On the Approximation of Convex, Closed Plane Curves by Multifocal Ellipses

P. ERDÖS
I. VINCZE

Abstract

The question whether a convex closed curve can be approximated by ellipses having a large number of foci is considered. It is shown that the limiting, convex figure of multifocal ellipses may have only one single straight segment. This happens only in the case, when the foci tend partly to infinity and partly to points of the line through the straight segment. The approximations of certain 'distance integrals' are treated; the characterization of approximability remains an open problem.

CONVEX CURVES; MULTIFOCAL ELLIPSES; APPROXIMATION OF CONVEX CLOSED CURVES

Introduction

Many years ago Vázsonyi[†] [2] raised the question of whether a convex closed curve could be approximated arbitrarily by ellipses, possibly having a large number of foci. Let $F_1, F_2, \ldots, F_n$ be n points in the plane; we shall call the set of points P satisfying the relation

$$\sum_{i=1}^{n} \overline{PF_i} = C(> C_0)$$

a multifocal ellipse or a W_n-curve. Here

$$C_0 = \min_{(P)} \sum_{i=1}^{n} \overline{PF_i}.$$

It is well known that the function $f(P) = \sum_{i=1}^{n} \overline{PF_i}$ achieves its minimum C_0 in one and only one point in the plane, when the set $\{F_1, F_2, \cdots, F_n\}$ does not lie on a straight line, while the set of minimum points of $f(P)$ may be a linear set when

 0021-9002/82/050089-08 \$01.05

[†] E. Vázsonyi earlier published under the name E. Weiszfeld.

the foci lie on a line. For $C > C_0$ the W_n-curves are convex, closed curves filling out the whole plane.

The answer to the problem of Vázsonyi was in the negative. In [1], we proved that *an equilateral triangle cannot be approximated arbitrarily by W_n-curves.* We constructed a convex, closed curve containing a straight segment, which could be approximated by W_n-curves when one focus tended to ∞. But the question remained open as to whether a convex, closed curve containing two straight segments could be approximated by W_n-curves. We turn now to the justification of the statement that: *the limiting, convex figure of W_n-curves may have only one single straight segment.* Some related questions will also be considered.

1. Some simple general statements

1.1. Let P_1, P_2 and F be three points in the plane. Denote by P_0 the middle point of the segment P_1P_2. We have

$$\overline{P_1F}+\overline{P_2F}-\overline{2P_0F}=\frac{\overline{P_1P}_2^2-(\overline{P_1F}-\overline{P_2F})^2}{\overline{P_1F}+\overline{P_2F}+\overline{2P_0F}}\geqq 0 \tag{1.1}$$

with the equality holding in the cases when F lies on the straight line through P_1 and P_2, or when F lies at ∞.

The equality in (1.1) is a consequence of the elementary relation

$$4\overline{P_0F}^2+\overline{P_1P}_2^2=2(\overline{P_1F}^2+\overline{P_2F}^2),$$

from which we obtain

$$(\overline{P_1F}+\overline{P_2F})^2=\overline{2P_0F}^2+\tfrac{1}{2}\overline{P_1F}_2^2+\overline{2P_1F}\cdot\overline{P_2F}$$

and

$$(\overline{P_1F}-\overline{P_2F})^2=2\overline{P_0F}^2+\tfrac{1}{2}\overline{P_1P}_2^2-2\overline{P_1F}\cdot\overline{P_2F}.$$

From these, the equality in (1.1) follows immediately. The inequality (1.1) is a simple consequence of the triangle inequality.

1.2. We shall consider point sets $F_t=(F_{1t},F_{2t},\cdots,F_{nt})$ depending on a real parameter t $(0\leqq t<\infty)$. Let (a_{it},b_{it}) and (r_{it},φ_{it}) be the cartesian and polar coordinates respectively of the point F_{it}. We shall assume that when $t\to\infty$, $F_{it}\to F_i$, where the point F_{it} and all of its coordinates tend *continuously* to the limit points with the coordinates (a_i,b_i) and (r_i,φ_i). The number of points n may be allowed to increase, but this will not play an essential role in our considerations, except in Section 3. The multifocal ellipse with focus-system F_t will be denoted by W_{nt} and $\lim W_{nt}=W_n$ or W according to whether n remains fixed or itself tends to ∞.

A system of W_{nt}-curves $(0 \leqq t < \infty)$, i.e. the system of foci and the corresponding constants c_t, for which

$$\sum_{i=1}^{n} \overline{PF_{it}} = C_t \tag{1.2}$$

will usually be determined by the requirement that they pass through two fixed points P_1 and P_2 of the plane for all values of t. This does not determine the single W_{nt}-curves and the system uniquely but what we need is a continuous change of the curves, tending to a finite (bounded) curve. Because of the broad possibilities in the choice of the points, F_{it} when $n \geqq 3$, this requirement can always be fulfilled.

1.3. Let us consider W_{nt}-curves with focus system F_t going through the points P_1 and P_2 for each t:

$$\sum_{i=1}^{n} \overline{P_1F_{it}} = \sum_{i=1}^{n} \overline{P_2F_{it}}. \tag{1.3}$$

Then the following result holds.

Proposition 1.1. *The necessary condition for a system of W_{nt}-curves $(0 \leqq t < \infty)$ satisfying* (1.3) *to tend to a limiting convex figure containing the linear segment P_1P_2, is the validity of the relation*

$$\lim_{t \to \infty} \sum_{i=1}^{n} (\overline{P_1F_{it}} + \overline{P_2F_{it}} - 2\overline{P_0F_{it}}) = 0, \tag{1.4}$$

where P_0 denotes the middle point of the straight segment P_1P_2.

This statement is a consequence of the definition and of the convexity of the limiting curve.

The relation (1.4) is not sufficient. This can be seen as follows: suppose that the F_{it} (for $i = 1, 2, \cdots, n$) tend to infinity in different directions in such a way that for any value of t, the W_{nt}-curve passes through P_1 and P_2. Then according to (1.1) the relation (1.4) holds. But, as we shall see in Section 2, if we consider the curves with $F_t\,(F_{1t}\,(t, 0), F_{2t}\,(-t, 0), F_{3t}\,(0, t), F_{4t}\,(0, -t))$ going through the points $P_1(1, 1)$, $P_2(1, -1)$ (and by symmetry also through $P_3(-1, 1)$ and $P_4(-1, -1)$), these will tend to a circle pasing through the four points.

1.4. As a result of the above considerations, namely relations (1.1) and (1.4), we can conclude that the following theorem holds.

Theorem 1.1. *The W_{nt}-curve passing through the points P_1 and P_2 for all values of t, can approach a limiting figure containing the segment P_1P_2 only in the case, where the F_t tend partly to ∞ and partly to points on the line through P_1 and P_2 outside the open segment P_1P_2.*

In (1.4) all terms are non-negative, hence the F_{it} must tend to ∞ or to points on the line P_1P_2; but when F_j lies inside the segment P_1P_2 we do not obtain 0 for the corresponding term.

We can also state the following result.

Theorem 1.2. *When $F_t \to F(F_1, F_2, \cdots, F_n)$ with $|F_i| = (a_i^2 + b_i^2)^{1/2} < R < \infty$, $i = 1, \cdots, n$, then the W_n-curves corresponding to different values of $C(> C_0)$, cannot have any linear segment.*

Proof. By (1.1) and the fact that not all the points F_j lie on the straight line P_1P_2, at least one of the relations

$$\overline{P_1F_i} + \overline{P_2F_i} - 2\overline{P_0F_i} > k/R$$

with k a positive constant, holds. Thus (1.4) cannot be satisfied (even if n tends to ∞), and from Proposition 1.1 our theorem is proved.

Starting from the above results, we should be able to conclude the statement given in our introduction. Let us now turn to a method based on the determination of the limiting figures.

2. Limiting figures of the W_n-curves

2.1. We turn now to the representation of the limits of the W_{nt}-curves when n is fixed and $t \to \infty$. Suppose that $F_{it} \to F_i$, $i = 1, 2, \cdots, k$ and $r_{it} \to \infty$, $\varphi_{it} \to \varphi_i$ for $i = k+1, k+2, \cdots, n$. We assume that the curves pass through the points $P_j(\xi_j, \eta_j)$ $j = 1, 2$, and furthermore that the curves for all t remain bounded. As a result of continuity, these conditions can be satisfied.

Proposition 2.1. *Under the given conditions and notations, the limiting position of the W_{nt}-curve system passing through the points P_1 and P_2, when $t \to \infty$ and $1 \leqq k \leqq n$ satisfies the equations*

$$(2.1) \sum_{i=1}^{k} \sqrt{(x-a_i)^2 + (y-b_i)^2} - S_1^{(k)} - (x-\xi_1) \sum_{i=k+1}^{n} \cos\varphi_i - (y-\eta_1) \sum_{i=k+1}^{n} \sin\varphi_i = 0$$

and

$$(2.2) \qquad S_2^{(k)} - S_1^{(k)} = (\xi_2 - \xi_1) \sum_{i=k+1}^{n} \cos\varphi_i + (\eta_2 - \eta_1) \sum_{i=k+1}^{n} \sin\varphi_i = 0,$$

where

$$S_j^{(k)} = \sum_{i=1}^{k} \sqrt{(\xi_j - a_i)^2 + (\eta_j - b_i)^2}, \qquad j = 1, 2.$$

Proof. The equation of the W_{nt}-curve passing through the points $P_1(\xi_1, \eta_1)$ and $P_2(\xi_2, \eta_2)$ has the form

$$\sum_{i=1}^{n}\sqrt{(x-a_{it})^2+(y-b_{it})^2}=\sum_{i=1}^{n}\sqrt{(\xi_j-a_{it})^2+(\eta_j-b_{it})^2},\qquad j=1,2.$$

Transforming to polar coordinates for the points (a_{it}, b_{it}), $i = k+1, k+2, \cdots, n$, we obtain for $j = 1, 2$

$$\sum_{i=1}^{k}\sqrt{(x-a_{it})^2+(y-b_{it})^2}+\sum_{i=k+1}^{n}\sqrt{r_{it}^2-2r_{it}(x\cos\varphi_{it}+y\sin\varphi_{it})+x^2+y^2}$$
$$=S_j^{(k)}+\sum_{i=k+1}^{n}\sqrt{r_{it}^2-2r_{it}(\xi_j\cos\varphi_{it}+y\sin\varphi_{it})+\xi_j^2+\eta_j^2}.$$

Using the simple asymptotic relation

$$\sqrt{z^2-2\alpha z+\beta}=z\left(1-\frac{\alpha}{z}+\frac{\beta-\alpha^2}{z^2}+\sigma\left(\frac{1}{z^2}\right)\right),\qquad z\to\infty$$

this can be rewritten in the form

$$\sum_{i=1}^{k}\sqrt{(x-a_{it})^2+(y-b_{it})^2}$$
$$(2.3)\quad +\sum_{i=k+1}^{n}r_{it}\left[1-\frac{x\cos\varphi_{it}+y\sin\varphi_{it}}{r_{it}}+\frac{(x\sin\varphi_{it}+y\cos\varphi_{it})^2}{r_{it}^2}+\sigma\left(\frac{1}{r_{it}^2}\right)\right]$$
$$=S_j^{(k)}+\sum_{i=k+1}^{n}r_{it}\left[1-\frac{\xi_j\cos\varphi_{it}+\eta_j\sin\varphi_{it}}{r_{it}}+\frac{(\xi_j\sin\varphi_{it}+\eta_j\cos\varphi_{it})^2}{r_{it}^2}+\sigma\left(\frac{1}{r_{it}^2}\right)\right].$$

Letting t tend to ∞, $r_{it}\to\infty$ for $i = k+1, k+2, \cdots, n$, we obtain (2.1) with (2.2). Since for $k = 1$ our equations lead in general to an ellipse, we assume in the next paragraph that $k \geqq 2$.

2.2. Without loss of generality we may assume that the curves pass through the points $P_1(0,-1)$ and $P_2(0,1)$. Then we have for $x = 0$, $k \geqq 2$

$$(2.4)\qquad \sum_{i=1}^{k}\sqrt{a_i^2+(y-b_i)^2}-S_1^{(k)}-(y+1)\sum_{i=k+1}^{n}\sin\varphi_i=0$$

$$(2.5)\qquad \sum_{i=1}^{k}\sqrt{a_i^2+(y-b_i)^2}-S_2^{(k)}-(y-1)\sum_{i=k+1}^{n}\sin\varphi_i=0$$

or

$$(2.5')\qquad S_2^{(k)}-S_1^{(k)}=2\sum_{i=k+1}^{n}\sin\varphi_i.$$

Equations (2.4) or (2.5) can be satisfied for $-1\leqq y\leqq 1$ only in the case where $a_i = 0$, $i = 1, 2, \cdots, k$, i.e. when

$$\sum_{i=1}^{k}|y-b_i|=\sum_{i=1}^{k}|1+b_i|+(y+1)\sum_{i=k+1}^{n}\sin\varphi_i,$$
$$\sum_{i=1}^{k}|1-b_i|-\sum_{i=1}^{k}|1+b_i|=2\sum_{i=k+1}^{n}\sin\varphi_i,$$

which hold for $y = -1$ and for $y = +1$. The left-hand side can be linear in the interval $(-1, 1)$, when $b_i \leqq -1$ or $b_i \geqq 1$ for $i = 1, 2, \cdots, k$, otherwise the derivative has a discontinuity in the interval $(-1, 1)$; this does not occur on the right side. This means that when the limiting curve contains two linear segments, then the foci must tend to ∞ or to the intersection of the two lines. This is essentially the case when $k = 1$, which corresponds to an ellipse, or when $k = 0$.

As we shall see in the next paragraph, when $k = 0$ so that all the foci tend to ∞, the limiting figure is either an infinite line or an ellipse. Restricting ourselves to bounded limiting curves, we have the following result.

Theorem 2.1. *The limiting figure of the W_{nt}-curves can have one linear segment only.*

2.3. Considering our Proposition 2.1, we can see that a third point $P_3(\xi_3, \eta_3)$ can still be prescribed through which the W_{nt}-curve system and the limiting figure may pass. Then we have the additional term

$$S_3^{(k)} - S_1^{(k)} = (\xi_3 - \xi_1) \sum_{i=k+1}^{n} \cos \varphi_i + (\eta_3 - \eta_1) \sum_{i=k+1}^{n} \sin \varphi_i . \tag{2.2'}$$

In (2.1) and (2.2) at the same time the prescription of the three points P_1, P_2 and P_3 determines completely the 'average' direction of the $(n-k)$ foci, and the approach to ∞; i.e. the $\Sigma_{i=k+1}^{n} \cos \varphi_i$ and $\Sigma_{i=k+1}^{n} \sin \varphi_i$ are determined uniquely.

2.4. For the case when the foci F_{it} all tend to ∞, we have the following proposition.

Proposition 2.2. *Using our earlier notation, when $t \to \infty$, $r_{it} = O(t)$, $\varphi_{it} \to \varphi_i$, $i = 1, 2, \cdots, n$, then the W_{nt}-curves, passing through two prescribed points, tend to an ellipse.*

To prove this statement we use (2.3) for $k = 0$:

$$\sum_{i=1}^{n} r_{it} \left[1 - \frac{x \cos \varphi_{it} + y \sin \varphi_{it}}{r_{it}} + \frac{(x \sin \varphi_{it} + y \cos \varphi_{it})^2}{r_{it}^2} + \sigma\left(\frac{1}{r_{it}^2}\right) \right]$$
$$= \sum_{i=1}^{n} \left[1 - \frac{\xi_j \cos \varphi_{it} + \eta_j \sin \varphi_{it}}{r_{it}} + \frac{(\xi_j \sin \varphi_{it} + \eta_j \cos \varphi_{it})^2}{r_{it}^2} + \sigma\left(\frac{1}{r_{it}^2}\right) \right] \tag{2.6}$$
$$j = 1, 2.$$

As $r_{it} = C_i t$ $(0 < C_i < \infty)$, from (2.6) we obtain for $j = 1, 2$

$$x \sum_{i=1}^{n} \cos \varphi_{it} + y \sum_{i=1}^{n} \sin \varphi_{it} + \frac{1}{t} \left[\sum_{i=1}^{n} \frac{(x \sin \varphi_{it} + y \cos \varphi_{it})^2}{C_i} + \sigma(1) \right]$$
$$= \xi_j \sum_{i=1}^{n} \cos \varphi_{it} + \eta_j \sum_{i=1}^{n} \sin \varphi_{it} + \frac{1}{t} \left[\sum_{i=1}^{n} \frac{(\xi_j \sin \varphi_{it} + \eta_j \cos \varphi_{it})^2}{C_i} + \sigma(1) \right] . \tag{2.7}$$

Now whenever $\Sigma_{i=1}^{n} \cos \varphi_{it}$ and $\Sigma_{i=1}^{n} \sin \varphi_{it}$ do not both tend to 0, then for $t \to \infty$ we obtain an infinite line, which is of little interest to us. Let us assume that $\varphi_{it} \to \varphi_i$, $i = 1, 2, \cdots, n$ and

$$\sum_{i=1}^{n} \cos \varphi_i = 0, \qquad \sum_{i=1}^{n} \sin \varphi_i = 0. \tag{2.8}$$

Then for finite t we obtain for $j = 1, 2$, the relation

$$\sum_{i=1}^{n} \frac{(x \sin \varphi_{it} + y \cos \varphi_{it})^2}{C_i} = \sum_{i=1}^{n} \frac{(\xi_j \sin \varphi_{it} + \eta_j \cos \varphi_{it})^2}{C_i} + \sigma(1).$$

Finally we have for the equation of the limiting figure

$$x^2 \sum_{i=1}^{n} \frac{\sin^2 \varphi_i}{C_i} + xy \sum_{i=1}^{n} \frac{\sin 2\varphi_i}{C_i} + y^2 \sum_{i=1}^{n} \frac{\cos^2 \varphi_i}{C_i} = C, \tag{2.9}$$

with the conditions (2.8) and for $j = 1, 2, \cdots$,

$$\xi_j^2 \sum_{i=1}^{n} \frac{\sin^2 \varphi_i}{C_i} + \xi_j \eta_j \sum_{i=1}^{n} \frac{\sin 2\varphi_i}{C_i} + \eta_j^2 \sum_{i=1}^{n} \frac{\cos^2 \varphi_i}{C_i} = C. \tag{2.10}$$

A simple analysis of the case where the r_{it} tend to ∞ in a different order leads to the same result, namely infinite lines, or if more than one focus tends to ∞ in the same order, an ellipse.

One can easily see from (2.1) that, for example, a circle cut by a circular chord segment cannot be approximated by W_n-curves.

3. Cases when n tends to ∞

3.1. It is obvious that one can approximate 'distance integrals' by W_n-curves. Let G be the closure of a bounded open set, and let Γ be a rectifiable, finite plane curve (open or closed). Then for a given point of the plane, the integrals

$$g(P) = \int_G \overline{PQ} df_Q, \tag{3.1}$$

$$\gamma(P) = \int_\Gamma \overline{PQ} ds_Q, \tag{3.1'}$$

where df_Q and ds_Q denote the area and arc elements repectively, are continuous functions of P. They have the known properties

$$g(P) = C(> C_0), \tag{3.2}$$

$$\gamma(P) = C'(> C_0') \tag{3.2'}$$

which represent smooth, convex, closed curves. Here

$$C_0 = \min_{(P)} g(P) = g(P_0,) \qquad C'_0 = \min_{(P)} \gamma(P) = \gamma(P'_0)$$

and P_0 and P'_0 are unique (see e.g. [3]).

Taking a uniform partition on G and Γ the sums

$$\sum_{i=1}^{n} \overline{PQ_i}\Delta f \quad \text{and} \quad \sum_{i=1}^{n} \overline{PQ}\Delta s, \quad \text{for } n \to \infty$$

will approximate the corresponding integrals on the whole curves $g(P) = C$ and $\gamma(P) = C'$. Thus we have the following result.

Theorem 3.1. *Any curve with Equations* (3.1) *and* (3.1′) *can be approximated by W_n-curves, when $n \to \infty$.*

The same statement holds for curves with the equation

$$g(P) + \gamma(P) + \sum_{(i)} \overline{PF_i} = C;$$

but whether these types of curves (integrating perhaps on more general domains), including those given in Section 2, are the only ones which can be approximated by means of W_n-curves remains an open problem.

Acknowledgement

The authors are indebted to G. Fejes Tóth for his valuable remarks.

References

[1] ERDÖS, P. AND VINCZE, I. (1958) On the approximation of convex, closed plane curves. *Math. Lapok* **9**, 1–2.

[2] WEISZFELD, E. (1937) Sur le point pour lequel la somme des distances de *n* points donnés est minimum. *Tohoku Math. J.* **43**, 355–386, and personal communication.

[3] VINCZE, ST. (1938) Über die Schwerpunkte der konvexen Kurven bei speziellen Belegungen. *Acta Lit. Acad. Sci. Szeged* **9**, 52–59.