Appendix 9D. Descent and the quadratics

9.13. Further solutions through linear algebra

We wish to determine all pairs of positive integers $a \geq b$ for which $ab + 1$ divides $a^2 + b^2$. If we have a solution with $a^2 + b^2 = k(ab + 1)$ for some positive integer k, then a is a root of the quadratic polynomial

$$x^2 - kbx + (b^2 - k).$$

If c is the other root, then $a + c = kb$, so $c = kb - a$ is another integer for which $b^2 + c^2 = k(bc + 1)$. Now $c \geq 0$ or else $bc + 1 \leq 0$, in which case $b^2 + c^2 \leq 0$, which is impossible. If $c = 0$, then, looking back at our equations, we see that $k = b^2$ and $a = b^3$. Otherwise $c > 0$ so that $b^2 - k = ac > 0$ and therefore $c = (b^2 - k)/a < b^2/b = b$.

We have proved that $0 \leq c < b$ which means that (b, c) is a smaller solution than the original solution (a, b), with the same quotient. We can iterate this map, $(a, b) \to (b, kb - a)$, to eventually descend, after a finite number of steps, to a basic solution. We only stop descending when $c = 0$, which means that k must be a square, a fact that is far from obvious in the formulation of the problem.

To obtain all solutions, we simply invert our map: Writing $k = m^2$ for some integer $m \geq 1$ we begin with the solution $(m, 0)$ to $m^2 + 0^2 = k(m \cdot 0 + 1)$ and obtain all others by iterating the map

$$(b, c) \to (kb - c, b).$$

This map is better understood through matrices: We have

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} kb - c \\ b \end{pmatrix} = \begin{pmatrix} k & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b \\ c \end{pmatrix},$$

the matrix representing a transformation of determinant 1. Therefore, if $k = m^2$, then all solutions to $a^2 + b^2 = k(ab+1)$ in non-negative integers a, b are given by

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} m^2 & -1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} m \\ 0 \end{pmatrix},$$

for some integers $m \geq 1$ and $n \geq 0$.

Exercise 9.13.1. Fix integer $k \geq 1$. Let $x_0 = 0$, $x_1 = 1$, and $x_n = kx_{n-1} - x_{n-2}$ for all $n \geq 2$. Prove that all solutions to $a^2 + b^2 = k(ab+1)$ in non-negative integers a, b with $k = m^2$ are given by $a = mx_n,\ b = mx_{n-1}$ for some integer $n \geq 1$.

Exercise 9.13.2. Prove that if A is any 2-by-2 matrix and the vector $u_n = A^n u_0$ for some given matrix u_0, then u_n satisfies a second-order linear recurrence.

Other quadratic equations can also be understood by recursions. Perhaps the most famous is the Markov equation.

9.14. The Markov equation

Here we seek all triples of positive integers x, y, z for which

$$x^2 + y^2 + z^2 = 3xyz.$$

One can find many solutions: $(1,1,1), (1,1,2), (1,2,5), (1,5,13), (2,5,29), \ldots$. If (a,b,c) is a solution, then a is a root of the quadratic, $x^2 - 3bcx + (b^2 + c^2)$, the other root being $3bc - a$, and so we obtain a new solution $(3bc - a, b, c)$. One can perform this same procedure singling out b or c instead of a, and get different solutions. For example, starting from $(1,2,5)$ one obtains the solutions $(29,2,5)$, $(1,13,5)$, and $(1,2,1)$, respectively.

If we fix one coordinate, say $z = c$, then we can get a new solution from an old one via the map

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 3c & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \end{pmatrix},$$

and then this map can be repeated arbitrarily often, as in the previous section, to obtain infinitely many solutions.

Despite knowing there are infinitely many, the solutions to the Markov equation remain mysterious. For example, one open question is to determine all of the integers that appear in a Markov triple. The first few are

$$1, 2, 5, 13, 29, 34, 89, 169, 194, 233, 433, 610, 985, 1325, \ldots.$$

It is believed that they are quite sparse.

Exercise 9.14.1.† Determine what solutions are obtained from (1,1,1) by using the maps $(x,y) \to (3y - x, y)$ and $(x,y) \to (x, 3x - y)$.

9.15. Apollonian circle packing

To my taste, the most beautiful such problem is the Apollonian circle packing problem.[5] Take three circles that touch each other (for example, take three coins

[5] Apollonius lived in Perga, 262–190 B.C.

and push them together):

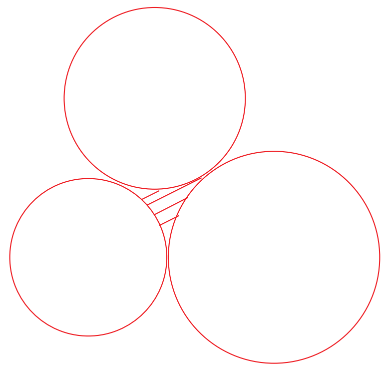


Figure 9.3. Three mutually tangent circles with a shaded crescent shape in between.

In between the circles one has a crescent-type shape (*a hyperbolic triangle*).

There are two circles that are tangent to each of these three circles: Inside that crescent shape one can inscribe a (unique) circle that touches all three of the original circles. There is also a unique circle that contains all of the original cycles and touches each of them.

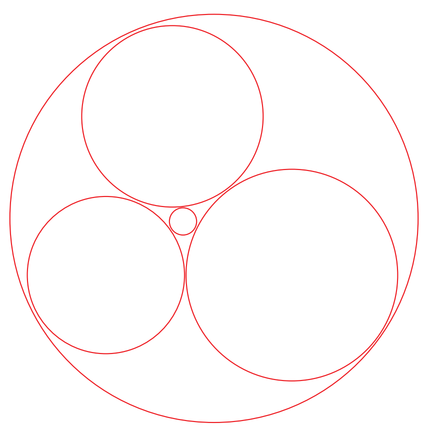


Figure 9.4. Two new circles, each tangent to all three of the old circles.

What is the relationship between the radii of the new circles and the radii of the original circles? Define the *curvature* of a circle to be C/r where r is the radius, for some appropriately selected constant $C > 0$. In 1643 Descartes, in a letter to Princess Elisabeth of Bohemia, noted that if b, c, and d are the curvatures of the original three circles, then the curvatures of the two new circles both satisfy the quadratic equation

$$2(a^2 + b^2 + c^2 + d^2) = (a + b + c + d)^2,$$

that is,

$$a^2 + b^2 + c^2 + d^2 - 2(ab + bc + cd + da + ac + bd) = 0.$$

Therefore given b, c, and d, there are two possibilities for a, the roots of the quadratic equation,

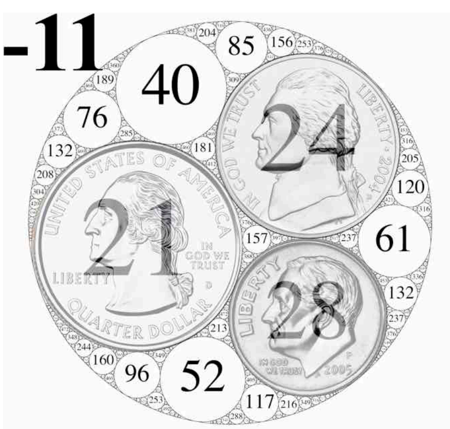
$$x^2 - 2(b + c + d)x + (b^2 + c^2 + d^2 - 2bc - 2cd - 2bd) = 0.$$

We select C so that the first three curvatures, b, c, and d are integers with $\gcd(b, c, d) = 1$. We will focus on the case that a is also an integer; for example if we start with $b = c = 2$ and $d = 3$, we have $a^2 - 14a - 15 = 0$ so that $a = -1$ or $a = 15$. Evidently $a = -1$ corresponds to the outer circle,[6] and $a = 15$ the inner one.

If we have one solution (a, b, c, d), then we also have another solution (A, b, c, d) for which

$$A = 2(b + c + d) - a.$$

Moreover if a, b, c, and d are integers, then so is A. We can iterate this (using any of the variables b, c, or d, in place of a) to obtain infinitely many Apollonian circles. These eventually *tile* the whole of the original circle, as each new circle fills in part of the crescent in between three existing circles.



In this example, we take the three most common American coins, a quarter, a nickel, and a dime, which have radii 24, 21, and 18 mms, respectively, to the nearest millimeter. In this case we define the curvature of a circle of radius r mm to be $504/r$, yielding curvatures of 21, 24, and 28, respectively. We proceed by filling in each successive crescent shape with a mutually tangent circle. What emerges is a tiling of the whole outer circle (which has curvature -11) by circles with larger and larger positive integer curvatures.[7]

There are many questions that can be asked: What integers appear as curvatures in a given packing? There are some integers that cannot appear because of congruence restrictions. For example if a, b, c, d are all odd, then all integers that arise as curvatures in this packing will be odd. The conjecture is that all sufficiently large integers that satisfy these congruence constraints, which can all be described mod 24, will appear as curvatures in the given packing. Although this is an open question, we do know that a positive proportion of the integers appear

[6] The negative sign is intriguing. The mathematics gives a negative integer which surely makes no sense; but can the mathematics lie? A more in-depth analysis indicates that the negative sign should be interpreted as meaning that whereas the *interior* of a circle usually means all points inside the circle, when we have negative curvature the interior is to be interpreted as all points *outside the circle*, going off to ∞. It is best to think of the circles as being drawn on a sphere because, for a circle on a sphere, the circle partitions the sphere into two parts, and there are two choices as to what is the interior and what is the exterior.

[7] Tiled circle defined by U.S. coins, reproduced here with the kind permission of Alex Kontorovich.

in any such packing, and that there are a surprisingly large number of circles in any packing with curvature $\leq T$, far more than T (so that many circles in the packing have the same curvature, and thus the same radius).[8] Since so many different integers appear as curvatures in any given packing, Peter Sarnak asked (and resolved) whether there are infinitely many pairs of mutually tangent circles, whose curvatures are both prime numbers, the *Apollonian twin prime conjecture*.

This last question is accessible because we see that any given solution $v = (a, b, c, d)$ is mapped to another solution by any permutation of the four elements, as well as the matrix $\begin{pmatrix} -1 & 2 & 2 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. These (linear) transformations generate a subgroup, G, of $\mathrm{SL}(4, \mathbb{Z})$,[9] and one can proceed by considering *orbits* (that is, the set $\{Av : A \in G\}$ for some starting vector v) under the actions of G.

Sarnak's approach to studying the curvatures brings us back to quadratic equations: In the American coins example, we begin with the circle of curvature 28 that is tangent to the circles of curvature 21 and 24. The circle inside the crescent that is tangent to these three circles has curvature 157. Next we determine the circle that is tangent to those of curvature 21, 24, and 157, and then the next one, always using the two circles of curvature 21 and 24. So if x_n is the curvature of the nth circle in this procedure, then $x_0 = 28$, $x_1 = 157$, and

$$x_{n+1} = 2(21 + 24 + x_n) - x_{n-1} \text{ for all } n \geq 1,$$

by Descartes's equation. We can prove by induction that

$$x_n = 45n^2 + 84n + 28 \text{ for all } n \geq 0;$$

so the circles in our circle problem, tangent to the original circles of curvatures 21 and 24, have curvature x_n for each $n \geq 0$.

Exercise 9.15.1. Suppose that you are given three mutually tangent circles A, B, and C_0 of curvatures a, b, and x_0 in an Apollonian circle packing. For each $n \geq 0$ let C_{n+1} be the circle tangent to the circles A, B, and C_n that lies in the crescent between these three circles, and let x_{n+1} be its curvature. Prove that

$$x_n = (a + b)(n^2 - n) + (x_1 - x_0)n + x_0 \text{ for all } n \geq 0.$$

Peter Sarnak developed this idea further, which we return to in appendix 12G.

Further reading on Apollonian packings

[1] Dana Mackenzie, *A tisket, a tasket, an Apollonian gasket*, American Scientist **98** (Jan–Feb 2010), 10–14.

[2] Peter Sarnak, *Integral Apollonian packings*, Amer. Math. Monthly **118** (2011), 291–306.

[8] The total number of circles in any given packing with curvature $\leq T$ is about cT^{α} where $\alpha = 1.30568\ldots$, and c is a positive constant that depends on the packing.

[9] $\mathrm{SL}(4, \mathbb{Z})$ is the set of 4-by-4 matrices of determinant 1 with integer entries.