Aaron Lin, Mehdi Makhul, Hossein Nassajian Mojarrad, Josef Schicho, Konrad Swanepoel and Frank de Zeeuw

On sets defining few ordinary circles

Article (Published version) (Refereed)

Discrete Comput Geom
DOI 10.1007/s00454-017-9885-8

On Sets Defining Few Ordinary Circles

Aaron Lin[1] · Mehdi Makhul[2] · Hossein Nassajian Mojarrad[3] · Josef Schicho[2] · Konrad Swanepoel[1] · Frank de Zeeuw[3]

Received: 14 September 2016 / Revised: 14 February 2017 / Accepted: 3 March 2017

Abstract An ordinary circle of a set P of n points in the plane is defined as a circle that contains exactly three points of P. We show that if P is not contained in a line or a circle, then P spans at least $n^2/4 - O(n)$ ordinary circles. Moreover, we determine the exact minimum number of ordinary circles for all sufficiently large n and describe all point sets that come close to this minimum. We also consider the circle variant of the orchard problem. We prove that P spans at most $n^3/24 - O(n^2)$ circles passing through exactly four points of P. Here we determine the exact maximum and the extremal configurations for all sufficiently large n. These results are based on the

Editor in Charge: János Pach

Aaron Lin
aaronlinhk@gmail.com

Mehdi Makhul
mmakhul@risc.uni-linz.ac.at

Hossein Nassajian Mojarrad
hossein.mojarrad@epfl.ch

Josef Schicho
Josef.Schicho@risc.jku.at

Konrad Swanepoel
k.swanepoel@lse.ac.uk

Frank de Zeeuw
fdezeeuw@gmail.com

1 Department of Mathematics, London School of Economics and Political Science, London WC2A 2AE, United Kingdom

2 Research Institute for Symbolic Computation, Johannes Kepler University, 4040 Linz, Austria

3 Department of Mathematics, EPFL, 1015 Lausanne, Switzerland

Published online: 20 March 2017

following structure theorem. If n is sufficiently large depending on K, and P is a set of n points spanning at most Kn^2 ordinary circles, then all but $O(K)$ points of P lie on an algebraic curve of degree at most four. Our proofs rely on a recent result of Green and Tao on ordinary lines, combined with circular inversion and some classical results regarding algebraic curves.

Keywords Ordinary circles · Circular curves · Inversion · Sylvester–Gallai · Green–Tao

Mathematics Subject Classification 52C35 (52C10 14N10 14H50)

1 Introduction

1.1 Background

The classical Sylvester–Gallai theorem states that any finite non-collinear point set in $\mathbb{R}^2$ spans at least one *ordinary line* (a line containing exactly two of the points). A more sophisticated statement is the so-called Dirac–Motzkin conjecture, according to which every non-collinear set of $n > 13$ points in $\mathbb{R}^2$ determines at least $n/2$ ordinary lines. This conjecture was proved by Green and Tao [13] for all sufficiently large n. Their proof was based on a structure theorem, which roughly states that any point set with a linear number of ordinary lines must lie mostly on a cubic curve (see Theorem 5.1 for a precise statement).

It is natural to ask the corresponding question for *ordinary circles* (circles that contain exactly three of the given points); see for instance [8, Sect. 7.2] or [17, Chap. 6]. Elliott [12] introduced this question in 1967, and proved that any n points, not all on a line or a circle, determine at least $2n^2/63 - O(n)$ ordinary circles. (Throughout the paper, by $O(f(n))$ we mean a function $g(n)$ such that $0 \leqslant g(n) \leqslant Cf(n)$ for some constant $C > 0$ and all sufficiently large n. Thus, $-O(n)$ is a function $g(n)$ satisfying $-Cn \leqslant g(n) \leqslant 0$ for sufficiently large n.) He suggested, cautiously, that the optimal bound is $n^2/6 - O(n)$. Elliott's result was improved by Bálintová and Bálint [1, Rem., p. 288] to $11n^2/247 - O(n)$, and Zhang [26] obtained $n^2/18 - O(n)$. Zhang also gave constructions of point sets on two concentric circles with $n^2/4 - O(n)$ ordinary circles.

We will use the results of Green and Tao to prove that $n^2/4 - O(n)$ is asymptotically the right answer, thus disproving the bound suggested by Elliott [12]. Nassajian Mojarrad and de Zeeuw proved this bound in an earlier preprint [19], which is subsumed by this paper, and will not be published independently. We will find the exact minimum number of ordinary circles, for sufficiently large n, and we will determine which configurations attain or come close to that minimum. We make no attempt to specify the threshold implicit in the phrase 'for sufficiently large n'; any improvement would depend on an improvement of the threshold in the result of Green and Tao [13]. For small n, the bound $\frac{1}{9}\binom{n}{2}$ due to Zhang [26] remains the best known lower bound on the number of ordinary circles.

Green and Tao [13] also solved (for large n) the even older *orchard problem*, which asks for the exact maximum number of lines passing through exactly three points of a set of n points in the plane. We refer to [13] for the history of this problem. The upper bound $\frac{1}{3}\binom{n}{2}$ is easily proved by double counting, but it is not the exact maximum. Using group laws on certain cubic curves, one can construct n non-collinear points with $\lfloor n(n-3)/6+1 \rfloor$ 3-point lines, and Green and Tao [13] proved (for large n) that this is optimal. This does not follow directly from the Dirac–Motzkin conjecture, but it does follow from the above-mentioned structure theorem of Green and Tao for sets with few ordinary lines (Theorem 5.1).

The analogous orchard problem for circles asks for the maximum number of circles passing through exactly four points from a set of n points. As far as we know, this question has not been asked before. We determine the exact maximum and the extremal sets for all sufficiently large n.

Although we do not consider other related problems, we remark that similar questions have been asked for ordinary conics [7,10,25], ordinary planes [2], and ordinary hyperplanes [3].

1.2 Results

Our first main result concerns the minimum number of ordinary circles spanned by a set of n points, not all lying on a line or a circle, and the structure of sets of points that come close to the minimum. The first part of the theorem solves Problem 6 in [8, Sect. 7.2].

Theorem 1.1 (Ordinary circles)

(i) *If n is sufficiently large, the minimum number of ordinary circles determined by n points in $\mathbb{R}^2$, not all on a line or a circle, equals*

$$\begin{cases} \frac{1}{4}n^2 - \frac{3}{2}n & \text{if } n \equiv 0 \pmod 4, \\ \frac{1}{4}n^2 - \frac{3}{4}n + \frac{1}{2} & \text{if } n \equiv 1 \pmod 4, \\ \frac{1}{4}n^2 - n & \text{if } n \equiv 2 \pmod 4, \\ \frac{1}{4}n^2 - \frac{5}{4}n + \frac{3}{2} & \text{if } n \equiv 3 \pmod 4. \end{cases}$$

(ii) *Let C be sufficiently large. If a set P of n points in $\mathbb{R}^2$ determines fewer than $n^2/2 - Cn$ ordinary circles, then P lies on the disjoint union of two circles, or the disjoint union of a line and a circle.*

In Sect. 4, we will describe constructions that meet the lower bound in part (i) of Theorem 1.1. For even n, the bound in part (i) is attained by certain constructions on the disjoint union of two circles, while for odd n, the bound is attained by constructions on the disjoint union of a line and a circle. The main tools in our proof are circle

inversion and the structure theorem of Green and Tao [13] for sets with few ordinary lines, together with some classical results about algebraic curves and their interaction with inversion.

Let us define a *generalised circle* to be either a circle or a line. Because inversion maps circles and lines to circles and lines, it turns out that in our proof it is more natural to work with generalised circles. Alternatively, we could phrase our results in terms of the *inversive plane* (or *Riemann sphere*) $\mathbb{R}^2 \cup \{\infty\}$, where ∞ is a single point that lies on all lines, which can then also be considered as circles. Yet another equivalent view would be to identify the inversive plane with the sphere $\mathbb{S}^2$ via stereographic projection, and consider circles on $\mathbb{S}^2$, which are in bijection with generalised circles. All our statements about generalised circles in $\mathbb{R}^2$ could thus be formulated in terms of circles in $\mathbb{R}^2 \cup \{\infty\}$ or on $\mathbb{S}^2$.

We define an *ordinary generalised circle* to be one that contains three points from a given set. Our proof of Theorem 1.1 proceeds via an analogous theorem for ordinary generalised circles, which turns out to be somewhat easier to obtain.

Theorem 1.2 (Ordinary generalised circles)

(i) *If n is sufficiently large, the minimum number of ordinary generalised circles determined by n points in $\mathbb{R}^2$, not all on a generalised circle, equals*

$$\begin{cases} \frac{1}{4}n^2 - n & \text{if } n \equiv 0 \pmod 4, \\ \frac{3}{8}n^2 - n + \frac{5}{8} & \text{if } n \equiv 1 \pmod 4, \\ \frac{1}{4}n^2 - \frac{1}{2}n & \text{if } n \equiv 2 \pmod 4, \\ \frac{3}{8}n^2 - \frac{3}{2}n + \frac{17}{8} & \text{if } n \equiv 3 \pmod 4. \end{cases}$$

(ii) *Let C be sufficiently large. If a set P of n points in $\mathbb{R}^2$ determines fewer than $n^2/2 - Cn$ ordinary generalised circles, then P lies on two disjoint generalised circles.*

We also solve the analogue of the orchard problem for circles (for sufficiently large n). We define a 4-*point* (*generalised*) *circle* to be a (generalised) circle that passes through exactly four points of a given set of n points. The 'circular cubics' in part (ii) will be defined in Sect. 2.

Theorem 1.3 (4-Point generalised circles)

(i) *If n is sufficiently large, the maximum number of 4-point generalised circles determined by a set of n points in $\mathbb{R}^2$ is equal to*

$$\begin{cases} \frac{1}{24}n^3 - \frac{1}{4}n^2 + \frac{5}{6}n - 2 & \text{if } n \equiv 0 \pmod 8, \\ \frac{1}{24}n^3 - \frac{1}{4}n^2 + \frac{11}{24}n - \frac{1}{4} & \text{if } n \equiv 1, 3, 5, 7 \pmod 8, \\ \frac{1}{24}n^3 - \frac{1}{4}n^2 + \frac{7}{12}n - \frac{1}{2} & \text{if } n \equiv 2, 6 \pmod 8, \\ \frac{1}{24}n^3 - \frac{1}{4}n^2 + \frac{5}{6}n - 1 & \text{if } n \equiv 4 \pmod 8. \end{cases}$$

(ii) *Let C be sufficiently large. If a set P of n points in $\mathbb{R}^2$ determines more than $n^3/24 - 7n^2/24 + Cn$ 4-point generalised circles, then up to inversions, P lies on an ellipse or a smooth circular cubic.*

Theorem 1.3 remains true if we replace 'generalised circles' by 'circles'. This is because we can apply an inversion to any set of n points with a maximum number of generalised circles in such a way that all straight-line generalised circles become circles. Therefore, the maximum is also attained by circles only.

The proofs of the above theorems are based on the following structure theorems in the style of Green and Tao [13]. The first gives a rough picture, by stating that a point set with relatively few ordinary generalised circles must lie on a bicircular quartic, a specific type of algebraic curve of degree four that we introduce in Sect. 2.

Theorem 1.4 (Weak structure theorem) *Let $K > 0$ and let n be sufficiently large depending on K. If a set P of n points in $\mathbb{R}^2$ spans at most Kn^2 ordinary generalised circles, then all but at most $O(K)$ points of P lie on a bicircular quartic.*

Ball [2] concurrently obtained a similar result as a consequence of a structure theorem for ordinary planes in $\mathbb{R}^3$. He shows that n points with $O(n^{2+1/6})$ ordinary circles must lie mostly on a quartic curve.

We define bicircular quartics in Sect. 2; they can be reducible, so in Theorem 1.4 the set P may also lie mostly on a lower-degree curve contained in a bicircular quartic. Our proof actually gives a more precise list of possibilities. The curve that P mostly lies on can be: a line; a circle; an ellipse; a line and a disjoint circle; two disjoint circles; a circular cubic that is acnodal or smooth; or a bicircular quartic that is an inverse of an acnodal or smooth circular cubic.

A more precise characterisation of the possible configurations with few ordinary generalised circles is given in the following theorem. The group structures referred to in the theorem are defined in Sect. 3; the circular points at infinity (α and β) referred to in Case (iii) are introduced in Sect. 2; and the 'aligned' and 'offset' double polygons are defined in Sect. 4.

Theorem 1.5 (Strong structure theorem) *Let $K > 0$ and let n be sufficiently large depending on K. If a set P of n points in $\mathbb{R}^2$ spans at most Kn^2 ordinary generalised circles, then up to inversions and similarities, P differs in at most $O(K)$ points from a configuration of one of the following types*:

(i) *a subset of a line*;

(ii) *a subgroup of an ellipse*;
(iii) *a coset $H \oplus x$ of a subgroup H of a smooth circular cubic, for some x such that $4x \in H \oplus \alpha \oplus \beta$, where α and β are the two circular points at infinity*;
(iv) *a double polygon that is 'aligned' or 'offset'.*

Conversely, every set of these types defines at most $O(Kn^2)$ ordinary generalised circles.

In Sect. 2, we carefully introduce circular cubics and bicircular quartics, and show their connection to inversion. In Sect. 3, we define group laws on these curves, which help us construct point sets with few ordinary (generalised) circles in Sect. 4. In Sect. 5, which forms the core of our proof, we derive Theorems 1.4 and 1.5 from the structure theorem of Green and Tao [13]. In Sect. 6, we combine the structure theorems with our analysis of the constructions from Sect. 4 to establish the precise statements in Theorems 1.1, 1.2, and 1.3.

2 Circular Curves and Inversion

The key tool in our proof is *circle inversion*, as it was in the earlier papers [1,12,26] on the ordinary circles problem; the first to use circle inversion in Sylvester–Gallai problems was Motzkin [18]. The simple reason for the relevance of circle inversion is that if we invert in a point of the given set, an ordinary circle through that point is turned into an ordinary line. Thus we can use results on ordinary lines, like those of Green and Tao [13], to deduce results about ordinary circles. To do this successfully, we need a thorough understanding of the effect of inversion on algebraic curves, and in particular we need to introduce the special class of *circular curves*.

2.1 Circular Curves and Circular Degree

In this subsection, we work in the real projective plane $\mathbb{RP}^2$, and partly in the complex projective plane $\mathbb{CP}^2$. See for instance [22, App. A] for an appropriate introduction to projective geometry. We use the homogeneous coordinates $[x : y : z]$ for points in $\mathbb{RP}^2$ or $\mathbb{CP}^2$, and we think of the line with equation $z = 0$ as the line at infinity. An affine algebraic curve in $\mathbb{R}^2$, defined by a polynomial $f \in \mathbb{R}[x, y]$, can be naturally extended to a projective algebraic curve, by taking the zero set of the homogenisation of f. This curve in $\mathbb{RP}^2$ then extends to $\mathbb{CP}^2$, by taking the complex zero set of the homogenised polynomial.

We define the *circular points* to be the points

$$\alpha = [i : 1 : 0], \quad \beta = [-i : 1 : 0]$$

on the line at infinity in $\mathbb{CP}^2$. The circular points play a key role in this paper, due to the fact that every circle contains both circular points. Moreover, any real conic containing α and β is either a circle, or a union of a line with the line at infinity. We could thus consider a generalised circle to be a conic that contains both circular points.

Definition 2.1 An algebraic curve in $\mathbb{RP}^2$ is *circular* if it contains α and β. For $k \geqslant 2$, an algebraic curve in $\mathbb{RP}^2$ is *k-circular* if it has singularities of multiplicity at least k at both α and β.

A classical reference for circular curves is Johnson [16], while a more modern one is Werner [24]. Let us make the definition more explicit in three concrete cases.

A *generalised circle* is an algebraic curve of degree two that contains α and β; equivalently, it is a curve in $\mathbb{RP}^2$ defined by a homogeneous polynomial of the form

$$t(x^2 + y^2) + \ell(x, y, z)z,$$

where $t \in \mathbb{R}$, and $\ell \in \mathbb{R}[x, y, z]$ is a non-trivial linear form. If $t \neq 0$, then the curve is a circle, while if $t = 0$, the curve is the union of a line with the line at infinity.

A *circular cubic* is an algebraic curve of degree three that contains α and β; equivalently, it is any curve in $\mathbb{RP}^2$ defined by a homogeneous polynomial of the form

$$(ux + vy)(x^2 + y^2) + q(x, y, z)z, \tag{1}$$

where $u, v \in \mathbb{R}$, and $q \in \mathbb{R}[x, y, z]$ is a non-trivial quadratic homogeneous polynomial. Note that we do not require a circular cubic to be irreducible or smooth. For instance, the union of a line and a circle is a circular cubic, and so is the union of any conic with the line at infinity (take $u = v = 0$ in (1)).

A *bicircular quartic* is an algebraic curve of degree four that is 2-circular; equivalently, it is any curve in $\mathbb{RP}^2$ defined by a homogeneous polynomial of the form

$$t(x^2 + y^2)^2 + (ux + vy)(x^2 + y^2)z + q(x, y, z)z^2, \tag{2}$$

where $t, u, v \in \mathbb{R}$, and $q \in \mathbb{R}[x, y, z]$ is a non-trivial homogeneous quadratic polynomial (see [24, Sect. 8.2] for a proof that a quartic is 2-circular if and only if its equation has the form (2)). A noteworthy example of a bicircular quartic is a union of two circles, for which it is easy to see that the curve has double points at α and β, since both circles contain those points.

Every circular cubic is contained in a bicircular quartic, since for $t = 0$ in (2) we get a union of a circular cubic and the line at infinity. A non-circular conic is also contained in a bicircular quartic, since for $t = u = v = 0$ in (2) we get a union of a conic and $z^2 = 0$, which is a double line at infinity.

Definition 2.2 The *circular degree* of an algebraic curve γ in $\mathbb{RP}^2$ is the smallest k such that γ is contained in a k-circular curve of degree $2k$.

The circular degree is well-defined, since given any curve γ of degree k, we can add k copies of the line at infinity, to get a k-circular curve of degree $2k$.

For example, a line has circular degree one, since its union with the line at infinity is a 1-circular curve of degree two. A conic that is not a circle has circular degree two, since its union with two copies of the line at infinity is a 2-circular curve of degree four. Similarly, a circular cubic has circular degree two, since its union with the line at infinity is a 2-circular curve of degree four. We can thus classify curves of low circular degree as follows:

- *Circular degree one*: lines and circles (that is, generalised circles).
- *Circular degree two*: non-circular conics, circular cubics, and bicircular quartics.
- *Circular degree three*: non-circular cubics, circular quartics, 2-circular quintics, and 3-circular sextics.

This classification is important to us, because we will see that circular degree is invariant under inversion.

We have defined circular curves and circular degrees in the projective plane, because that is their most natural setting. In the rest of the paper, to avoid confusion between the projective and inversive planes, we will use these notions for curves in $\mathbb{R}^2$, with the understanding that to inspect the definitions we should consider $\mathbb{RP}^2$ and $\mathbb{CP}^2$.

2.2 Inversion

Circular curves are intimately related to circle inversion, which we now introduce. A general reference is [6].

Definition 2.3 Let $C(p, r)$ be the circle with centre $p = (x_p, y_p) \in \mathbb{R}^2$ and radius $r > 0$. The *circle inversion* with respect to $C(p, r)$ is the mapping $I_{p,r} : \mathbb{R}^2\backslash\{p\} \to \mathbb{R}^2\backslash\{p\}$ defined by

$$I_{p,r}(x, y) = \left(\frac{r^2(x - x_p)}{(x - x_p)^2 + (y - y_p)^2} + x_p, \frac{r^2(y - y_p)}{(x - x_p)^2 + (y - y_p)^2} + y_p \right)$$

for $(x, y) \neq p$. We write I_p for $I_{p,1}$. We call p the *centre* of the inversion $I_{p,r}$.

In the inversive plane $\mathbb{R}^2 \cup \{\infty\}$, the inversion map can be completed by setting $I_{p,r}(p) = \infty$ and $I_{p,r}(\infty) = p$, so that inversions take generalised circles to generalised circles. The group of transformations of the inversive plane generated by the inversions and the similarities is called the inversive group. It is known that a bijection of the inversive plane that takes generalised circles to generalised circles has to be an element of this group, and that any element of this group is either a similarity or an inversion followed by an isometry [9, Thm. 6.71].

The image of an algebraic curve in $\mathbb{R}^2$ under an inversion is also an algebraic curve, in the following sense.

Definition 2.4 For any algebraic curve γ there is an algebraic curve γ' such that

$$I_{p,r}(\gamma\backslash\{p\}) = \gamma'\backslash\{p\}.$$

We refer to γ' as the *inverse of γ with respect to the circle $C(p, r)$*, and abuse notation slightly by writing $\gamma' = I_{p,r}(\gamma)$. Also, since for different choices of radius r, $I_{p,r}(\gamma)$ differs only by a dilatation in p, we will often only consider the inverse $I_p(\gamma) = I_{p,1}(\gamma)$ and refer to it as the *inverse of γ in the point p*.

If a curve has degree d, then its inverse has degree at most $2d$ [24, Thm. 4.14]. If γ is irreducible, then its inverse is also irreducible. Note that inverses of algebraic

curves can behave somewhat unintuitively; for instance, Proposition 2.6 states that the inverse of an ellipse has an isolated point, which is surprising if one thinks of an ellipse as just a closed continuous curve.

It is well known that the inverses of generalised circles are again generalised circles. It turns out that, more generally, circular degree is preserved under inversion. We now make precise what this means for curves of low circular degree. A proof can be found in the classical paper [16]; for a more modern reference, see [24, Sect. 9.2].

Lemma 2.5 (Inversion and circular degree) *Let C_k be a curve of circular degree k. Then:*

(i) *The inverse of C_1 in a point on C_1 is a line; the inverse of C_1 in a point not on C_1 is a circle.*
(ii) *The inverse of C_2 in a singular point on C_2 is a non-circular conic; the inverse of C_2 in a regular point on C_2 is a circular cubic; the inverse of C_2 in a point not on C_2 is a bicircular quartic.*
(iii) *The inverse of C_3 in a singularity of multiplicity three is a non-circular cubic*; *the inverse of C_3 in a singularity of multiplicity two is a circular quartic*; *the inverse of C_3 in a regular point on C_3 is a 2-circular quintic; the inverse of C_3 in a point not on C_3 is a 3-circular sextic.*

One particular subcase of Case (ii) will play an important role in our paper, and we state it separately in Proposition 2.6. A proof can be found in [14, p. 202]. Let us recall that an *acnodal cubic* is a singular cubic with a singularity that is an isolated point; for example, $(2x-1)(x^2+y^2)-y^2=0$ is an acnodal circular cubic with a singularity at the origin.

Proposition 2.6 *The inverse of an ellipse in a point on the ellipse is an acnodal circular cubic with the centre of inversion as its singularity*; *the inverse of an acnodal circular cubic in its singularity is an ellipse through the singularity.*

For example, the inverse of the cubic $(2x-1)(x^2+y^2)-y^2=0$ in its singularity at the origin is the ellipse $(x-1)^2+2y^2=1$.

3 Groups on Circular Curves

3.1 Groups on Irreducible Circular Cubics

The extremal configurations in our main theorems are all based on group laws on certain circular curves. It is well known that irreducible smooth cubics (elliptic curves) have a group law (see for instance [22]). These groups play a crucial role in the work of Green and Tao [13]. The reason that these groups are relevant to ordinary lines is the following collinearity property of this group (when defined in the standard way). Three points on the curve are collinear if and only if in the group they sum to the identity element. For this property to hold, the identity element must be an inflection point. Here we will define a group in a slightly different way (described for instance in [22, Sect. 1.2]), in which the identity element is not necessarily an inflection point, and the same collinearity property does not hold. However, for circular cubics, we

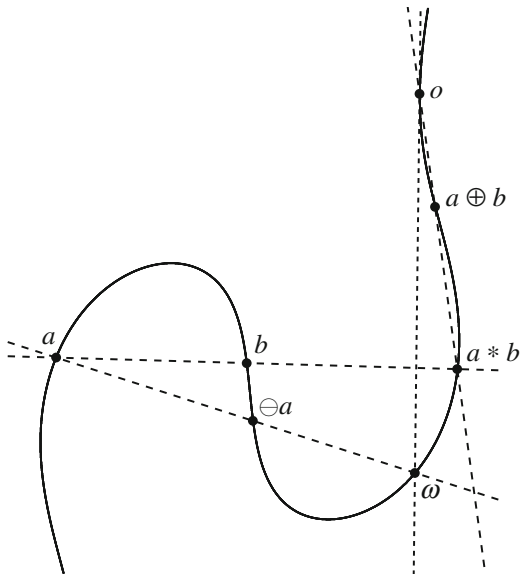

Fig. 1 Group law on a smooth circular cubic curve

show that we can choose the identity element so that we get a similar property for concyclicity.

First let γ be any irreducible cubic, write γ^* for its set of regular points, and pick an arbitrary point $o \in \gamma^*$. We describe an additive group operation $\oplus$ on the set γ^* for which o is the identity element. The construction is depicted in Fig. 1. Given $a, b \in \gamma^*$, let $a * b$ be the third intersection point of γ and the line ab, and define $a \oplus b$ to be $(a * b) * o$, the third intersection point of γ and the line through $a * b$ and o. When $a = b$, the line ab should be interpreted as the tangent line at a; when $a * b = o$, the line through $a * b$ and o should be interpreted as the tangent line to γ at o. We refer to [22] for a more careful definition and a proof that this operation really does give a group.

Now consider a circular cubic γ. Since the circular points α and β lying on it are conjugate, γ has a unique real point on the line at infinity, which we choose as our identity element o. We define the point ω to be the third intersection point of the tangent line to γ at o (if there is no third intersection point, then o is an inflection point, and we consider o itself to be the third point). Throughout this paper we will use ω to denote this special point on a circular cubic; note that ω is not fixed like α and β, but depends on γ. Also note that ω is real, since it corresponds to the third root of a real cubic polynomial whose other two roots correspond to the real point o. Observe that

$$\omega = \alpha \oplus \beta,$$

since $\alpha * \beta = o$, and by definition $o * o = \omega$.

With this group law, we no longer have the property that three points are collinear if and only if they sum to o (unless o happens to be an inflection point). Nevertheless, one can check that three points $a, b, c \in \gamma^*$ are collinear if and only if $a \oplus b \oplus c = \omega$.

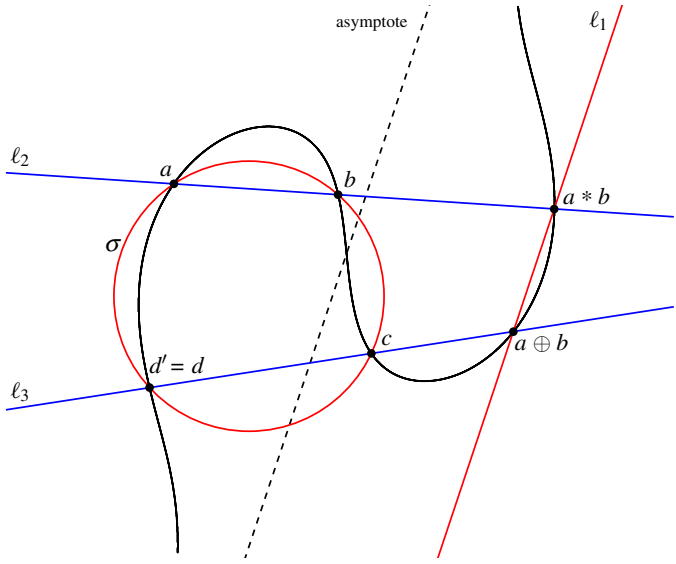

Fig. 2 Concyclicity of four regular points on a circular cubic

More important for us, four points of γ^* lie on a generalised circle if and only if they sum to ω. This amounts to a classical fact (see [4, Art. 225] for an equivalent statement), but we include a proof for completeness. We use the following version of the Cayley–Bacharach Theorem, due to Chasles (see [11]).

Theorem 3.1 (Chasles) *Suppose two cubic curves in $\mathbb{CP}^2$ with no common component intersect in nine points, counting multiplicities. If γ is another cubic curve containing eight of these intersection points, counting multiplicities, then γ also contains the ninth.*

Recall from Sect. 2 that a generalised circle, viewed projectively, is either a circle, or the union a line with the line at infinity.

Proposition 3.2 *Let γ be an irreducible circular cubic in $\mathbb{RP}^2$, and let $a, b, c, d \in \gamma^*$ be points that are not necessarily distinct. A generalised circle intersects γ in the points a, b, c, d (taking into account multiplicity) if and only if $a \oplus b \oplus c \oplus d = \omega$.*

Proof We consider the cubic γ extended to $\mathbb{CP}^2$. We first show the forward direction. All statements in the proof should be considered with multiplicity.

If the generalised circle is the union of a line ℓ and the line at infinity ℓ_∞, then $\ell \cup \ell_\infty$ intersects γ in $a, b, c, d, \alpha, \beta$. Since ℓ intersects γ in at most three points, one of the points a, b, c, d must equal o, say $d = o$. Since ℓ_∞ also intersects γ in at most three points, we must have $a, b, c \in \ell$. Thus a, b, c are collinear, and we have $a \oplus b \oplus c = \omega$, by the definition of the group law. It then follows from $d = o$ that $a \oplus b \oplus c \oplus d = \omega$.

Suppose next that the generalised circle is a circle σ, and intersects γ in $a, b, c, d, \alpha, \beta$. The construction that follows is depicted in Fig. 2. Let ℓ_1 be the line

through o and $a * b$ (and thus through $a \oplus b$), ℓ_2 the line through a and b (and thus through $a * b$), and ℓ_3 the line through c and $a \oplus b$. Note that σ and ℓ_∞ intersect in α and β. Then $\gamma_1 = \sigma \cup \ell_1$ and $\gamma_2 = \ell_2 \cup \ell_3 \cup \ell_\infty$ are two cubic curves that intersect in nine points, of which the eight points $a, b, c, a * b, a \oplus b, o, \alpha$, and β certainly lie on γ; the remaining point is the third intersection point of γ_1 and ℓ_3 beside c and $a \oplus b$, which we denote by d'. By Theorem 3.1, γ contains d'. By the group law on γ, we have $d' = (a \oplus b) * c$. Moreover, d' must be the sixth intersection point of γ and σ beside a, b, c, α, β, which is d, so $d = d' = (a \oplus b) * c$. By the definition of the group law, this implies $a \oplus b \oplus c = o * d$, so $(a \oplus b \oplus c) * d = (o * d) * d = o$, and finally $a \oplus b \oplus c \oplus d = o * o = \omega$.

For the converse, suppose that $a \oplus b \oplus c \oplus d = \omega$, and let d' be the fourth point where the generalised circle σ through a, b, c intersects γ. Then, by what we have just shown, $a \oplus b \oplus c \oplus d' = \omega$, and it follows that $d = d'$, and a, b, c, d lie on σ. □

This proposition is a consequence of the more general fact that six points on a circular cubic lie on a conic if and only if they sum to 2ω. (In the standard group structure on a cubic, where the identity o is chosen as an inflection point, they would sum to o; see [23, Thm. 9.2].) Since a generalised circle in $\mathbb{RP}^2$ is a conic containing α and β, and $\alpha \oplus \beta = \omega$, it follows that four points a, b, c, d (possibly including o) lie on a generalised circle if and only if they sum to ω.

3.2 Groups on Other Circular Curves

We now define group laws on two other types of curves of circular degree two, and observe that they satisfy similar concyclicity properties. Let us note at this point that most bicircular quartics can also be given a group structure (if an irreducible bicircular quartic has no singularities besides α and β, then it is a curve of genus one, and thus has a group law by [21, Sect. III.3]). However, in our proofs we will handle bicircular quartics by inverting in a point on the curve, which by Lemma 2.5 transforms a bicircular quartic into a circular cubic. For that reason, we do not need to study the group law on bicircular quartics separately.

Ellipses We discuss a group law on ellipses, although we do not actually need it in our proof, because inversion lets us transform an ellipse into an acnodal cubic (Proposition 2.6), which we have already given a group structure in the previous subsection. Nevertheless, we treat the group law on ellipses here because it is especially elementary, and it would be strange not to mention it.

Consider the ellipse σ given by the equation $x^2 + (y/s)^2 = 1$, with $s \neq 0, 1$. For any point $a \in \sigma$, we project a vertically to the point a' on the unit circle around the origin, as in Fig. 3, and call the angle θ_a the *eccentric angle* of a. We define the sum of two points $a, b \in \sigma$ to be the point $c = a \oplus b$ whose eccentric angle is $\theta_c = \theta_a + \theta_b$. This gives σ a group structure isomorphic to $\mathbb{R}/\mathbb{Z}$. The identity element is $o = (1, 0)$, and the inverse of a point is its reflection in the x-axis. We have the following classical fact that describes when four points on an ellipse are concyclic (see [15] for the oldest reference we could find, and [5, Problem 17.2] for two detailed proofs).

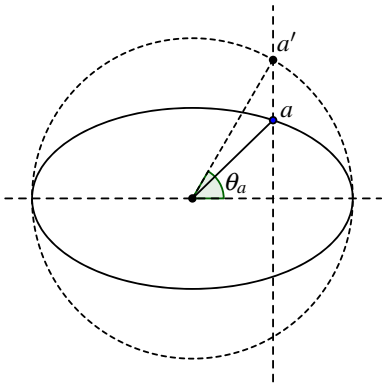

Fig. 3 Eccentric angle of a point on an ellipse

Proposition 3.3 *Four points $a, b, c, d \in \sigma$ are concyclic if and only if $a \oplus b \oplus c \oplus d = o$. We may allow two of the points to be equal, in which case the circle through the three distinct points is tangent to the ellipse at the repeated point.*

Another way to look at this group law is that we are parametrising the ellipse using lines through $o = (1, 0)$ (see for instance [22, Sect. 1.1]). More precisely, each point $a \in \sigma$ corresponds to the line oa; oa makes an angle $\pi - \theta_a/2$ with the x-axis, and the set of lines through o thus has a group structure equivalent to the one above. This view lets us relate the group on the ellipse to the group on the acnodal cubic. By Proposition 2.6, inverting in o maps the ellipse to an acnodal circular cubic γ, with o becoming the isolated point of the cubic. The lines through o now parametrise the cubic, and this parametrisation gives the same group on γ as the line construction that we gave in Sect. 3.1 (see [22, Sect. 3.7]).

Concentric circles We now define a group on the union of two disjoint circles. For notational convenience, we identify $\mathbb{R}^2$ with $\mathbb{C}$. After an appropriate inversion, we can assume the circles to be

$$\left\{ e^{2\pi i t} \;\middle|\; t \in [0, 1) \right\}, \qquad \left\{ r e^{-2\pi i t} \;\middle|\; t \in [0, 1) \right\},$$

with $r > 1$, and we represent each element of $\sigma_1 \cup \sigma_2$ as $r^{\varepsilon} e^{2\pi i t}$ with $\varepsilon \in \mathbb{Z}_2$ (with the obvious convention $r^0 = 1$ and $r^1 = r$). We define a group operation on $\sigma_1 \cup \sigma_2$ by

$$r^{\varepsilon_1} e^{2\pi i t_1} \oplus r^{\varepsilon_2} e^{2\pi i t_2} = r^{(\varepsilon_1 + \varepsilon_2) \bmod 2} e^{2\pi i (t_1 + t_2)},$$

which turns $\sigma_1 \cup \sigma_2$ into a group isomorphic to $\mathbb{R}/\mathbb{Z} \times \mathbb{Z}_2$, with identity element $o = 1 = r^0 e^{2\pi i \cdot 0}$. We again have the following concyclicity property, which is easily seen using symmetry.

Proposition 3.4 *Points $a, b \in \sigma_1$ and $c, d \in \sigma_2$ lie on a generalised circle if and only if $a \oplus b \oplus c \oplus d = o$. If $a = b$ or $c = d$, then the generalised circle is tangent at that point.*

4 Constructions

4.1 Ellipse

Let σ be the ellipse defined by $x^2 + (y/s)^2 = 1$, with the group structure introduced in Sect. 3.2. Let $n \geqslant 5$. We have a finite subgroup of size n given by

$$S = \left\{ \left(\cos\left(\frac{2\pi k}{n}\right), s \sin\left(\frac{2\pi k}{n}\right)\right) \;\middle|\; k = 0, \ldots, n-1 \right\} \subset \sigma.$$

By Proposition 3.3, the circle through any three points $a, b, c \in S$ passes through the point $d = \ominus a \ominus b \ominus c \in S$. Therefore, the only way S spans an ordinary circle is when d coincides with one of the points a, b, c (which occurs if the circle is tangent to σ at that point). It follows that the number of ordinary circles is equal to

$$\frac{1}{2}\left| \left\{ (k_1, k_2, k_3) \in \mathbb{Z}_n^3 \;\middle|\; 2k_1 + k_2 + k_3 \equiv 0 \,(\mathrm{mod}\, n),\ k_1, k_2, k_3 \text{ distinct} \right\} \right|,$$

which is $n^2/2 - O(n)$.

Similarly, the number of 4-point circles is equal to

$$\frac{1}{4!}\left| \left\{ (k_1, k_2, k_3, k_4) \in \mathbb{Z}_n^4 \;\middle|\; k_1 + k_2 + k_3 + k_4 \equiv 0 \,(\mathrm{mod}\, n),\ k_1, k_2, k_3, k_4 \text{ distinct} \right\} \right|,$$

which is, by inclusion–exclusion, equal to $(n^3 - 6n^2 + (8 + 3\delta_n)n - 6\varepsilon_n)/24$, where δ_n is the number of solutions in $\mathbb{Z}_n$ to the equation $2k = 0$ and ε_n is the number of solutions in $\mathbb{Z}_n$ to the equation $4k = 0$. This works out to

$$\begin{cases} \dfrac{1}{24}n^3 - \dfrac{1}{4}n^2 + \dfrac{7}{12}n - 1 & \text{if } n \equiv 0 \pmod 4, \\ \dfrac{1}{24}n^3 - \dfrac{1}{4}n^2 + \dfrac{11}{24}n - \dfrac{1}{4} & \text{if } n \equiv 1, 3 \pmod 4, \\ \dfrac{1}{24}n^3 - \dfrac{1}{4}n^2 + \dfrac{7}{12}n - \dfrac{1}{2} & \text{if } n \equiv 2 \pmod 4. \end{cases}$$

4.2 Circular Cubic Curve

Let γ be an irreducible circular cubic, and let $\oplus$ be the group operation defined in Sect. 3.1. It is well known (see for instance [13]) that the group $(\gamma^*, \oplus)$ is isomorphic to the circle $\mathbb{R}/\mathbb{Z}$ if γ is acnodal or if γ is smooth and has one connected component, and is isomorphic to $\mathbb{R}/\mathbb{Z} \times \mathbb{Z}_2$ if γ is smooth and has two connected components. Let H_n be a subgroup of order n of γ^*, and let $x \in \gamma^*$ be such that $4x = \omega \ominus h$ for some $h \in H_n$. By Proposition 3.2, the number of ordinary generalised circles in the coset $S = H_n \oplus x$ equals

$$\frac{1}{2}\left| \left\{ (a, b, c) \in H_n^3 \;\middle|\; 2a \oplus b \oplus c = h,\ a, b, c \text{ distinct} \right\} \right|,$$

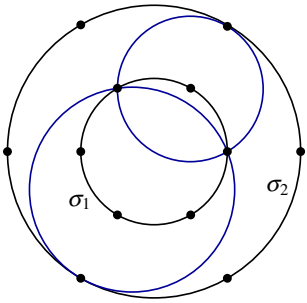

Fig. 4 'Aligned' double hexagon

which is easily seen to equal $n^2/2 - O(n)$. Similarly, the number of ordinary circles in $S = H_n \oplus x$ equals

$$\frac{1}{2}\left|\left\{(a,b,c) \in H_n^3 \;\middle|\; 2a \oplus b \oplus c = h,\ a,b,c \neq \ominus x \text{ and distinct}\right\}\right|,$$

which also equals $n^2/2 - O(n)$.

As in the previous construction, if $o \notin S$ (equivalently, $x \notin H_n$) then the number of 4-point circles is equal to $(n^3 - 6n^2 + (8 + 3\delta_n)n - 6\varepsilon_n)/24$, where δ_n is the number of solutions in H_n to the equation $2k = h$ and ε_n is the number of solutions in H_n to the equation $4k = h$. If H_n is cyclic, then we get the same numbers as in the previous construction. Otherwise, $n = 0 \pmod 4$, $H_n \cong \mathbb{Z}_{n/2} \times \mathbb{Z}_2$, and the number of 4-point circles equals

$$\begin{cases} \dfrac{1}{24}n^3 - \dfrac{1}{4}n^2 + \dfrac{5}{6}n - 2 & \text{if } n \equiv 0 \pmod 8, \\ \dfrac{1}{24}n^3 - \dfrac{1}{4}n^2 + \dfrac{5}{6}n - 1 & \text{if } n \equiv 4 \pmod 8, \end{cases}$$

which is greater than the corresponding number in the previous construction.

4.3 'Aligned' Double Polygons

Let $n \geqslant 6$ be even and set $m = n/2$. We identify $\mathbb{R}^2$ with $\mathbb{C}$. Let σ_1 be the circle with centre the origin and radius one, and σ_2 the circle with centre the origin and radius $r > 1$. Let $S_1 = \left\{e^{2\pi ik/m} \;\middle|\; k = 0, \ldots, m-1\right\} \subset \sigma_1$ and $S_2 = \left\{re^{2\pi ik/m} \;\middle|\; k = 0, \ldots, m-1\right\} \subset \sigma_2$. Thus, S_1 and S_2 are the vertex sets of regular m-gons on σ_1 and σ_2 that are 'aligned' in the sense that their points lie at the same set of angles from the common centre (see Fig. 4).

Let $S = S_1 \cup S_2$. By Proposition 3.4, the points $a, b \in \sigma_1, c, d \in \sigma_2$ are collinear or concyclic if and only if $a \oplus b \oplus c \oplus d = o$. In particular, if $a = b$, then the generalised circle through the three points is tangent to σ_1. It follows that if $n \geqslant 8$, the ordinary generalised circles of S are exactly those through $e^{2\pi ik_1/m}, re^{-2\pi ik_2/m}, re^{-2\pi ik_3/m}$ or through $re^{-2\pi ik_1/m}, e^{2\pi ik_2/m}, e^{2\pi ik_3/m}$ where $2k_1 + k_2 + k_3 \equiv 0 \pmod m$, with $k_2 \not\equiv k_3 \pmod m$.

For generic $r > 1$, we then obtain that the number of ordinary generalised circles equals

$$\left| \left\{ (k_1, k_2, k_3) \in \mathbb{Z}_m^3 \;\middle|\; 2k_1 + k_2 + k_3 \equiv 0 \,(\mathrm{mod}\, n),\ k_2, k_3 \text{ distinct} \right\} \right|$$

(although k_2 and k_3 are not ordered, we either have two points on σ_1 or two points on σ_2). This equals $m(m-2)$ if m is even and $m(m-1)$ if m is odd. That is, for generic r, we obtain $n^2/4 - n$ ordinary generalised circles if $n \equiv 0 \pmod 4$ and $n^2/4 - n/2$ ordinary generalised circles if $n \equiv 2 \pmod 4$.

If we choose $r = (\cos(2\pi k/m))^{-1}$ (there are $\lfloor m/4 \rfloor$ choices for r), then the tangent lines at points of S_1 pass through two points of S_2, so are ordinary generalised circles. Thus, for these choices of r we lose m ordinary circles, and obtain $n^2/4 - 3n/2$ ordinary circles if $n \equiv 0 \pmod 4$ and $n^2/4 - n$ ordinary circles if $n \equiv 2 \pmod 4$. Note that this is much less than the number of ordinary circles given by Constructions 4.1 and 4.2.

Similarly, the number of 4-point generalised circles spanned by S equals

$$\frac{1}{4} \left| \left\{ (k_1, k_2, k_3, k_4) \in \mathbb{Z}_m^4 \;\middle|\; k_1 + k_2 + k_3 + k_4 \equiv 0 \,(\mathrm{mod}\, m),\ k_1 \neq k_2 \text{ and } k_3 \neq k_4 \right\} \right|,$$

which is $m^3/4 - O(m^2) = n^3/32 - O(n^2)$, also much less than the number in Constructions 4.1 and 4.2.

4.4 'Offset' Double Polygons

We modify the previous construction by rotating S_2 around the origin by an angle of $\pi k/m$. This results in $S_2' = \left\{ re^{-i\pi(2k-1)/m} \;\middle|\; k = 0, \ldots, m-1 \right\}$ and $S' = S_1 \cup S_2'$ (see Fig. 5). As before, if $n \geqslant 8$, the ordinary generalised circles of S' are exactly those through $e^{2\pi i k_1/m}, re^{-i\pi(2k_2-1)/m}, re^{-i\pi(2k_3-1)/m}$ or through $re^{-i\pi(2k_1-1)/m}, e^{2\pi i k_2/m}, e^{2\pi i k_3/m}$, where $2k_1 + k_2 + k_3 \equiv 1 \pmod m$ with $k_2 \not\equiv k_3 \pmod m$.

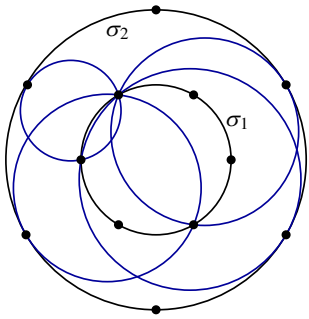

Fig. 5 'Offset' double hexagon

For generic $r > 1$, we now have to count the number of ordered triples in the set

$$\left\{(k_1, k_2, k_3) \in \mathbb{Z}_m^3 \;\middle|\; 2k_1 + k_2 + k_3 \equiv 1 \,(\mathrm{mod}\, n),\ k_2, k_3 \text{ distinct}\right\}.$$

This equals m^2 if m is even and $m(m-1)$ if m is odd. That is, for generic r, we obtain $n^2/4$ ordinary generalised circles if $n \equiv 0 \pmod 4$, worse than Construction 4.3, and $n^2/4 - n/2$ ordinary generalised circles if $n \equiv 2 \pmod 4$, the same number as in Construction 4.3.

Again, if we choose $r = (\cos(2\pi k/m))^{-1}$ (there are $\lfloor m/4 \rfloor$ choices for r), we lose m ordinary circles. Thus, we obtain $n^2/4 - n$ ordinary circles if $n \equiv 2 \pmod 4$, the same number as in Construction 4.3.

As in Construction 4.3, we get $n^3/32 - O(n^2)$ 4-point circles.

4.5 Punctured Double Polygons

Let $n = 2m - 1 \geqslant 11$ be odd. Take Construction 4.3 with $n + 1 = 2m$ points and remove an arbitrary point $p \in S_1$.

First assume that m is odd. Before we remove p, there are $m(m-1)$ ordinary generalised circles. Of these, there are $(m-1)/2$ tangent at p. There are also $m - 1$ ordinary generalised circles through p tangent at some point of S_2. Thus, by removing p, we destroy $3(m-1)/2$ ordinary generalised circles and create $\binom{m}{2} - (m-1)/2$ new ones. Therefore, $S \backslash \{p\}$ has

$$m(m-1) - \frac{3}{2}(m-1) + \binom{m}{2} - \frac{1}{2}(m-1) = \frac{3}{2}m^2 - \frac{7}{2}m + 2$$

ordinary generalised circles. That is, there are $3n^2/8 - n + 5/8$ ordinary generalised circles if $n \equiv 1 \pmod 4$.

Next assume that m is even. Before we remove p, there are $m(m-2)$ ordinary generalised circles, of which there are $(m-2)/2$ through two different points of S_2 tangent at p, and there are also $m - 2$ ordinary generalised circles through p tangent at a point of S_2. As before, we obtain

$$m(m-2) - \frac{3}{2}(m-2) + \binom{m}{2} - \frac{1}{2}(m-2) = \frac{3}{2}m^2 - \frac{9}{2}m + 4$$

ordinary generalised circles. Thus, we obtain $3n^2/8 - 3n/2 + 17/8$ ordinary generalised circles if $n \equiv 3 \pmod 4$.

Instead of starting with Construction 4.3, we can take the 'offset' Construction 4.4 and remove a point. It is easy to see that when $n \equiv 1 \pmod 4$ we obtain the same number of ordinary generalised circles, while if $n \equiv 3 \pmod 4$ we obtain more.

Since there are no 5-point circles in Constructions 4.3 and 4.4 when $m \geqslant 6$, removing a point does not add any 4-point circle, but destroys $O(n^2)$ of them. We thus get $n^3/32 - O(n^2)$ 4-point generalised circles, which is asymptotically the same as in Constructions 4.3 and 4.4.

4.6 Inverted Double Polygons

We can use inversion to make new constructions out of old ones.

Invert Construction 4.5 in the removed point p. The resulting point set has m points on a circle and $m-1$ points on a line disjoint from the circle. Every ordinary circle after the inversion corresponds to an ordinary generalised circle not passing through p before the inversion. If m is odd, there are $(m-1)/2$ ordinary generalised circles tangent at p and a further $m-1$ ordinary generalised circles through p tangent to σ_2, so we obtain $m(m-1)-3(m-1)/2 = (m-1)(2m-3)/2$ ordinary circles. For even m we similarly obtain $m(m-2)-3(m-2)/2 = (m-2)(2m-3)/2$ ordinary circles. That is, we have $(n-1)(n-2)/4 = n^2/4-3n/4+1/2$ ordinary circles when $n \equiv 1 \pmod 4$ and $(n-3)(n-2)/4 = n^2/4-5n/4+3/2$ ordinary circles when $n \equiv 3 \pmod 4$.

If we remove another point from this inverted construction, we obtain a set of n points where n is even, with $3n^2/8 - O(n)$ ordinary circles.

4.7 Other Inverted Examples

If we invert Construction 4.1 in a point on the ellipse that is not in the set S, then by Proposition 2.6, we obtain points on an acnodal circular cubic (without its acnode) as in Construction 4.2, with the same number of ordinary and 4-point generalised circles.

If we invert a circular cubic in a point not on the curve, then we obtain a bicircular quartic by Lemma 2.5. There will again be $n^2/2 - O(n)$ ordinary circles (or ordinary generalised circles) and $n^3/24 - O(n^2)$ 4-point circles among the inverted points.

5 The Structure Theorems

5.1 Proof of the Weak Structure Theorem

The proofs of our structure theorems for sets with few ordinary circles crucially rely on the following structure theorem for sets with few ordinary lines due to Green and Tao [13]. Recall that an *ordinary line* is a line containing exactly two points of the given point set.

Theorem 5.1 (Green–Tao) *Let $K > 0$ and let n be sufficiently large depending on K. If a set P of n points in $\mathbb{R}^2$ spans at most Kn ordinary lines, then P differs in at most $O(K)$ points from an example of one of the following types*:

(i) *$n - O(K)$ points on a line*;
(ii) *m points each on a line and a disjoint conic, for some $m = n/2 \pm O(K)$*;
(iii) *$n \pm O(K)$ points on an acnodal or smooth cubic.*

We commence the proof of Theorem 1.4. Let P be a set of n points spanning at most Kn^2 ordinary generalised circles. We wish to show that P lies mostly on a bicircular quartic (we will repeatedly use 'mostly' to mean 'for all but $O(K)$ points').

Note that for at least $2n/3$ points p of P, there are at most $9Kn$ ordinary circles through p, hence the set $I_p(P\backslash\{p\})$ spans at most $9Kn$ ordinary lines. Let P' be the set of such points. For n sufficiently large depending on K, applying Theorem 5.1 to $I_p(P\backslash\{p\})$ for any $p \in P'$ gives that $I_p(P\backslash\{p\})$ lies mostly on a line, a line and a conic, an acnodal cubic, or a smooth cubic.

If there exists $p \in P'$ such that $I_p(P\backslash\{p\})$ lies mostly on a line, then inverting again in p, we see that P must lie mostly on a line or a circle.

If there exists $p \in P'$ such that $I_p(P\backslash\{p\})$ lies mostly on a line ℓ and a disjoint conic σ, we have two cases, depending on whether p lies on ℓ or not.

If $p \in \ell$, we invert again in p to find that P lies mostly on the union of ℓ and $I_p(\sigma)$. By Lemma 2.5, $I_p(\sigma)$ is either a circle (if σ is a circle) or an irreducible bicircular quartic (if σ is a non-circular conic). Furthermore, p is the only point that could possibly lie on both ℓ and $I_p(\sigma)$. Since roughly $n/2$ points of P lie on ℓ, there must be another point $q \in \ell \cap P'$ that does not lie on $I_p(\sigma)$. In $I_q(P\backslash\{q\})$, the line ℓ remains a line, and by definition of P' the set $I_q(P\backslash\{q\})$ spans few ordinary lines, so Theorem 5.1 tells us $I_q(I_p(\sigma))$ is a conic. It follows from Lemma 2.5 that $I_p(\sigma)$ cannot be a quartic, since we inverted in the point q outside $I_p(\sigma)$ and did not obtain a quartic. That means $I_p(\sigma)$ has to be a circle, and it is disjoint from ℓ. Thus, P lies mostly on the union of a line and a disjoint circle.

If $p \notin \ell$, we invert in p to see that P lies mostly on the union of the circle $I_p(\ell)$ and the curve $I_p(\sigma)$, which is either a circle or a quartic. Again p is the only point that can lie on both curves. Inverting in another point $q \in I_p(\ell) \cap P'$, $I_q(I_p(\ell))$ becomes a line, so Theorem 5.1 tells us that $I_q(I_p(\sigma))$ is a conic, so that $I_p(\sigma)$ must be a circle disjoint from $I_p(\ell)$ as before. Thus, P lies mostly on the union of two disjoint circles.

The case that remains is when for all $p \in P'$, the set $I_p(P\backslash\{p\})$ lies mostly on an acnodal or smooth cubic γ. Fix such a p, and consider $I_p(\gamma)$, which mostly contains P. If γ is not a circular cubic, then by the classification in Sect. 2 it has circular degree three, so $I_p(\gamma)$ has circular degree three as well. For any $q \in I_p(\gamma) \cap P'$ other than p, the curve $I_q(I_p(\gamma))$ is also a cubic curve, by the definition of P' and Theorem 5.1. By Case (iii) of Lemma 2.5, this can only happen if q is a singularity of $I_p(\gamma)$. But $I_p(\gamma)$ is an irreducible curve of degree at most six, and so has at most ten singularities by [23, Thm. 4.4], which is a contradiction. So γ must be a circular cubic that is acnodal or smooth. If γ is acnodal, then $I_p(\gamma)$ is either a bicircular quartic (if $p \notin \gamma$), an acnodal circular cubic (if p is a regular point of γ), or a non-circular conic (if p is the singularity of γ). In the last case, the conic is an ellipse by Proposition 2.6. If γ is smooth, then $I_p(\gamma)$ is either a bicircular quartic or a smooth circular cubic.

We have encountered the following curves that P could mostly lie on: a line, a circle, an ellipse, a disjoint union of a line and a circle, a disjoint union of two circles, a circular cubic, or a bicircular quartic. All of these are subsets of bicircular quartics, which proves the statement of Theorem 1.4. □

5.2 Proof of the Strong Structure Theorem

We now prove Theorem 1.5. First of all, as explained in Sect. 4, a subgroup of an ellipse and an appropriate coset of a subgroup of a smooth circular cubic both have

at most $n^2/2$ ordinary generalised circles, and a double polygon has at most $n^2/4$ ordinary generalised circles. It follows from Lemma 5.2 below that if we add and/or remove $O(K)$ points, then there will be at most $O(Kn^2)$ ordinary generalised circles.

Lemma 5.2 *Let S be a set of n points in $\mathbb{R}^2$ with s ordinary generalised circles. Let T be a set that differs from S in at most K points: $|S \triangle T| \leqslant K$. Then T has at most $s + O(Kn^2 + K^2n + K^3)$ ordinary generalised circles.*

Proof First note that if we add a point to any set of n points, we create at most $\binom{n}{2}$ ordinary generalised circles. Secondly, since two circles intersect in at most two points, the number of 4-point circles through a fixed point in a set of n points is at most $\frac{1}{3}\binom{n-1}{2}$, so by removing a point we create at most $\frac{1}{3}\binom{n-1}{2} < \binom{n}{2}$ ordinary generalised circles. It follows that by adding and removing $O(K)$ points, we create at most

$$\binom{n}{2} + \binom{n+1}{2} + \cdots + \binom{n+K-1}{2} = O(Kn^2 + K^2n + K^3)$$

ordinary generalised circles. □

Next, let P be a set of n points with at most Kn^2 ordinary generalised circles. From the proof of Theorem 1.4 above, we see that P differs in at most $O(K)$ points from a line, a circle, an ellipse, a disjoint union of a line and a circle, a disjoint union of two circles, a circular cubic, or a bicircular quartic. Moreover, in the proof we saw that the circular cubic must be acnodal or smooth, and that the bicircular quartic has the property that if we invert in a point on the curve, the resulting circular cubic is acnodal or smooth.

Using inversions, we can reduce the number of types of curves that we need to analyse further.

- If P lies mostly on a line, then we are in Case (i) of Theorem 1.5, so we are done.
- If P lies mostly on a circle, then inverting in a point on the circle puts us in Case (i) again.
- If P lies mostly on an ellipse, then inverting in a point of the ellipse places P mostly on an acnodal circular cubic.
- If P lies mostly on a bicircular quartic, then inverting in any regular point on the curve gives us a circular cubic. As mentioned above, this cubic is acnodal or smooth.
- If P lies mostly on a line and a disjoint circle, then an inversion in a point not on the line or circle places P mostly on two disjoint circles.
- If P lies mostly on the disjoint union of two circles, we can apply an inversion that maps the two disjoint circles to two concentric circles [6, Thm. 1.7].

So, up to inversions, we need only consider the cases when P lies mostly on an acnodal or smooth circular cubic, or on two concentric circles. We do this in Lemmas 5.5 and 5.6 below, which will complete the proof of Theorem 1.5.

To determine the structure of P, we use a variant of a lemma from additive combinatorics that was used by Green and Tao [13]. It captures the principle that if a finite subset of a group is almost closed under addition, then it is close to a subgroup. The following statement is Proposition A.5 in [13].

Proposition 5.3 *Let $K > 0$ and let n be sufficiently large depending on K. Let A, B, C be three subsets of some abelian group $(G, \oplus)$, all of cardinality within K of n. Suppose there are at most Kn pairs $(a, b) \in A \times B$ for which $a \oplus b \notin C$. Then there is a subgroup $H \leqslant G$ and cosets $H \oplus x$, $H \oplus y$ such that*

$$|A \triangle (H \oplus x)|, |B \triangle (H \oplus y)|, |C \triangle (H \oplus x \oplus y)| = O(K).$$

The variant that we need is a simple corollary of Proposition 5.3.

Corollary 5.4 *Let $K > 0$ and let n be sufficiently large depending on K. Let A, B, C, D be four subsets of some abelian group $(G, \oplus)$, all of cardinality within K of n. Suppose there are at most Kn^2 triples $(a, b, c) \in A \times B \times C$ for which $a \oplus b \oplus c \notin D$. Then there is a subgroup $H \leqslant G$ and cosets $H \oplus x$, $H \oplus y$, $H \oplus z$ such that*

$$|A \triangle (H \oplus x)|, |B \triangle (H \oplus y)|, |C \triangle (H \oplus z)|, |D \triangle (H \oplus x \oplus y \oplus z)| = O(K).$$

Proof By the pigeonhole principle, there exists an $a_0 \in A$ such that there are at most $K'n$ (where $K' = O(K)$) pairs $(b, c) \in B \times C$ for which $a_0 \oplus b \oplus c \notin D$, or equivalently $b \oplus c \notin D \ominus a_0$. Applying Proposition 5.3, we have a subgroup $H \leqslant G$ and cosets $H \oplus y$, $H \oplus z$ such that

$$|B \triangle (H \oplus y)|, |C \triangle (H \oplus z)|, |(D \ominus a_0) \triangle (H \oplus y \oplus z)| = O(K).$$

Since $|B \cap (H \oplus y)| \geqslant n - O(K)$, we repeat the argument above to obtain $b_0 \in B \cap (H \oplus y)$ such that there are at most $O(Kn)$ pairs $(a, c) \in A \times C$ with $a \oplus b_0 \oplus c \notin D$, and Proposition 5.3 gives a subgroup $H' \leqslant G$ and cosets $H' \oplus x$, $H' \oplus z'$ such that

$$|A \triangle (H' \oplus x)|, |C \triangle (H' \oplus z')|, |(D \ominus b_0) \triangle (H' \oplus x \oplus z')| = O(K).$$

From this, it follows that $|(H \oplus z) \triangle (H' \oplus z')| = O(K)$, hence $|(H \oplus z) \cap (H' \oplus z')| \geqslant n - O(K)$. Since $(H \oplus z) \cap (H' \oplus z')$ is not empty, it has to be a coset of $H' \cap H$. If $H' \neq H$, then $|H' \cap H| \leqslant n/2 + O(K)$, a contradiction. Therefore, $H = H'$ and $H \oplus z = H' \oplus z'$. So we have $|A \triangle (H \oplus x)|, |B \triangle (H \oplus y)|, |C \triangle (H \oplus z)|, |D \triangle (H \oplus x \oplus b_0 \oplus z)| = O(K)$. Since $b_0 \in H \oplus y$, we obtain $|D \triangle (H \oplus x \oplus y \oplus z)| = O(K)$ as well. □

Lemma 5.5 (Circular cubic) *Let $K > 0$ and let n be sufficiently large depending on K. Suppose P is a set of n points in $\mathbb{R}^2$ spanning at most Kn^2 ordinary generalised circles, and all but at most K points of P lie on an acnodal or smooth circular cubic γ. Then there is a coset $H \oplus x$ of a subgroup $H \leqslant \gamma^*$, with $4x \in H \oplus \omega$, such that $|P \triangle (H \oplus x)| = O(K)$.*

Proof Let $P' = P \cap \gamma^*$. Then $|P \triangle P'| = O(K)$, and by Lemma 5.2, P' spans at most $O(Kn^2)$ ordinary circles. If $a, b, c \in \gamma$ are distinct, then by Proposition 3.2, the generalised circle through a, b, c meets γ again in the unique point $d = \omega \ominus (a \oplus b \oplus c)$. This implies that $d \in P'$ for all but at most $O(Kn^2)$ triples $a, b, c \in P'$, or equivalently

$a \oplus b \oplus c \in \omega \ominus P'$. Applying Corollary 5.4 with $A = B = C = P'$ and $D = \omega \ominus P'$, we obtain $H \leqslant \gamma^*$ and a coset $H \oplus x$ such that $|P \triangle (H \oplus x)| = O(K)$ and $|(\omega \ominus P') \triangle (H \oplus 3x)| = O(K)$, which is equivalent to $|P \triangle (H \ominus 3x \oplus \omega)| = O(K)$. Thus we have $|(H \oplus x) \triangle (H \ominus 3x \oplus \omega)| = O(K)$, which implies $4x \in H \oplus \omega$. □

Lemma 5.6 (Concentric circles) *Let $K > 0$ and let n be sufficiently large depending on K. Suppose P is a set of n points in $\mathbb{R}^2$ spanning at most Kn^2 ordinary generalised circles. Suppose all but at most K of the points of P lie on two concentric circles, and that P has $n/2 \pm O(K)$ points on each. Then, up to similarity, P differs in at most $O(K)$ points from an 'aligned' or 'offset' double polygon.*

Proof By scaling and rotating, we can assume that P lies mostly on the two concentric circles $\{e^{2\pi it} \mid t \in [0, 1)\}$ and $\{re^{-2\pi it} \mid t \in [0, 1)\}$, $r > 1$, which we gave a group structure in Sect. 3.2.

Let $P_1 = P \cap \sigma_1$ and $P_2 = P \cap \sigma_2$. Then $|P \triangle (P_1 \cup P_2)| = O(K)$, and by Lemma 5.2, $P_1 \cup P_2$ spans at most $O(Kn^2)$ ordinary circles. If $a, b \in \sigma_1$ and $c \in \sigma_2$ with $a \neq b$, then by Lemma 3.4, the generalised circle through a, b, c meets $\sigma_1 \cup \sigma_2$ again in the unique point $d = \ominus(a \oplus b \oplus c)$. This implies $d \in P_2$ for all but at most $O(Kn^2)$ triples (a, b, c) with $a, b \in P_1$ and $c \in P_2$. Applying Corollary 5.4 with $A = B = P_1, C = P_2$ and $D = \ominus P_2$, we get cosets $H \oplus x$ and $H \oplus y$ of $\sigma_1 \cup \sigma_2$ such that $|P_1 \triangle (H \oplus x)|, |P_2 \triangle (H \oplus y)| = O(K)$ and $2x \oplus 2y \in H$, where $x \in \sigma_1$ and $y \in \sigma_2$. It follows that $H \leqslant \sigma_1$, hence H is a cyclic group of order $m = n/2 \pm O(K)$, and $H \oplus x$ and $H \oplus y$ are the vertex sets of regular m-gons inscribed in σ_1 and σ_2, respectively, either 'aligned' or 'offset' depending on whether $x \oplus y \in H$ or not. □

Together these lemmas prove Theorem 1.5. It just remains to remark that if P differs in $O(K)$ points from a coset on an acnodal circular cubic, then we apply inversion in its singularity. By Proposition 2.6, we obtain that P differs in $O(K)$ points from a coset $H \oplus x$ of a finite subgroup H of an ellipse, where $4x = o$. Thus, x is a point of the ellipse with eccentric angle a multiple of $\pi/2$. After a rotation, we can assume that $x = o$, which is Case (ii) of Theorem 1.5. □

6 Extremal Configurations

In this section we prove Theorems 1.1, 1.2, and 1.3. We first consider generalised circles.

6.1 Ordinary Generalised Circles

Suppose P is an n-point set in $\mathbb{R}^2$ spanning fewer than $n^2/2$ ordinary generalised circles, and that P is not contained in a generalised circle. Applying Theorem 1.5, we can conclude that, up to inversions, P differs in $O(1)$ points from one of the following examples: points on a line, a coset of a subgroup of an acnodal or smooth circular cubic, or a double polygon.

The first type of set is very easy to handle. Note that the lower bound is on the number of ordinary circles, not counting 3-point lines.

Lemma 6.1 *Let $K \geqslant 1$ and $n \geqslant 2K+4$. If all except K points of a set $P \subset \mathbb{R}^2$ of n points lie on a line, then P spans at least $\binom{n-1}{2}$ ordinary circles.*

Proof Let ℓ be a line such that $|P \cap \ell| = n - K$. For any $p \in P \cap \ell$ and $q \in P \backslash \ell$ there are at most $K-1$ non-ordinary circles through p, q, another point on $P \cap \ell$, and another point in $P \backslash \ell$. Therefore, there are at least $K(n-2K)$ ordinary circles through p. This holds for any of the $n-K$ points $p \in P \cap \ell$, and we obtain at least $K(n-2K)(n-K)/2$ ordinary circles. It is easy to see that when $1 \leqslant K \leqslant (n-4)/2$, $K(n-2K)(n-K)/2$ is minimised when $K = 1$. □

Cosets on cubics are also relatively easy to handle. We again obtain a lower bound on the number of ordinary circles, not including 3-point lines.

Lemma 6.2 *Suppose $P \subset \mathbb{R}^2$ differs in K points from a coset $H \oplus x$ of an acnodal or smooth circular cubic, where $|H| = n \pm O(K)$ and $4x \ominus \omega \in H$. Then P spans at least $n^2/2 - O(Kn)$ ordinary circles.*

Proof Suppose that P differs in K points from $H \oplus x$. We know from Construction 4.2 that $H \oplus x$ spans $n^2/2 - O(n)$ ordinary circles, all of which are tangent to γ. We show that adding or removing K points destroys no more than $O(Kn)$ of these ordinary circles, so that the resulting set P still spans at least $n^2/2 - O(Kn)$ ordinary circles.

Suppose we add a point $q \notin H \oplus x$. For $p \in H \oplus x$, at most one circle tangent to γ at p can pass through q. Thus, adding q destroys at most n ordinary circles. Now suppose we remove a point $p \in H \oplus x$. Since ordinary circles of $H \oplus x$ correspond to solutions of $2p \oplus q \oplus r = \omega$ or $p \oplus 2q \oplus r = \omega$, there are at most $O(n)$ solutions for a fixed p. Thus removing p destroys at most $O(n)$ ordinary circles.

Repeating K times, we see that adding or removing K points to or from $H \oplus x$ destroys at most $O(Kn)$ ordinary generalised circles out of the $n^2/2 - O(n)$ spanned by $H \oplus x$. This proves that P spans at least $n^2/2 - O(Kn)$ ordinary circles. □

From the two lemmas above we know that there is an absolute constant C such that a set of n points, not all collinear or concyclic, spanning at most $n^2/2 - Cn$ ordinary generalised circles, differs in $O(1)$ points from Case (iv) in Theorem 1.5. This case, where P is close to the vertex set of a double polygon, requires a more careful analysis of the effect of adding or removing points.

We use the following special case of a result due to Raz et al. [20].

Proposition 6.3 *If $P \subset \mathbb{R}^2$ is a set of n points contained in two circles, then the number of lines with at least three points of P is at most $O(n^{11/6})$.*

Proof Denote the two circles by σ_1 and σ_2. We use [20, Thm. 6.1], which states that for (not necessarily distinct) algebraic curves C_1, C_2, C_3 of constant degree, and finite sets $S_i \subset C_i$, the number of collinear triples $(p_1, p_2, p_3) \in S_1 \times S_2 \times S_3$, with p_1, p_2, p_3 distinct, is bounded by $O(|S_1|^{1/2}|S_2|^{2/3}|S_3|^{2/3} + |S_1| + |S_1|^{1/2}|S_2| + |S_1|^{1/2}|S_3|)$, unless $C_1 \cup C_2 \cup C_3$ is a line or a cubic. Let $C_1 = \sigma_1$ and $C_2 = C_3 = \sigma_2$. Set $S_i = P \cap C_i$ for $i = 1, 2, 3$. Every line with at least one point of S_1 and two points of $S_2 = S_3$ corresponds to a collinear triple in $S_1 \times S_2 \times S_3$. Since the union of two circles is not a line or a cubic, we can apply the theorem to get the bound $O(n^{11/6})$

for the number of collinear triples in P with one point in σ_1 and two points in σ_2. Similarly, the number of collinear triples in P with one point in σ_2 and two points in σ_1 is also $O(n^{11/6})$. Since a line intersects $\sigma_1 \cup \sigma_2$ in at most four points, we also obtain the bound $O(n^{11/6})$ for the number of lines with at least three points. □

Lemma 6.4 *Let S be a double polygon with m points on each circle. Let $P = (S\backslash A) \cup B$ be a set of n points, where A is a subset of S with $a = O(1)$ points and B is a set disjoint from S with $b = O(1)$ points. Then P spans at least $(2 + a + 4b)n^2/8 - O(n^{11/6})$ ordinary generalised circles.*

Proof We know from Constructions 4.3 and 4.4 that S spans $n^2/4 - O(n)$ ordinary generalised circles.

Consider first the number of ordinary generalised circles spanned by $S\backslash A$. As we saw in Construction 4.5, removing a point $p \in S$ destroys at most $3m/2$ ordinary generalised circles spanned by S, and adds $m^2/2 - O(m) = n^2/8 - O(n)$ ordinary generalised circles. Noting that there are at most m 4-point generalised circles spanned by S that go through any two given points of A, we thus have by inclusion–exclusion that $S\backslash A$ determines at least $(1/4 + a/8)n^2 - O(n)$ ordinary generalised circles.

Now consider adding $q \in B$ to S. For any pair of points from $S\backslash A$, adding $q \in B$ creates a new ordinary generalised circle, unless the generalised circle through the pair and q contains three or four points of $S\backslash A$. We already saw that the number of ordinary generalised circles hitting a fixed point is $O(n)$, so it remains to bound the number of 4-point generalised circles of S that hit q. If q lies on one of the concentric circles, then no 4-point generalised circles hit q, so we can assume that q does not. Applying inversion in q reduces the problem to bounding the number of 4-point lines determined by a subset of two circles. By Proposition 6.3, this number is bounded by $O(n^{11/6})$, so p lies on at most $O(n^{11/6})$ of the 4-point generalised circles spanned by S. Adding q to S thus creates at least $\binom{n}{2} - O(n^{11/6})$ ordinary generalised circles. Note that each $p \in A$ that was removed destroys at most n of these circles.

Adding q to $S\backslash A$ also destroys at most $O(n)$ ordinary circles, since for each $p \in S$ there is only one circle tangent at p and going through q, and for each $p \in A$, at most m ordinary circles spanned by $S\backslash A$ go through p. Finally, since there are at most $2m$ circles through two points of B that also go through two points of $S\backslash A$, $P = (S\backslash A) \cup B$ spans at least $(1/4 + a/8 + b/2)n^2 - O(n^{11/6})$ ordinary generalised circles. □

Theorem 1.2 then follows easily from the lemmas above.

Proof of Theorem 1.2 Suppose that P is a set of n points in $\mathbb{R}^2$ with fewer than $n^2/2 - Cn$ ordinary generalised circles, where C is sufficiently large. Without loss of generality, n is also sufficiently large. By Lemmas 6.1 and 6.2, we need only consider the case where P differs by $O(1)$ points from a double polygon. In the notation of Lemma 6.4, we have $P = (S\backslash A) \cup B$ and $(2 + a + 4b)/8 < 1/2$, which implies that $a \leqslant 1$ and $b = 0$. So P is either equal to S, or is obtained from S by removing one point, which are exactly the cases in Constructions 4.3, 4.4, and 4.5 . In particular, the minimum number of ordinary generalised circles occurs in Construction 4.3 when $n \equiv 0 \pmod 4$, in Construction 4.5 when $n \equiv 1, 3 \pmod 4$, and in Constructions 4.3 and 4.4 when $n \equiv 2 \pmod 4$. □

6.2 Ordinary Circles

We now consider what happens if we do not count generalised circles that are lines, and prove Theorem 1.1.

Proof of Theorem 1.1 Let P be a set of n points not all on a line or a circle, with at most $n^2/2 - Cn$ ordinary circles, for a sufficiently large C. By a simple double counting argument, there are at most $n^2/6$ 3-point lines, so there are at most $2n^2/3 - O(n)$ ordinary generalised circles. By Theorem 1.5, up to inversions and up to $O(1)$ points, P lies on a line, an ellipse, a smooth circular cubic, or two concentric circles. By Lemmas 6.1 and 6.2, the first three cases give us at least $n^2/2 - O(n)$ ordinary circles, contrary to assumption. Therefore, we only need to consider the case where, when P is transformed by an inversion to P', we have $P' = (S \backslash A) \cup B$, where S is a double polygon ('aligned' or 'offset'), and $|A| = a$, $|B| = b$.

By Lemma 6.4, P' has at least $(2 + a + 4b)n^2/8 - O(n^{11/6})$ ordinary generalised circles, which gives us the inequality $(2 + a + 4b)/8 < 2/3$, which in turn gives us $a \leqslant 3$ and $b = 0$. Therefore, P' lies on two concentric circles, and P lies on the disjoint union of two circles or the disjoint union of a line and a circle.

Suppose that $a = 3$ (and $b = 0$). Then P' has $5n^2/8 - O(n)$ ordinary generalised circles. Those passing through the centre of the inversion that transforms P to P', are inverted back to straight lines passing through three points of P. As in the proof of Lemma 6.4, there are $n^2/8 - O(n)$ ordinary generalised circles that pass through any point of A. Also, we can use Lemma 6.5 below to show that there are at most $O(n)$ ordinary generalised circles spanned by $S \backslash A$ that intersect in the same point not in S. Indeed, by Lemma 6.5, there are at most $n/2$ ordinary generalised circles of S that intersect in the same point $p \notin S$. Furthermore, for each point $q \in A$ there are $O(n)$ generalised circles through p, q, and two more points of S. It follows that there are $O(n)$ ordinary generalised circles spanned by $S \backslash A$ through p.

Thus, if the centre of inversion is in A, P has $n^2/2 - O(n)$ ordinary circles, which is a contradiction if C is chosen large enough. On the other hand, if the centre of inversion is not in A, then P has $5n^2/8 - O(n)$ ordinary circles, also a contradiction.

Therefore, we have $a \leqslant 2$, which means that P' is a set of n points as in Constructions 4.3, 4.4, 4.5, or 4.6.

Next, suppose that n is even. If $a = 2$, then there are $n^2/2 - O(n)$ ordinary generalised circles and through both points of A there are $n^2/8 - O(n)$ ordinary generalised circles. If we invert in one of these points in A, we obtain a set with $3n^2/8 - O(n)$ ordinary circles (as in Construction 4.6), which is not extremal. Otherwise, $a = 0$, P' is as in Constructions 4.3 or 4.4, and there are at least $n^2/4 - n$ ordinary generalised circles if $n \equiv 0 \pmod 4$ and $n^2/4 - n/2$ if $n \equiv 2 \pmod 4$. Let p be the centre of the inversion that transforms P to P'. Then all the 3-point lines of P are inverted to ordinary circles in the double polygon P', all passing through p. By Lemma 6.5 below, there are at most $n/2$ ordinary circles that intersect in the same point not in P'. Thus, in P there at most $n/2$ 3-point lines, and the number of ordinary circles (not including lines) is at least $n^2/4 - 3n/2$ if $n \equiv 0 \pmod 4$ and $n^2/4 - n$ if $n \equiv 2 \pmod 4$, which match Construction 4.3 (and Construction 4.4 if $n \equiv 2 \pmod 4$), if

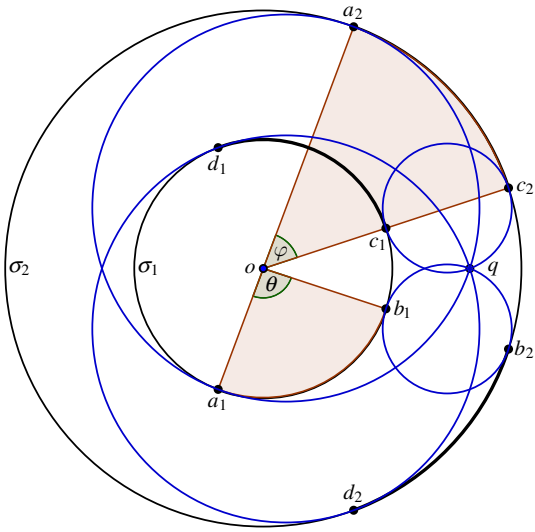

Fig. 6 Bitangent circles through q

the radii are chosen so that each vertex of the inner polygon has an ordinary generalised circle that is a straight line tangent to it.

Finally, suppose that n is odd. Then $a = 1$ and P' is as in Construction 4.5, with $3n^2/8 - O(n)$ ordinary generalised circles. It follows that P must be as in Construction 4.6, with $n^2/4 - 3n/4 + 1/2$ ordinary circles if $n \equiv 1 \pmod 4$ and $n^2/4 - 5n/4 + 3/2$ ordinary circles if $n \equiv 3 \pmod 4$. This finishes the proof. □

Lemma 6.5 *Let S be a double polygon ('aligned' or 'offset') with m points on each circle. Then a point $q \notin S$ lies on at most m ordinary generalised circles spanned by S.*

Proof Denote the inner circle by σ_1 and the outer circle by σ_2, both with centre o. We proceed by case analysis on the position of q with respect to σ_1 and σ_2. Note that for each point $p \in S$, at most one of the ordinary generalised circles tangent at p can go through q.

If q lies on either σ_1 or σ_2, then q does not lie on any ordinary generalised circle spanned by S.

If q lies inside σ_1, then q lies on at most m ordinary generalised circles spanned by S, since ordinary generalised circles tangent to σ_1 cannot pass through q. Similarly, if q lies outside σ_2, it lies on at most m ordinary generalised circles, since ordinary generalised circles tangent to σ_2 lie inside σ_2.

The remaining case to consider is when q lies in the annulus bounded by σ_1 and σ_2. Consider the subset $S' \subset S$ of points p such that there exists an ordinary generalised circle tangent at p going through q. Consider the four circles passing through q and tangent to both σ_1 and σ_2. They touch σ_1 at a_1, b_1, c_1, d_1 and σ_2 at a_2, b_2, c_2, d_2 as in Fig. 6. Any circle through q tangent to σ_1 and intersecting σ_2 in two points, must touch σ_1 on one of the open arcs a_1b_1 or c_1d_1. Similarly, any circle through q tangent to σ_2 and intersecting σ_1 in two points, must touch σ_2 on one of the open arcs a_2c_2 or

b_2d_2. It follows that S' must be contained in the relative interiors of one of these four arcs. Since S consists of m equally spaced points on each of σ_1 and σ_2,

$$|S'| < \left\lceil \frac{2m(\angle\, a_1ob_1 + \angle\, c_1od_1 + \angle\, b_2od_2 + \angle\, a_2oc_2)}{4\pi} \right\rceil = \left\lceil \frac{m(\theta+\varphi)}{\pi} \right\rceil,$$

where θ and φ are as indicated in Fig. 6. In order to show that $|S'| \leqslant m$, it suffices to show that the angle sum $\theta + \varphi$ is strictly less than π. This is clear from Fig. 6 (note that a_1, o, a_2 are collinear with a_1 and a_2 on opposite sides of o). □

6.3 Four-Point Circles

Proof of Theorem 1.3 Let P be a set of n points in $\mathbb{R}^2$ with at least $n^3/24 - 7n^2/24 + O(n)$ 4-point generalised circles. Let t_i denote the number of i-point lines ($i \geqslant 2$) and s_i the number of i-point circles ($i \geqslant 3$) in P. By counting unordered triples of points, we have

$$\binom{n}{3} = \sum_{i \geqslant 3} \binom{i}{3}(t_i + s_i) \geqslant t_3 + s_3 + 4(t_4 + s_4),$$

hence

$$\frac{1}{6}n^3 - O(n^2) \geqslant t_3 + s_3 + 4\left(\frac{1}{24}n^3 - O(n^2)\right)$$

and $t_3 + s_3 = O(n^2)$, so we can apply Theorem 1.5. We next consider each of the cases of that theorem in turn.

If all except $O(1)$ points of P lie on a straight line, it is easy to see that P determines only $O(n^2)$ generalised circles, contrary to assumption.

If all except $O(1)$ are vertices of two regular m-gons on concentric circles where $m = n/2 \pm O(1)$, then we know from Constructions 4.3, 4.4, and 4.5 that P determines at most $n^3/32 + O(n^2)$ 4-point generalised circles, again contrary to assumption.

Suppose next that $P = ((H \oplus x)\backslash A) \cup B$, where H is a finite subgroup of order $m = n \pm O(1)$ of a smooth circular cubic, A is a subset of $H \oplus x$ with $a = O(1)$ points, and B is a set disjoint from $H \oplus x$ with $b = O(1)$ points. Then $n = m - a + b$. The number of 4-point generalised circles in $H \oplus x$ is $m^3/24 - m^2/4 + O(m)$. We next determine an upper bound for the number of 4-point generalised circles in P.

For each $p \in A$, let C_p be the set of 4-point generalised circles of $H \oplus x$ that pass through p. Then $|C_p| = m^2/6 - O(m)$ and $|C_p \cap C_q| = O(m)$ for distinct $p, q \in A$. By inclusion–exclusion, we destroy at least $\left|\bigcup_{p\in A} C_p\right| \geqslant am^2/6 - O(m)$ 4-point generalised circles by removing A, and we still have at most $m^3/24 - m^2/4 - am^2/6 + O(m)$ 4-point generalised circles in $(H \oplus x)\backslash A$.

For each $p \in B$, the number of ordinary generalised circles spanned by $H \oplus x$ passing through p is at most $O(m)$. This is because each such generalised circle is tangent to the cubic at one of the points of $H \oplus x$, and there is only one generalised

circle through p and tangent at a given point of $H \oplus x$. Also, for each pair of distinct $p, q \in B$, there are at most $O(m)$ generalised circles through p and q and two points of $H \oplus x$; and for any three $p, q, r \in B$ there are at most $O(1)$ generalised circles through p, q, r and one point of $H \oplus x$. Therefore, again by inclusion–exclusion, by adding B we gain at most $O(m)$ 4-point generalised circles.

It follows that the number of 4-point generalised circles determined by P is

$$t_4 + s_4 \leqslant \frac{1}{24}m^3 - \frac{1}{4}m^2 - \frac{1}{6}am^2 + O(m) = \frac{n^3 - (a+3b+6)n^2 + O(n)}{24}.$$

Since we assumed that

$$t_4 + s_4 \geqslant \frac{n^3 - 7n^2 + O(n)}{24},$$

we obtain $a + 3b < 1$. Therefore, $a = b = 0$ and $P = H \oplus x$. The maximum number of 4-point circles in a coset has been determined in Constructions 4.1 and 4.2.

The final case, when all but $O(1)$ points of P lie on an ellipse, can be reduced to the previous case. Indeed, by Lemma 2.6, if we invert the ellipse in a point on the ellipse, we obtain an acnodal circular cubic, and then the above analysis holds verbatim for the group of regular points on this cubic. □

Acknowledgements M. Makhul and J. Schicho were supported by the Austrian Science Fund (FWF): DKW1214, subproject DK9. H.N. Mojarrad and F. de Zeeuw were partially supported by Swiss National Science Foundation grants 200020-165977 and 200021-162884.

References

1. Bálintová, A., Bálint, V.: On the number of circles determined by n points in the Euclidean plane. Acta Math. Hung. **63**(3), 283–289 (1994)
2. Ball, S.: On sets defining few ordinary planes. arXiv:1606.02138 (2016)
3. Ball, S., Monserrat, J.: A generalisation of Sylvester's problem to higher dimensions. J. Geom. doi:10.1007/s00022-016-0357-8
4. Basset, A.B.: An Elementary Treatise on Cubic and Quartic Curves. Cambridge Deighton Bell, London (1901)
5. Berger, M., Pansu, R., Berry, J.-P., Saint-Raymond, X.: Problems in Geometry. Problem Books in Mathematics. Springer, New York (1984)
6. Blair, D.E.: Inversion Theory and Conformal Mappings. Student Mathematical Library, vol. 9. American Mathematical Society, Providence (2000)
7. Boys, T., Valculescu, C., de Zeeuw, F.: On the number of ordinary conics. SIAM J. Discrete Math. **30**(3), 1644–1659 (2016)
8. Brass, P., Moser, W., Pach, J.: Research Problems in Discrete Geometry. Springer, New York (2005)
9. Coxeter, H.S.M.: Introduction to Geometry. Wiley Classics Library. Wiley, New York (1969)
10. Czapliński, A., Dumnicki, M., Farnik, Ł., Gwoździewicz, J., Lampa-Baczyńska, M., Malara, G., Szemberg, T., Szpond, J., Tutaj-Gasińska, H.: On the Sylvester–Gallai theorem for conics. Rend. Sem. Mat. Univ. Padova (in press). arXiv:1411.2648

11. Eisenbud, D., Green, M., Harris, J.: Cayley–Bacharach theorems and conjectures. Bull. Am. Math. Soc. (N.S.) **33**(3), 295–324 (1996)
12. Elliott, P.D.T.A.: On the number of circles determined by n points. Acta Math. Acad. Sci. Hung. **18**(1–2), 181–188 (1967)
13. Green, B., Tao, T.: On sets defining few ordinary lines. Discrete Comput. Geom. **50**(2), 409–468 (2013)
14. Hilton, H.: Plane Algebraic Curves. The Clarendon Press, Oxford (1920)
15. Joachimsthal, F.: Démonstration d'un théorème de Mr. Steiner. J. Reine Angew. Math. **36**, 95–96 (1848)
16. Johnson, W.W.: Classification of plane curves with reference to inversion. The Analyst **4**, 42–47 (1877)
17. Klee, V., Wagon, S.: Old and New Unsolved Problems in Plane Geometry and Number Theory. The Dolciani Mathematical Expositions, vol. 11. Mathematical Association of America, Washington, DC (1991)
18. Motzkin, T.: The lines and planes connecting the points of a finite set. Trans. Am. Math. Soc. **70**(3), 451–464 (1951)
19. Nassajian Mojarrad, H., de Zeeuw, F.: On the number of ordinary circles. arXiv:1412.8314 (2014)
20. Raz, O.E., Sharir, M., de Zeeuw, F.: Polynomials vanishing on Cartesian products: The Elekes–Szabó theorem revisited. Duke Math. J. **165**(18), 3517–3566 (2016)
21. Silverman, J.H.: The Arithmetic of Elliptic Curves. Graduate Texts in Mathematics, vol. 106, 2nd edn. Springer, Dordrecht (2009)
22. Silverman, J.H., Tate, J.: Rational Points on Elliptic Curves. Undergraduate Texts in Mathematics, 2nd edn. Springer, New York (1992)
23. Walker, R.J.: Algebraic Curves. Springer, New York (1978)
24. Werner, T.R.: Rational Families of Circles and Bicircular Quartics. Ph.D. Thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (2011). https://opus4.kobv.de/opus4-fau/files/2301/ThomasWernerDissertation.pdf
25. Wiseman, J.A., Wilson, P.R.: A Sylvester theorem for conic sections. Discrete Comput. Geom. **3**(4), 295–305 (1988)
26. Zhang, R.: On the number of ordinary circles determined by n points. Discrete Comput. Geom. **46**(2), 205–211 (2011)