A Brief History of Impossibility

JEFF SUZUKI
Brooklyn College
Brooklyn, NY
jeff_suzuki@yahoo.com

Sooner or later every student of geometry learns of three "impossible" problems:

1. Trisecting the angle: Given an arbitrary angle, construct an angle exactly one-third as great.
2. Duplicating the cube: Given a cube of arbitrary volume, find a cube with exactly twice the volume.
3. Squaring the circle: Given an arbitrary circle, find a square with the same area.

These problems originated around 430 BC at a time when Greek geometry was advancing rapidly. We might add a fourth problem: inscribing a regular heptagon in a circle. Within two centuries, all these problems had been solved (see [**3**, Vol. I, p. 218–270] and [**1**] for some of these solutions).

So if these problems were all solved, why are they said to be impossible? The "impossibility" stems from a restriction, allegedly imposed by Plato (427–347 BC), that geometers use no instruments besides the compass and straightedge. This restriction requires further explanation. For that, we turn to Euclid (fl. 300 BC), who collected and systematized much of the plane geometry of the Greeks in his *Elements*.

Euclid's goal was to develop geometry in a deductive manner from as few basic assumptions as possible. The first three postulates in the *Elements* are (in modernized form):

1. Between any two points, there exists a unique straight line.
2. A straight line may be extended indefinitely.
3. Given any point and any length, a circle may be constructed centered at the point with radius equal to the given length.

These three postulates correspond to the allowable uses of compass and straightedge: to draw a line that passes through two given points; to extend a given line segment indefinitely; and to draw a circle about any given point with any given radius. To solve a problem using compass and straightedge means to use only these operations, repeated a finite number of times. The construction's validity can then be proven using only the postulates of Euclidean geometry.

For example, consider the problem of duplicating the cube. In order to duplicate a cube with a side length of a, it is necessary to construct a line segment of length $\sqrt[3]{2}a$. One of the simpler solutions, presented by Menaechmus around 350 BC, is equivalent to locating the intersection point of the parabola $ay = x^2$ and the hyperbola $xy = 2a^2$; these two curves intersect at the point $\left(\sqrt[3]{2}a, \sqrt[3]{4}a\right)$. Since this solution requires the use of the hyperbola and parabola, it is not a compass and straightedge solution.

A more subtle problem occurs with the trisection problem. Suppose we wish to trisect $\angle BOE$, which we may assume to be the central angle of arc BE in a circle (see Figure 1). There are several *neusis* ("verging") solutions, one of which is the following. Draw BC parallel to OE and then draw CA with the property that $DA = OB$ (the radius of the circle). It is relatively easy to prove that $\angle DOA = \frac{1}{3}\angle BOE$ (a proof we will leave to the reader). We can accomplish this construction with compass and

straightedge as follows: Open the compass to fixed length equal to the radius OB. Using C as a pivot, swing the straightedge around, using the compass to measure out a length OB from the point where the straightedge crosses the circle, until you find the point D where the length $DA = OB$.

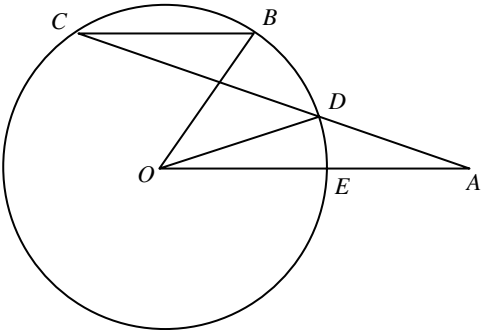

Figure 1 Neusis trisection of an angle.

There are at least two objections that can be raised to this "compass and straightedge" solution. First, the postulates only guarantee the existence of a line between two points, or the extension of an existing line; hence there is no guarantee that the line like CA, specified by a point C and a length DA, even exists. Second, the postulates only allow us to measure out a length by means of a circle of known center. This means we cannot measure the length DA equal to OB until we locate D. Thus, even though this solution uses compass and straightedge, it is not a compass and straightedge solution.

Even if we restrict ourselves to the canonical uses of the compass and straightedge, how can we distinguish between constructions that have never been done and those that are actually impossible? Before 1796, no compass and straightedge construction for a regular heptadecagon was known, but in that year Gauss discovered how to inscribe one in a circle. Might there be some as-yet-undiscovered means of trisecting an angle or duplicating the cube using compass and straightedge? In 1837 an obscure French mathematician named Pierre Wantzel (1814–1848) proved this could not be: cube duplication and angle trisection are in fact impossible, as is constructing a regular heptagon or squaring the circle. In the following we'll trace the steps leading up to Gauss's construction of the heptadecagon and Wantzel's proof of impossibility.

Descartes

The first important step towards proving certain constructions impossible was taken by René Descartes (1596–1650) in his *The Geometry* (1637). Descartes's key insight was that by identifying the lengths of line segments with real numbers, one could restate a geometric problem as an algebraic one, express the solution symbolically, then convert the algebraic expression into a geometric construction procedure.

In order to take this last step, we must develop an arithmetic of lines. Let AB and CD be two line segments (where we will assume CD is shorter than AB). Compass and straightedge techniques from the *Elements* allow us to find line segments that correspond to the sum $AB + CD$, difference $AB - CD$, and $q \cdot AB$ (for any positive rational q). The problem arises when trying to interpret the product $AB \cdot CD$. Euclid and others identified this product as the rectangle whose adjacent sides were equal in length to AB and CD. This would mean the arithmetic of line segments was not

closed under multiplication; moreover, it would make the division of two line segments impossible to define.

Descartes realized that the theory of proportions could be used to identify the product of two line segments with another line segment, provided we had a line segment of unit length. Imagine two lines intersecting at B at any angle whatsoever, and say we wish to multiply BD by BC. Mark off BA equal to the unit, and join AC (see Figure 2). Draw DE parallel to AC. Then triangles BAC, BDE are similar, and we have the ratio $BE : BD = BC : BA$. This corresponds to the equality of the two products $BE \cdot BA = BC \cdot BD$. Since BA is equal to the unit, we can thus identify the line segment BE with the product $BC \cdot BD$. Thus the product of two line segments is another line segment. Division can be handled in virtually the same way.

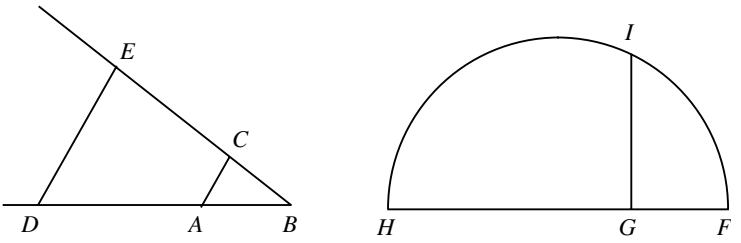

Figure 2 Multiplication and roots.

Proposition 14 of Book II of the *Elements* gives the construction technique for finding square roots (literally the side of a square equal in area to a given rectangle), which Descartes modified to extract square roots [**7**, p. 5]. Suppose we wish to find the square root of GH. Extend GH by GF equal to the unit, then draw the circle with FH as its diameter. The perpendicular GI will equal the square root of GH (see Figure 2).

Suppose we begin with a line segment AB (which we can take to be our unit). If we can construct a line of length $k \cdot AB$ using the above techniques, we say that k is a constructible number (and $k \cdot AB$ is a constructible line segment). In general, k is a constructible number if it is rational, or the root of a quadratic equation with constructible coefficients. A figure is constructible if all the line segments required for its construction are constructible. Moreover, given a constructible figure, any line segment we can obtain from it (e.g., the diagonal of a square) is constructible. For example, if we could square the circle, then $\sqrt{\pi}$ would be constructible; equivalently, if $\sqrt{\pi}$ is inconstructible, squaring the circle is impossible.

This identification of a geometric problem with an algebraic problem allows us to phrase the problem of constructibility in terms of the roots of a specific equation: if the root is a constructible number, the corresponding geometric problem can be solved using compass and straightedge alone. Duplicating the cube would allow us to find a line of length $\sqrt[3]{2}$, which is a root of the equation $x^3 - 2 = 0$. Constructing a regular n-gon would allow us to find a line of length $\sin\frac{2\pi}{n}$, which is the imaginary part of one of the roots of $x^n - 1 = 0$ (because of this, the problem of finding the roots of $x^n - 1 = 0$ is also known as the cyclotomy problem).

Trisection of an angle corresponds to a cubic equation as follows. Given a circle with center O and unit radius, with central angle AOC equal to 3θ. We wish to find point B on the circle where angle BOC is equal to θ. If we drop AD and BE perpendicular to OC, we have $AD = \sin 3\theta$, $BE = \sin\theta$. These quantities are related through the identity:

$$\sin 3\theta = 3\sin\theta - 4\sin^3\theta$$

Since angle AOC is the given angle, then $\sin 3\theta$ is a known quantity which we can designate as l. Thus if the real roots of $l = 3x - 4x^3$ are not constructible, trisection of the corresponding angle is impossible.

Vandermonde and Lagrange

The next step towards answering the constructibility problem came from the work of Alexandre-Théophile Vandermonde (1735–1796) and Joseph Louis Lagrange (1736–1813). Vandermonde's [**8**], presented to the Paris Academy in 1770, and Lagrange's [**6**], presented to the Berlin Academy in 1771, examined why general solutions to equations of degree 3 and 4 existed. Both came to the same conclusion independently: Our ability to solve these equations is due to the fact that we can find the value of certain expressions of the roots without knowing the roots themselves.

To understand their methods, consider the quadratic equation $x^2 - px + q = 0$, with roots $x = a$ and $x = b$. Thus $p = a + b$ and $q = ab$. Next, take any function of the roots of this equation. Some functions, such as $f(r_1, r_2) = r_1 + r_2$, have the same value regardless of which root we regard as r_1 and which root we regard as r_2; these are called symmetric functions. It was widely believed (though not proven until the middle of the nineteenth century) that every symmetric function of the roots of a polynomial could be expressed as a rational function of the coefficients. In this case, $f(a, b) = f(b, a) = p$.

On the other hand, consider a function like $g(r_1, r_2) = r_1 - r_2$. Depending on which root we call r_1 and which root we call r_2, g might take on one of two possible values, $a - b$ or $b - a$. In order to find the values of this non-symmetric function of the roots, Lagrange let the k distinct values be the roots of a kth degree equation. In our example, the two values of g would be the roots of:

$$(y - (a - b))(y - (b - a)) = y^2 - (a - b)^2$$

A little algebra shows us that $(a - b)^2 = (a + b)^2 - 4ab$. Since we know the values of $a + b$ and ab, we can determine, even without knowing the roots, that

$$(a - b)^2 = p^2 - 4q.$$

Thus the two different values of $a - b$ will be the two roots of $y^2 - (p^2 - 4q) = 0$. Hence $a - b = \sqrt{p^2 - 4q}$ or $-\sqrt{p^2 - 4q}$. It will make no difference which we choose; for example, we might let $a - b = \sqrt{p^2 - 4q}$. To solve for a and b separately, we need a second equation, which we can obtain from the coefficients: $a + b = p$. These two equations give us the system:

$$a + b = p \qquad a - b = \sqrt{p^2 - 4q}$$

Hence $a = \frac{p+\sqrt{p^2-4q}}{2}$, and $b = \frac{p-\sqrt{p^2-4q}}{2}$.

Both Vandermonde and Lagrange considered the problem of finding the nth roots of unity, which would be the roots of $x^n - 1 = 0$. Lagrange noted the correspondence between the roots of $x^n - 1 = 0$ and the cyclotomy problem; further, he observed that if n is prime, all of the roots can be generated by the successive powers of any root except $x = 1$. This allowed him to write equations relating the roots; solving the equations would give the roots of unity. Lagrange used his method to find the roots of unity for $n = 3$ through $n = 6$ (all of which can be found using only square roots), while Vandermonde found the roots of unity up to $n = 11$ using similar methods.

Gauss

According to legend, Carl Friedrich Gauss (1777–1855) discovered the constructibility of the regular heptadecagon in 1796; this inspired him to choose mathematics as his future field of study, despite the indifferent reception of his discovery by A. G. Kästner at Göttingen. Gauss's main contribution to the problem of cyclotomy was inventing a method of splitting the roots of unity into sets where the sum of the roots in each set was the root of an equation with determinable coefficients. He described his method in [**2**], where the solution to the cyclotomy problem appeared as one application of the theory of quadratic residues. While Gauss's discovery was unprecedented, it was a straightforward, albeit clever, application of the ideas of Lagrange and Vandermonde.

The nth roots of unity are solutions to the equation $x^n - 1 = 0$. Obviously any root r must satisfy $r^n = 1$. If n is the least power of r that is equal to 1, then r is said to be a primitive nth root of unity. For example, the roots of $x^4 - 1 = 0$ are $\pm 1, \pm i$. Since $1^1 = 1$ and $(-1)^2 = 1$, then neither 1 nor -1 is primitive. On the other hand, the least power of i or $-i$ that gives 1 is the fourth power; thus i and $-i$ are primitive roots, and their powers will generate all the roots; for example:

$$i, \quad i^2 = -1, \quad i^3 = -i, \quad i^4 = 1$$

In general, as Lagrange noted, if n is prime, then there are $n - 1$ primitive roots of unity.

As we note above, constructibility of the regular n-gon corresponds to constructibility of the roots of $x^n - 1 = 0$. We'll illustrate Gauss's general method by finding the 5th roots of unity. These would be solutions to the equation $x^5 - 1 = 0$. There is one (non-primitive) root $x = 1$. Removing a factor of $x - 1$ we obtain the equation

$$x^4 + x^3 + x^2 + x + 1 = 0$$

which is called the cyclotomic equation. All primitive fifth roots r must satisfy this equation.

Gauss considered a sequence whose first term is a primitive root, and where each term is some (constant) power of the previous term. For example, if we take r and cube it repeatedly, we obtain:

$$r, r^3, r^9, r^{27}, r^{81} \ldots$$

Since r is a root of $x^5 - 1 = 0$, then $r^5 = 1$. Hence the above sequence simplifies to r, r^3, r^4, r^2, $r, \ldots$, and all the roots appear in this sequence. On the other hand, suppose we take r and repeatedly raise it to the fourth power, obtaining the sequence:

$$r, r^4, r^{16}, r^{64}, \ldots$$

In this case, the only distinct members of the sequence are r and r^4.

Note that the remaining roots, r^2 and r^3, are the squares of the two distinct terms of this last sequence: $(r)^2 = r^2$, and $(r^4)^2 = r^8 = r^3$. More generally, suppose n is prime and r is a primitive nth root of unity. Gauss showed that our sequence of powers will have k distinct elements, where k is a divisor of $n - 1$. Moreover the remaining roots (if k is not equal to $n - 1$) can be separated into sets of k distinct elements, each of which is a power of a root of the original set.

For example, consider the $n = 7$ case, and a primitive root p. The sequence

$$p, p^6, p^{36}, p^{216}, \ldots$$

contains only two distinct roots, p and p^6. The squares of these are p^2, $p^{12} = p^5$, and the cubes are p^3, p^4. Thus the six roots have been partitioned into three sets, $\{p, p^6\}$, $\{p^2, p^5\}$, and $\{p^3, p^4\}$.

Note that the decomposition is not unique; for example, the sequence

$$p, p^2, p^4, p^8, \ldots$$

contains three distinct roots, p, p^2, and p^4; the remaining roots are the cubes of these roots and the six roots will be partitioned into two sets, $\{p, p^2, p^4\}$, and $\{p^3, p^5, p^6\}$.

Returning to the $n = 5$ case, we have split the roots into two sets: $\{r, r^4\}$ and $\{r^2, r^3\}$. Gauss then considered the sum of the roots in each set (designating these sums as "periods"), and let the sums be the roots of an equation:

$$\begin{aligned}(y - (r + r^4))(y - (r^2 + r^3)) &= y^2 - (r^4 + r^3 + r^2 + r)y + (r + r^4)(r^2 + r^3) \\ &= y^2 - (r^4 + r^3 + r^2 + r)y + (r^3 + r^4 + r^6 + r^7) \\ &= y^2 - (r^4 + r^3 + r^2 + r)y + (r^4 + r^3 + r^2 + r) \\ &= y^2 + y - 1\end{aligned}$$

where we made use of the fact that r satisfies the equation $x^4 + x^3 + x^2 + x + 1 = 0$. Hence the two periods $r + r^4$ and $r^2 + r^3$ correspond to the two roots of the quadratic equation $y^2 + y - 1 = 0$. We find the roots are $y = \frac{-1 \pm \sqrt{5}}{2}$.

One of these roots corresponds to $r + r^4$, and the other corresponds to $r^2 + r^3$. In principle it makes no difference which we assign to $r + r^4$, though in practice it is convenient if r is the principal fifth root of unity $\cos \frac{2\pi}{5} + i \sin \frac{2\pi}{5}$. Gauss noted that we could find this root numerically and see which of the two roots of $y^2 + y - 1 = 0$ was equal to $r + r^4$. Alternatively, we might note that $r + r^4$ will have a positive real component; hence $r + r^4 = \frac{-1 + \sqrt{5}}{2}$.

To find r, we can construct a quadratic equation with r and r^4 as roots:

$$(z - r)(z - r^4) = z^2 - (r + r^4)z + r^5 = z^2 - \left(\frac{-1 + \sqrt{5}}{2}\right)z + 1$$

Note that the coefficients of this equation are constructible numbers; hence its roots will also be constructible. These roots are:

$$z = \frac{\left(\frac{-1+\sqrt{5}}{2}\right) \pm \sqrt{\left(\frac{-1+\sqrt{5}}{2}\right)^2 - 4}}{2}$$

One of these will be the principal fifth root of unity, and the other will be its fourth power. Since the principal fifth root of unity is equal to $\cos \frac{2\pi}{5} + i \sin \frac{2\pi}{5}$, we can, as Gauss suggested, approximate the sine and cosine values and determine which of the two roots corresponds to the principal root (which will tell us, among other things, $\cos \frac{2\pi}{5} = \frac{-1+\sqrt{5}}{4}$ and $\sin \frac{2\pi}{5} = \frac{1}{4}\sqrt{10 + 2\sqrt{5}}$). Since the $\sin \frac{2\pi}{5}$ is constructible, so is the regular pentagon, a fact known to the ancients: Euclid's construction appears as Proposition 11 of Book IV (though Ptolemy gave a much easier construction in the *Almagest*).

On the other hand, consider the regular heptagon. In the above, we found that the roots can be separated into two periods, $p + p^2 + p^4$ and $p^3 + p^5 + p^6$. Letting the sum of the roots in the set be the roots of a quadratic equation and reducing as before

we obtain:

$$(y-(p+p^2+p^4))(y-(p^3+p^6+p^5)) = y^2+y+2$$

with roots $y = \frac{-1\pm\sqrt{-7}}{2}$; if p is the principal root, then $p+p^2+p^4$ has a positive imaginary component so $\frac{-1+\sqrt{-7}}{2} = p+p^2+p^4$ and $\frac{-1-\sqrt{-7}}{2} = p^3+p^5+p^6$.

The next step would be letting p, p^2, and p^4 be the three roots of a cubic equation:

$$\begin{aligned}(z-p)(z-p^2)(z-p^4) &= z^3-(p+p^2+p^4)z^2+(p^3+p^6+p^5)z-p^7\\ &= z^3-\left(\frac{-1+\sqrt{-7}}{2}\right)z^2+\left(\frac{-1-\sqrt{-7}}{2}\right)z-1\end{aligned}$$

While we can solve the cubic equation, we cannot do so by means of basic arithmetic operations and square roots alone; we must extract a cube root. Hence it would appear that the primitive seventh root of unity (and consequently the regular heptagon) is inconstructible.

The preceding example suggests the following: Suppose we wish to construct a regular n-gon, where n is prime. If $n-1$ has any prime factors other than 2, then at some point in separating the roots, we will have to solve an equation of a degree higher than 2. Hence constructing a regular n-gon using this method requires $n = 2^k+1$.

We can go a little further. If k has any odd factors, then 2^k+1 is composite; this follows because if $k = pq$ and q is odd, $x^{pq}+1$ has a factor of x^p+1. Thus a regular n-gon, where n is prime, *might* be constructible if n is a so-called Fermat prime, with $F_m = 2^{2^m}+1$. The known Fermat primes are 3, 5, 17, 257, 65537; it is unknown if higher Fermat primes exist.

In any case, consider $n = 17$. The corresponding cyclotomic equation has 16 roots. Gauss split these into two sets of eight roots apiece; hence, a quadratic equation could be used to find the sum of eight of the roots. Each set of eight could in turn be split into two sets of four; again, a quadratic equation could be used to find the sum of four of the roots. Each set of four could be split into two sets of two, and the sum of two of the roots could be found. Finally the sets of two could be broken down into their individual roots, so a primitive 17th root of unity could be found. Since none of the equations has a degree higher than 2, the roots are constructible; hence a regular heptadecagon can be constructed using only compass and straightedge.

Wantzel

Gauss's method suggests but does not prove the constructibility of the 257- and 65,537-gons (we need Sylow's Theorem to guarantee constructibility); likewise, it suggests but does not prove the impossibility of constructing a regular heptagon.

The first proof of the impossibility of certain geometric constructions came from Pierre Wantzel (1814–1848) in [**9**] (1837). Wantzel began by considering a system of quadratic equations (which for brevity we will call a Wantzel System):

$$\begin{aligned}x_1^2+Ax_1+B &= 0\\ x_2^2+A_1x_2+B_1 &= 0\\ x_3^2+A_2x_2+B_2 &= 0\\ &\vdots\\ x_n^2+A_{n-1}x_2+B_{n-1} &= 0\end{aligned}$$

where A, B are rational functions of some given quantities; A_1, B_1 are rational functions of the given quantities together with x_1 (and hence the coefficients of the second equation are constructible numbers); A_2, B_2 are rational functions of the given quantities, together with x_1, x_2, and in general A_m, B_m are rational functions of the given quantities and the variables $x_1, x_2, \ldots, x_m$. Note that Gauss's method of showing the constructibility of a pentagon or heptadecagon made use of precisely such a system; in the case of the pentagon the Wantzel system is:

$$y^2 + y - 1 = 0$$
$$z^2 - yz + 1 = 0$$

More generally, every constructible number r corresponds to some Wantzel system.

Consider any of these equations $x_{m+1}^2 + A_m x_{m+1} + B_m = 0$. Remarkably, the rational functions A_m, B_m can always be reduced to a *linear* function of the form $A'_{m-1}x_m + B'_{m-1}$, where A'_{m-1} and B'_{m-1} are rational functions of the given quantities and the variables $x_1, x_2, \ldots, x_{m-1}$. This reduction can be performed in two steps. First, the preceding equation $x_m^2 + A_{m-1}x_m + B_{m-1} = 0$ can be used to eliminate the higher powers of x_m in the expression for A_m and B_m, reducing them to the form $\frac{C_m x_m + D_m}{E_m x_m + F_m}$. Then the numerator and denominator can be multiplied by a constant quantity to reduce the rational function to a linear one.

For example, suppose we have the system of equations:

$$x^2 - 5x + 2 = 0$$
$$y^2 + \left(\frac{x^3 + 3x + 1}{2x - 1}\right) y + \left(\frac{1}{x^2 + 7x + 5}\right) = 0$$

From the first equation we have $x^2 = 5x - 2$. Hence $x^3 = 5x^2 - 2x = 23x - 10$. Thus the second equation can be reduced to:

$$y^2 + \left(\frac{26x - 9}{2x - 1}\right) y + \left(\frac{1}{12x + 3}\right) = 0$$

How can we eliminate the rational functions? Consider the first rational function. Suppose we multiply numerator and denominator by some constant C so that

$$C(26x - 9) = (2x - 1)(\alpha x + \beta)$$

for some values of α, β; then the common factor of $2x - 1$ can be removed and the rational expression simplified to a linear one. Expanding gives us:

$$26Cx - 9C = 2\alpha x^2 + (2\beta - \alpha)x - \beta$$

We can make use of the substitution $x^2 = 5x - 2$ to eliminate the square term:

$$26Cx - 9C = (2\beta + 9\alpha)x - (\beta + 4\alpha)$$

Comparing coefficients gives us a system of 2 linear equations in 3 unknowns:

$$26C = 2\beta + 9\alpha \qquad 9C = \beta + 4\alpha$$

Because this system is underdetermined, we may express two of the variables in terms of the third. For example, one solution is $\alpha = 2$, $\beta = -23/4$, and $C = 1/4$; in other words $\frac{1}{4}(26x - 9) = (2x - 1)(2x - \frac{23}{4})$. Thus:

$$\frac{26x-9}{2x-1} = \frac{\frac{1}{4}}{\frac{1}{4}}\frac{26x-9}{2x-1} = \frac{(2x-1)(2x-23/4)}{\frac{1}{4}(2x-1)} = 8x-23$$

In this way the final equation $x_n^2 + A_{n-1}x_n + B_{n-1} = 0$ can be converted into an equation where the coefficients A_{n-1} and B_{n-1} are linear functions of x_{n-1}.

Next, consider that x_{n-1} is one of the solutions to a quadratic equation. If we allow x_{n-1} to take on its two possible values, we obtain two different expressions for A_{n-1} and B_{n-1}, and consequently two different equations quadratic in x_n. Multiplying these two equations together will give us a fourth degree equation in x_n whose coefficients are functions of the given quantities and the variables $x_1, x_2, \ldots, x_{n-2}$. As before we can reduce these coefficients to linear functions of x_{n-2}; letting x_{n-2} take on its two possible values and multiplying the corresponding expressions will give us an eighth degree equation in x_n whose coefficients can be reduced to linear functions of x_{n-3}. Eventually we will end with an equation in x_n of degree 2^n whose coefficients are rational functions of the given quantities. This leads us to a preliminary theorem:

THEOREM. *Any Wantzel system of n equations corresponds to an equation of degree 2^n whose coefficients are rational functions of the given quantities; consequently, any constructible number is a root of an equation of degree 2^n whose coefficients are rational functions of the given quantities.*

For example, the Wantzel system corresponding to the construction of the pentagon was:

$$y^2 + y - 1 = 0$$
$$z^2 - yz + 1 = 0$$

Let the two roots of the first equation be $y = a$ and $y = b$. In the above we found these two roots, and used them to form a quadratic equation in z to find the principal fifth root of unity.

On the other hand, we can also write a single expression (which we will call the Wantzel polynomial) which contains all the roots. In this case, we can substitute the two roots $y = a$ and $y = b$ in to the left hand side of the second equation, then multiply the two expressions to obtain:

$$(z^2 - az + 1)(z^2 - bz + 1) = z^4 - (a+b)z^3 + (2+ab)z^2 - (a+b)z + 1$$

Since a and b are the two roots of $y^2 + y - 1$, we have $a + b = -1$ and $ab = -1$. Thus the equation

$$z^4 + z^3 + z^2 + z + 1 = 0$$

contains all solutions to the Wantzel system.

Next, suppose $x_n = r$ is a root of the Wantzel polynomial corresponding to a Wantzel system of n equations; further suppose that no Wantzel system of fewer than n equations exists with $x_n = r$ as a root. Wantzel then proved that no variable x_k could be expressed as a rational function of $x_1, x_2, \ldots, x_{k-1}$; equivalently, the quadratic equations are irreducible. This is because if one of the equations can be factored, then the preceding equation can be eliminated and we would obtain two Wantzel system of $n - 1$ equations, which would contain all the roots of the original system (and in particular r could be found by a Wantzel system of $n - 1$ equations). For example, consider the system:

$$x^2 - 3x - 7 = 0$$
$$y^2 - (4x - 1)y + 8x = 0$$
$$z^2 - (4y)z + (4y^2 - 1) = 0$$

and let $z = r$ be one of the roots. Note that the last equation factors, so we may write two separate Wantzel systems where the third equation differs, namely

$$x^2 - 3x - 7 = 0$$
$$y^2 - (4x - 1)y + 8x = 0$$
$$z - (2y + 1) = 0$$

and

$$x^2 - 3x - 7 = 0$$
$$y^2 - (4x - 1)y + 8x = 0$$
$$z - (2y - 1) = 0$$

where z can be expressed as a rational function of the preceding variables.

Consider the first system. Let the roots $y^2 - (4x - 1)y + 8x = 0$ be $y = a$ and $y = b$; letting y take on these two values in the third equation and multiplying the factors gives us the expression

$$(z - (2a + 1))\,(z - (2b + 1) = z^2 - (2a + 2a + 2)z + (4ab + 2a + 2b + 1)$$

But if the roots of $y^2 - (4x - 1)y + 8x = 0$ are $y = a$ and $y = b$, then $a + b = 4x - 1$, and $ab = 8x$; hence the second and third equation can be combined to form the single equation $z^2 - 8xz + (40x - 1) = 0$. Thus in place of the three equations, we have two equations:

$$x^2 - 3x - 7 = 0$$
$$z^2 - 8xz + (40x - 1) = 0$$

The reader can verify that the second Wantzel system would have $z^2 - (8x - 4)z + (24x - 4) = 0$ as its second equation. Thus in place a Wantzel system containing n equations, we would have two systems containing $n - 1$ equations, which between them contain all the roots z of the original system; hence $z = r$ would be the root of a Wantzel system containing $n - 1$ equations, which contradicts our original assumption.

Note that any solution x_n of the Wantzel polynomial $f(x)$ is a solution of $x_n^2 + A_{n-1}x_n + B_{n-1} = 0$, where A_{n-1}, B_{n-1} are found by substituting some set of solutions $\{x_1, x_2, \ldots, x_{n-1}\}$ to the equations of the Wantzel system. For example, the primitive fifth root of unity $z = \cos\frac{2\pi}{5} + i\sin\frac{2\pi}{5}$ of $z^4 + z^3 + z^2 + z + 1 = 0$ corresponds to a root of $z^2 - yz + 1 = 0$ where y is a solution to $y^2 + y - 1 = 0$.

Wantzel used this idea to prove that if another polynomial $F(x)$ had any root $x_n = a$ in common with $f(x)$, then it must have all roots in common; hence $f(x)$ is irreducible. Let $x_n = a$ be the root corresponding to the set $\{x_1, x_2, \ldots, x_{n-1}\}$, and let $F(x)$ be a polynomial with rational coefficients with $F(a) = 0$. As before we can reduce $F(x)$ to an expression of the form $A'_{n-1}x_n + B'_{n-1}$, where A'_{n-1}, B'_{n-1} are functions of the given quantities and the variables $x_1, x_2, \ldots, x_{n-1}$. Moreover, A'_{n-1} and B'_{n-1} must be equal to zero (since if they were not, x_n could be expressed as a rational

function of $x_1, x_2, \ldots, x_{n-1}$); hence we have $A'_{n-1} = 0$ (and likewise, $B'_{n-1} = 0$). But A'_{n-1} can be reduced as before to a linear function of x'_{n-1}. Thus the equation $A'_{n-1} = 0$ gives us an equation of the form $A'_{n-2}x_{n-1} + B'_{n-2} = 0$, where A'_{n-2} and B'_{n-2} are functions of the given quantities and the variables $x_1, x_2, \ldots, x_{n-2}$.

As before A'_{n-2} and B'_{n-2} must be both equal to zero; from $A'_{n-2} = 0$ we can obtain an equation of the form $A'_{n-3}x_{n-2} + B'_{n-3} = 0$. Continuing in this fashion we will eventually arrive at an equation of the form $A'x_1 + B' = 0$, where A' and B' are functions of the given quantities only. Again, x_1 cannot be a rational function of the given quantities only, so A' and B' must both equal zero; since they contain no variables, they are identically zero. Thus the two roots of $x_1^2 + Ax_1 + B = 0$ satisfy $A'x_1 + B' = 0$.

Now consider the equation $A'_1x_2 + B'_1 = 0$. A'_1 and B'_1 have both been reduced to linear functions of x_1 that will equal zero for any value of x_1 that satisfies $x_1^2 + Ax_1 + B = 0$. Thus the two possible values of x_1 will make both A'_1 and B'_1 equal to zero; consequently the four possible values of x_2 will make $A'_1x_2 + B'_1 = 0$. In a like manner, the eight possible values of $x_3^2 + A_2x_3 + B_1 = 0$ will satisfy the equation $A'_2x_3 + B'_2 = 0$, and so on, and so the 2^n possible roots of $x_n^2 + A_{n-1}x_n + B_{n-1} = 0$ will satisfy $F(x)$. Hence if $F(x)$ shares any root with $f(x)$, it will share all the roots of $f(x)$.

For example, consider our system

$$y^2 + y - 1 = 0$$
$$z^2 - yz + 1 = 0$$

which corresponded to the single equation $z^4 + z^3 + z^2 + z + 1 = 0$. Let $z = z_1$ be the root corresponding to one of the roots $y = y_1$ of the first equation, and suppose there was another polynomial $F(z)$ with rational coefficients that also had $z = z_1$ as a root.

First, we can eliminate the higher powers of z in $F(z)$ by the equation $z^2 - yz + 1 = 0$. This allows us to write $F(z)$ as a polynomial in y and z of the form

$$(z^2 - yz + 1)f(y, z) + A_1z + B_1,$$

where A_1 and B_1 are functions of y and $f(y, z)$ is some polynomial in y and z. Since $z = z_1$ satisfies (by assumption) the equation $z^2 - yz + 1 = 0$ when $y = y_1$, then substituting in these values gives us $A_1z_1 + B_1$, which (since z_1 is a root of F) must equal zero. Since the system is minimal, z cannot be expressed as a rational function of y, so A_1 and B_1 must both equal zero when $y = y_1$.

Next take (for example) the expression A_1, which we can write as

$$(y^2 + y - 1)g(y) + A'y + B',$$

where A', B' are rational functions of the given quantities only. Since $y = y_1$ satisfies $y^2 + y - 1 = 0$, and (by the above) satisfies $A_1 = 0$, then $A'y_1 + B' = 0$. But y_1 (by assumption) cannot be written as a rational function of the given quantities; hence A' and B' are simultaneously equal to zero. Since they contain no variable terms at all, then A' and B' must be identically zero and $A_1 = (y^2 + y - 1)g(y)$. Hence any solution to $y^2 + y - 1 = 0$ will make $A_1 = 0$. The same reasoning applies to B_1.

Since $F(z)$ can be written as $(z^2 - yz + 1)f(y, z) + A_1z + B_1$, and $A_1 = 0$, $B_1 = 0$ when y is equal to either root of $y^2 - y + 1 = 0$, then any of the four roots of $z^4 + z^3 + z^2 + z + 1 = 0$ will satisfy $F(z) = 0$. Hence $F(z)$ must contain all the roots.

At last this gives us a necessary condition for constructibility:

WANTZEL'S THEOREM. *If r is a constructible number, it must be the root of an irreducible polynomial of degree 2^n.*

Equivalently, let r be the root of an irreducible polynomial $f(x)$. If the degree of f is not equal to 2^n, then r is not constructible. This proves the impossibility of duplicating the cube or trisecting an arbitrary angle. In the first case, $\sqrt[3]{2}$ is the root of $x^3 - 2 = 0$, which is irreducible but not of degree 2^n; the same reasoning proves that arbitrary nth roots cannot be found, unless n is a power of 2. Likewise trisecting an arbitrary angle requires finding a root of $l = 3x - 4x^3$, which will in general be irreducible and not of degree 2^n.

What about the cyclotomy problem? If n is prime, the corresponding cyclotomic equation is irreducible, but if n is not a Fermat prime, then the degree of this equation is not a power of 2 and so the regular n-gon will not be constructible. Thus it is impossible to construct regular polygons of 7, 11, 13, etc. sides using only compass and straightedge.

Wantzel's Theorem alone is insufficient to prove the impossibility of squaring the circle, though it does lay the groundwork for a proof. If $\sqrt{\pi}$ is a constructible number, it must be the root of an irreducible equation of degree 2^n. In 1882 Ferdinand Lindemann (1852–1939) proved that π is transcendental: hence no equation of any degree with rational coefficients can have π as a root. Consequently squaring the circle is impossible.

REFERENCES

1. A. Aaboe, *Episodes from the Early History of Mathematics*, MAA New Mathematical Library, 1964.
2. C. F. Gauss, *Disquisitiones Arithmeticae*, Yale University Press, 1966. Translated by A. A. Clarke.
3. T. L. Heath, *A History of Greek Mathematics*, Dover, 1981.
4. T. L. Heath, *The Thirteen Books of the Elements*, Dover, 1956.
5. J. L. Lagrange, *Œuvres*, Gauthiers-Villars, Paris, 1869.
6. ———, Réflexions sur la résolution algébrique des équations, *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres* (1770) 134–215; (1771) 138–253.
7. D. E. Smith, *The Geometry of Rene Descartes*, Dover, 1954.
8. A. T. Vandermonde, Mémoire sur la Résolution des Équations, *Histoire de l'Académie Royale des Sciences*, 1771, 365–416.
9. P. Wantzel, Recherches sur les moyens de reconnaître si un Problème de Géométrie peut se résoudre avec la règle et le compas. *Journal de Mathématiques Pures et Appliquées*, 1837, 366–372.