

# THE ISOPERIMETRIC INEQUALITY: THE ALGEBRAIC VIEWPOINT

H. WU

*This is an expanded version of the handout that accompanied my lecture to a group of high school teachers on April 25, 1998.*

## 1. THE LITERATURE

The amount of education that takes place in a two hour lecture is very small. The hope is rather that this lecture will make you feel sufficiently intrigued afterwards to pursue the subject on your own. With this in mind, I want to begin with a discussion of the literature. Two books can be recommended regardless of the subject of this lecture. The first is the following classic:

[CR] Richard Courant and Herbert Robbins, *What is mathematics?*, 2nd ed., revised by Ian Stewart, Oxford University Press, 1996

This book has stayed in print for almost sixty years since its publication date of 1941, and the revised edition is even in paperback! I urge you to buy a copy right away if you don't already own one. It is an approachable and erudite discourse on the nature of mathematics, one that mathematics researchers and high school students alike would find enlightening. Steiner's attempted proof of the isoperimetric inequality is on pp. 373–376.

A second book I recommend is of a far more recent vintage:

[R] Joseph Rotman, *Journey into Mathematics*, Prentice Hall, 1998.

The arithmetic-geometric-mean inequality is given two proofs on pp. 12–16, and both are different from the one discussed in the lecture. Heron's formula and the fact that among triangles with a fixed perimeter the equilateral one encloses the largest area are proved on pp. 79–81. In general, this book is one of the few that give detailed exposition of the kind mathematics that is both beneficial and relevant to high school teaching.

The next three are elementary monographs on inequalities. The first two are still in print, but all three should be available in most libraries.

[**BB**] E. Beckenbach and R. Bellman, *An Introduction to Inequalities*, New Mathematical Library, Volume 3, Mathematical Association of America, 1961.

[**K1**] N. D. Kazarinoff, *Geometric Inequalities*, New Mathematical Library, Volume 4, Mathematical Association of America, 1961.

[**K2**] N. D. Kazarinoff, *Analytic Inequalities*, Holt, Rinehart and Winston, 1964.

All three contain a careful discussion of the arithmetic-geometric-mean inequality, and the last two also treat the isoperimetric inequalities for triangles, quadrilaterals and polygons. Each only assumes high school mathematics and the first one, [**BB**], is especially suitable for enrichment programs in high school.

A rigorous proof of the isoperimetric inequality can be found on pp. 31-35 of

[**dC**] Manfredo do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice Hall, 1976.

A good knowledge of calculus of two variables, including Green's Theorem is required. This is probably the most elementary known proof of this inequality but is by no means the most natural. An outline of a more natural argument is on pp. 283-284 of

[**OP**] John Oprea, *Differential Geometry and Its Applications*, Prentice Hall, 1997.

What is missing in the argument of [**OP**] is the fact that the circle actually furnishes the absolute minimum.

It may be superfluous to say so, but let it be said nonetheless, that far from being an end in itself, the isoperimetric inequality for curves is only the first step of an unending journey. The following survey article gives an entry into the current literature.

[**OS**] R. Osserman, The isoperimetric inequality, *Bulletin of the American Mathematical Society*, 84 (1978), 1182-1238.

## 2. THE INEQUALITY

Let  $\mathcal{R}$  be a region in the plane, and let  $\partial\mathcal{R}$  be its boundary. We will henceforth assume that the area of  $\mathcal{R}$  and the length of  $\partial\mathcal{R}$  both make sense. Denote them by  $A$  and  $C$ , respectively. By tradition, we shall refer to  $C$  as the *perimeter* of  $\mathcal{R}$ . At first glance, it does not look promising that there would be any relationship between  $A$  and  $C$ . But in fact there is, and this surprising fact constitutes

**The isoperimetric inequality.** *If the area and perimeter of a region in the plane are  $A$  and  $C$  respectively, then*

$$(1) \quad A \leq \frac{1}{4\pi} C^2,$$

*and equality holds exactly when the region is a round disk.*

Note that this theorem asserts three things:

- (a)  $A$  and  $C$  always satisfy inequality (1),
- (b) if the region is a round disk, then inequality (1) is actually an equality, and
- (c) if inequality (1) is an equality for a region, then the region is a round disk.

Of these three, (b) can be immediately verified. Indeed, if  $\mathcal{R}$  is a round disk of radius  $r$ , then  $A = \pi r^2$  and  $\frac{1}{4\pi} C^2 = \frac{1}{4\pi} (2\pi r)^2 = \pi r^2$  so that equality prevails in (1). However, (a) and (c) are far from obvious. One can paraphrase this theorem by saying that among all regions with a fixed perimeter  $C$ , the round disk has the largest area, namely,  $\frac{1}{4\pi} C^2$ . (Note that when we say “the” round disk, we consider all round disks with equal radii to be the same.) Or, in self-explanatory language, the round disk is the solution to the “isoperimetric” problem.

From the point of view of isoperimetry, (1) is an *optimal* inequality in a sense that we now explain. Fix the perimeter  $C$ . Let  $\beta$  be any positive constant and let us consider in general an inequality of the type  $A \leq \beta C^2$ . From (1), we know that such an inequality is valid provided  $\beta \geq \frac{1}{4\pi}$ . Note the obvious fact: since (1) is true, then so is  $A \leq \frac{1}{\pi} C^2$  or  $A \leq C^2$  (corresponding to  $\beta = \frac{1}{\pi}$  and  $\beta = 1$ , respectively). But (1) assures us that it *suffices* to take  $\beta = \frac{1}{4\pi}$ , and observation (b) above says that it is *necessary* to require  $\beta$  to be at least  $\frac{1}{4\pi}$ , because if  $\beta_0 < \frac{1}{4\pi}$ , the round disk with perimeter  $C$  would contradict  $A \leq \beta_0 C^2$ . With the perimeter  $C$  fixed then, (1) gives the smallest possible constant  $\beta$  to make the inequality  $A \leq \beta C^2$  valid in general. This is the meaning of (1) being an optimal inequality.

The intuitive reason why such a theorem should be true is the subject of another lecture.<sup>1</sup> What we hope to accomplish here is, first, to closely examine why the *inequality* (1) is true when the region  $\mathcal{R}$  is triangular (in which case,  $\partial\mathcal{R}$  is just an ordinary triangle). Then we shall go on to investigate whether, when  $\mathcal{R}$  is restricted to be a triangular region (which then excludes the possibility of its being a round disk), there is an optimal analog of (1). It will be seen that indeed there is a stronger isoperimetric inequality for triangular regions, and along the way, we shall encounter a classical inequality among positive numbers.

From now on, let  $\mathcal{R}$  be a region bounded by a triangle  $\partial\mathcal{R}$ . The first problem facing us is to express the area  $A$  in terms of the perimeter  $C$ .

<sup>1</sup>See the author’s *The isoperimetric inequality: the geometric story*.

The usual area formula for the triangle  $\mathcal{R}$  involves the height of  $\mathcal{R}$  on a base and does not seem to be related to  $C$ . It turns out that there is a formula expressing  $A$  directly in terms of  $C$ . Moreover, such a formula was most likely already known to Archimedes (287-212 B.C.), the greatest mathematician of antiquity, although it usually goes under the name of *Heron's formula* because it was first clearly stated and proved by Heron (75 A.D.? Very uncertain). Let

$$s = \frac{1}{2}C.$$

Heron's formula states that if  $a$ ,  $b$  and  $c$  denote the lengths of the three sides of a triangle, then:

$$(2) \quad A = \sqrt{s(s-a)(s-b)(s-c)}.$$

A proof can be found on pp. 79–80 of [R] or pp. 35–36 of [K1].

Now we want to show that  $A \leq \frac{1}{4\pi}C^2 = \frac{1}{\pi}s^2$ . According to (2), we may restate (1) in an equivalent form:

$$(3) \quad \sqrt{s(s-a)(s-b)(s-c)} \leq \frac{1}{\pi}s^2$$

Our next problem is: how to approach (3)? It would help to know that there is a simple inequality:

**Theorem AGM<sub>2</sub>.** *For any two positive numbers  $a$  and  $b$ ,*

$$(4) \quad ab \leq \left(\frac{a+b}{2}\right)^2,$$

*and equality holds exactly when  $a = b$ .*

We shall give three different proofs of Theorem AGM<sub>2</sub>. As in the case of the isoperimetric inequality, the condition for the inequality (4) to be an equality will be seen to be as interesting as the inequality itself. This lecture will pay as much attention to inequalities as the condition for equality in each case. The meaning of “AGM<sub>2</sub>” will become clear presently. Our immediate

concern is to use (4) to prove (3), and it goes as follow:

$$\begin{aligned}
\sqrt{s(s-a)(s-b)(s-c)} &= \sqrt{s(s-a)} \cdot \sqrt{(s-b)(s-c)} \\
&\leq \frac{1}{2}[s+(s-a)] \cdot \frac{1}{2}[(s-b)+(s-c)] \quad (\text{by(4)}) \\
&= \frac{1}{4}[2s-a][2s-(b+c)] \\
&\leq \frac{1}{4} \left( \frac{(2s-a)+(2s-(b+c))}{2} \right)^2 \quad (\text{by(4)again}) \\
&= \frac{1}{16} (4s-(a+b+c))^2 \\
&= \frac{1}{16} (4s-2s)^2 \\
&= \frac{1}{4} s^2,
\end{aligned}$$

where we have made use of the fact that  $2s = a + b + c$ . In view of (2), this implies that for a triangle with area  $A$  and perimeter  $2s$ ,

$$(5) \quad A \leq \frac{1}{4} s^2.$$

Since  $4 > \pi$ ,  $\frac{1}{4} < \frac{1}{\pi}$ . Thus (3) is proved, and therewith, the isoperimetric inequality for a triangular region.

We pause to make some remarks about the preceding proof of inequality (3). Recall that for triangular regions, (1) is equivalent to (3), and we have just seen that (5) implies (3). Thus with (5) in mind, we should ask if we can similarly assert as we did with (1) that

for a *triangular* region with area  $A$  and perimeter  $2s$ , (5) holds, and the inequality is an equality exactly when the region is *special* in a sense to be determined.

Unfortunately, the proof of inequality (5) shows that even for triangular regions, equality in  $A \leq \frac{1}{4} s^2$  can never take place. The reason is simple: if equality is achieved in (5), then the chain of equalities and inequalities preceding (5) must all be *equalities*. In particular, the first inequality in the proof of (5) comes from the application of (4); let us look at the part concerning

$$\sqrt{s(s-a)} \leq \frac{1}{2}(s+(s-a)).$$

According to (4), equality is possible only when  $s = s - a$ , which is equivalent to  $a = 0$ . Since  $a$  is the length of a side of the triangle, this is impossible. Therefore, it is *intrinsic to the proof itself* that this way of using  $\text{AGM}_2$  cannot lead to an optimal isoperimetric inequality for triangular regions.

Thus there remains the task of getting the smallest possible constant  $\beta_0$  so that, for all triangular regions,  $A \leq \beta_0 s^2$  and, furthermore, we need to characterize those triangular regions which satisfy  $A = \beta_0 s^2$ . We shall do that, but let us first prove Theorem  $\text{AGM}_2$ .

As mentioned earlier, we are going to give three proofs of Theorem  $\text{AGM}_2$  and each sheds a different light on it. For convenience, we shall use the standard symbol  $\iff$  to denote “is exactly the same as”.

**First Proof.** Without asking whether (4) is true or not, we shall re-interpret it several times as an equivalent inequality and at the end, its truth will be obvious.

$$\begin{aligned} ab \leq \left(\frac{a+b}{2}\right)^2 &\iff 4ab \leq (a+b)^2 \iff 4ab \leq a^2 + 2ab + b^2 \\ (6) \qquad \qquad \qquad &\iff 0 \leq a^2 - 2ab + b^2. \end{aligned}$$

So (4) is true *exactly when*  $0 \leq a^2 - 2ab + b^2$ . However the latter is true because  $(a^2 - 2ab + b^2) = (a - b)^2 \geq 0$ . So (4) is proved. Now if  $a = b$ , clearly equality holds in (4). We have to also show, however, that if equality holds in (4) so that  $ab = \left(\frac{a+b}{2}\right)^2$ , then in fact  $a = b$ . Now  $ab = \left(\frac{a+b}{2}\right)^2$  means that equality holds in (6), so that  $a^2 - 2ab + b^2 = 0$ . Thus  $(a - b)^2 = 0$ , and this is possible only if  $a - b = 0$ , which means  $a = b$  as desired. Q.E.D.

**Second Proof.** If  $a = b$ , then  $ab = \left(\frac{a+b}{2}\right)^2 = a^2$  and (4) certainly holds. Let then  $a \neq b$ , say  $a < b$ . Then  $\left(\frac{a+b}{2}\right) > \frac{a+a}{2} = a$  so that

$$a = \left(\frac{a+b}{2}\right) - s \quad \text{for some } s > 0.$$

Then

$$b = (b+a) - a = (b+a) - \left[\left(\frac{a+b}{2}\right) - s\right] = \left(\frac{a+b}{2}\right) + s,$$

so that

$$ab = \left(\frac{a+b}{2} - s\right) \left(\frac{a+b}{2} + s\right) = \left(\frac{a+b}{2}\right)^2 - s^2 < \left(\frac{a+b}{2}\right)^2.$$

This proves (4) in the form of a *strict* inequality. In particular, if  $a < b$ , then only strict inequality in (1) is possible. Therefore, if equality holds in (4), we must have  $a = b$ . Q.E.D.

**Third Proof.** Write  $P = a + b$ ; then  $ab = (P - b)b$ . In terms of  $P$ , the theorem may be rephrased as:

with  $P$  held fixed, the function  $f(b)$  given by  $f(b) = b(P - b)$  is  
at most  $\left(\frac{P}{2}\right)^2$ , and equals this maximum exactly when  $b = \frac{P}{2}$ .

To see this, notice that  $f(b) = -b^2 + Pb$ , so that its graph is an inverted parabola which crosses the  $x$ -axis at 0 and  $b$  because  $f(0) = f(P) = 0$ . Pictorially, we'd expect the highest point of the parabola to be over the midpoint  $\frac{P}{2}$  of the interval  $[0, P]$ , and since  $f\left(\frac{P}{2}\right) = \left(\frac{P}{2}\right)^2$ , the theorem is proved. But we don't need to rely on pictures because we know by completing the square that  $f(b) = -\left(b - \frac{P}{2}\right)^2 + \left(\frac{P}{2}\right)^2$ . It follows that for all

$b \in [0, P]$ ,  $f(b) \leq (\frac{P}{2})^2$ , and equality holds exactly when  $-(b - \frac{P}{2})^2 = 0$ , *i.e.* when  $b = \frac{P}{2}$ . Q.E.D.

It may be of interest to point out that  $\text{AGM}_2$  itself carries information about isoperimetry. Indeed, consider a rectangle with sides of length  $a$  and  $b$ . Then its area is  $A_0 \equiv ab$  and its perimeter is  $C_0 \equiv 2(a + b)$ . Inequality (4) is now equivalent to  $A_0 \leq \frac{1}{16}C_0^2$ , and  $\text{AGM}_2$  asserts that a rectangle with a fixed perimeter  $C_0$  achieves the maximum area  $\frac{1}{16}C_0^2$  exactly when it is a square.

For two positive numbers  $a$  and  $b$ ,  $\sqrt{ab}$  is called the *geometric mean* of  $a$  and  $b$ , and  $\frac{a+b}{2}$  is called their *arithmetic mean*. Theorem  $\text{AGM}_2$  is therefore the statement that the geometric mean of two positive numbers is less than or equal to their arithmetic mean, and equals the latter exactly when the two numbers are themselves equal. For this reason, Theorem  $\text{AGM}_2$  is called the *arithmetic-geometric-mean inequality* for two numbers, which then also explains the meaning of  $\text{AGM}_2$ . What about three positive numbers?

There are two reasons why we raise the question about three positive numbers. The first reason is that it is natural to do so. But there is a second reason: we have seen that one reason the inequality (5) is not optimal (the number  $\frac{1}{4}$  is too big) is that its proof breaks up the product  $s(s-a)(s-b)(s-c)$  into two groups of two factors each, and then (4) is applied to each group separately. An obvious alternative is to consider  $(s-a)(s-b)(s-c)$  together and hope to apply some version of  $\text{AGM}_2$  to this product of three numbers.

In any case, we would guess that the analogue of  $\text{AGM}_2$  for three numbers is:

$$(7) \quad (a_1 a_2 a_3)^{1/3} \leq \left( \frac{a_1 + a_2 + a_3}{3} \right),$$

and equality holds exactly when  $a_1 = a_2 = a_3$ ,

and for four positive numbers,

$$(8) \quad (a_1 a_2 a_3 a_4)^{1/4} \leq \left( \frac{a_1 + a_2 + a_3 + a_4}{4} \right),$$

and equality holds exactly when  $a_1 = \dots = a_4$ .

Assertions (7) and (8) will be referred to as  $\text{AGM}_3$  and  $\text{AGM}_4$ , respectively. In general, we define  $(a_1 a_2 \dots a_n)^{1/n}$  to be the *geometric mean* of the positive numbers  $a_1, \dots, a_n$ , and  $\left( \frac{a_1 + a_2 + \dots + a_n}{n} \right)$  to be their *arithmetic mean*. The obvious conjecture is that the geometric mean is always less than or equal to the arithmetic mean. We shall deal with this conjecture after we have taken a closer look at (7) and (8).

It is true that one can prove  $\text{AGM}_3$ , *i.e.*, (7), by a direct algebraic computation, but this turns out to be neither enlightening nor productive, in the

sense that it does not make us understand (7) any better and the method does not generalize to a proof of (8). A good proof is one that makes us understand what it proves, so such a computational proof is no good and will therefore not be pursued here. There is apparently no direct computational proof of (8). What is surprising, however, is that we can quite easily prove (8) on the basis of  $\text{AGM}_2$  and then *go backwards* to prove (7). This is the insight of Auguste Cauchy (1789-1857), who used this idea to prove the arithmetic-geometric-mean inequality in general. Let us first show the  $\text{AGM}_2$  implies  $\text{AGM}_4$ . So let  $a_1, a_2, a_3, a_4$  be positive numbers and we shall prove:  $a_1 a_2 a_3 a_4 \leq \left( \frac{a_1 + a_2 + a_3 + a_4}{4} \right)^4$ .

$$\begin{aligned}
a_1 a_2 a_3 a_4 &= (a_1 a_2)(a_3 a_4) \\
&\leq \left( \frac{a_1 + a_2}{2} \right)^2 \left( \frac{a_3 + a_4}{2} \right)^2 && \text{(by AGM}_2\text{)} \\
&= \left[ \left( \frac{a_1 + a_2}{2} \right) \left( \frac{a_3 + a_4}{2} \right) \right]^2 \\
&\leq \left[ \left\{ \frac{\left( \frac{a_1 + a_2}{2} \right) + \left( \frac{a_3 + a_4}{2} \right)}{2} \right\}^2 \right]^2 && \text{(by AGM}_2\text{ again)} \\
&= \left[ \left\{ \frac{a_1 + a_2 + a_3 + a_4}{4} \right\}^2 \right]^2 \\
&= \left( \frac{a_1 + a_2 + a_3 + a_4}{4} \right)^4.
\end{aligned}$$

Moreover, suppose equality holds in (8), then the first and last terms in the preceding chain of equalities-inequalities are equal, and therefore every inequality in this chain is necessarily an equality. But both inequalities above are a result of  $\text{AGM}_2$ , and we know from  $\text{AGM}_2$  that equality holds on the second line exactly when  $a_1 = a_2$  and  $a_3 = a_4$ , and equality holds on the fourth line exactly when  $\left( \frac{a_1 + a_2}{2} \right) = \left( \frac{a_3 + a_4}{2} \right)$ . Together, these imply that equality holds in (4) exactly when  $a_1 = a_2 = a_3 = a_4$ . We have thus completely proved  $\text{AGM}_4$ .

Now we prove  $\text{AGM}_3$  ((7)), using Cauchy's idea by *adding* an extra number to both sides of (7). Let  $a_1, a_2$ , and  $a_3$  be given positive numbers, and let  $A = \left( \frac{a_1 + a_2 + a_3}{3} \right)$ . This  $A$  is the arithmetic mean of  $a_1, a_2$ , and  $a_3$ ; equivalently,  $a_1 + a_2 + a_3 = 3A$ . By (8),

$$(9) \quad a_1 a_2 a_3 A \leq \left( \frac{a_1 + a_2 + a_3 + A}{4} \right)^4.$$

However,  $a_1 + a_2 + a_3 + A = 3A + A = 4A$ , so the right side of (9) is just  $A^4$ , and we have  $a_1 a_2 a_3 A \leq A^4$ . Since  $A$  is positive, multiplying both sides

by  $\frac{1}{A}$  does not change the inequality. So we obtain

$$a_1 a_2 a_3 \leq A^3 = \left( \frac{a_1 + a_2 + a_3}{3} \right)^3.$$

Moreover, suppose equality holds, *i.e.*,  $a_1 a_2 a_3 = \left( \frac{a_1 + a_2 + a_3}{3} \right)^3$ . Then  $a_1 a_2 a_3 = A^3$ , so that the left side of (9) becomes  $A^4$ , which then equals the right side of (9) which as we observed is also equal to  $A^4$ . So equality holds in (9). By  $\text{AGM}_4$ , this happens exactly when  $a_1 = a_2 = a_3 = A$ . So (7) is now also completely proved!

This idea of Cauchy's immediately lends itself to the general proof of the arithmetic-geometric-mean inequality,  $\text{AGM}_n$ , namely: if  $a_1, a_2, \dots, a_n$  are positive numbers, then

$$(10) \quad (a_1 a_2 \cdots a_n)^{\frac{1}{n}} \leq \frac{a_1 + a_2 + \cdots + a_n}{n} \quad \text{and equality holds exactly when } a_1 = \cdots = a_n$$

The proof can be easily described: In the same way  $\text{AGM}_2$  was used to prove  $\text{AGM}_4$ , we can use  $\text{AGM}_4$  to prove  $\text{AGM}_8$ , as follows.

$$\begin{aligned} a_1 a_2 \cdots a_8 &= (a_1 \cdots a_4)(a_5 \cdots a_8) \\ &\leq \left( \frac{a_1 + \cdots + a_4}{4} \right)^4 \left( \frac{a_5 + \cdots + a_8}{4} \right)^4 \quad (\text{by } \text{AGM}_4) \\ &= \left[ \left( \frac{a_1 + \cdots + a_4}{4} \right) \left( \frac{a_5 + \cdots + a_8}{4} \right) \right]^4 \\ &\leq \left[ \left\{ \frac{\left( \frac{a_1 + \cdots + a_4}{4} \right) + \left( \frac{a_5 + \cdots + a_8}{4} \right)}{2} \right\}^2 \right]^4 \quad (\text{by } \text{AGM}_2) \\ &= \left( \frac{a_1 + \cdots + a_8}{8} \right)^8, \end{aligned}$$

as desired.

This suggests a somewhat spectacular induction argument. For each positive integer  $k$ , let  $\mathcal{P}_k$  be the statement: if  $a_1, \dots, a_{2^k}$  are  $2^k$  positive numbers, then

$$(11) \quad (a_1 a_2 \cdots a_{2^k})^{\frac{1}{2^k}} \leq \frac{a_1 + \cdots + a_{2^k}}{2^k}, \quad \text{and equality holds exactly when } a_1 = \cdots = a_{2^k}$$

We claim that  $\mathcal{P}_k$  is true for all  $k = 1, 2, 3, \dots$ . Since  $\mathcal{P}_1$  is just  $\text{AGM}_2$ ,  $\mathcal{P}_1$  is true. Then we can show  $\mathcal{P}_k$  implies  $\mathcal{P}_{k+1}$  in exactly the same way that  $\text{AGM}_4$  implies  $\text{AGM}_8$ . The assertion about the condition of equality is also proved in the same way that the equality assertion in  $\text{AGM}_4$  was proved. (If you really want to see the details, they are in [BB] and [K2].) Thus we have proved  $\text{AGM}_{2^k}$ .

What about  $\text{AGM}_n$  when  $n$  is not a power of 2? We shall handle it in exactly the same way that we did  $\text{AGM}_3$  above. To illustrate, suppose we want to prove  $\text{AGM}_{29}$ . The power of 2 after 29 is 32, and we already know that  $\text{AGM}_{32}$  is true ( $32 = 2^5$ ). So in order to prove

$$(12) \quad (a_1 a_2 \cdots a_{29}) \leq \left( \frac{a_1 + \cdots + a_{29}}{29} \right)^{29}$$

we introduce as before  $A = \left( \frac{a_1 + \cdots + a_{29}}{29} \right)$  and use  $\text{AGM}_{32}$  to conclude that

$$(13) \quad (a_1 a_2 \cdots a_{29} A A A) \leq \left( \frac{a_1 + \cdots + a_{29} + A + A + A}{29} \right)^{32}$$

But  $a_1 + \cdots + a_{29} = 29A$ , so  $(a_1 + \cdots + a_{29} + A + A + A)/32 = (32A)/32 = A$ , and the right side of (13) is  $A^{32}$ . Thus  $(a_1 \cdots a_{29}) A A A \leq A^{32}$ . Since  $A$  is positive, we may multiply both sides by  $A^{-3}$  and get  $(a_1 \cdots a_{29}) \leq A^{29}$ , which is the inequality in (12). The case of equality in (12) can be argued exactly as in the case of  $\text{AGM}_3$ .

We have therefore proved (10), which is  $\text{AGM}_n$ .

We can now return to our roots: why we got started on the AGM inequality in the first place. We want to prove the isoperimetric inequality for triangular regions. We have seen that by Heron's formula, the area  $A$  is given by (2). We now apply  $\text{AGM}_3$  to (2) and make use of  $2s = a + b + c$ :

$$\begin{aligned} A &= \sqrt{s} [(s-a)(s-b)(s-c)]^{\frac{1}{2}} \\ &\leq \sqrt{s} \left[ \left( \frac{(s-a) + (s-b) + (s-c)}{3} \right)^3 \right]^{\frac{1}{2}} \\ &= \sqrt{s} \left[ \left( \frac{3s - (a+b+c)}{3} \right)^3 \right]^{\frac{1}{2}} \\ &= \sqrt{s} \left[ \left( \frac{s}{3} \right)^3 \right]^{\frac{1}{2}} \\ &= \frac{s^2}{3\sqrt{3}} \end{aligned}$$

Thus,

$$(14) \quad A \leq \frac{s^2}{3\sqrt{3}}.$$

Now  $\text{AGM}_3$  also says that the chain of inequalities would be equalities (and hence equality holds in (14)) exactly when  $s-a = s-b = s-c$ , which is the same as  $a = b = c$ . Thus the area attains its maximum value  $\frac{s^2}{3\sqrt{3}}$  exactly when the triangle is equilateral. Thus we have completely proved

the isoperimetric inequality for triangular regions. To bring it to the same form as the general isoperimetric inequality announced earlier, recall that the perimeter  $C$  equals  $2s$ . Thus we have:

**The isoperimetric inequality for triangular regions.** If  $A$  and  $C$  are the area and perimeter of a triangle, then:

$$(15) \quad A \leq \frac{1}{12\sqrt{3}}C^2,$$

and equality holds exactly when the triangle is equilateral.

By the way, it is instructive to directly check that the area of an equilateral triangle with each side equal to  $\frac{2s}{3}$  is indeed  $\frac{s^2}{3\sqrt{3}}$ . Also note the obvious fact that  $4\pi = 12.566\dots$ , whereas  $12\sqrt{3} = 20.78\dots$ . This gives a comparison between (1) and (15).

It goes without saying that the same conclusion could be obtained by the method of calculus: just maximize the function

$$F(a, b, c) = (s - a)(s - b)(s - c),$$

where  $s$  is a constant, within the region  $0 < a, b, c < s$ . However, the use of  $\text{AGM}_3$  gives an entirely different perspective on this problem, and serves to underline the importance of inequalities in mathematics.

We have used the isoperimetric inequality to introduce the arithmetic-geometric-mean inequality, but it would be wrong to suggest that the isoperimetric inequality somehow *justifies* the latter. The fact is that the arithmetic-geometric-mean inequality is one of the truly basic inequalities in mathematics and it comes up frequently in advanced mathematics. In addition, it is both surprising and beautiful. If you have any doubts of the latter fact, ask a friend to give you three very big numbers (e.g., 51244585, 60231463, 44428791) and ask her which of the two is bigger: their product or the cube of their average? Try this a few more times with other big numbers and watch her struggle before letting her in on the secret. What do you think would be her reaction?

**EXERCISE.** Of all the triangles with the same base and area, prove that the isosceles triangle has the smallest perimeter.

**EXERCISE.** Let  $r_1$  and  $r_2$  be the two roots of the quadratic polynomial  $x^2 - bx + c$ , where  $b$  and  $c$  are real numbers.

- (1) Prove that  $r_1r_2 = c$  and  $r_1 + r_2 = b$ .
- (2) Consider the following statement: "Using (1) and  $\text{AGM}_2$ , we have  $c \leq (\frac{b}{2})^2$ . Since  $b$  and  $c$  are arbitrary, this is a contradiction." Explain what is wrong with the preceding argument. What is a correct statement about the relationship

between  $b$  and  $c$ ?

DEPARTMENT OF MATHEMATICS #3840, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720-3840 WU@MATH.BERKELEY.EDU