Jan van de Craats
Korteweg-de Vries Institute for Mathematics
University of Amsterdam
j.vandecraats@uva.nl

Jan Brinkhuis
Econometric Institute, Erasmus School of Economics
Erasmus University, Rotterdam
brinkhuis@ese.eur.nl

Research

Cardioids and Morley's trisector theorem

At the end of the nineteenth century the algebraic geometer Frank Morley discovered a nice little theorem on the trisectors of a triangle, known as 'Morley's trisector theorem'. Since the beginning of the twentieth century, many proofs of this theorem have been published. In this article Jan van de Craats and Jan Brinkhuis elucidate the role of cardioids in Morley's discovery of his theorem.

The algebraic geometer Frank Morley (1860–1937) came from England to the United States in 1887. He was the editor of the *American Journal of Mathematics* from 1900 to 1921 and served as President of the American Mathematical Society from to 1919 to 1920. In 1899, while studying properties of general configurations of n lines in the Euclidean plane by means of complex numbers, Morley discovered a nice little theorem on the trisectors of a triangle. He mentioned it to friends, who spread it over the world in the form of mathematical gossip (Coxeter [1, p. 23]). The simple version of Morley's trisector theorem, as it later became known, reads as follows (see Figure 1): *in any triangle, the three points of intersection of the adjacent angle trisectors form an equilateral triangle.*

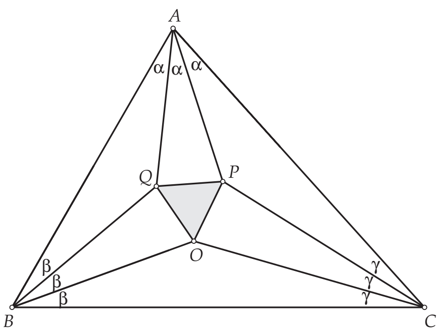

Figure 1 Morley's trisector theorem.

Morley never bothered to publish an elementary proof. In the book *Inversive Geometry* from 1933, which he wrote with his son Frank Vigor Morley, the authors, after stating this version of the theorem as a corollary of some rather intricate results on cardioids (certain heart-shaped curves, see Figure 5) simply ask: *Exercise 10: Verify this by trigonometry* [2, p. 244]. Indeed, it is straightforward to verify that, in the notation of Figure 1, $OP = 8r \sin\alpha \sin\beta \sin\gamma$, where r equals the circumradius of triangle ABC (see [3, p. 740]). This, by symmetry, proves that $OP = PQ = QO$, as desired.

Since the beginning of the twentieth century, many other proofs of Morley's theorem have been published. An article in the November 1978 issue of *The American Mathematical Monthly* by Cletus Oakley and Justine C. Baker (with supplements by Charles W. Trigg), see [3], lists no less than 150 references. Some only give a proof of the simple version of the theorem. But many proofs not only consider inner trisectors, but also their outer counterparts. In fact, if all trisectors are extended to full lines, there are precisely eighteen trisectors: six for each vertex of ABC, the outer ones making angles of $\pm\pi/3$ with the inner trisectors through the same vertex. Note that the directed angle from one line to another is determined modulo π, so trisectors are determined modulo $\pi/3$. From these eighteen trisectors many more equilateral Morley triangles can be constructed, as will be indicated in the next section. Since 1978, numerous other articles and notes on Morley's theorem have appeared in print.

A preview of Morley's analysis

At this point, the reader might like to get a preview of Morley's analysis. Let a triangle ABC be given. Extend its sides to full

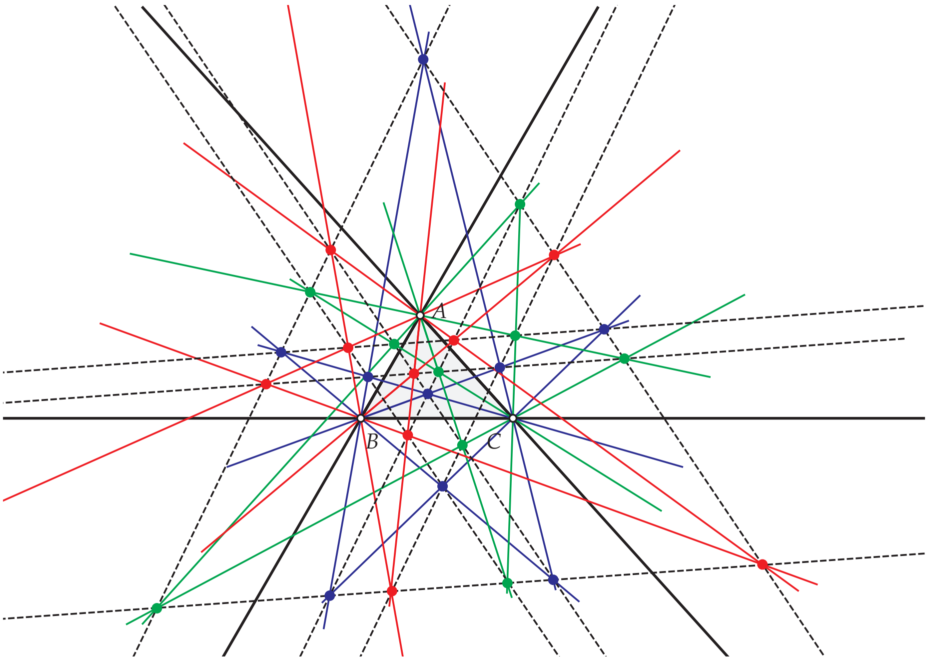

Figure 2 The full Morley trisector configuration.

lines. All eighteen trisectors of the triangle ABC from Figure 1, extended to full lines, are shown in Figure 2, colored in red, blue and green. To be more precise, the two inner trisectors through the endpoints of side AB and adjacent to AB are red, for BC this is blue and for CA this is green; each other trisector has the same color as the inner trisector through the same vertex of triangle ABC with which it makes an angle $\pm\pi/3$. Thus, six trisectors are red, six are blue and six are green. Trisectors of the same color but not at the same vertex intersect in points of the same color indicated by small dots: nine red, nine blue and nine green intersection points. The reader is invited to inspect the resulting collection of 27 colored intersection points of pairs of trisectors in Figure 2.

Morley considered cardioids that are tangent to each of the extended sides of triangle ABC. He discovered that the set of the so-called centers of these cardioids is very special: it consists of nine lines in three directions, three in each direction, intersecting in 27 points at angles $\pm\pi/3$. These lines are called the *axes* of the triangle. In Figure 2, the nine axes are drawn as dashed black lines. Furthermore, Morley proved the astonishing fact that *the 27 intersection points of the axes are precisely the 27 colored intersection points of pairs of trisectors constructed above.*

Now we are ready to explain the construction of the Morley triangles. Choose an axis in each of the three directions. This can be done in 27 ways. Each of these creates an equilateral triangle, called a *Morley triangle*. From the Morley triangles, 18 are *proper*, having vertices in three different colors, and the remaining 9 triangles turn out to be *monochromatic*, having vertices in one color only. Triangle OPQ from Figure 1 is one of the proper Morley triangles.

In the book *Inversive Geometry* [2], it takes less than four pages (§137, §138 and §140 on pp. 239–244) to elucidate the role of cardioids in the discovery of Morley's results, but this is not easy reading because of the idiosyncratic notation and style. The present paper offers self-contained explanations of this role of cardioids with full details of the proofs and with figures to illustrate the arguments. In this way, we hope to make this elegant, and almost forgotten fragment of analytic Euclidean geometry more accessible to modern readers.

Clinants, bisectors and trisectors

Throughout this paper, the points in the plane will be viewed as complex numbers in the Argand plane (the Euclidean plane co-ordinated by complex numbers). Complex numbers will be represented by lower case letters. We will always use the letter t for points on the unit circle, so $t\bar{t} = 1$, in other words, $\bar{t} = 1/t$. Sometimes, we will also use τ (the Greek letter 'tau') instead of t. Recall that, for given lines l and m, the directed angle from l to m is the angle through which a variable line has to be turned in the counterclockwise sense, in order to pass from the position l to the position m. This angle is considered modulo π. For parallel lines, this angle can be taken 0.

The *clinant* of a line is the complex number on the unit circle that has as its argument *twice* the directed angle from the real axis to the line. The clinant uniquely describes the direction of the line: two lines are parallel if and only if they have the same clinant. Furthermore, two lines are perpendicular if and only if their clinants differ by a factor -1.

The clinant will always be calculated — and arises in a natural way — as the quotient $\tau = (z-x)/(\bar{z}-\bar{x})$, where x and z are two distinct points on the line. Indeed, the vectors $z-x$ and $x-z$ both indicate the direction of the line, but in opposite sense. The quotient $(z-x)/(\bar{z}-\bar{x})$, however, remains the same if we interchange z and x. In fact, it is a complex number on the unit circle that only depends on the line, and not on the particular choice of z and x on this line. For any fixed point x on a line with clinant τ, the equation $(z-x)/(\bar{z}-\bar{x}) = \tau$, or

$$z - x = \tau(\bar{z} - \bar{x})$$

represents, for variable z, the points on the line.

Clinants are very convenient to define trisectors. First the idea is explained for the simpler case of bisectors.

Definition of bisectors. Let a, b and c be distinct points, determining lines ac and bc. Let the clinants of the lines ac and bc be t_a and t_b, respectively. Then any line through c with clinant t satisfying $t^2 = t_a t_b$ is called a *bisector* at c of the lines ac and bc.

Note that, by this definition, the directed angle at c from the line bc to the line ac

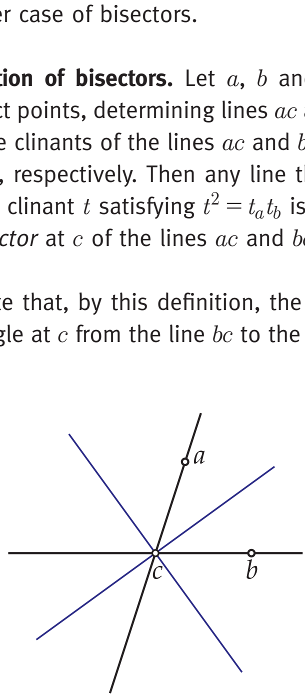

Figure 3 Bisectors of lines ac and bc.

has precisely two bisectors, intersecting at right angles, see Figure 3. Their clinants are $\pm\sqrt{t_a t_b}$. Recall that each complex number different from 0 has two complex square roots, differing by a factor -1. Multiplying the directed angle from bc to either one of the two bisectors by a factor 2, yields the directed angle from bc to ac, modulo π. This justifies the terminology *bisector*.

In relation to Morley's theorem, trisectors are important. They are defined in a similar manner.

Definition of trisectors. Let a, b and c be distinct points, determining lines ac and bc. Let the clinants of the lines ac and bc be t_a and t_b, respectively. Then any line through c with clinant t satisfying $t^3 = t_a t_b^2$ is called a *trisector* at c of the lines ac and bc, adjacent to the line bc.

See Figure 4. The intuition behind this definition is that the direction of the line bc pulls twice as hard as the direction of the line ac. Note that there are *three* trisectors adjacent to bc, intersecting at angles $\pm\pi/3$. If $\sqrt[3]{t_a t_b^2}$ is one of the (complex) cube roots of $t_a t_b^2$, the others are $\omega\sqrt[3]{t_a t_b^2}$ and $\omega^2\sqrt[3]{t_a t_b^2}$, where $\omega = e^{2\pi i/3}$. These cube roots are the clinants of the trisectors adjacent to bc.

Multiplying the directed angle from bc to any trisector adjacent to bc by a factor 3, yields the directed angle from bc to ac, modulo π. This justifies the terminology *trisector*. By the definition above, any angle determined by two distinct intersecting lines, has *six* trisectors, three adjacent to one leg and three adjacent to the other. Note that the use of the word 'adjacent' in the definition of trisectors agrees with its use in the formulation of the simple version of Morley's theorem. Also note that in our definition the case that a, b and c are collinear is not excluded: even in this case there are three trisectors at c, one of these being the line through a, b and c itself.

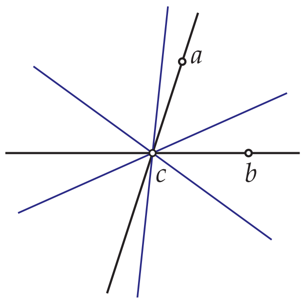

Figure 4 Trisectors adjacent to bc.

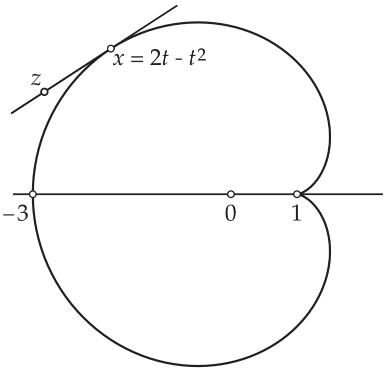

Figure 5 The standard cardioid.

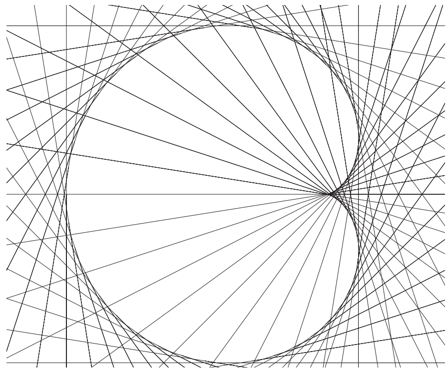

Figure 6 Tangent lines to the cardioid.

The standard cardioid and its tangent lines

When t runs through the unit circle, the equation

$$x = 2t - t^2 \tag{1}$$

is a parametric representation of a closed curve, called the *standard cardioid* (see Figure 5). More generally, a *cardioid* is a curve that is similar to the standard cardioid. Its name is derived from its heart-like shape. The point 0 is called the *center* of the cardioid. Centers of cardioids will play an important role in the sequel. The cardioid has a *cusp* when $dx/dt = 0$, which occurs for parameter value $t = 1$ at the point $x = 1$. The real axis, characterized by $z = \bar{z}$, will be called the cusp tangent line. The point $x = -3$, taken for parameter value $t = -1$, is called the *apse* of the cardioid. Its tangent line is perpendicular to the cusp tangent line.

Now we will derive an equation for the tangent line at the point on the cardioid with parameter value $t \neq 1$. Let us first find its clinant. Taking differentials from the equation $x = 2t - t^2$ yields

$$dx = 2(1-t)\,dt = 2(1-t)\,\mathrm{i}t\,d\theta$$

where the real number θ is the argument of $t = e^{\mathrm{i}\theta}$, always taken, as usual, modulo 2π. It follows that

$$\begin{aligned}\overline{dx} &= 2(1-\bar{t})\,\overline{dt} = 2(1-\frac{1}{t})(-\mathrm{i})\frac{1}{t}d\theta\\ &= \frac{1}{t^3}2(1-t)\,\mathrm{i}t\,d\theta = \frac{1}{t^3}dx.\end{aligned}$$

Therefore, $dx/\overline{dx} = t^3$. Since in terms of differentials the clinant of the tangent line equals $dx/\overline{dx}$, we have proved: *the clinant of the tangent line at the point with parameter value t of the standard cardioid $x = 2t - t^2$ equals t^3*. Moreover, for any point z on the tangent line, we have

$$t^3 = \frac{z - x}{\bar{z} - \bar{x}} = \frac{z - (2t - t^2)}{\bar{z} - (2\bar{t} - \bar{t}^2)}$$

which, using $\bar{t} = 1/t$, can be simplified to

$$z - 3t + 3t^2 - \bar{z}t^3 = 0. \tag{2}$$

This is the promised equation for the tangent line for parameter value t. It also holds for $t = 1$: the tangent line at the cusp is $z - \bar{z} = 0$ with clinant $t^3 = 1$.

As above, we will use the Greek letter ω ('omega') to denote the complex number $\omega = e^{2\pi i/3}$. Note that the tangent lines to the cardioid for parameter values t, ωt and $\omega^2 t$ are parallel. Indeed, their clinants are t^3, $(\omega t)^3$ and $(\omega^2 t)^3$, and so are all equal, as $\omega^3 = 1$.

Figure 6 shows the cardioid as the envelope of its tangent lines. The curve divides the plane into two regions: an outer region and an inner region. For each point z in the outer region, there are precisely three distinct points on the cardioid for which the tangent line runs through z. This is obvious from Figure 6, but it can also be explained as follows. Note that for any solution $t = u$ of the cubic equation (2), conjugating the equation shows that also $t = 1/\bar{u}$ is a solution. Since $u = 1/\bar{u}$ precisely when u is on the unit circle, (2) has either one or three solutions on the unit circle, counted with multiplicity. Multiple roots can be found by differentiating (2) with respect to t and conjugating, yielding $z = 2t - t^2$, so they occur precisely if z is on the cardioid. If z is in the outer region of the cardioid, then there are at least two distinct points on the cardioid with tangent lines through z. Hence, then there must always be a third point on the cardioid with tangent line through z.

Taking, as in Figure 7, a point z in the outer region of the cardioid, the cubic equation (2) has three distinct solutions t_1, t_2 and t_3, all situated on the unit circle.

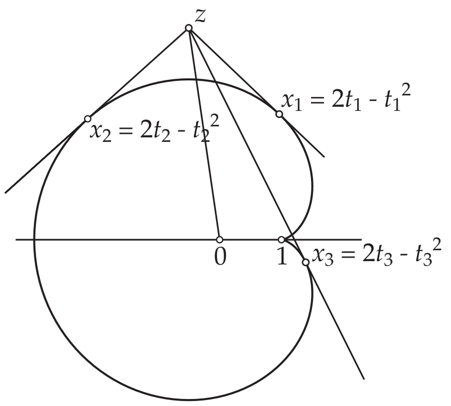

Figure 7 Three tangent lines through z.

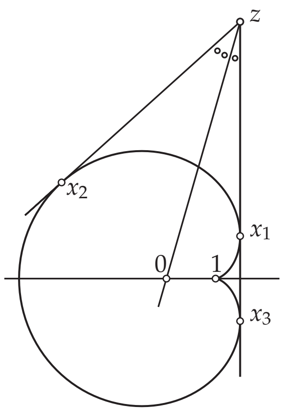

Figure 8 z on the double tangent line.

After dividing (2) by $-\bar{z}$, we see that the constant term is $-z/\bar{z}$, so

$$z/\bar{z} = t_1 t_2 t_3. \tag{3}$$

Taking cubes gives $(z/\bar{z})^3 = t_1^3 t_2^3 t_3^3$. This can be seen as a relation between clinants of certain lines. Indeed, $z/\bar{z}$ is the clinant of the line connecting z to the center 0 of the cardioid, and t_k^3 is the clinant of the tangent line for parameter value t_k for $k = 1, 2, 3$. Therefore, we get that the cube of the clinant of the line connecting z to the center 0 of the cardioid equals the product of the clinants of the three tangent lines from z. In other words, *the clinant of the line $z0$ is one of the cube roots of the product of the clinants of the three tangent lines from z*. Note that the cube roots of a complex number different from zero differ by a factor ω^k, where $k = 0, 1$ or 2.

Equation (3) will play an important role in the next section, where the relation between tangent lines and trisectors will be explained. To help the reader to get an intuitive feeling for equation (3) in the case depicted in Figure 7, we reformulate it in terms of the arguments of the vectors $0 - z$, $x_1 - z$, $x_2 - z$ and $x_3 - z$, pointing from z to 0, x_1, x_2 and x_3, respectively. Since the argument of a direction vector of a line is, modulo π, half the argument of the clinant of the line, we get

$$\arg(0-z) = \frac{1}{3}(\arg(x_1-z) + \arg(x_2-z) + \arg(x_2-z)) \pmod{\pi/3} \tag{4}$$

so the argument of $0 - z$ equals the *arithmetic mean* of the arguments of the three complex numbers $x_k - z$ $(k = 1, 2, 3)$, modulo $\pi/3$. The reader is invited to check this in Figure 7 by verifying, using a protractor, that $\angle x_2 z 0 = \angle 0 z x_1 + \angle 0 z x_3$.

Trisectors and double lines

There is a unique line that is tangent to the standard cardioid in more than one point (see Figure 8): the double tangent line, or *double line* for short. It is vertical, so its clinant is -1. The parameter values t yielding a vertical tangent line satisfy $t^3 = -1$. The value $t = -1$ gives the tangent line at the apse, so each of the other two possibilities, $t = -\omega$ and $t = -\omega^2$, must give the double line. Indeed, substitution into (2) in both cases, using $\omega^2 + \omega + 1 = 0$, yields the same equation:

$$z + \bar{z} = 3.$$

Therefore, this equation represents the double line and the two tangency points x_3 and x_1 on the double line have parameter values $-\omega$ and $-\omega^2$, respectively.

From any point z on the double line different from the two tangency points, there is precisely one other tangent line to the cardioid. Take z on the double line, but not on the closed interval between x_1 and x_3, as in Figure 8. Then it follows from (4) that $\angle x_2 z 0 = \angle 0 z x_1 + \angle 0 z x_3 = 2\angle 0 z x_1$. Hence, the line connecting z to 0 is the *inner trisector* of the angle $\angle x_1 z x_2$ adjacent to the double line, i.e., $\angle x_1 z x_2 = 3\angle x_1 z 0$. This is a first hint of the connection between cardioids and Morley's theorem.

To further explore this connection, let $z_1 z_2 z_3$ be any triangle. From now on, we assume, as we may without loss of generality, that the standard cardioid $x = 2t - t^2$ is situated inside triangle $z_1 z_2 z_3$, touching each of its sides, and that $z_2 z_3$ is the double line, as in Figure 9. Note that this implies that the given triangle $z_1 z_2 z_3$ is drawn in a way that is different from the triangle in the Figures 1 and 2, where it was denoted ABC. However, the choice of the coloring in Figure 9 (and later figures) for triangle $z_1 z_2 z_3$ has been chosen according to the coloring in Figure 2 for triangle ABC.

The lines $z_2 0$ and $z_3 0$ (drawn in blue) are inner trisectors adjacent to $z_2 z_3$. They have been given the same color, as they are adjacent to the same side of the triangle. The other inner trisectors of the angles at z_3 and z_2 in triangle $z_1 z_2 z_3$ are drawn in green and red, respectively.

Clearly, there are precisely two other cardioids situated inside triangle $z_1 z_2 z_3$ that touch each side and have one side of the triangle as a double line. In Figure 10, the three cardioids with their accompanying inner trisectors are drawn in blue, green and red. Their centers are $c_1 = 0$, c_2 and c_3. Therefore, in order to prove the simple version of Morley's theorem, it suffices to show that the centers of the three cardioids form an equilateral triangle. However, from now on, we will aim directly at proving Morley's results in their general form. To this end, we first have to consider cardioids in general position.

Cardioids in general position

In this section, the formulas (2) and (3), the clinant interpretation of (3) and the

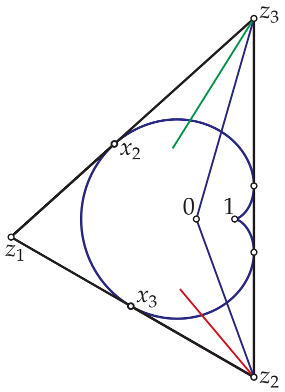

Figure 9 Inner trisectors adjacent to $z_2 z_3$.

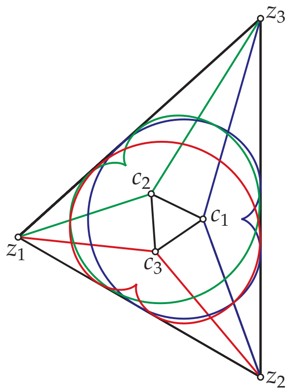

Figure 10 Inner trisectors and cardioids.

relation between centers and trisectors for the standard cardioid are extended to cardioids in general position. For given complex numbers c and a with $a \neq 0$, the equation

$$x = c + 2a\tau - \bar{a}\tau^2 \qquad (5)$$

represents, for parameter τ running through the unit circle, a cardioid with center c. Indeed, with $a = |a|e^{\mathrm{i}\alpha}$, $b = |a|e^{3\mathrm{i}\alpha}$, $\tau = e^{2\mathrm{i}\alpha}t$, equation (5) can be rewritten as

$$\frac{x-c}{b} = 2\frac{a}{b}\tau - \frac{\bar{a}}{b}\tau^2 = 2t - t^2. \qquad (6)$$

Therefore, equations (5) and (6) both represent a cardioid with center c of any size in any orientation.

To find an equation of the tangent line to a cardioid with equation (5) for parameter value τ with $\tau \neq a/\bar{a}$ (the cusp value, given by $dx/d\tau = 0$), we first determine its clinant by means of differentials, as we did for the standard cardioid. With $\tau = e^{\mathrm{i}\theta}$, we have

$$dx = (2a - 2\bar{a}\tau)\,d\tau = (2a - 2\bar{a}\tau)\,\mathrm{i}\tau d\theta$$

so

$$\begin{aligned}\overline{dx} &= (2\bar{a} - 2a\bar{\tau})\,\bar{\mathrm{i}}\bar{\tau}\,d\theta \\ &= (2\bar{a} - 2a\frac{1}{\tau})(-\mathrm{i})\frac{1}{\tau}d\theta \\ &= \frac{1}{\tau^3}(2a - 2\bar{a}\tau)\,\mathrm{i}\tau d\theta \\ &= \frac{1}{\tau^3}dx.\end{aligned}$$

Therefore, *the clinant* $dx/\overline{dx}$ *of the tangent line to the cardioid in general position* $x = c + 2a\tau - \bar{a}\tau^2$ *at the point with parameter value* τ *equals* τ^3. It follows that the tangent line is given by the equation $z - x = \tau^3(\bar{z} - \bar{x})$ or, using $x = c + 2a\tau - \bar{a}\tau^2$,

$$z - c - 2a\tau + \bar{a}\tau^2 = \tau^3(\bar{z} - \bar{c} - 2\bar{a}\frac{1}{\tau} + a\frac{1}{\tau^2})$$

which can be simplified to

$$(z - c) - 3a\tau + 3\bar{a}\tau^2 - \tau^3(\bar{z} - \bar{c}) = 0. \quad (7)$$

This is the equation of the tangent line to (5) at the point with parameter value τ. It also holds for the cusp parameter $\tau = a/\bar{a}$. Then the equation of the tangent line is $(z - c) - (a/\bar{a})^3(\bar{z} - \bar{c}) = 0$ with clinant $\tau^3 = (a/\bar{a})^3$. This completes the promised extension of equation (2) to cardioids in general position.

Special attention deserves the case that the double line of the cardioid is vertical. This occurs if and only if the complex number b in equation (6) is *real*, as this equation can be written as $x = c + b(2t - t^2)$ and, moreover, the double line of the standard cardioid is vertical. However, if b is real (and $b \neq 0$), we can write a instead of b, thus getting $x = c + 2at - at^2$, which is of the form (5) with a real and $\tau = t$ (if b is not real, this is not the case!).

If a is real, substitution of the parameter values $\tau = -\omega$ and $\tau = -\omega^2$ in (7) yields the same equation, which therefore must be the equation of the double line. Using $\omega^3 = 1$ and $\omega^2 + \omega + 1 = 0$, this equation can be simplified to

$$z + \bar{z} = c + \bar{c} + 3a \ \ (a \text{ real}).$$

We thus have proved: *if a cardioid has a vertical double line, then its equation can be taken as* $x = c + 2a\tau - a\tau^2$ *with* a *real. The tangency points of the cardioid on the double line then occur for parameter values* $\tau = -\omega$ and $\tau = -\omega^2$.

Returning to the case of a cardioid Γ in general position, take a point z in the outer region of Γ. Then there are three distinct parameter values τ_1, τ_2, τ_3 for which z is on the tangent line. Hence, the cubic equation (7) in τ has the three solutions τ_1, τ_2 and τ_3. After dividing (7) by $-(\bar{z} - \bar{c})$, we see that the constant term is $-(z - c)/(\bar{z} - \bar{c})$, so

$$(z - c)/(\bar{z} - \bar{c}) = \tau_1\tau_2\tau_3. \qquad (8)$$

This is the extension of equation (3) to cardioids in general position. It follows from (8) that *the clinant of the line connecting* z *to the center* c *is one of the cube roots of the product* $\tau_1^3\tau_2^3\tau_3^3$ *of the clinants of the three tangent lines through* z.

An important consequence of (8) is the following lemma, which describes the relation between centers and trisectors for cardioids in general position.

Lemma 1. *If* z *is a point on the double line of a cardioid with center* c*, then the line* zc *is a trisector at* z*, adjacent to the double line, of the angle from the double line to the other tangent line through* z *to the cardioid.*

Proof. Let again τ_1, τ_2, τ_3 be the parameter values of the three points on the cardioid for which the tangent line runs through the point z. Two of these values are the parameter values of the tangency points on the double line. Let these values be τ_1 and τ_3. Then $\tau_1^3 = \tau_3^3$ is the clinant of the double line, while τ_2^3 is the clinant of the other tangent line through z. It follows from (8) that the cube of the clinant of the line zc equals $\tau_2^3(\tau_1^3)^2$, so by the definition of trisectors, zc is one of the trisectors adjacent to the double line of the angle from the double line to the other tangent line through z. □

Note that Lemma 1 holds for all points z on the double line, even for the two tangency points. Then the other tangent line through z coincides with the double line, but also in that case it is true that zc is one of the three trisectors of the (zero) angle at z. Indeed, then the directed angle from the double line to zc is either $\pi/3$ or $-\pi/3$, so multiplication by a factor 3 gives 0 modulo π.

The converse of Lemma 1 also holds:

Lemma 2. *If a cardioid* Γ *is tangent to each of two distinct intersecting lines* za *and* zb *and if the center* c *of* Γ *is on one of the trisectors adjacent to* zb *of the angle at* z *from* zb *to* za*, then* zb *is the double line of* Γ.

Proof. Let t_a, t_b and t_c be the clinants of za, zb and zc, respectively. Then $t_c^3 = t_a t_b^2$. Let zd, with clinant t_d, be the third tangent line from z to Γ. It then follows from (8) that $t_c^3 = t_a t_b t_d$, so $t_b = t_d$. Hence, the lines zb and zd coincide, in other words, zb is the double line of Γ. □

Trisectors and doubly inscribed cardioids

Morley discovered his results by studying the infinitely many cardioids touching the sides, extended to full lines, of a given triangle $z_1z_2z_3$. That is, he considered all cardioids touching each one of the three given lines z_2z_3, z_3z_1 and z_1z_2. We will call such cardioids *inscribed cardioids*. Note that inscribed cardioids need not be situated inside the triangle. Morley paid special attention to inscribed cardioids for which, moreover, one of the given lines is the double line (such as, for example, the three cardioids in Figure 10 and the nine cardioids in Figure 11). We will call such cardioids *doubly inscribed cardioids*.

Lemma 3. *For each line of triangle* $z_1z_2z_3$ *there are nine doubly inscribed cardioids with that line as double line. The centers of these cardioids are precisely the nine points of intersection of pairs of trisectors adjacent to the double line at the two vertices on the double line.*

Proof. We assume without loss of generality that the chosen line is z_2z_3 and that

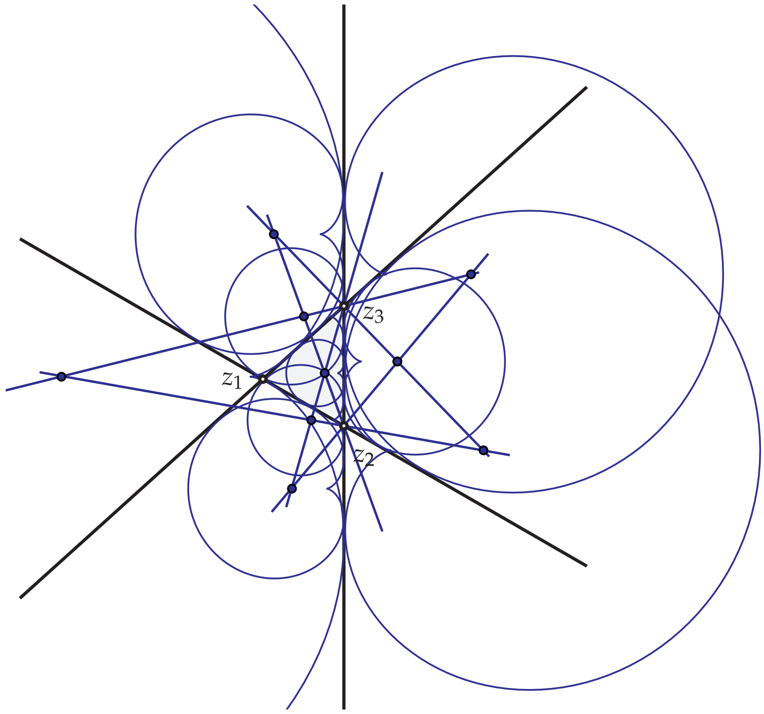

Figure 11 The 6 trisectors adjacent to z_2z_3 and the 9 inscribed cardioids with double line z_2z_3.

the doubly inscribed cardioid with z_2z_3 as double line that is situated inside triangle $z_1z_2z_3$, is the standard cardioid $x = 2t - t^2$.

We start our proof by investigating pairs of trisectors adjacent to z_2z_3, of course with one trisector at z_2 and the other at z_3. At each of the vertices z_2 and z_3 there are three trisectors adjacent to z_2z_3, as shown in Figure 11, where they have been drawn in blue. Note that no pair of these trisectors is parallel. Indeed, in the notation of Figure 9, for the *inner* trisectors z_20 and z_30, we have $0 < \angle 0z_3z_2 + \angle 0z_2z_3 < \pi/3$. Furthermore, for any other pair of trisectors adjacent to z_2z_3, we have to add integer multiples of $\pi/3$ to these angles, but this never leads to an angle sum that is an integer multiple of π, so we never get parallel trisectors. Therefore, the $2 \times 3 = 6$ trisectors give $3 \times 3 = 9$ pairs of trisectors adjacent to z_2z_3 and each of these pairs intersects in a point, which is indicated in Figure 11 by a small blue dot.

Any cardioid Γ with double line z_2z_3 is determined as soon as its center is given. This follows from the definition of a cardioid as a curve in the plane that is similar to the standard cardioid. Therefore, each of the nine intersection points c of pairs of trisectors mentioned above defines a unique cardioid Γ with center c and double line z_2z_3. Since z_2c is a trisector adjacent to z_2z_3 of the angle from z_2z_3 to z_1z_2, Γ is also tangent to z_1z_2, and since z_3c is a trisector adjacent to z_2z_3 of the angle from z_2z_3 to z_1z_3, Γ is also tangent to z_3z_1. Hence Γ is a doubly inscribed cardioid and Lemma 3 is proved. □

The axes of triangle $z_1z_2z_3$

Perhaps the most surprising aspect of Morley's results on inscribed cardioids is his discovery that the set of their centers consists of nine lines in three directions, three in each direction, intersecting at angles $\pm\pi/3$ in 27 points. These lines are the *axes* of triangle $z_1z_2z_3$. In Figure 2, the axes have been drawn as dashed black lines. The axes of a triangle are defined as follows.

Definition of axes. Let Γ, given by $x = c + 2at - \bar{a}t^2$, be an inscribed cardioid of a triangle $z_1z_2z_3$, and let t_1, t_2, t_3 be parameter values of points where Γ touches the lines of the triangle. Then the line through the center c of Γ with clinant $t_1t_2t_3$ is called an *axis* of the triangle.

Note that t_1, t_2, t_3 are not uniquely defined for a doubly inscribed cardioid Γ. In that case there are two possibilities for the parameter value of the point where Γ touches the double line, giving two axes through its center.

Lemma 4. *The clinant of any axis l of a triangle $z_1z_2z_3$ is a cube root of the product of the clinants of its sides. Each point of l is the center of an inscribed cardioid.*

Proof. The first statement follows from the definition of axes and the property of cardioids that the clinant of the tangent line for parameter value t is t^3.

Let c be the center of an inscribed cardioid Γ given by $x = c + 2at - \bar{a}t^2$. Suppose that t_1, t_2 and t_3 are parameter values for which Γ touches the lines z_2z_3, z_3z_1 and z_1z_2, respectively, of triangle $z_1z_2z_3$. Let l be the axis through the center c of Γ with clinant $t_1t_2t_3$.

On account of (7), the tangent line at Γ for parameter value t is given by

$$(z - c) - 3at + 3\bar{a}t^2 - t^3(\bar{z} - \bar{c}) = 0. \quad (9)$$

For $t = t_1$, $t = t_2$, $t = t_3$ this equation represents the lines of triangle $z_1z_2z_3$.

We will prove next that every point of the axis l is the center of an inscribed cardioid. But first, define the following symmetric expressions in t_1, t_2 and t_3:

$$s_1 = t_1 + t_2 + t_3,$$
$$s_2 = t_1t_2 + t_2t_3 + t_3t_1,$$
$$s_3 = t_1t_2t_3.$$

Then

$$\overline{s_1} = \frac{1}{t_1} + \frac{1}{t_2} + \frac{1}{t_3} = \frac{s_2}{s_3}.$$

Note that $s_1 = 0$ would imply, as t_1, t_2 and t_3 are on the unit circle, that t_1, t_2 and t_3 are the vertices of an equilateral triangle. Consequently, $t_1^3 = t_2^3 = t_3^3$ would hold, so the lines of triangle $z_1z_2z_3$, with equal clinants, would be parallel, contradiction. Hence, we may suppose that $s_1 \neq 0$.

Let c_1 be any point on the axis l other than c. Then

$$(c_1 - c)/(\overline{c_1} - \bar{c}) = t_1t_2t_3 = s_3.$$

Define

$$a_1 = a - \frac{1}{3}(c_1 - c)\overline{s_1}. \quad (10)$$

We claim that $a_1 \neq 0$ holds and that the cardioid Γ_1, given by the equation $x = c_1 + 2a_1t - \overline{a_1}t^2$, is an inscribed cardioid. Let, for each parameter value t, a line be given by the equation

$$(z - c_1) - 3a_1t + 3\overline{a_1}t^2 - t^3(\bar{z} - \overline{c_1}) = 0. \quad (11)$$

If $a_1 \neq 0$, this is the equation of the tangent line of Γ_1 at the point with parameter value t. We will prove that, without any assump-

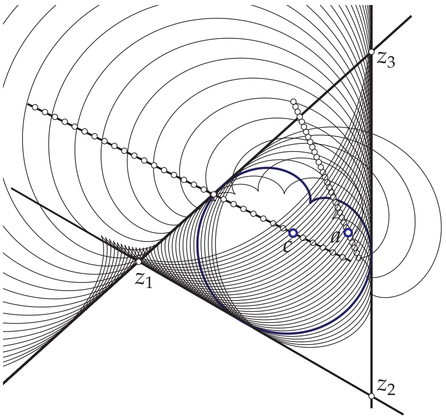

Figure 12 A sequence of inscribed cardioids with their centers on an axis of $z_1z_2z_3$, together with the sequence of corresponding collinear a-values.

tion on a_1, equation (11) for $t=t_1$, $t=t_2$, $t=t_3$ represents the lines z_2z_3, z_3z_1 and z_1z_2, respectively. Then, as a consequence, $a_1 \neq 0$ must hold, since otherwise, by (11), the point $z=c_1$ would be on each of the three lines of $z_1z_2z_3$, which is impossible.

To prove our claims, note that, with the notations $\Delta c = c_1 - c$ (so $\Delta c/\overline{\Delta c} = s_3$) and $\Delta a = a_1 - a$, we have, on account of (10)

$$\Delta a = -\frac{1}{3}\Delta c\overline{s_1} = -\frac{1}{3}\Delta c\frac{s_2}{s_3} = -\frac{1}{3}\overline{\Delta c}\, s_2$$

so

$$\overline{\Delta c}\, s_2 = -3\Delta a$$

and, since $\overline{\Delta a} = -\frac{1}{3}\overline{\Delta c}\, s_1$,

$$\overline{\Delta c}\, s_1 = -3\overline{\Delta a}.$$

Hence, for $t=t_1$, $t=t_2$ or $t=t_3$, we have

$$\begin{aligned} 0 &= \overline{\Delta c}(t-t_1)(t-t_2)(t-t_3) \\ &= \overline{\Delta c}(t^3 - t^2 s_1 + t s_2 - s_3) \\ &= \overline{\Delta c}\, t^3 + 3\overline{\Delta a}\, t^2 - 3(\Delta a)t - \Delta c. \end{aligned}$$

Expanding and rearranging the last expression yields that, for $t=t_1$, $t=t_2$ or $t=t_3$,

$$(c_1 - c) + 3(a_1 - a)t - 3(\overline{a_1} - \bar{a})t^2 - (\overline{c_1} - \bar{c})t^3 = 0. \qquad (12)$$

Subtracting equation (12) from equation (9) gives equation (11). Hence, for $t=t_1$, $t=t_2$ and $t=t_3$ equation (11), just like (9), represents the lines z_2z_3, z_3z_1 and z_1z_2, respectively. Therefore, $a_1 \neq 0$ holds and the cardioid Γ_1 is an inscribed cardioid, as desired. □

The proof above shows that, as c_1 runs over the axis l, the inscribed cardioid Γ_1 with equation $x = c_1 + 2a_1t - \overline{a_1}t^2$ that we have defined in the proof of Lemma 4 has the property that the parameters of the tangency points to the triangle $z_1z_2z_3$ do not depend on c_1. In particular, they are equal to the parameters t_1, t_2, t_3 of the tangency points of Γ to $z_1z_2z_3$. Moreover, by (10), the point a_1 also runs over a line. We will call this line the a-line.

As an illustration of Lemma 4 and the a-line, we have drawn in Figure 12 an inscribed but not doubly inscribed cardioid $x = c + 2at - \bar{a}t^2$ with center c and point a (in blue), together with a sequence of inscribed cardioids with centers on the (uniquely determined) axis through c. The corresponding part of the a-line is also marked.

Lemma 5. *If a cardioid is doubly inscribed, then its center lies on precisely two axes.*

Proof. If Γ is a doubly inscribed cardioid with, say, z_2z_3 as double line, then there are two parameter values t_1 and t_1' for which Γ touches the double line. The clinants of the corresponding axes through the center c of Γ then are $t_1t_2t_3$ and $t_1't_2t_3$. By Lemma 4, a third axis through c would only be possible if c would also be the center of an inscribed cardioid $\Gamma_1 \neq \Gamma$. On account of Lemma 3, z_2c and z_3c are trisectors adjacent to z_2z_3, so on account of Lemma 2 also Γ_1 must be doubly inscribed with z_2z_3 as double line. It follows that Γ and Γ_1, as doubly inscribed cardioids with the same center and the same double line, must be the same. Contradiction. □

Monochromatic Morley triangles

In this section, we will prove the following lemma on monochromatic Morley triangles, i.e., triangles with their sides along axes of triangle $z_1z_2z_3$ and with vertices that are centers of doubly inscribed cardioids with the same double line.

Lemma 6. *For each center c_1 of a doubly inscribed cardioid there exist two more centers c_2 and c_3 of doubly inscribed cardioids with the same double line such that $c_1c_2c_3$ is a monochromatic Morley triangle.*

Proof. First, we will define c_2 and c_3. Suppose that c_1 is the center of a doubly inscribed cardioid Γ_1 and that z_2z_3 is its double line, which, as before, we assume to be vertical. By Lemma 3, its center c_1 is the intersection of two trisectors z_3c_1 and z_2c_1 adjacent to z_2z_3. Let c_2 and c_3 be the intersections of the pairs of trisectors adjacent to z_2z_3 that are obtained from the former pair by multiplying their clinants simultaneously by ω or ω^2, respectively (see Figure 13). We will prove that the (elongated) sides of triangle $c_1c_2c_3$ are axes of triangle $z_1z_2z_3$. This will establish the lemma.

Let Γ_1 be given by the equation $x = c_1 + 2a_1t - a_1t^2$ for some *real* number a_1 (cf. section 'Cardioids in general position'). As a_1 is real, the parameter values of the tangency points of Γ_1 on the double line z_2z_3 are $t = -\omega$ and $t = -\omega^2$. Let t_2 and t_3 be the parameter values of the points of tangency of Γ_1 on z_3z_1 and z_1z_2, respectively. Then, on account of equation (8), the clinants of the trisectors z_3c_1 and z_2c_1 are $(-\omega)t_2(-\omega^2) = t_2$ and $(-\omega)t_3(-\omega^2) = t_3$, respectively. Take $t_1 = -\omega$.

Let Γ_2 be the doubly inscribed cardioid with center c_2, double line z_2z_3 and equation $x = c_2 + 2a_2\tau - a_2\tau^2$ for some real a_2. Then, as above, the parameter values for the points of tangency on the double line are $\tau = -\omega$ and $\tau = -\omega^2$. Let τ_2 and τ_3 be the parameter values of the points of tangency on z_3z_1 and z_1z_2, respectively. Then, again by (8), the parameter values τ_2 and τ_3 are also the clinants of the trisectors z_3c_2

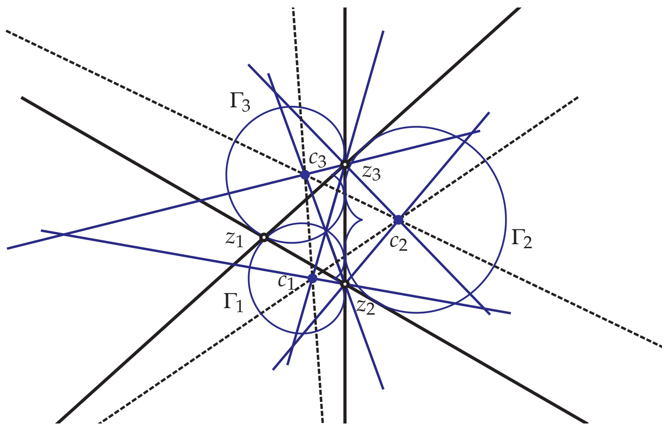

Figure 13 The monochromatic Morley triangle $c_1c_2c_3$.

and z_2c_2, which are equal to ωt_2 and ωt_3 by the definition of c_2. Therefore, $\tau_2 = \omega t_2$ and $\tau_3 = \omega t_3$. This time, we take $\tau_1 = -\omega^2 = \omega t_1$, so $\tau_k = \omega t_k$ for all $k = 1,2,3$.

The equation of the tangent line to the cardioid Γ_1 for parameter value t (cf. (7)) is given by

$$(z-c_1)-3a_1t+3a_1t^2-t^3(\bar{z}-\overline{c_1})=0. \quad (13)$$

The tangent line to Γ_2 for $\tau = \omega t$ is given by

$$(z-c_2)-3a_2\omega t+3a_2\omega^2t^2 -\omega^3t^3(\bar{z}-\overline{c_2})=0. \quad (14)$$

For $t = t_1$, $t = t_2$ or $t = t_3$, equations (13) and (14) both represent the lines z_2z_3, z_3z_1 or z_1z_2, respectively. Subtracting (13) from (14) and using $\omega^3 = 1$ yields the following cubic equation in t:

$$(c_1-c_2)+3(a_1-\omega a_2)t-3(a_1-\omega^2a_2)t^2 -(\overline{c_1}-\overline{c_2})t^3=0. \quad (15)$$

Its three solutions are t_1, t_2 and t_3, so $(c_1 - c_2)/(\overline{c_1} - \overline{c_2}) = t_1t_2t_3$. Hence, the clinant of the line c_1c_2 equals $t_1t_2t_3$, which implies that the line c_1c_2 is an axis of triangle $z_1z_2z_3$.

By cyclically permuting c_1, c_2 and c_3, it follows that also the lines c_2c_3 and c_3c_1 are axes of triangle $z_1z_2z_3$. Indeed, starting with c_2 instead of c_1, in the calculation above we only have to replace t_2 and t_3 by ωt_2 and ωt_3, respectively, while keeping $t_1 = -\omega$. Then it follows that the clinant of c_2c_3 equals $t_1(\omega t_2)(\omega t_3) = \omega^2 t_1t_2t_3$. Similarly, starting with c_3 yields $t_1(\omega^2 t_2)(\omega^2 t_3) = \omega t_1t_2t_3$ as the clinant of c_3c_1. Therefore, $c_1c_2c_3$ is a monochromatic Morley triangle. □

Figure 11 shows the nine centers of doubly inscribed cardioids with z_2z_3 as double line as small blue dots. We have shown that three of these centers form a blue monochromatic Morley triangle. From the remaining six blue centers in a similar way two more blue monochromatic Morley triangles can be constructed. The reader is invited to identify them in Figure 11.

Axes and doubly inscribed cardioids

In this section, we will prove the following lemma:

Lemma 7. *Each axis of triangle $z_1z_2z_3$ contains precisely six centers of doubly inscribed cardioids, two for each side as double line.*

Proof. Choose an inscribed cardioid and let τ_1, τ_2 and τ_3 be parameter values of its tangency points with the lines z_2z_3, z_3z_1 and z_1z_2, respectively. Let l be the axis of $z_1z_2z_3$ with clinant $\tau_1\tau_2\tau_3$ through the center of the cardioid. In the section 'The axes of triangle $z_1z_2z_3$' we have shown that for each point c on l it is possible to choose an inscribed cardioid Γ with center c and equation $x = c + 2a\tau - \bar{a}\tau^2$ such that the parameter values of the tangency points of Γ with the lines z_2z_3, z_3z_1 and z_1z_2 are τ_1, τ_2 and τ_3, respectively, independently of the choice of c on l. Moreover, using equation (10), it was shown that, as c runs over the axis l, the point a also runs over a line, which we called the a-line.

As observed in the section 'Cardioids in general position', the a-line cannot pass through the origin, since for $a = 0$ the cardioid degenerates into a point, and a point cannot touch all lines of triangle $z_1z_2z_3$. Therefore, if the center c runs over the axis l while a runs over the a-line, the argument of a runs over an open interval of length π. To be more precise, let $a = |a|e^{i\alpha}$, then we may take $\alpha_0 < \alpha < \alpha_0 + \pi$ for a certain α_0.

Now we will investigate for which points c on the axis l the chosen inscribed cardioid Γ with c as its center is doubly inscribed. In the section 'Cardioids in general position', we have rewritten the equation of Γ as $x = c + b(2t - t^2)$ using $b = |a|e^{3i\alpha}$ and $\tau = e^{2i\alpha}t$, thus showing in an explicit way its similarity to the standard cardioid $x = 2t - t^2$ (cf. equation (6)). For the standard cardioid, the two points of contact with the double line occur for parameter values $t = -\omega$ and $t = -\omega^2$, so for the cardioid Γ the two points of contact with its double line occur for parameter values $\tau = -e^{2i\alpha}\omega$ and $\tau = -e^{2i\alpha}\omega^2$. If for some α one of the two parameter values $-e^{2i\alpha}\omega$ and $-e^{2i\alpha}\omega^2$ equals τ_1, τ_2 or τ_3, then Γ is a doubly inscribed cardioid with z_2z_3, z_3z_1 or z_1z_2, respectively, as double line. To be more precise, let $\tau_k = e^{i\theta_k}$ for $k = 1,2,3$, and suppose that for a certain α we have $-e^{2i\alpha}\omega = \tau_k = e^{i\theta_k}$. Then $e^{2i\alpha} = e^{i\pi}e^{-2\pi i/3}e^{i\theta_k}$ so $\alpha = \frac{1}{2}(\pi/3 + \theta_k)$ modulo π. Similarly, if we have $-e^{2i\alpha}\omega^2 = \tau_k = e^{i\theta_k}$ then $e^{2i\alpha} = e^{i\pi}e^{-4\pi i/3}e^{i\theta_k}$ so $\alpha = \frac{1}{2}(-\pi/3 + \theta_k)$ modulo π. Therefore, as a rule, there are precisely six values of α in the open interval $\langle \alpha_0, \alpha_0 + \pi \rangle$, given by $\alpha = \frac{1}{2}(\pm\pi/3 + \theta_k)$ modulo π, corresponding to six centers c on the axis l for which Γ is doubly inscribed, two centers for each line of triangle $z_1z_2z_3$ as double line.

The only exception would occur if one of the six α-values would be equal to α_0 modulo π. In that case, for one of the lines of triangle $z_1z_2z_3$, say for z_2z_3, there would be only one center c_1 of a double inscribed cardioid with z_2z_3 as double line on l. However, on account of Lemma 6, c_1 is a vertex of a monochromatic Morley triangle $c_1c_2c_3$. The axes through c_1 then are the lines c_1c_2 and c_1c_3, and since by Lemma 5 no other axes through c_1 are possible, l must be one of these, contradicting our assumption that c_1 is the only center of a doubly inscribed cardioid with z_2z_3 as double line on l. □

The full Morley trisector configuration

In the section 'Monochromatic Morley triangles', we identified three monochromatic blue Morley triangles. Their vertices are the nine centers of doubly inscribed cardioids with z_2z_3 as double line. On account of Lemma 7, the nine axes along their sides must all be different. We claim that this set of nine axes is complete: there can be no other axes. Indeed, choose any axis. Then by Lemma 7 it contains a blue point. By Lemma 6 this point is a vertex of a monochromatic blue Morley triangle, so it lies on two of the nine axes, and by Lemma 5 it cannot lie on any other axis. Therefore the chosen axis is one of the nine.

The nine axes occur in three directions, three in each direction, intersecting in $3 \times 3 \times 3 = 27$ points, nine blue points, nine red points and nine green points, with two points of each color on each axis and two axes through each colored point. Note that two points of the same color on an axis are vertices of a unique monochromatic Morley triangle. Since there are $3 \times 3 = 9$ monochromatic Morley triangles, the remaining $27 - 9 = 18$ Morley triangles are 'proper', having vertices in three different colors.

It would be nice to produce an animation based on Figure 2 in which one could trace the nine axes, while for each center c on an axis the corresponding cardioid touching the sides of triangle $z_1z_2z_3$ were shown (cf. Figure 12). For reasons of clarity, we have abstained from adding cardioids to Figure 2.

This concludes our elaboration of the text on cardioids and Morley's full trisector configuration on pages 239–244 in [2]. We summarize our lemmas and other results in the following theorem.

Theorem. *Let $z_1z_2z_3$ be an arbitrary triangle with sides extended to full lines. Let t_0 be the product of the clinants of these lines. Then*:

1. *The set of all centers of* inscribed cardioids — *cardioids touching each of the three lines of triangle $z_1z_2z_3$ — consists of nine lines, called the* axes *of the triangle, in three directions under angles $\pm\pi/3$, three axes in each direction. The clinants of the axes are the three cube roots of t_0.*
2. *The 27 intersection points of the axes are the centers of the 27* doubly inscribed cardioids, *i.e.*, *the inscribed cardioids having one of the lines of $z_1z_2z_3$ as double line, nine cardioids for each line of $z_1z_2z_3$.*
3. *Each axis contains exactly six centers of doubly inscribed cardioids, two for each line of $z_1z_2z_3$ as double line.*
4. *For each side of the triangle, the nine centers of the doubly inscribed cardioids with this side as double line are the nine points of intersection of pairs of trisectors adjacent to this side, at different endpoints.*
   Consider the 27 equilateral triangles that are created by making a choice of three axes, one in each direction.
5. *For nine of these equilateral triangles, the vertices are intersection points of pairs of trisectors adjacent to the same side of triangle $z_1z_2z_3$, three triangles for each side of $z_1z_2z_3$.*
6. *For the remaining 18 equilateral triangles, the vertices are intersection points of pairs of trisectors adjacent to the three sides of $z_1z_2z_3$, one vertex for each side.*

It seems appropriate to end with a quotation from a letter by Frank Vigor Morley to Cletus Oakley, who had asked him about the origins of his father's theorem [3, p. 741]:

"Now my father did not lack warmth for any geometric property so simple and startling as this one. I never asked him outright the question, though it is a proper one, that Professor Oakley now asks me, namely, why at the time of the discovery my father kept his cool about promoting the 'gem' — there might have been some bit of hoo-ha if he had removed the cover and sent it to the showroom as a separate static cut stone. I think the way the theorem is presented in the book *Inversive Geometry* [2] may answer the question. Attention to the detached theorem was not, for him, to interfere with the pleasure of watching his 'mobile' of cardioids and their tangents: it was the cardioids which led him to, and provided for him the most elegant proof of, the trisector theorem. Proof and theorem were pleasing in their togetherness."

References

1 H. S. M. Coxeter, *Introduction to Geometry*, 2nd ed., Wiley, 1969.

2 Frank Morley and F. V. Morley, *Inversive Geometry*, Ginn, 1933. [Reissued by Chelsea, 1954 and by Dover Books on Mathematics, 2013.]

3 Cletus O. Oakley and Justine C. Baker, The Morley trisector theorem, *The American Mathematical Monthly* 85(9) (1978), 737–745.