On cardioids and Morley's theorem

JOHN R. SILVESTER

1. *Introduction*

Morley's trisector theorem says that the three intersections of the trisectors of the angles of a triangle, lying near the three sides respectively, form an equilateral triangle. See Figure 1. Morley discovered his theorem in 1899, and news of it quickly spread. Over the years many proofs have been published, by trigonometry or by geometry, but a simple angle-chasing argument is elusive. See [1] for a list up to 1978. Perhaps the easiest proof is that of John Conway [2], who assembles a triangle similar to the given triangle by starting with an equilateral triangle and surrounding it by triangles with very carefully chosen angles.

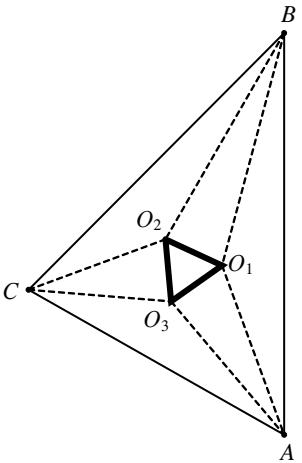


FIGURE 1

Morley's proof, however, came as a by-product of his study of cardioids, using complex numbers. He published this in the book [3] he wrote with his son. The book is a difficult read, because so much is left to the reader. Recently Jan van de Craats and Jan Brinkhuis published a much more readable version of Morley's argument [4], in which the missing details of Morley's proof are all supplied. Their article covers all the cases of Morley's theorem (outlined at the end of Section 6, below); here, in the interests of simplicity, we concentrate solely on the most elementary case, and we have also shortened one or two of the proofs.

I wish to thank Jan van de Craats and Jan Brinkhuis for helpful correspondence, and to acknowledge my debt to their article. I also wish to thank Michael Fox, whose Sketchpad animation of an inscribed cardioid bouncing around a triangle first aroused my interest in this topic.

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press

2. *Turns and clinants*

Given a line ℓ: $ax + by + c = 0$ ($a, b, c \in \mathbb{R}$. a, b not both zero), consider the parallel line through the origin, $\ell_0 : ax + by = 0$. Identifying $\mathbb{R}^2$ with the complex numbers, $\mathbb{C}$, we see that ℓ_0 meets the unit circle $|z| = 1$ in two points, $\pm t_0$. Here $t_0 = e^{i\theta}$, where θ is the angle between ℓ_0 (or ℓ) and the x-, or real, axis. We call such a number, of modulus 1, a *turn*, and we shall reserve the letter t for turns. Now we could use the turn t_0 (or $-t_0$) as a measure of the slope of ℓ, but the ambiguity of sign is a nuisance, so instead we use its (their) square, $t = t_0^2 = e^{2i\theta} = t_0/\overline{t_0}$, which we call the *clinant* of ℓ.

Now let z, z_0 be two arbitrary points of ℓ. Then $z - z_0$ is a point of ℓ_0, so that $z - z_0 = \lambda t_0$ for some $\lambda \in \mathbb{R}$. So

$$\frac{z - z_0}{\bar{z} - \bar{z}_0} = \frac{\lambda t_0}{\lambda \bar{t}_0} = t,$$

the clinant of ℓ. If we fix z_0 and regard z as a variable, then the equation ℓ of becomes $\dfrac{z - z_0}{\bar{z} - \bar{z}_0} = t$, or $z - t\bar{z} = w$, where $w = z_0 - t\bar{z}_0$. This can be obtained directly from the equation $ax + by + c = 0$ by means of the substitutions $x = \frac{1}{2}(z + \bar{z})$, $y = \frac{1}{2i}(z - \bar{z})$, yielding $t = -\left(\dfrac{a + ib}{a - ib}\right)$ and $w = -\left(\dfrac{2c}{a - ib}\right)$. Note that in the equation $z - t\bar{z} = w$, the constants t and w cannot be assigned arbitrarily to give a line; for example, if we were to put $t = w = 1$, the equation would become the nonsensical $z - \bar{z} = 1$, with the left-hand side pure imaginary but the right-hand side real. Indeed, t and w are not independent, and we leave it to the reader to show that $t = -w/\overline{w}$, and also that w is the reflection of 0 in ℓ.

Let ℓ_1, ℓ_2 be two lines with respective clinants t_1, t_2. Then

(i) $\ell_1 \parallel \ell_2$ if, and only if, $t_2 = t_1$;

(ii) $\ell_1 \perp \ell_2$ if, and only if, $t_2 = -t_1$;

(iii) ℓ_1 and ℓ_2 meet at an angle of $\pi/3$ (or $2\pi/3$) if, and only if, $t_2 = \varepsilon t_1$, where ε is a complex cube root of unity, that is, $\varepsilon^3 = 1, \varepsilon \neq 1$.

(iv) more generally, if ℓ_1 and ℓ_2 meet at X, then a rotation through θ about X takes ℓ_1 to ℓ_2 if, and only if, $t_2 = e^{2i\theta} t_1$.

Suppose three lines ℓ_1, ℓ_2, ℓ_3, with respective clinants t_1, t_2, t_3, meet at X. Then by (iv), ℓ_2 bisects the angle between ℓ_1 and ℓ_3 if, and only if, $t_3/t_1 = (t_2/t_1)^2$, that is, if, and only if, $t_1 t_3 = t_1^2$. By (ii), there are, as expected, two perpendicular solutions for ℓ_2, given ℓ_1 and ℓ_3. Again, ℓ_2 is a trisector of the angle between ℓ_1 and ℓ_3 if $t_3/t_1 = (t_2/t_1)^3$, that is, if $t_1^2 t_3 = t_2^3$. By (iii) there are three solutions for ℓ_2, given ℓ_1 and ℓ_3, which are at angles of $\pi/3$ to each other, that is, with clinants t_2, εt_2 and $\varepsilon^2 t_2$. In [4] these are called the trisectors *nearer to* ℓ_1. There are also three trisectors of the angle between ℓ_1 and ℓ_3 *nearer to* ℓ_3, so six trisectors in all.

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press

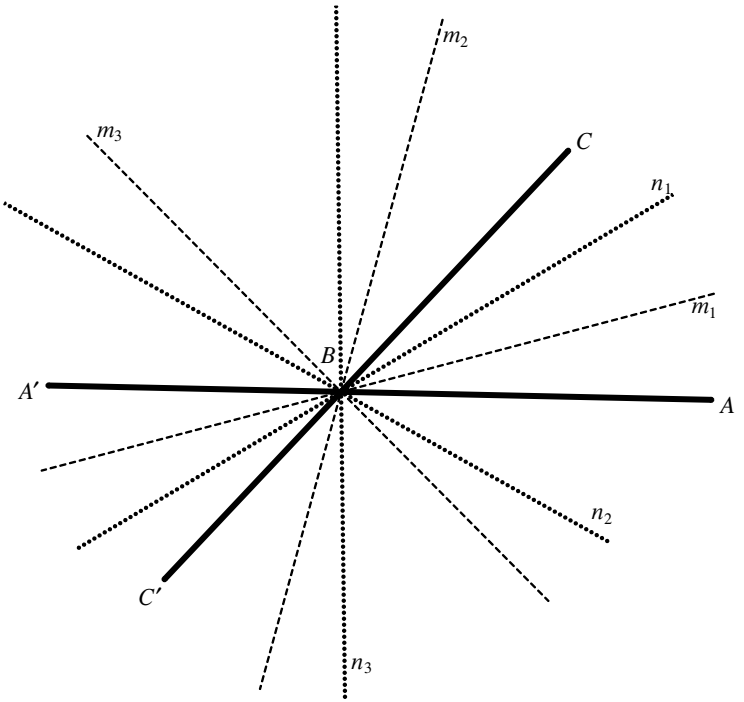


FIGURE 2

Suppose two lines AA' and CC' meet at B. See Figure 2, where the three trisectors nearer to AA' are labelled m_1, m_2 and m_3, and the three trisectors nearer to CC' are labelled n_1, n_2 and n_3. If $\angle ABC = \alpha$, that is, AA' is rotated through the angle α about B to give CC', then AA' is rotated about B through $\theta = \frac{1}{3}\alpha$ to give m_1, through $\theta + \frac{1}{3}\pi$ to give m_2, and through $\theta + \frac{2}{3}\pi$ to give m_3. Likewise, CC' is rotated through $-\theta$ about B to give n_1, through $-\theta - \frac{1}{3}\pi$ to give n_2, and through $-\theta - \frac{2}{3}\pi$ to give n_3. We need a better terminology, as m_2 is actually physically nearer to CC' than to AA', in spite of being defined as one of the trisectors 'nearer to' AA' — indeed, if $\alpha = \frac{1}{2}\pi$, then $m_2 = CC'$! So let us call m_1 the *internal trisector* of $\angle ABC$, so that n_1 is the internal trisector of $\angle CBA$. Then, m_3 is the *external trisector* of $\angle ABC$; it is also the internal trisector of $\angle A'BC$; likewise, n_3 is the external trisector of $\angle CBA$, or the internal trisector of $\angle C'BA$. Finally, if we need it, m_2 is the *mid-trisector* of $\angle ABC$ (or of $\angle A'BC$), and n_2 is the mid-trisector of $\angle CBA$ (or of $\angle C'BA$). So now the three trisectors m_1, m_2, m_3 of $\angle ABC$ nearer to AA' are simply the three trisectors (internal, mid-, and external) of $\angle ABC$, and furthermore, we can distinguish between them.

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press

3. *Cardioids*

A cardioid is the curve traced out by a fixed point on the circumference of a circle rolling, without slipping, around an equal but fixed circle. See Figure 3. Take the fixed circle S as $|z| = 1$, centre O, the *centre* of the cardioid, and the rolling circle S' with its centre initially at $z = 2$, and the fixed point P initially at $A : z = 1$. Then, when S' has moved so that its point of contact B with S is at $t = e^{i\theta}$, the centre of S' is at C: $z = 2t$, and S' has turned through an angle 2θ, so that P is now at $z = 2t - t^2$. This, then, is the equation of the *standard cardioid*, parameterized by the turn t. It is a heart-shaped curve, as indicated by its name, with a cusp at A.

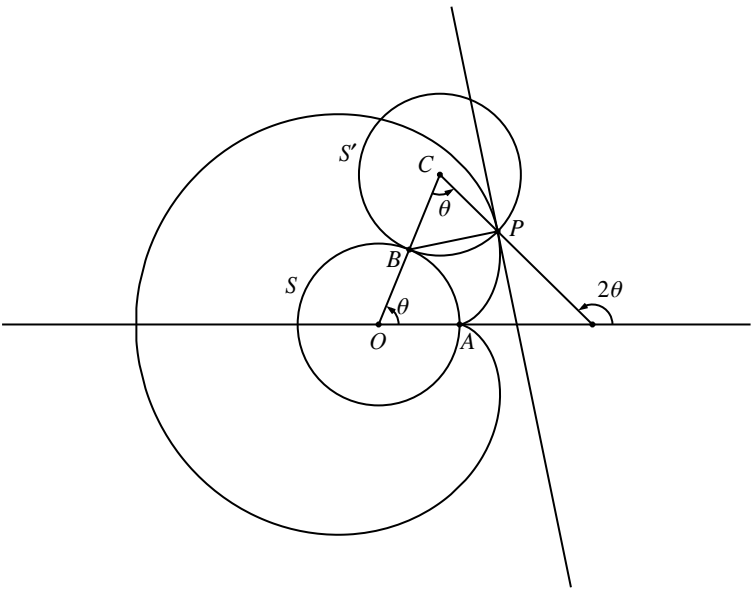


FIGURE 3

At any moment the point P is rotating instantaneously about B, so that PB is normal to the cardioid at P (when $P \neq B$, that is, $t \neq 1$). Its clinant is

$$\frac{(2t - t^2) - t}{(2\bar{t} - \bar{t}^2) - \bar{t}} = \frac{t - t^2}{t^{-1} - t^{-2}} = -t^3,$$

so the clinant of the tangent to the cardioid at P is t^3. Consequently, the equation of the tangent is

$$\frac{z - (2t - t^2)}{\bar{z} - (2t^{-1} - t^{-2})} = t^3,$$

or

$$z - 3t + 3t^2 - t^3\bar{z} = 0. \tag{1}$$

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press

This holds also for $t = 1$, as the cuspidal tangent is the real axis, $z - \bar{z} = 0$. We note in passing that the tangents at the points with parameters t, εt, $\varepsilon^2 t$ all have clinant t^3, and so are all parallel.

4. *The double tangent*

The tangent to the standard cardioid is vertical when $t^3 = -1$, and this happens when $t = -1$, $-\varepsilon$ or $-\varepsilon^2$. In the first case, the tangent is $z + \bar{z} = -6$, the tangent at the *apse* of the cardioid, $z = -3$. In the second and third cases, the tangent is $z + \bar{z} = 3$, the *double tangent* of the cardioid, touching it at the two points $R = \frac{3}{2} - \frac{\sqrt{3}}{2}i$, $R' = \frac{3}{2} + \frac{\sqrt{3}}{2}i$. See Figure 4.

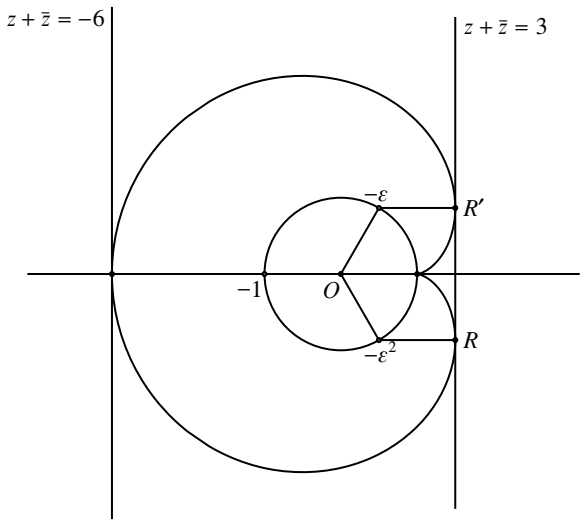


FIGURE 4

For any chosen z, (1) is a cubic in t, so there are three roots, and consequently for z outside the cardioid, there are in general three tangents to the cardioid through z. If we choose Q on the double tangent RR' as our point z, where $Q \neq R, R'$, then there is one further tangent PQ through Q, and (1) has zeros $-\varepsilon$, $-\varepsilon^2$, t, where t is the parameter of P, and $t \neq -\varepsilon$, $-\varepsilon^2$, that is, $1 - t + t^2 \neq 0$. See Figure 5. Eliminating z between (1) and $z + \bar{z} = 3$, we have $z - 3t + 3t^2 - t^2(3 - z) = 0$, or $(1 + t^3)z = 3t(1 - t + t^2)$, or $(1 + t)z = 3t$. So $(1 + t^{-1})\bar{z} = 3t^{-1}$, or $(1 + t)\bar{z} = 3$, whence $z/\bar{z} = t$, so the clinant of OQ is t.

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press

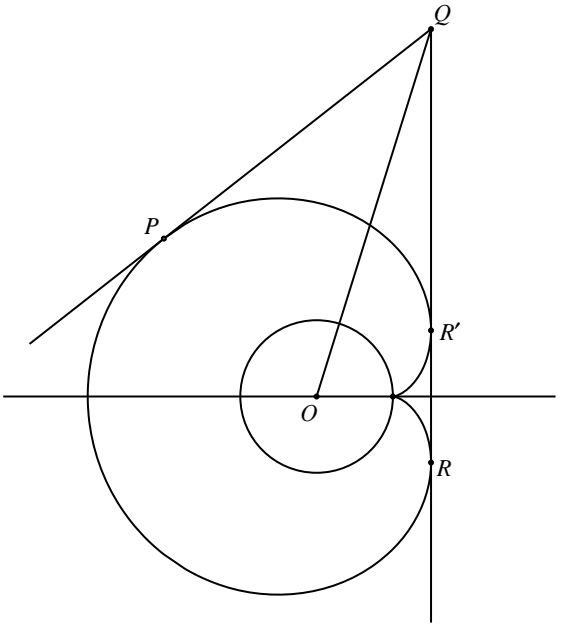


FIGURE 5

We have three lines RQ, OQ, PQ meeting at Q, with respective clinants $t_1 = -1$, $t_2 = t$ and $t_s = t^3$. Since $t_1^2 t_3 = t_2^3$, then OQ is one of the trisectors of $\angle RQP$. Which one? In Figure 6, A is the apse, and B is the mid-point of RR'. We have chosen three points Q_i on the double tangent so that the tangents P_iQ_i, $i = 1, 2, 3$, are parallel. Necessarily, one of the Q_i will be above R', one between R and R', and one below R, and we have labelled these Q_1, Q_2, Q_3 in that order. Then, with C on BQ_3 produced, from the parallels we have $\angle CQ_1P_1 = \angle CQ_2P_2 = \angle CQ_3P_3 = \alpha$, say. Let $\theta = \frac{1}{3}\alpha$. If we start with R at A, then Q is at infinity and PQ, OQ are vertical. As P moves to P_1, the tangent PQ turns through $\angle CQ_1P_1 = \alpha$, and since the line OQ is turning at one third of the rate, $\angle CQ_1O = \theta$, and thus OQ is the internal trisector of $\angle CQ_1P_1 = \angle BQ_1P_1$. Then, as P moves around to P_2, the tangent PQ turns through a further π, so that $\angle CQ_2O = \theta + \frac{\pi}{3}$, and thus OQ_2 is the mid-trisector of $\angle CQ_2P_2$. Finally, as P moves further around to P_3, the tangent OQ turns through another π, so that $\angle CQ_3O = \theta + \frac{2\pi}{3}$, and thus OQ_3 is the external trisector of $\angle CQ_3P_3$, or the internal trisector of $\angle BQ_3P_3$. Thus, for any Q not between R and R', OQ is the internal trisector of $\angle BQP$.

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press

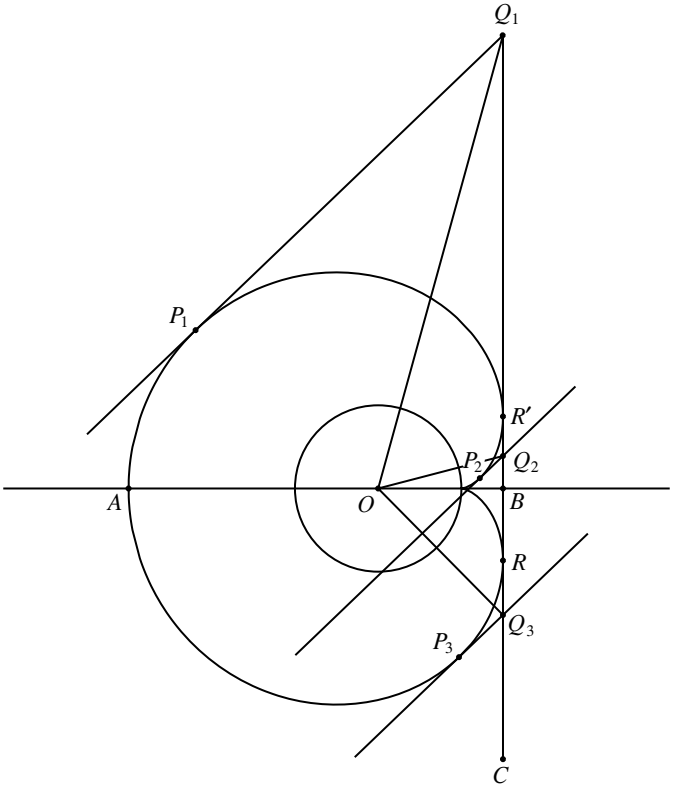


FIGURE 6

5. *The axes of a triangle*

Consider ΔABC formed by the tangents (1) to the standard cardioid, for $t = t_1, t_2, t_3$. If we replace t_I by ut_i for some turn u, then the tangents

$$z - 3ut_i + 3u^2t_i^2 - u^3t_i^3\bar{z} = 0 \tag{2}$$

form $\Delta A'B'C'$. See Figure 7. Now the clinants of AB, BC and CA are, respectively, t_1^3, t_2^3 and t_3^3; and the clinants of $A'B'$, $B'C'$ and $C'A'$ are, respectively, $u^3t_1^3$, $u^3t_2^3$ and $u^3t_3^3$.

It follows from (iv) at the end of Section 2 that the angle between AB and $A'B'$, the angle between BC and $B'C'$, and the angle between CA and $C'A'$ are all equal, so that ΔABC and $\Delta A'B'C'$ are similar. Therefore ΔABC can be obtained from $\Delta A'B'C'$ by applying a similarity $f : z \to bz + c$ (where $b \neq 0$) to (2) at $t = t_i$:

$$\frac{z - c}{b} - 3ut_i + 3u^2t_i^2 - u^3t_i^3\frac{\bar{z} - \bar{c}}{\bar{b}} = 0,$$

or

$$z - c - 3but_i + 3bu^2t_i^2 - \frac{bu^3t_i^3\bar{z}}{\bar{b}} + \frac{bu^3t_i^3\bar{c}}{\bar{b}} = 0. \tag{3}$$

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press

Now (1), at $t = t_i$, and (3) are the same line, so comparing $\bar{z}$-terms, we have $\frac{bu^3}{\bar{b}} = 1$. Subtracting (1) from (3), we now have

$$t^3\bar{c} - 3(bu^2 - 1)t^2 + 3(bu - 1)t - c = 0 \tag{4}$$

for $t = t_1, t_2, t_3$. Clearly $c \neq 0$. Put $s_1 = t_1 + t_2 + t_3$, $s_2 = t_2t_3 + t_3t_1 + t_1t_2$ and $s_3 = t_1t_2t_3$. Note that $\overline{s_1}s_3 = s_2$, and also $\overline{s_2}s_3 = s_1$.

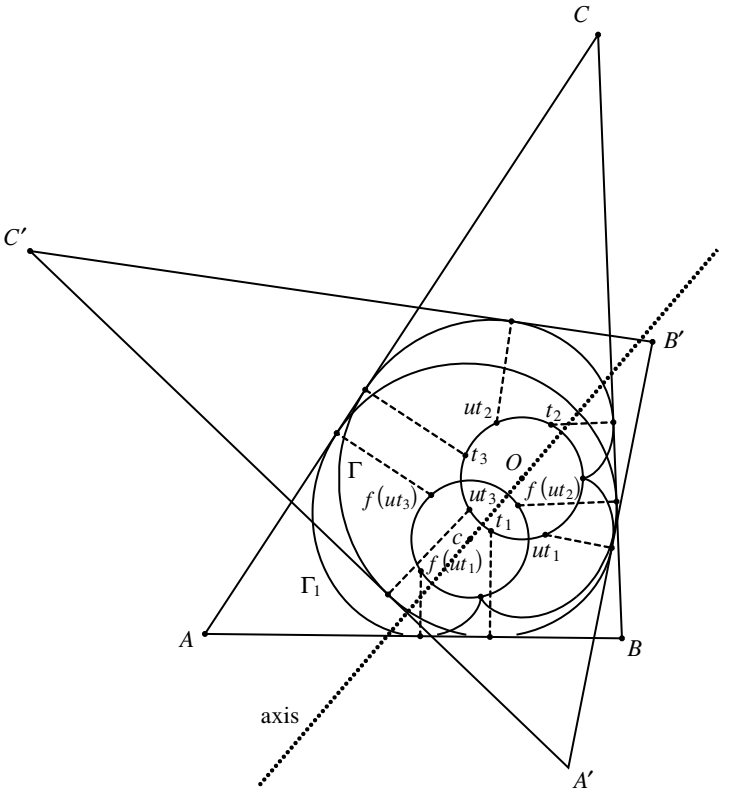


FIGURE 7

Then from (4), $\bar{c}s_1 = 3(bu^2 - 1)$, $\bar{c}s_2 = 3(bu - 1)$, and $\bar{c}s_3 = c$. From these, $s_1 \neq s_2$, since $u \neq 1$, $b \neq 0$ and $c \neq 0$. Then

$$\bar{c} = \frac{3(u - 1)}{s_1 - us_2}$$

$$\text{whence } c = \frac{3s_3(u - 1)}{s_1 - us_2}. \tag{5}$$

Applying f to the cardioid $\Gamma : z = 2t - t^2$, which is inscribed in both ΔABC and $\Delta A'B'C'$, gives the cardioid $\Gamma_1 : \frac{z - c}{b} = 2t - t^2$, a cardioid

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press

with centre c, inscribed in ΔABC. So the line joining 0, the centre of Γ, to c, the centre of Γ_1, has clinant $\dfrac{c}{\bar{c}} = s_3$, which is independent of u. We deduce that the locus of the centre c of Γ_1, as u varies, is a straight line, an *axis* of ΔABC. To see that every point of this line is the centre of a cardioid Γ_1 for suitable u, choose c_1 on this line, that is, choose $c_1 \neq 0$ with $c_1 = s_3\overline{c_1}$, whence, by (5) (with c_1, u_1 in place of c, u),

$$u_1 = \frac{c_1 s_1 + 3s_3}{c_1 s_2 + 3s_3}. \tag{6}$$

We must check that $c_1 s_2 + 3s_3 \neq 0$, and that $|u_1| = 1$. Now

$$\begin{aligned} |c_1 s_1 + 3s_3|^2 &= (c_1 s_1 + 3s_3)\left(\overline{c_1}\,\overline{s_1} + 3\overline{s_3}\right) \\ &= c_1\overline{c_1}s_1\overline{s_1} + 3c_1 s_1\overline{s_3} + 3\overline{c_1}\,\overline{s_1}s_3 + 9s_3\overline{s_3} \\ &= \overline{c_1}^{\,2} s_1 s_2 + 3\overline{c_1}s_1 + 3\overline{c_1}s_2 + 9. \end{aligned}$$

A similar calculation shows this is also equal to $|c_1 s_2 + 3s_3|^2$, so that $|c_1 s_1 + 3s_3| = |c_1 s_2 + 3s_3|$. If $c_1 s_2 + 3s_3 = 0$, then $c_1 s_1 + 3s_3 = 0$ also, whence $s_1 = s_2$, a contradiction; and it now follows that $|u_1| = 1$, as required. So the value (6) of u_1 will yield the chosen value of c_1. (Note, from (5), that $u_1 = s_1 / s_2$ will give the point at infinity on the axis.)

What about parameters? The parameter of a point of contact of Γ with $\Delta A'B'C'$ is ut_i, so the point of contact itself is $z_i = 2ut_i - u^2 t_i^2$. Then the corresponding point of contact of Γ' with ΔABC is

$$f(z_i) = b\left(2ut_i - u^2 t_i^2\right) + c = c + 2but_i - bu^2 t_i^2.$$

Put $a = bu$, so that $\overline{a} = \overline{b}\,\overline{u} = (bu^3)u^{-1} = bu^2$. So $f(z_i) = c + 2at_i - \bar{a}t_i^2$, and the parameters of the points of contact of $\Gamma : z = 2t - t^2$ and $\Gamma_1 : z = c + 2at - \bar{a}t^2$ with ΔABC are t_1, t_2, t_3 in both cases.

As an exercise, the reader might like to show that, as u varies, a moves on the line through 1 with clinant s_2 / s_1, that is, the line $s_1 z - s_2 \bar{z} = s_1 - s_2$.

6. *Morley's trisector theorem*

Given ΔABC, let the internal trisectors of $\angle ABC$ and $\angle BAC$ meet at O_1. There is a cardioid Γ_1 with centre O_1 for which AB is the double tangent, and by Section 4 this cardioid also touches BC and CA, so it is inscribed inside ΔABC. See Figure 8.

Likewise, there are two more cardioids Γ_2, Γ_3 inscribed inside ΔABC with BC, CA as respective double tangents, and centres O_2 (the intersection of the internal trisectors of $\angle BCA$ and $\angle CBA$), and O_3 (the intersection of the internal trisectors of $\angle CAB$ and $\angle ACB$), respectively. See Figure 8, again.

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press

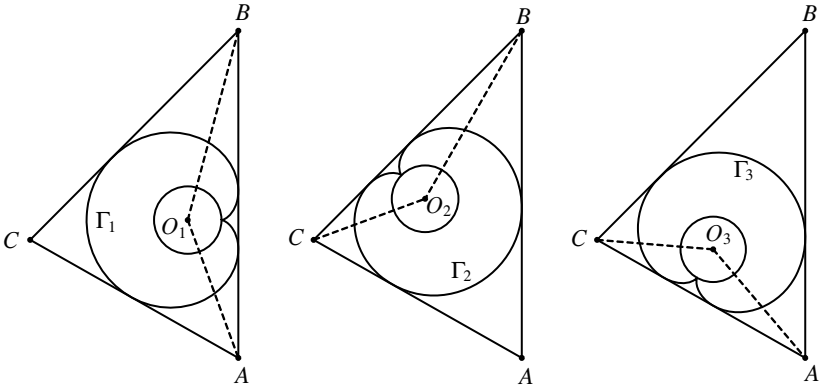


FIGURE 8

Now, looking at Figure 9, we take t_1, t_2, t_3 as the parameters of the points of contact of Γ_1 with the sides of ΔABC, where t_1 is the parameter of R. We take u, as in Section 5, so that ut_2 is the parameter of R', and Γ_1 has double contact with $B'C'$ and therefore, with f defined as in Section 5, $f(\Gamma_1) = \Gamma_2$, and $f(O_1) = O_2$. Therefore O_1O_2 is an axis of ΔABC, with clinant $t_1t_2t_3$.

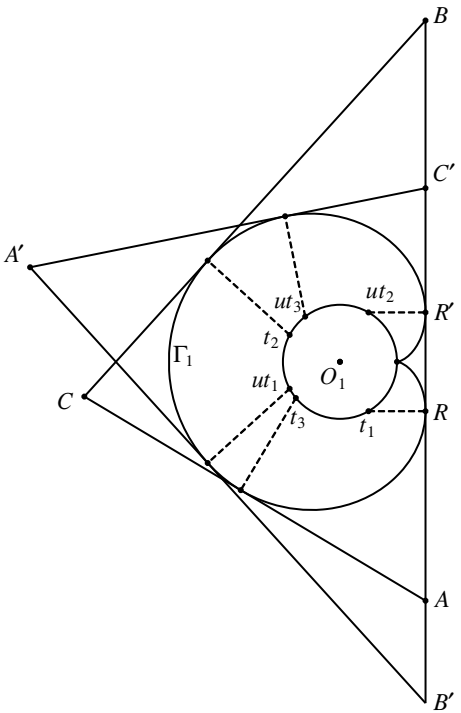


FIGURE 9

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press

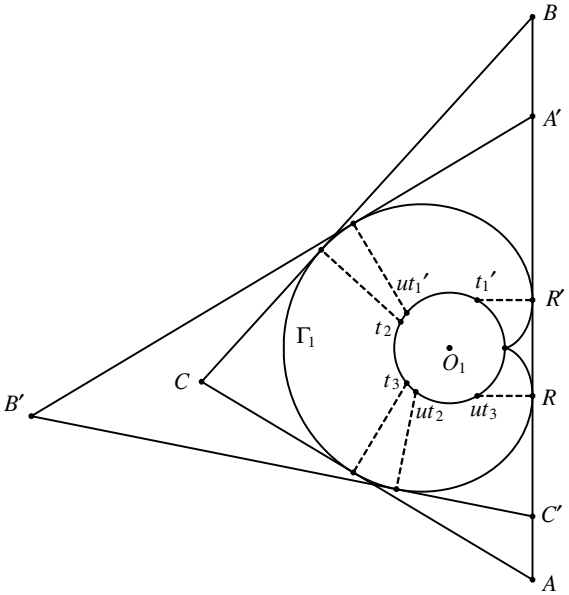


FIGURE 10

Next, looking at Figure 10, we take t_1' as the parameter of R' and t_2, t_3 as before. We take u, as in Section 5, so that ut_3 is the parameter of R, and Γ_1 has double contact with $C'A'$ and therefore, with f defined as in Section 5, $f(\Gamma_1) = \Gamma_3$, and $f(O_1) = O_3$. Therefore O_1O_3 is another axis of ΔABC, this time with clinant $t_1't_2t_3$.

But $t_1' = \varepsilon t_1$, and therefore $\angle O_2O_1O_3 = \frac{1}{3}\pi$. Similarly, $\angle O_3O_2O_1 = \frac{1}{3}\pi$, and therefore $\Delta O_1O_2O_3$ is equilateral, which is Morley's trisector theorem. See Figure 1.

The preceding result is just the tip of an iceberg. The three trisectors of $\angle ABC$ meet the three trisectors of $\angle BAC$ in nine points, each the centre of a cardioid touching the lines AB (twice), BC (once), and CA (once). Likewise there are nine centres of cardioids touching the lines AB (once), BC (twice), and CA (once); and nine centres of cardioids touching AB (once), BC (once), and CA (twice). The triangle has nine axes, two through each of the twenty seven centres just described and meeting at an angle of $\frac{1}{3}\pi$. There are six centres on each axis, and the axes are parallel in threes. For all the details, including coloured diagrams, see [4].

7. *References*

1. Cletus O. Oakley and Justine C. Baker, The Morley trisector theorem, *Amer. Math. Monthly* **85**(9) (November 1978) pp. 737-745.
2. John Conway, On Morley's trisector theorem, *Math. Intelligencer* **36**(3) (Septenber 2014) p. 3, available online at https://link.springer.com/content/pdf/10.1007/s00283-014-9463-3.pdf

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press

3. Frank Morley and F. V. Morley, *Inversive geometry*, Bell, London (1933) pp. 239-244.
4. Jan van de Craats and Jan Brinkhuis, *Cardioids and Morley's trisector theorem*, Nieuw Arch. Wiskd. (5) 18, No. 1 (March 2017) pp. 25-33, available online at http://www.nieuwarchief.nl/serie5/pdf/naw5-2017-18-1-025.pdf

10.1017/mag.2021.6
JOHN R. SILVESTER
© The Mathematical Association 2021
Department of Mathematics, King's College, Strand, London WC2R 2LS
e-mail: *jrs@kcl.ac.uk*

The answers to the *Nemo* page from November 2020 on learning algebra were:

Jack London	Martin Eden	Chapter 12
Stephen Leacock	Literary Lapses: how to avoid getting married	
Edith Nesbitt	The Book of Dragons	(The Island of the Nine Whirlpools)
DH Lawrence	The Rainbow	Chapter 12
Marcel Proust	Sodome et Gormorrhe	(Cities of the Plain)
Samuel Johnson	Miscellanies	Vol I, p 301

Congratulations to Martin Lukarevski on tracking all of these down. The theme this time is parallelograms. The quotations are to be identified by reference to author and work. Solutions are invited to the Editor by 23rd May 2021

1. He stood before me, a thin parallelogram of black with a mosaic of white about the throat.

2. What Grecian, what Roman, what Turk, what Italian would endure, or would ever have endured, to use a room with a monstrous cantle in the form of a parallelogram cut sheerly out of one corner of it ?

3. Every passenger train along the lines on a great parallelogram between Southampton, Manchester, Brighton and Horsham, travelled with locked doors.

4. Stephen lay watching the Great Bear; Elfride was regarding a monotonous parallelogram of window blind.

5. The snowy parallelogram was here presented and accepted with a mutual bow.

6. I love the parallelogram of grass marked out with chalk or tape, as if its limits were the frontiers of my sacred Fatherland, the four seas of Britain.

https://doi.org/10.1017/mag.2021.6 Published online by Cambridge University Press