lie in that interval. Thus it suffices to determine which of the n_a are almost-squares.

Furthermore, suppose that n_a is an almost-square for some $a \geq 1$. Then $F(n_a) \geq F(n)$ for all $n < n_a$ by the definition of $\mathcal{A}$, while $F(n_a) > F(n)$ for all $n_a < n < n_{a-1}$ since we've already concluded that no such n can be an almost-square. Moreover, $n_{a-1} > n_a$ and $s(n_{a-1}) = 2m - 1 = s(n_a)$, so $F(n_{a-1}) > F(n_a)$, and thus n_{a-1} is an almost-square as well. Therefore it suffices to find the largest value of a (corresponding to the smallest n_a) such that n_a is an almost-square.

By Lemma 3, we know that $(m - 1)^2$ is an almost-square, and so we need to find the largest a such that $F(n_a) \geq F((m - 1)^2)$, i.e.,

$$\frac{(m - a - 1)(m + a)}{2m - 1} \geq \frac{m - 1}{2},$$

which is the same as $2a(a + 1) + 1 \leq m$. By completing the square and solving for a, we find that this inequality is equivalent to

$$\frac{-\sqrt{2m - 1} - 1}{2} \leq a \leq \frac{\sqrt{2m - 1} - 1}{2}, \tag{2}$$

and so the largest integer a satisfying the inequality is exactly $a = \lfloor (\sqrt{2m-1}-1)/2 \rfloor = a_m$, as defined in the statement of the Main Theorem. This establishes the first part of the theorem.

By the same reasoning, it suffices to find the largest value of b such that $F(m^2 - b^2) \geq F(m(m-1))$, i.e.,

$$\frac{m^2 - b^2}{2m} \geq \frac{m(m-1)}{2m-1},$$

which is the same as

$$b^2 \leq m^2/(2m-1) \tag{3}$$

or $b \leq \lfloor m/\sqrt{2m-1} \rfloor$. But by Lemma 5(b), $\lfloor m/\sqrt{2m-1} \rfloor = b_m$ for $m \geq 2$, and so the second part of the theorem is established. $\qquad\square$

With the Main Theorem now proven, we remark that Lemma 5(c) implies that for any integer $m \geq 2$, the number of almost-squares in the two flocks between $(m-1)^2 + 1$ and m^2 is exactly $(1 + a_m) + (1 + b_m) = 1 + \lfloor \sqrt{2m} \rfloor$, while Lemma 5(a) implies that there are either equally many in the two flocks or else one more in the second flock than in the first.

5. Taking Notice of Triangular Numbers

Our next goal is to derive Corollary 1 from the Main Theorem. First we establish a quick lemma giving a closed-form expression for $T(x)$, the number of triangular numbers not exceeding x.

Lemma 6. *For all $x \geq 0$, we have $T(x) = \lfloor \sqrt{2x + 1/4} + 1/2 \rfloor$.*

Proof: $T(x)$ is the number of positive integers n such that $t_n \leq x$, or $n(n-1)/2 \leq x$. This inequality is equivalent to $(n - 1/2)^2 \leq 2x + 1/4$, or $-\sqrt{2x + 1/4} + 1/2 \leq n \leq \sqrt{2x + 1/4} + 1/2$. The left-hand expression never exceeds $1/2$, and so $T(x)$ is simply the number of positive integers n such that $n \leq \sqrt{2x + 1/4} + 1/2$; in other words, $T(x) = \lfloor \sqrt{2x + 1/4} + 1/2 \rfloor$ as desired. $\qquad\square$

Proof of Corollary 1: Suppose first that $n = k(k+h)$ for some integers $k \geq 1$ and $h \leq T(k)$. Let $k' = k + h$, and define

$$\begin{cases} m = k + (h+1)/2 \text{ and } a = (h-1)/2, & \text{if } h \text{ is odd,} \\ m = k + h/2 \text{ and } b = h/2, & \text{if } h \text{ is even,} \end{cases}$$

so that

$$\begin{cases} k = m - a - 1 \text{ and } k' = m + a, & \text{if } h \text{ is odd,} \\ k = m - b \text{ and } k' = m + b, & \text{if } h \text{ is even.} \end{cases}$$

We claim that

$$\begin{cases} (m-1)^2 < (m-a-1)(m+a) \leq (m-1/2)^2, & \text{if } h \text{ is odd,} \\ (m-1/2)^2 < m^2 - b^2 \leq m^2, & \text{if } h \text{ is even.} \end{cases} \tag{4}$$

To see this, note that in terms of k and h, these inequalities become

$$\left(k + \frac{h-1}{2}\right)^2 < k(k+h) \leq \left(k + \frac{h}{2}\right)^2.$$

A little bit of algebra reveals that the right-hand inequality is trivially satisfied while the left-hand inequality is true provided that $h < 2\sqrt{k} + 1$. However, from Lemma 6 we see that

$$T(k) = \lfloor \sqrt{2k + 1/4} + 1/2 \rfloor \leq \sqrt{2k + 1/4} + 1/2 < 2\sqrt{k} + 1$$

for $k \geq 1$. Since we are assuming that $h \leq T(k)$, this shows that the inequalities (4) do indeed hold.

Because of these inequalities, we may apply Lemma 2 (see the remarks following the proof of the lemma) and conclude that

$$\begin{cases} s(n) = 2m - 1, \ d(n) = m - a - 1, \ \text{and} \ d'(n) = m + a, & \text{if } h \text{ is odd,} \\ s(n) = 2m, \ d(n) = m - b, \ \text{and} \ d'(n) = m + b, & \text{if } h \text{ is even.} \end{cases}$$

Consequently, the Main Theorem asserts that n is an almost-square if and only if

$$\begin{cases} a \leq a_m, & \text{if } h \text{ is odd,} \\ b \leq b_m, & \text{if } h \text{ is even,} \end{cases} \tag{5}$$

which by the definitions of a, b, and m is the same as

$$\begin{cases} (h-1)/2 \leq \lfloor (\sqrt{2m-1} - 1)/2 \rfloor = \lfloor (\sqrt{2k+h} - 1)/2 \rfloor, & \text{if } h \text{ is odd,} \\ h/2 \leq \lfloor \sqrt{m/2} \rfloor = \lfloor \sqrt{k/2 + h/4} \rfloor, & \text{if } h \text{ is even.} \end{cases}$$

Since in either case, the left-hand side is an integer, the greatest-integer brackets can be removed from the right-hand side, whence both cases reduce to $h \leq \sqrt{2k + h}$. From here, more algebra reveals that this inequality is equivalent to $h \leq \sqrt{2k + 1/4} + 1/2$; and since h is an integer, we can add greatest-integer brackets to the right-hand side, thus showing that the inequality (5) is equivalent to $h \leq T(k)$ (again using Lemma 6). In particular, n is indeed an almost-square.

This establishes one half of the characterization asserted by Corollary 1. Conversely, suppose we are given an almost-square n, which we can suppose to be greater than 1 since 1 can obviously be written as $1(1 + 0)$. If we let $h = d'(n) - d(n)$, then the Main Theorem tells us that

$$\begin{cases} n = (m - a - 1)(m + a), \ d(n) = m - a - 1, \ \text{and} \ d'(n) = m + a, & \text{if } h \text{ is odd,} \\ n = m^2 - b^2, \ d(n) = m - b, \ \text{and} \ d'(n) = m + b, & \text{if } h \text{ is even} \end{cases}$$

for some integers $m \geq 2$ and either a with $0 \leq a \leq a_m$ or b with $0 \leq b \leq b_m$. If we set $k = d(n)$, then certainly $n = k(k + h)$. Moreover, the algebraic steps showing that the inequality (5) is equivalent to $h \leq T(k)$ are all reversible; and (5) does in fact hold, since we are assuming that n is an almost-square. Therefore n does indeed have a representation of the form $k(k + h)$ with $0 \leq h \leq T(k)$. This establishes the corollary. □

We take a slight detour at this point to single out some special almost-squares. Let us make the convention that the kth flock refers to the flock of almost-squares with semiperimeter k, so that the first flock is actually empty, the second and third poor flocks contain only $1 = 1 \times 1$ and $2 = 1 \times 2$, respectively, the fourth flock contains $3 = 1 \times 3$ and $4 = 2 \times 2$, and so on. The Main Theorem tells us that a_m and b_m control the number of almost-squares in the odd-numbered and even-numbered flocks, respectively; thus every so often, a flock will have one more almost-square than the preceding flock of the same "parity". We'll let a *pioneer* be an almost-square that begins one of these suddenly-longer flocks.

For instance, from the division of $\mathcal{A}$ into flocks on page 3, we see that the 4th flock $\{1 \times 3, 2 \times 2\}$ is longer than the preceding even-numbered flock $\{1 \times 1\}$, so $1 \times 3 = 3$ is the first pioneer; the 9th flock $\{3 \times 6, 4 \times 5\}$ is longer than the preceding odd-numbered flock $\{3 \times 4\}$, so $3 \times 6 = 18$ is the second pioneer; and so on, the next two pioneers being 6×10 in the 16th flock and 10×15 in the 25th flock. Now if this isn't a pattern waiting for a proof, nothing is! The following lemma shows another elegant connection between the almost-squares and the squares and triangular numbers.

Corollary 4. *For any positive integer j, the jth pioneer equals $t_{j+1} \times t_{j+2}$ (where t_i is the ith triangular number), which begins the $(j+1)^2$-th flock. Furthermore, the "record-tying" almost-squares (those whose F-values are equal to the F-values of their immediate predecessors in $\mathcal{A}$) are precisely the even-numbered pioneers.*

Proof: First, Lemma 5(a) tells us that the odd- and even-numbered flocks undergo their length increases in alternation, so that the pioneers alternately appear in the flocks of each parity. The first pioneer $3 = 1 \times 3$ appears in the 4th flock, and corresponds to $m = 2$ and the first appearance of $b_m = 1$ in the notation of the Main Theorem. Thus the $(2k-1)$-st pioneer will equal $m^2 - k^2$, where m corresponds to the first appearance of $b_m = k$. It is easy to see that the first appearance of $b_m = k$ occurs when $m = 2k^2$, in which case the $(2k-1)$-st pioneer is

$$m^2 - k^2 = (2k^2)^2 - k^2 = (2k^2 - k)(2k^2 + k) = \frac{2k(2k-1)}{2} \frac{(2k+1)2k}{2} = t_{2k}t_{2k+1}.$$

Moreover, the flock in which this pioneer appears is the $2m$-th or $(2k)^2$-th flock.

Similarly, the $2k$-th pioneer will equal $(m - k - 1)(m + k)$, where m corresponds to the first appearance of $a_m = k$. Again one can show that the first appearance of $a_m = k$ occurs when $m = 2k^2 + 2k + 1$, in which case the $2k$-th pioneer is

$$(m - k - 1)(m + k) = (2k^2 + k)(2k^2 + 3k + 1) = \frac{(2k+1)2k}{2} \frac{(2k+2)(2k+1)}{2} = t_{2k+1}t_{2k+2}.$$

Moreover, the flock in which this pioneer appears is the $(2m - 1)$-st or $(2k + 1)^2$-th flock. This establishes the first assertion of the corollary.

Since the F-values of the almost-squares form a nondecreasing sequence by the definition of almost-square, to look for almost-squares with equal F-values we only need to examine consecutive almost-squares. Furthermore, two consecutive almost-squares in the same flock never have equal F-values, since they are distinct numbers but by Lemma 4 their semiperimeters are the same. Therefore we only need to determine when the last almost-square in a flock can have the same F-value as the first almost-square in the following flock.

The relationship between the F-values of these pairs of almost-squares was determined in the proof of the Main Theorem. Specifically, the equality $F((m-1)^2) = F((m-a-1)(m+a))$ holds if and only if the right-hand inequality in (2) is actually an equality; this happens precisely when $m = 2a^2 + 2a + 1$, which corresponds to the even-numbered pioneers as was determined above. On the other hand, the equality $F(m(m-1)) = F(m^2 - b^2)$ holds if and only if the inequality (3) is actually an equality; but m^2 and $2m - 1$ are always relatively prime (any prime factor of m^2 must divide m and thus divides into $2m-1$ with a "remainder" of -1), implying that $m^2/(2m-1)$ is never an integer for $m \geq 2$, and so the inequality (3) can never be an equality. This establishes the second assertion of the corollary. $\square$

We know that all squares are almost-squares, and so t_j^2 is certainly an almost-square for any triangular number t_j; also, Corollary 4 tells us that the product $t_j t_{j+1}$ of two consecutive triangular numbers is always an almost-square. This led the author to wonder which numbers of the form $t_m t_n$ are almost-squares. If m and n differ by more than 1, it would seem that the rectangle of dimensions $t_m \times t_n$ is not the most cost-effective rectangle of area $t_m t_n$, and so the author expected that these products of two triangular numbers would behave randomly with respect to being almost-squares—that is, a few of them might be but most of them wouldn't. After some computations, however, Figure 3 emerged, where a point has been plotted in the (m, n) position if and only if $t_m t_n$ is an almost-square; and the table exhibited a totally unexpected regularity.
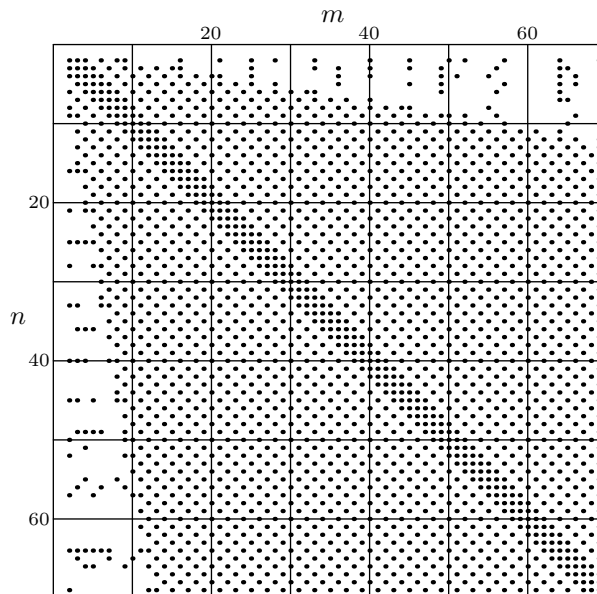

FIGURE 3. Amazing almost-square patterns in products of two triangles

Of course the symmetry of the table across the main diagonal is to be expected since $t_m t_n = t_n t_m$. The main diagonal and the first off-diagonals are filled with plotted points, corresponding to the almost-squares t_m^2 and $t_m t_{m+1}$; and in hindsight, the second off-diagonals correspond to

$$t_m t_{m+2} = \frac{m(m-1)}{2} \frac{(m+2)(m+1)}{2} = \frac{m^2 + m - 2}{2} \frac{m^2 + m}{2},$$

which is the product of two consecutive integers (since $m^2 + m$ is always even) and is thus an almost-square as well. But apart from these central diagonals and some garbage along the edges of the table where m and n are quite different in size, the checkerboard-like pattern in the kite-shaped region of the table seems to be telling us that the only thing that matters in determining whether $t_m t_n$ is an almost-square is whether m and n have the same parity!

Once this phenomenon had been discovered, it turned out that the following corollary could be derived from the prior results in this paper. We leave the proof of this corollary as a challenge to the reader.

Corollary 5. *Let m and n be positive integers with $n - 1 > m > 3n - \sqrt{8n(n-1)} - 1$. Then $t_m t_n$ is an almost-square if and only if $n - m$ is even.*

We remark that the function $3n - \sqrt{8n(n-1)} - 1$ is asymptotic to $(3 - 2\sqrt{2})n + (\sqrt{2} - 1)$, which explains the straight lines of slope $-(3 - 2\sqrt{2}) \approx -0.17$ and $-1/(3 - 2\sqrt{2}) \approx -5.83$ that seem to the eye to separate the orderly central region in Figure 3 from the garbage along the edges.

6. COUNTING AND COMPUTING

In this section we establish Corollaries 2 and 3. We begin by defining a function $B(x)$ that will serve as the backbone of our investigation of the almost-square counting function $A(x)$. Let $\{x\} = x - \lfloor x \rfloor$ denote the fractional part of x, and define the quantities $\gamma = \gamma(x) = \{\sqrt{2}x^{1/4}\}$ and $\delta = \delta(x) = \{x^{1/4}/\sqrt{2}\}$. Let $B(x) = B_0(x) + B_1(x)$, where

$$B_0(x) = \frac{2\sqrt{2}}{3}x^{3/4} + \frac{1}{2}x^{1/2} + \Big(\frac{2\sqrt{2}}{3} + \frac{\gamma(1-\gamma)}{\sqrt{2}}\Big)x^{1/4} \tag{6}$$

and

$$B_1(x) = \frac{\gamma^3}{6} - \frac{\gamma^2}{4} - \frac{5\gamma}{12} - \frac{\delta}{2} - 1.$$

We remark that $\gamma = \{2\delta\}$ and that $B_1(x^4)$ is a periodic function of x with period $\sqrt{2}$, and so it is easy to check that the inequalities $-2 \leq B_1(x) \leq -1$ always hold. The following lemma shows how the strange function $B(x)$ arises in connection with the almost-squares.

Lemma 7. *For any integer $M \geq 1$, we have $A(M^2) = B(M^2)$.*

Proof: As remarked at the end of Section 4, the number of almost-squares between $(m-1)^2 + 1$ and m^2 is $\lfloor \sqrt{2m} \rfloor + 1$ for $m \geq 2$. Therefore

$$A(M^2) = 1 + \sum_{m=2}^{M}(\lfloor \sqrt{2m} \rfloor + 1) = M - 1 + \sum_{m=1}^{M}\lfloor \sqrt{2m} \rfloor.$$

It's almost always a good idea to interchange orders of summation whenever possible—and if there aren't enough summation signs, find a way to create some more! In this case, we convert the greatest-integer function into a sum of 1s over the appropriate range of integers:

$$A(M^2) = M - 1 + \sum_{m=1}^{M} \sum_{1 \leq k \leq \sqrt{2m}} 1$$

$$= M - 1 + \sum_{1 \leq k \leq \sqrt{2M}} \sum_{k^2/2 \leq m \leq M} 1$$

$$= M - 1 + \sum_{\substack{1 \leq k \leq \sqrt{2M} \\ k \text{ odd}}} \Big(M - \frac{k^2 - 1}{2}\Big) + \sum_{\substack{1 \leq k \leq \sqrt{2M} \\ k \text{ even}}} \Big(M - \big(\frac{k^2}{2} - 1\big)\Big).$$

If we temporarily write μ for $\lfloor \sqrt{2M} \rfloor$, then

$$A(M^2) = M - 1 + \mu\big(M + \frac{1}{2}\big) - \frac{1}{2}\sum_{k=1}^{\mu} k^2 + \sum_{\substack{k=1 \\ k \text{ even}}}^{\mu} \frac{1}{2}$$

$$= M(\mu + 1) + \frac{\mu}{2} - 1 - \frac{1}{2}\frac{\mu(\mu + 1)(2\mu + 1)}{6} - \frac{1}{2}\lfloor \frac{\mu}{2} \rfloor, \tag{7}$$

using the well-known formula for the sum of the first μ squares. Since $\lfloor \lfloor x \rfloor / n \rfloor = \lfloor x/n \rfloor$ for any real number $x \geq 0$ and any positive integer n, the last term can be written as

$$\frac{1}{2}\Big\lfloor \frac{\mu}{2} \Big\rfloor = \frac{1}{2}\Big\lfloor \frac{\lfloor \sqrt{2M} \rfloor}{2} \Big\rfloor = \frac{1}{2}\Big\lfloor \sqrt{\frac{M}{2}} \Big\rfloor = \frac{1}{2}\sqrt{\frac{M}{2}} - \frac{1}{2}\Big\{ \sqrt{\frac{M}{2}} \Big\} = \frac{1}{2}\sqrt{\frac{M}{2}} - \frac{\delta(M^2)}{2},$$

while we can replace the other occurrences of μ in equation (7) by $\sqrt{2M} - \{\sqrt{2M}\} = \sqrt{2M} - \gamma(M^2)$. Writing γ for $\gamma(M^2)$ and δ for $\delta(M^2)$, we see that

$$A(M^2) = M(\sqrt{2M} - \gamma + 1) + \sqrt{2M} - \gamma - 1$$

$$- \frac{1}{2}\frac{(\sqrt{2M} - \gamma)(\sqrt{2M} - \gamma + 1)(2(\sqrt{2M} - \gamma) + 1)}{6} - \frac{1}{2}\sqrt{\frac{M}{2}} + \frac{\delta}{2}$$

$$= \frac{2\sqrt{2}}{3}M^{3/2} + \frac{M}{2} + \Big(\frac{2\sqrt{2}}{3} - \frac{\gamma(1-\gamma)}{\sqrt{2}}\Big)\sqrt{M} + \frac{\gamma^3}{6} - \frac{\gamma^2}{4} - \frac{5\gamma}{12} - 1 - \frac{\delta}{2} = B(M^2)$$

after much algebraic simplification. This establishes the lemma. $\square$

Now $B(x)$ is a rather complicated function of x, but the next lemma gives us a couple of ways to predict the behavior of $B(x)$. First, it tells us how to predict $B(x + y)$ from $B(x)$ if y is small compared to x (roughly speaking, their difference will be $y/\sqrt{2}x^{1/4}$); second, it tells us how to predict approximately when $B(x)$ assumes a given integer value.

Lemma 8. *There is a positive constant C such that:*

(a) for all real numbers $x \geq 1$ and $0 \leq y \leq \min\{x/2, 3\sqrt{x}\}$, we have

$$\Big|(B(x + y) - B(x)) - \frac{y}{\sqrt{2}x^{1/4}}\Big| < C; \tag{8}$$

(b) if we define $z_j = \frac{1}{2}(3j)^{2/3} - \frac{1}{4}(3j)^{1/3}$ for any positive integer j, then for all $j > C$ we have $z_j > 2$ and $B((z_j - 1)^2) < j < B(z_j^2)$.

If the proof of Lemma 5 was omitted due to its tediousness, the proof of this lemma should be omitted and then buried .... The idea of the proof is to rewrite $B_0(x)x^{-3/4}$ using the new variable $t = x^{-1/4}$, and then expand in a Taylor series in t (a slight but easily overcome difficulty being that the term $\gamma(x)(1 - \gamma(x))$ is not differentiable when $\sqrt{2}x^{1/4}$ is an integer). For the proof of part (b), we also need to rewrite $z_j j^{-2/3}$ using the new variable $u = j^{-1/3}$ and expand in a Taylor series in u. We remark that the constant C in Lemma 8 can be taken to be quite small—in fact, $C = 5$ will suffice.

With these last lemmas in hand, we can dispatch Corollaries 2 and 3 in quick succession.

Proof of Corollary 2: Let $x > 1$ be a real number and define $R(x) = A(x) - 2\sqrt{2}x^{3/4}/3 - \sqrt{x}/2$, as in the statement of the corollary. We will describe how to prove the following more precise statement:

$$R(x) = \Big(\frac{2\sqrt{2}}{3} + g(\sqrt{2}x^{1/4}) - h(2\sqrt{x})\Big)x^{1/4} + R_1(x), \tag{9}$$

where

$$g(t) = \frac{\{t\}(1 - \{t\})}{\sqrt{2}} \quad \text{and} \quad h(t) = \begin{cases} \dfrac{\{t\}}{\sqrt{2}}, & \text{if } 0 \leq \{t\} \leq \dfrac{1}{2}, \\[2mm] \sqrt{1 - \{t\}} - \dfrac{1 - \{t\}}{\sqrt{2}}, & \text{if } \dfrac{1}{2} \leq \{t\} \leq 1 \end{cases}$$

and $|R_1(x)| < D$ for some constant D. The functions g and h are continuous and periodic with period 1, and are the causes of the oscillations in the error term $R(x)$. The expression $g(\sqrt{2}x^{1/4})$ goes through a complete cycle when x increases by about $2\sqrt{2}x^{3/4}$ (one can show this using Taylor expansions yet again!), which causes the large-scale bounces in the normalized error term $R(x)x^{-1/4}$ shown in Figure 4 below. Similarly, the expression $h(2\sqrt{x})$ goes through a complete cycle when x increases by about $\sqrt{x}$, which causes the smaller-scale stutters shown in the (horizontally magnified) right-hand graph in Figure 4.
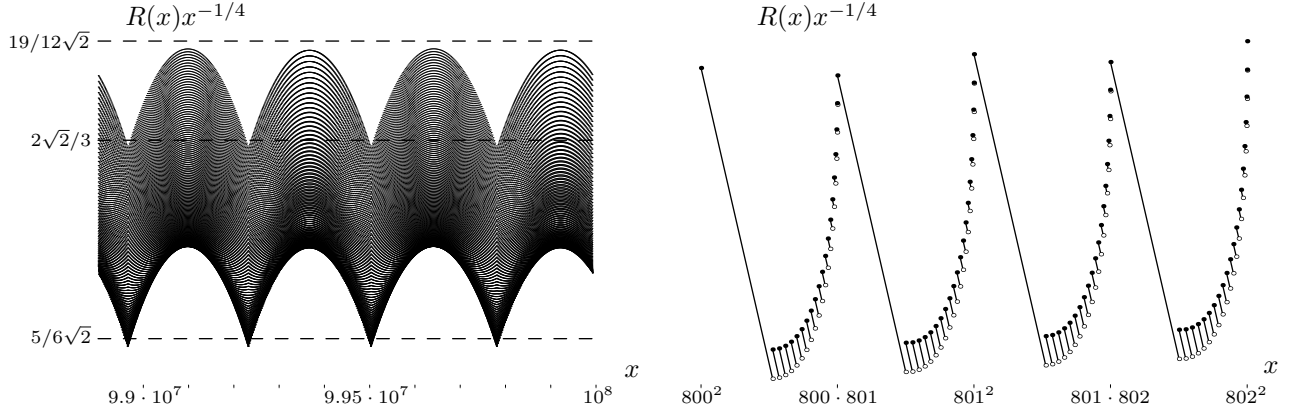

FIGURE 4. Big bounces and small stutters for $R(x)x^{-1/4}$

To establish the formula (9), we shrewdly add $B(x) - B(x)$ to the expression defining $R(x)$, which yields

$$R(x) = (A(x) - B(x)) + \Big(\frac{2\sqrt{2}}{3} + \frac{\gamma(x)(1-\gamma(x))}{\sqrt{2}}\Big)x^{1/4} + B_1(x)$$

from the definition (6) of $B_0(x)$. Now $B_1(x)$ is a bounded function; and since $\gamma(x) = \{\sqrt{2}x^{1/4}\}$, the expression $\gamma(x)(1-\gamma(x))/\sqrt{2}$ is precisely $g(\sqrt{2}x^{1/4})$. So what we need to show is that $B(x) - A(x) = h(2\sqrt{x})x^{1/4} + R_2(x)$, where $R_2(x)$ is another bounded function.

While we won't give all the details, the outline of showing this last fact is as follows: suppose first that $x \geq m^2$ but that x is less than the first almost-square $(m+1+a_{m+1})(m-a_{m+1})$ in the $(2m+1)$-st flock, so that $A(x) = A(m^2)$. Since $A(m^2) = B(m^2)$ by Lemma 7, we only need to show that $B(x) - B(m^2)$ is approximately $h(2\sqrt{x})x^{1/4}$; this we can accomplish with the help of Lemma 8(a).

Similarly, if $x < m^2$ but x is at least as large as the first almost-square $(m+b_m)(m-b_m)$ in the $2m$-th flock, the same method works as long as we take into account the difference between $A(m^2)$ and $A(x)$, which is $\lfloor\sqrt{m^2 - x}\rfloor$ by the Main Theorem. And if x is close to an almost-square of the form $m(m-1)$ rather than m^2, the same method applies; even though $A(m(m-1))$ and $B(m(m-1))$ are not exactly equal, they differ by a bounded amount.

Notice that the functions $g(t)$ and $h(t)$ take values in $[0, 1/4\sqrt{2}]$ and $[0, 1/2\sqrt{2}]$, respectively. From this and the formula (9) we can conclude that

$$\liminf_{x\to\infty} \frac{R(x)}{x^{1/4}} = \frac{5}{6\sqrt{2}} \text{ and } \limsup_{x\to\infty} \frac{R(x)}{x^{1/4}} = \frac{19}{12\sqrt{2}}.$$

The interested reader can check, for example, that the sequences $y_j = 4j^4 + j^2$ and $z_j = (2j^2 + j)^2$ satisfy $\lim_{j\to\infty} R(y_j)/y_j^{1/4} = 5/6\sqrt{2}$ and $\lim_{j\to\infty} R(z_j)/z_j^{1/4} = 19/12\sqrt{2}$. $\square$

Proof of Corollary 3: The algorithms we describe will involve only the following types of operations: performing ordinary arithmetical calculations $+$, $-$, $\times$, $\div$; computing the greatest-integer $\lfloor \cdot \rfloor$, least-integer $\lceil \cdot \rceil$, and fractional-part $\{\cdot\}$ functions; taking square roots; and comparing two numbers to see which is bigger. All of these operations can be easily performed in polynomial time. To get the ball rolling, we remark that the functions a_m, b_m, and $B(x)$ can all be computed in polynomial time, since their definitions only involve the types of operations just stated.

We first describe a polynomial-time algorithm for computing the number of almost-squares up to a given positive integer N. Let $M = \lceil \sqrt{N} \rceil$, so that $(M-1)^2 < N \leq M^2$. Lemma 7 tells us that the number of almost-squares up to M^2 is $B(M^2)$, and so we simply need to subtract from this the number of almost-squares larger than N but not exceeding M^2. This is easy to do by the characterization of almost-squares given in the Main Theorem. If $N > M(M-1)$, then we want to find the positive integer b such that $M^2 - b^2 \leq N < M^2 - (b-1)^2$, except that we want $b = 0$ if $N = M^2$. In other words, we set $b = \lceil M^2 - N \rceil$. Then, if $b \leq b_M$, the number of almost-squares up to N is $B(M^2) - b$, while if $b > b_M$, the number of almost-squares up to N is $B(M^2) - b_M - 1$.

In the other case, where $N \leq M(M-1)$, we want to find the positive integer a such that $(M-a-1)(M+a) \leq N < (M-a)(M+a-1)$, except that we want $a = 0$ if $N = M(M-1)$, In other words, we set $a = \lceil \sqrt{(M-1/2)^2 - N} + 1/2 \rceil$. Then, if $a \leq a_M$, the number of almost-squares up to N is $B(M^2) - b_m - 1 - a$, while if $a > a_M$, the number of almost-squares up to N is $B((M-1)^2)$. This shows that $A(N)$ can be computed in polynomial time, which establishes part (c) of the corollary.

Suppose now that we want to compute the Nth almost-square. We compute in any way we like the first C almost-squares, where C is as in Lemma 8; this only takes a constant amount of time (it doesn't change as N grows) which certainly qualifies as polynomial time. If $N \leq C$ then we are done, so assume that $N > C$. Let $M = \lceil z_N \rceil$, where z_N is defined as in Lemma 8(b), so that M is at least 3 by the definition of C. By Lemma 7,

$$A(M^2) = B(M^2) \geq B(z_N^2) > N \quad \text{and} \quad A((M-2)^2) = B((M-2)^2) < B((z_N-1)^2) < N,$$

where the last inequality in each case follows from Lemma 8(b). Therefore the Nth almost-square lies between $(M-2)^2$ and M^2, and so is either in the $2M$-th flock or one of the preceding three flocks. If $0 \leq B(M^2) - N \leq b_M$, then the Nth almost-square is in the $2M$-th flock, and by setting $b = B(M^2) - N$ we conclude that the Nth almost-square is $M^2 - b^2$ and the dimensions of the optimal rectangle are $(M-b) \times (M+b)$. If $1 + b_M \leq B(M^2) - N \leq b_M + 1 + a_M$, then the Nth almost-square is in the $(2M-1)$-st flock, and so on. This establishes part (d) of the corollary.

Finally, we can determine the greatest almost-square not exceeding N by computing $J = A(N)$ and then computing the Jth almost-square, both of which can be done in polynomial time by parts (c) and (d); and we can determine whether N is an almost-square simply by checking whether this result equals N. This establishes the corollary in its entirety.   □

7. Final Filibuster

We have toured some very pretty and precise properties of the almost-squares, and there are surely other natural questions that can be asked about them, some of which have already been noted. When Grantham posed this problem, he recalled the common variation on

the original calculus problem where the fence for one of the sides of the rectangle is more expensive for some reason (that side borders a road or something), and suggested the more general problem of finding the most cost-effective rectangle with integer side lengths and area at most N, where one of the sides must be fenced at a higher cost. This corresponds to replacing $s(n)$ with the more general function $s_\alpha(n) = \min_{d|n}(d + \alpha n/d)$, where α is some constant bigger than 1. While the elegance of the characterization of such "α-almost-squares" might not match that of Corollary 1, it seems reasonable to hope that an enumeration every bit as precise as the Main Theorem would be possible to establish.

How about generalizing this problem to higher dimensions? For example, given a positive integer N, find the dimensions of the rectangular box with integer side lengths and volume at most N whose volume-to-surface area ratio is largest among all such boxes. (It seems a little more natural to consider surface area rather than the sum of the box's length, width, and height, but who knows which problem has a more elegant solution?) Perhaps these "almost-cubes" have an attractive characterization analogous to Corollary 1; almost certainly a result like the Main Theorem, listing the almost-cubes in order, would be very complicated. And of course there is no reason to stop at dimension 3.

In another direction, intuitively it seems that numbers with many divisors are more likely to be almost-squares, and the author thought to test this theory with integers of the form $n!$. However, computations reveal that the only values of $n \le 500$ for which $n!$ is an almost-square are $n = 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 13, 15$. Is it the case that these are the only factorial almost-squares? This seems like quite a hard question to resolve. Perhaps a better intuition about the almost-squares is that only those numbers that lie at the right distance from a number of the form m^2 or $m(m-1)$ are almost-squares—more an issue of good fortune than of having enough divisors.

The reader is welcome to contact the author for the Mathematica code used to calculate the functions related to almost-squares described in this paper. With this code, for instance, one can verify that with 8,675,309 square feet of land at his disposal, it is most cost-effective for Farmer Ted to build a $2,919' \times 2,972'$ supercoop ... speaking of which, we almost forgot to finish the Farmer Ted story:

> After learning the ways of the almost-squares, Farmer Ted went back to Builders Square, where the salesman viewed the arrival of his $\mathbb{R}$-rival with trepidation. But Farmer Ted reassured him, "Don't worry—I no longer think it's inane to measure fences in $\mathbb{N}$." From that day onward, the two developed a flourishing business relationship, as Farmer Ted became an integral customer of the store.

And that, according to this paper, is that.

References

[1] J.H. Conway and R.K. Guy, *The Book of Numbers*, Copernicus, New York, 1996
[2] A.K. Dewdney, *The (new) Turing Omnibus*, Computer Science Press, New York, 1993
[3] S.W. Golomb, Problem E2491, *Amer. Math. Monthly* 82 (1975), 854–855
[4] C. Pomerance, A tale of two sieves, *Notices Amer. Math. Soc.* 43 (1996), no. 12, 1473–1485
[5] N.J.A. Sloane, An on-line version of the encyclopedia of integer sequences, *Electron. J. Combin.* 1 (1994), feature 1 (electronic)

[6] N.J.A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press Inc., San Diego, 1995

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, CANADA M5S 3G3
E-mail address: gerg@math.toronto.edu