Asymptotes, Cubic Curves, and the Projective Plane

Jeffrey Nunemacher

Asymptotes, Cubic Curves, and the Projective Plane

JEFFREY NUNEMACHER
Ohio Wesleyan University
Delaware, OH 43015

1. Introduction

Among the most beautiful and naturally appealing mathematical objects are the various plane curves. It is a pity that our undergraduates encounter so few of them. One extensive class of curves, which played a role in the recent proof of Fermat's Last Theorem, is the class of cubic curves, i.e., curves defined by an equation $P(x, y) = 0$, where P is a polynomial in x and y of total degree three. Famous ancient examples, which can be explored using simple analytic techniques (see, for example, [8]), are the folium of Descartes $x^3 + y^3 - 3xy = 0$, the witch of Maria Agnesi $y(1 + x^2) = 1$, the cissoid of Diocles $y^2(2 - x) = x^2$, and the Fermat curve $x^3 + y^3 = 1$. Using a classical formula to express the roots of a cubic equation in terms of its coefficients, it is possible to solve for y in terms of x. The resulting functions are usually not easy to sketch by hand using standard methods of calculus, but software such as *Derive* or *Mathematica* makes it possible to study cubic curves in a computer laboratory. Such a study requires knowledge and care, since the packages often use formulas that select complex branches; hence they can miss certain real branches of the curve.

Newton studied the general cubic equation in two variables and classified irreducible cubic curves into 72 different species. Here *irreducible* means that the polynomial defining the curve does not factor as a product of lower degree polynomials. For example, the curve defined by $x^3 - x^2 y - xy + y^2 = 0$ is reducible, since its defining polynomial factors as $(x^2 - y)(x - y)$; this curve is the union of a parabola and a straight line. In fact, Newton missed 6 species—according to his classification scheme (which allows affine coordinate changes), there are a total of 78 species. It makes a good project in a calculus course to explore the diversity of cubic curves and to reconsider Newton's classification. For suggestions on how this might be done making use of both classical algebra and modern technology, see [6].

Newton's classification begins by studying the asymptotic behavior of cubic curves. This approach is very natural, since the behavior "at infinity" is a dominant feature of the shape of any curve. But asymptotes can be far from obvious on a computer-generated graph. The folium of Descartes $x^3 + y^3 - 3xy = 0$ is shown in FIGURE 1 together with its asymptote $x + y + 1 = 0$. If the line were not drawn, would you be confident that the folium has an asymptote, or of the asymptote's exact location? It is an
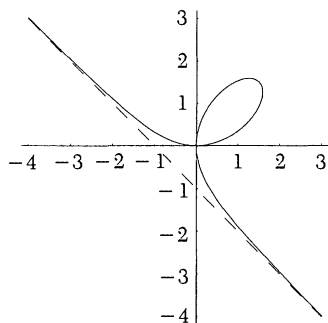

FIGURE 1

interesting and somewhat nonstandard exercise in calculus to verify that this line is asymptotic to the folium (see [8], p. 512). How does one find the asymptotes of a cubic curve or, more generally, of an algebraic curve of degree n? For such curves y is not given explicitly in terms of x so standard limiting techniques do not apply.

There is a theorem involving only polynomial algebra that almost answers this question. It provides a set of at most n lines that are the only possible asymptotes, and these lines will be asymptotic to the curve except in one rare situation. Special cases of this theorem used to be part of the standard repertoire of mathematical techniques that students learned when studying analytic geometry or the theory of equations. Fashions change, however, and this result is now encountered only in algebraic geometry, if at all. One reason for its disappearance is the unfortunate decline in interest and knowledge of geometry in high school and college. Another lies in the fact that the natural domain for thinking about asymptotes is the projective plane, which is often not studied at the undergraduate level. A third is that complications result from looking only at the real portion of an algebraic curve, which is best viewed as an object in complex space. In this article we present this theorem with some background in projective geometry and apply it to study cubic curves.

The theorem specifying the asymptotes of an algebraic curve is not easy to locate in current references, although at one time it must have been well known. A special case occurs explicitly in [7], pp. 8–10, but I have been unable to locate the general case in any reference. The nonsingular case of the theorem is treated nicely in [2]. An approach to finding asymptotes using only calculus can be found in [5], where it is applied to study various examples but not to find a theorem yielding all asymptotes. Classical methods and many examples with beautiful hand-drawn graphs can be found in [3]. Good modern references for many concrete facts in elementary algebraic geometry (but not this one!) are [1] and [4].

2. Statement of the theorem and applications

In this paper *asymptote* always refers to a line that is approached by points (x, y) on a branch of a curve as x or y becomes unbounded. This is the kind of asymptote encountered in a calculus course, but there it is almost always horizontal or vertical. A degenerate case, which we shall exclude from now on, occurs when the curve contains a line as a component, i.e., when the asymptote is actually part of the curve. Thus, in what follows, we assume that the defining polynomial $P(x, y)$ does not vanish identically on any line. The term *asymptotic direction* refers to a vector parallel to such a line so, in particular, the location of the line in the plane is not specified.

The following theorem specifies at most n candidate lines to be real asymptotes to a curve defined by $P(x, y) = 0$, where $P(x, y)$ is a polynomial of total degree n in the variables x and y. Such a curve is called an (affine) algebraic curve of degree n. Let us denote by $P_k(x, y)$ the sum of all terms occurring in $P(x, y)$ of total degree k. Then $P(x, y)$ can be expressed as $\sum_{k=0}^{n} P_k(x, y)$. The polynomials $P_k(x, y)$ are homogeneous of degree k; this means that, for any scalar λ, we have $P_k(\lambda x, \lambda y) = \lambda^k P_k(x, y)$. Polynomials such as $P_k(x, y)$, which are homogeneous of some degree k, are sometimes called *forms*.

Each vanishing form $P_k(x, y)$ factors over the complex numbers into a product of k linear factors, which are unique up to scalar multiple. The existence of such a factorization is a direct consequence of the Fundamental Theorem of Algebra. To see this, divide $P_k(x, y)$ by the term x^k, and replace the powers of y/x by powers of a new variable u. The resulting polynomial of one variable $p_k(u)$ has k complex roots and factors completely over the complex numbers. Suppose that $bu + a$ is one of the

factors of $p_k(u)$. When u is replaced by y/x, this term gives rise to the factor $ax + by$ of $P_k(x, y)$. If $bu + a$ occurs to exact multiplicity m as a factor of $p_k(u)$, then $P_k(x, y)$ can be expressed as $(ax + by)^m Q(x, y)$, where $Q(x, y)$ is a homogeneous polynomial of degree $k - m$ with $Q(b, -a) \neq 0$.

MAIN THEOREM. *Suppose that $ax + by$ is a factor of the top degree form $P_n(x, y)$ of multiplicity m with a and b real. Let $r \leq m$ denote the largest integer with the property that there exist polynomials $Q_j(x, y)$ for $n - r + 1 \leq j \leq n$ satisfying the conditions:*

$$P_n(x, y) = (ax + by)^r Q_n(x, y), \ P_{n-1}(x, y) = (ax + by)^{r-1} Q_{n-1}(x, y), \quad \text{(A)}$$
$$\dots, \text{and finally } P_{n-r+1}(x, y) = (ax + by)Q_{n-r+1}(x, y).$$

Then associated with the factor $ax + by$ is a set of at most r possible asymptotes $ax + by = t_0$, where t_0 is a real root of the equation

$$t^r Q_n(b, -a) + t^{r-1} Q_{n-1}(b, -a) + \cdots + t Q_{n-r+1}(b, -a) + P_{n-r}(b, -a) = 0. \quad \text{(B)}$$

All real asymptotes to the curve defined by $P(x, y) = 0$ arise in this way as $ax + by$ ranges over the real linear factors of $P(x, y)$. If $r > 1$ it may happen that some of these lines are spurious asymptotes.

Equation (B) has at most r roots, which may be complex or have multiplicity greater than one. Since $r \leq m$, the multiplicity of $ax + by$ as a factor of $P_n(x, y)$, the total number of possible asymptotes cannot exceed n. There is an actual asymptote associated with the factor $ax + by$ for each distinct real root except in the case discussed below.

The candidate asymptotes thus satisfy the equation

$$(ax + by)^r Q_n(b, -a) + (ax + by)^{r-1} Q_{n-1}(b, -a) + \cdots$$
$$+ (ax + by)Q_{n-r+1}(b, -a) + P_{n-r}(b, -a) = 0. \quad \text{(C)}$$

Condition (A) is a divisibility condition requiring that descending powers of $ax + by$ must be factors of the top r forms $P_k(x, y)$. Since r is the largest such integer, $ax + by$ does not further divide $P_{n-k}(x, y)$ or $Q_k(x, y)$ for some k between $n - r + 1$ and n, i.e., $P_{n-r}(b, -a)$ or at least one of these $Q_k(b, -a)$'s is nonzero. The most common situation is covered by the following simpler result in which the candidate line is guaranteed to be an asymptote. It is a special case of the Main Theorem.

COROLLARY. *If $ax + by$ is a simple factor of $P_n(x, y)$, i.e., if $P_n(x, y) = (ax + by)Q_n(x, y)$ with $Q_n(b, -a) \neq 0$, then associated with this factor is the single asymptote to $P(x, y) = 0$ defined by the equation*

$$(ax + by)Q_n(b, -a) + P_{n-1}(b, -a) = 0. \quad \text{(D)}$$

We give some examples of the application of this theorem and its corollary.

Example 1. For the folium of Descartes $x^3 + y^3 - 3xy = 0$, which is displayed in FIGURE 1, the sole real linear factor of $P_3(x, y) = x^3 + y^3$ is $x + y$ with $Q_3(x, y) = x^2 - xy + y^2$. Here we have $r = m = 1$ with $a = b = 1$ and $P_2(x, y) = -3xy$. Thus the single asymptote is given by (D), namely, $(x + y)Q_3(1, -1) + P_2(1, -1) = 0$, i.e., $3x + 3y + 3 = 0$.

Example 2. Consider the curve $xy^2 - 2.45y = x^3 - 5x^2 + 8.75x - 6.25$; the coefficients have been chosen to present a "typical" nontrivial cubic curve. Here $P_3(x, y) = xy^2 - x^3$ which factors as $x(y + x)(y - x)$ and $P_2(x, y) = 5 x^2$. Then the Corollary asserts that there are three asymptotes, namely, $x = 0$, $y = x - 2.5$, and $y = -x + 2.5$. See FIGURE 2 for a graph of this curve and its asymptotes. One
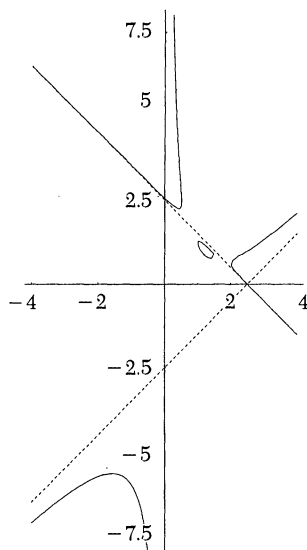
FIGURE 2

limitation of computer graphics is illustrated well by this figure. While the rightmost branch of this curve appears to be asymptotic to the line $x + y = 2.5$ from above, a more careful analysis shows that the curve actually crosses the line at the point $(2.5, 0)$ and is then asymptotic from below. This crossing becomes evident in the graph only with a tenfold increase in scale, so it would not be noticed unless it was deliberately sought.

Example 3. The quartic curve defined by $(x^3 - x)y = 1$ can be analyzed completely using techniques of calculus, since it is easy to solve for y in terms of x. But our theorem applies as well, with the following results. Here $P_4(x, y) = x^3 y$, so the factor x has multiplicity three. Also x^2 divides $P_3(x, y) = 0$ and x divides $P_2(x, y) = -xy$, so $r = 3$ with $a = 1$ and $b = 0$. From (B) we obtain the equation $t^3(-1) + t^2(0) + t(1) + 0 = 0$, i.e., $-t^3 + t = 0$. Solving this equation yields the three parallel candidate asymptotes $x = 0$, $x + 1$, and $x - 1 = 0$, which can easily be verified to be true asymptotes using limits. The other factor of $P_4(x, y) = x^3 y$ is y with multiplicity $m = 1$. Equation (D) now gives $y = 0$ as the only other asymptote. Thus this curve has three asymptotes in the direction $\langle 0, 1 \rangle$ and one in the direction $\langle 1, 0 \rangle$.

Example 4. Consider the parabola $x^2 - y = 0$, which we know has no asymptotes from basic analytic geometry. The factor x has multiplicity two in $P_2(x, y) = x^2$, but x does not divide $P_1(x, y) = -y$. Thus $r = 1$ with $a = 1$ and $b = 0$. We obtain from (C) the equation $x(0) - 1 = 0$. This linear equation does not describe a line, so we confirm there are no (finite) asymptotes to the parabola.

Example 5. Finally, let us analyze the curve $x^2 y^2 - y^2 + 1 = 0$. It is easy to solve this equation for y in terms of x. We find that the curve is the union of the graphs of the two functions $f_{\pm}(x) = \pm 1 / \sqrt{1 - x^2}$. Thus, using limits, we see that there are exactly two asymptotes, vertical ones at $x = \pm 1$. (See FIGURE 3 below.) Applying the theorem to this example, we see that the leading term $P_4(x, y) = x^2 y^2$ has the two factors x and y, each of which yields possible asymptotes with $r = 2$ (since $P_3 = 0$). The factor x gives rise to the two asymptotes already noticed. The factor y (with $a = 0$ and $b = 1$) produces an equation (B) of the form $t^2\ 1 = 0$. Thus we obtain the candidate line $y = 0$, which is clearly not an asymptote for the curve. This example
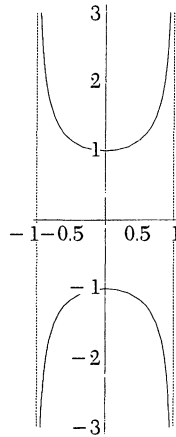
FIGURE 3

shows that the lines provided by the theorem may not be asymptotes if no real branch of the curve approaches them. We shall explore this situation in more detail below.

The conclusions of the theorem become simpler and more pleasing if we move to a space larger than $\mathbb{R}^2$. If we allow a and b to be complex numbers, equations (B) and (D) give a criterion for complex asymptotes, which are two-dimensional real planes in $\mathbb{C}^2$ (see [3], p. 44). For instance, the circle $x^2 + y^2 = 1$ has $x + iy = 0$ and $x - iy = 0$ as complex asymptotes in the four-dimensional space $\mathbb{C}^2$, but these planes intersect the real plane $\mathbb{R}^2$ only at the origin. A different extension is relevant in the situation of Example 4. There the asymptotic equation does not describe a real line. A line is actually present, but it is the line at infinity in the projective plane $\mathbb{RP}^2$ (which contains $\mathbb{R}^2$). We shall return to this situation below when the necessary definitions have been made. A combination of these two extensions is necessary to explain Example 5. In a suitable space there are always exactly n asymptotes, if we interpret "asymptote" correctly and assign multiplicities to asymptotes $ax + by = t_0$ according to the number of times that t_0 occurs as a root when (B) is factored over $\mathbb{C}$. In $\mathbb{R}^2$, however, n provides only an upper bound for the number of finite real asymptotes.

For cubic curves, therefore, there can be no more than three asymptotes. In fact, cubic curves exist with 0, 1, 2, or 3 real asymptotes. The curve $yx(x - 1) = 1$ has three asymptotes; $yx^2 = 1$ has two; the folium of Descartes has one, as we saw above; and the polynomial $y = x^3$ has no finite asymptotes.

Notice that for a curve $P(x, y) = 0$ of degree n the possible asymptotic directions $\langle b, -a \rangle$ are defined by the factors $ax + by$ of the top degree term $P_n(x, y)$. It is intuitively reasonable that the dominant term should determine the behavior of the curve at infinity, i.e., the asymptotic directions. However, where the asymptotes are situated in the plane is dependent on some of the lower order terms. In the case of a simple factor $ax + by$ only the next term $P_{n-1}(x, y)$ is relevant. This is the nonsingular case (in a sense to be defined below). For a factor of higher multiplicity $m > 1$, the next r terms in the homogeneous expansion are relevant to the asymptotic behavior, where r is defined by the theorem but is always bounded above by m.

3. Curves and asymptotes in the projective plane

It is natural to regard an asymptote to a curve as a line to which the curve is tangent "at infinity." This idea works very well for curves that are defined by the vanishing of a polynomial, since such curves extend in a simple fashion to a larger space

including points at infinity. In addition to simplifying the hunt for asymptotes, the addition of points at infinity to $\mathbb{R}^2$ to form the projective plane $\mathbb{RP}^2$ has various other benefits. Newton's original classification of irreducible cubic curves into 72 species was criticized by later authors as being too complicated to be useful. By regarding the curve as lying in the projective plane and enlarging the group of allowable coordinate changes to include all projective transformations, it can be simplified into a classification containing just 5 different species. The points at infinity, which are sometimes called ideal points, are rather intuitive—we simply introduce one additional point at which all parallel lines in a given direction meet. Thus railroad tracks meet at the horizon, but perhaps less intuitively, they meet at the *same* point in both directions. This point is specified by any nonzero vector $\langle m, n \rangle$ parallel to the line. Notice that for points satisfying $ax + by + c = 0$ with x or y large, only the ratio of a to b is important, rather than a, b, or c. This ratio is the information that a direction vector $\langle m, n \rangle$ contains.

There is a clean algebraic way to add points at infinity to $\mathbb{R}^2$. We consider the set of all 3-tuples (X, Y, Z) with real coordinates not all zero, and define an equivalence relation: $(X_1, Y_1, Z_1) \sim (X_2, Y_2, Z_2)$ if each triple is a scalar multiple of the other. These equivalence classes are defined to be the points in the real projective plane $\mathbb{RP}^2$. If the Z-coordinate of the 3-tuple (X, Y, Z) is nonzero, we may divide by it and obtain the equivalent 3-tuple $(X/Z, Y/Z, 1)$, which we identify with the Euclidean point (x, y), where $x = X/Z$ and $y = Y/Z$. Only when $Z = 0$ is this not possible, and it is these points which are the points at infinity. Notice that on the line $ax + by + c = 0$ the points escape to infinity as Z approaches 0, so the condition $Z = 0$ for the line at infinity is quite natural. Any nonzero vector $\langle x, y \rangle$ has exactly one point $(x, y, 0)$ associated with it, which we regard as the point at infinity in this direction. An ordinary point (x, y) in $\mathbb{R}^2$ is identified with the point given by the class containing the 3-tuple $(x, y, 1)$ in $\mathbb{RP}^2$.

This algebraic construction of $\mathbb{RP}^2$ has a simple geometric realization. Consider the closed upper unit hemisphere H in $\mathbb{R}^3$ and a plane T tangent to it at the north pole. (See FIGURE 4.) Each equivalence class of 3-tuples (X, Y, Z) in $\mathbb{RP}^2$ when $Z \neq 0$ has a
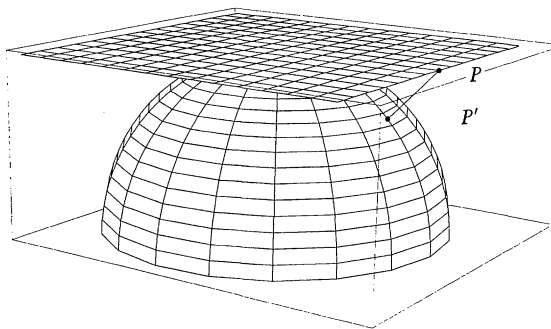

FIGURE 4

unique representative in H, obtained by dividing all components by $\sqrt{X^2 + Y^2 + Z^2}$ and multiplying by -1 (if necessary) to make the third component positive. When $Z = 0$ there are two representatives for (X, Y, Z) on the boundary circle B of H, which differ by a factor of -1. If we identify these antipodal points on the boundary unit circle, we have a model for $\mathbb{RP}^2$ that we can visualize easily, although we need to be careful to regard pairs of antipodal points as single entities. The correspondence with real points in $\mathbb{R}^2$ is obtained by regarding T as a copy of $\mathbb{R}^2$ and then projecting from T to H along rays from the origin of $\mathbb{R}^3$. Notice that this projection is

compatible with the equivalence relation $\sim$, i.e., each point P of T projects to a unique point P' on the open upper hemisphere contained in H.

Each line L in T projects to a great semi-circle in H which meets B in two antipodal points, which are identified as a single point in $\mathbb{RP}^2$. We regard this single point as the point at infinity on L. Lines parallel to L project to great semi-circles which meet at the same antipodal points on B; thus any two parallel lines meet at a single ideal point in $\mathbb{RP}^2$. Curves in T that are asymptotic to a line L approach the curve at infinity, i.e., they are tangent to the corresponding great semi-circle in H at the point at infinity which lies on it. There is an exceptional case that occurs when the tangent circle is the boundary circle B, which is not the image of any finite line in T. B consists of all points at infinity, so a curve in T whose projection is tangent to B at a point of B has no real asymptote. This happens, for instance, for the parabola $y = x^2$, as we shall see below. Thus we have a nice geometric criterion to detect asymptotes to curves: a curve has a real asymptote if and only if its image in H is tangent at a point of B to a great semicircle different from B.

The extended space $\mathbb{RP}^2$ is in many ways superior to the Euclidean plane $\mathbb{R}^2$. It has a natural topology, which is obtained by forming the quotient of $\mathbb{R}^3$ under the equivalence relation defined above. We obtain this same topology if we form the quotient space of H under the identification of the antipodal points of the boundary circle B. This second approach makes it clear that $\mathbb{RP}^2$ is compact, since it is the continuous image of the compact hemisphere H under the identification map. Each point in $\mathbb{RP}^2$ has a two-dimensional Euclidean neighborhood. This is obvious for all points that are images of the open upper hemisphere in H and is true for the points on B as well, either by noticing that the identification glues together two half discs to create a full Euclidean disc surrounding each such point, or by noticing that in the $\mathbb{R}^3$ construction of $\mathbb{RP}^2$ all points are created equal, so those points that are images of points on B cannot be topologically different from the other points.

The nonsingular linear transformations of $\mathbb{R}^3$ respect the defining equivalence relation, so they define a group of homeomorphisms of $\mathbb{RP}^2$. This group is transitive, since there is such a transformation mapping any nonzero point of $\mathbb{R}^3$ onto any other nonzero point. This transitivity make $\mathbb{RP}^2$ into a homogeneous space. For later work observe that these linear transformations are differentiable with nonvanishing Jacobian, since they are nonsingular. Thus they preserve the tangency of curves (even though they may change the angles at which nontangent curves meet). To establish the theorem specifying asymptotes, we shall use such a linear transformation to map a point at infinity to the origin, where calculations are more familiar. Finally, as noted above, the Euclidean plane $\mathbb{R}^2$ sits naturally in $\mathbb{RP}^2$ as those classes of 3-tuples containing a representative with Z coordinate equal to 1. In summary, $\mathbb{RP}^2$ is a homogeneous compact manifold in which $\mathbb{R}^2$ is naturally embedded.

Any curve in $\mathbb{R}^2$ defined by a polynomial equation $P(x, y) = 0$ of degree n extends naturally to a curve in $\mathbb{RP}^2$ as follows. Replace x by X/Z and y by Y/Z and multiply the entire equation by Z^n to clear the fractions. This procedure produces a homogeneous polynomial in the three variables X, Y, and Z. The resulting equation defines a curve in $\mathbb{RP}^2$, since a homogeneous polynomial has the same zero value at all scalar multiples of any 3-tuple at which it vanishes. For points with $Z \neq 0$ it restricts to the original curve in $\mathbb{R}^2$; thus it defines an extension of the curve $P(x, y) = 0$ into the projective plane. Such curves are called projective algebraic curves. They can be studied using a variety of classical and modern techniques, and form the basic objects of interest in algebraic geometry. The linear equation $ax + by + c = 0$ in $\mathbb{R}^2$, for example, extends to the homogeneous equation $aX + bY + cZ = 0$ in $\mathbb{RP}^2$. As long as not all of a, b, and c are zero, the original line extends to a line in the projective plane

by the addition of the single point $(b, -a, 0)$ at infinity. If $a = b = 0$ the line consists entirely of points at infinity and is called the line at infinity.

Our basic problem is to determine all the asymptotes to a general algebraic curve. As noted above, these are the finite lines that are tangent to the curve at a point of infinity in $\mathbb{RP}^2$. For algebraic curves there is an algebraic way to identify such tangent lines. It is based on the idea that a line is tangent to an algebraic curve at a finite point if it has higher "order of contact" at the point than do "generic" lines through the point. When the point is the origin and the curve is defined by a polynomial $P(x, y) = 0$ with $P(x, y) = \sum_{k=1}^{n} P_k(x, y)$ as in Section 2 (here we may start the summation at $k = 1$ since $P(0, 0) = 0$), the order of contact with a line is the least k for which $P_k(x, y)$ does not vanish identically on the line. Let l denote this order of contact. Since $P_l(x, y)$ is a form of degree l, at most l distinct lines will have order of contact greater than l, and all others will have order of contact l with the curve. The former set we declare to be the (algebraic) tangent lines to the curve at the origin. This approach is simple algebraically and allows us also to cope with curves that are singular at the origin.

To see the connection with the more familiar approach to tangent lines in calculus, notice that $P_l(x, y)$ is the lth degree Taylor expansion of $P(x, y)$ at $(0, 0)$. When $l = 1$ the curve is said to be nonsingular at $(0, 0)$; otherwise it is singular there. In the nonsingular case let us compute y' at $(0, 0)$ using implicit differentiation. Let $P_1(x, y) = ax + by$. Then we obtain $0 = P_x + P_y y' = a + by' +$ terms that evaluate to 0 at $(0, 0)$ because of the presence of x or y in the term. Thus $y' = -a/b$, so there is a unique tangent line at $(0, 0)$ given by $y = -ax/b$, i.e., by $P_1(x, y) = 0$. This argument justifies this method of finding tangent lines in the nonsingular case. For a discussion of the singular case see [3], pp. 22–24. In this situation the algebraic concept of tangent line does not always agree with our geometric one (because of the artificial restriction that we are looking only at the real portion of our curves). For example, the curve $x^2 y^2 = x^4 + y^2$ algebraically has the x-axis ($y^2 = 0$) as a tangent line at the origin, but the curve has the origin as an isolated point on the real graph, since $x^2 y^2 \geq y^2$ implies that $x^2 \geq 1$ unless $y = 0$ and the only point on the curve with $y = 0$ is the origin.

As asymptotes for real algebraic curves, we are interested in lines that are real, i.e., we work in $\mathbb{RP}^2$ and not in the corresponding complex projective space $\mathbb{CP}^2$, and in lines that are finite, i.e., not the line at infinity, $Z = 0$. The latter restriction explains the phenomenon that occurred in Example 4 above. It may happen in (B) that all the coefficients $Q_k(b, -a) = 0$ while $P_{n-r}(b, -a) \neq 0$. In this situation there is no finite asymptote in the direction $\langle b, -a \rangle$. If we projectivize the picture, this situation gives rise to the equation $P_{n-r}(b, -a)Z^r = 0$, which does define a line in $\mathbb{RP}^2$, namely, the line at infinity. So there *is* an asymptote in this case, just not a finite asymptote. This is the situation for all curves defined by $y = p(x)$, where $p(x)$ is a polynomial in x of degree greater than one. All such curves have the line at infinity as their only (ideal) asymptote.

We are now in a position to understand what happens in Example 5 above. As noted above, the curve $x^2 y^2 = x^4 + y^2$ has the origin as an isolated point, which gives rise to the spurious tangent $y = 0$, obtained by setting its lowest degree form y^2 equal to zero. The curve $x^2 y^2 - y^2 + 1 = 0$ of Example 5 was obtained from this curve by the transformation $[X, Y, Z] \to [Z, Y, X]$, which has the effect of mapping the origin $[0, 0, 1]$ to the point at infinity $[1, 0, 0]$. Thus this curve has a point at infinity as an isolated singular point and it has the x-axis as a spurious asymptote (as we saw in Example 5). Singularities of algebraic curves can be very complicated. But only in this rare situation of an isolated point at infinity in $\mathbb{RP}^2$, which is not isolated if one looks

at the entire curve in $\mathbb{CP}^2$, is there no portion of the real curve abutting onto the line that our method has identified. The fact that such candidate lines are not asymptotic will be evident from a machine-drawn graph.

4. Proof of the theorem

Consider now an algebraic curve $P(x, y) = 0$ of degree n, where the polynomial $P(x, y)$ is expressed as a sum of forms $P(x, y) = \sum_{k=0}^n P_k(x, y)$. Putting this equation into homogeneous coordinates, we obtain the equation

$$F(X, Y, Z) = P_n(X, Y) + ZP_{n-1}(X, Y) + Z^2 P_{n-2}(X, Y) + \cdots + Z^n P_0(X, Y) = 0,$$

which extends the affine curve to a projective curve in $\mathbb{RP}^2$. The extended curve, which we shall denote by A, contains points at infinity $(X, Y, 0)$ precisely when $F(X, Y, 0) = 0$, i.e., when $P_n(X, Y) = 0$. This occurs in those at most n directions $\langle b, -a \rangle$ for which $ax + by$ is a real linear factor of $P_n(x, y)$. We must now determine which of these directions yield lines in $\mathbb{R}^2$ that are possibly tangent to the curve at a point of infinity.

Let us fix a particular factor $ax + by$ of $P_n(x, y)$ to analyze. Without loss of generality we may assume that b is nonzero (since either a or b is nonzero). We shall use a linear transformation to map the point at infinity $[b, -a, 0]$ to $[0, 0, 1]$. This will enable us to do our calculations at the origin. Consider the transformation $T: \mathbb{RP}^2 \to \mathbb{RP}^2$ defined by $T[X, Y, Z] = [bZ, -aZ + Y, X]$. This transformation is invertible, with inverse $T^{-1}[X, Y, Z] = [bX, bY + aX, X]$, and takes the origin $[0, 0, 1]$ of the affine plane to the point at infinity $[b, -a, 0]$. This formula for the inverse was obtained by inverting the corresponding matrix and using homogeneity to simplify the expression. It is easy to check that $T^{-1} \circ T[X, Y, Z] = [bX, bY, bZ] \sim [X, Y, Z]$. The curve A defined by $F[X, Y, Z] = 0$ then "pulls back" under T to an associated curve A' defined by $F(T[X, Y, Z]) = F[bZ, -aZ + Y, X] = 0$. In terms of the homogeneous components of $P(X, Y, Z)$, the curve A' is defined by the equation $P_{n-j}(bZ, -aZ + Y)X^j = 0$. With r defined as in the statement of the theorem, we have

$$P_{n-j}(bZ, -aZ + Y) = (bY)^{r-j} Q_{n-j}(bZ, -aZ + Y) \quad \text{for} \quad j = 0, 1, \ldots, r,$$

(where for simplicity we have set Q_{n-r} equal to P_{n-r}). Our equation for A' now takes the form

$$\sum_{j=0}^r (bY)^{r-j} Q_{n-j}(bZ, -aZ + Y) X^j + \sum_{j=r+1}^n P_{n-j}(bZ, -aZ + Y) X^j = 0.$$

To study the tangents at the origin $[0, 0, 1]$ of A', we set $Z = 1$, $X = x$, and $Y = y$ to obtain the restriction of the curve to the affine plane $\mathbb{R}^2$. This yields the equation

$$\sum_{j=0}^r (by)^{r-j} Q_{n-j}(b, -a + y) x^j + \sum_{j=r+1}^n P_{n-j}(b, -a + y) x^j = 0.$$

This equation is a polynomial equation in x and y, and the tangent lines to the curve at the origin are given by those lines that satisfy the equation $L(x, y) = 0$, where $L(x, y)$ is the lowest order nonvanishing form in this polynomial. It is clear from this equation that the degree of $L(x, y)$ is r since at least one of the terms $Q_{n-j}(b, -a + y)$ has nonzero constant term $Q_{n-j}(b, -a)$ (by the definition of r in the statement of the theorem). Thus the lowest order nonvanishing form in the above

equation is given by

$$L(x, y) = \sum_{j=0}^{r} (by)^{r-j} x^j Q_{n-j}(b, -a). \tag{E}$$

The tangent lines to A' at the origin are the solutions to the equation $L(x, y) = 0$.

To find the tangent lines to our original curve A at the point at infinity $[b, -a, 0]$, we extend this equation to $\mathbb{RP}^2$ and pull back using T^{-1}. Notice that we are using the fact that T and T^{-1} preserve tangency of curves. Since L is homogeneous, the extension to $\mathbb{RP}^2$ is given by

$$H(X, Y, Z) = \sum_{j=0}^{r} (bY)^{r-j} X^j Q_{n-j}(b, -a) = 0.$$

Under T^{-1} this equation pulls back to the equation

$$0 = H(T^{-1}[X, Y, Z]) = H[bZ, bY + aX, X]$$

$$= \sum_{j=0}^{r} (b(bY + aX))^{r-j} (bZ)^j Q_{n-j}(b, -a)$$

$$= b^r \sum_{j=0}^{r} (aX + bY)^{r-j} Z^j Q_{n-j}(b, -a).$$

Since b is nonzero, this is equivalent to the equation

$$\sum_{j=0}^{r} (aX + bY)^{r-j} Z^j Q_{n-j}(b, -a) = 0.$$

Setting $Z = 1$, $X = x$, and $Y = y$, we obtain $\sum_{j=0}^{r}(ax + by)^{r-j} Q_{n-j}(b, -a) = 0$, which specifies the lines in the direction $\langle b, -a \rangle$ possibly asymptotic to the curve A. This equation is identical to (C) in the statement of the theorem.

Remarks. The ideas in [5] can be developed to yield a proof of the Main Theorem which is based entirely on limits and avoids the projective plane. But this route is somewhat circuitous, and does not emphasize the geometric interpretation of an asymptote as a line that is tangent to a curve at infinity. It is also possible to establish the Main Theorem by using an appropriate version of Taylor's theorem at the point at infinity $[b, -a, 0]$. This avoids the transformation T but introduces its own complications. The argument using T is essentially just linear algebra, and was suggested by one of the referees.

REFERENCES

1. Egbert Brieskorn and Horst Knörrer, *Plane Algebraic Curves*, Birkhäuser Verlag, Boston, MA, 1986.
2. Julian Lowell Coolidge, *A Treatise on Algebraic Plane Curves*, Clarendon Press, Oxford, UK, 1931.
3. Percival Frost, *An Elementary Treatise on Curve Tracing*, 2nd ed., Macmillan, London, UK, 1911.
4. Francis Kirwan, *Complex Algebraic Curves*, Cambridge University Press, Cambridge, UK, 1992.
5. Gabriel Klambauer, *Aspects of Calculus*, Springer-Verlag, New York, NY, 1986.
6. Jeffrey Nunemacher, Newton's investigation of cubic curves, in *Problems for Student Investigation*, MAA Notes Number 30, Math. Assoc. of America, Washington, DC, 1993.
7. E. J. F. Primrose, *Plane Algebraic Curves*, Macmillan, London, UK, 1955.
8. George Simmons, *Calculus with Analytic Geometry*, McGraw-Hill, New York, NY, 1985.