Geometric differentiation

for the *intelligence* of curves and surfaces

Second edition

I. R. Porteous

Senior Lecturer, Department of Pure Mathematics
University of Liverpool

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011–4211, USA
10 Stamford Road, Oakleigh, VIC 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

© Cambridge University Press 1994, 2001

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 1994
Second edition 2001

Printed in the United Kingdom at the University Press, Cambridge

Typeface Times 10/13. *System* 3B2 [KT]

A catalogue record for this book is available from the British Library

ISBN 0 521 81040 X hardback
ISBN 0 521 00264 8 paperback

Contents

1
Plane curves

1.0 Introduction

Sir Christopher Wren
Went to dine with some men.
'If anyone calls,
Say I'm designing St Paul's!'

St Paul's Cathedral was designed following the Great Fire of London in 1666. Six years earlier Wren, a mathematician as well as architect, was one of the founder members of the Royal Society. At that time one of the men that he might well have been dining with was the great Dutch Scientist, Christiaan Huygens (*natus* 1629, *denatus* 1695, as a late picture of him has it! (Figure 1.1)). At the time we are speaking of Newton (*natus* 1642) and Leibniz (*natus* 1646) were still teenagers, and the Calculus had yet to be invented. Indeed the first elementary calculus textbook was published only in 1696, the year after Huygens' death. This purported to be written by an aristocratic friend of the Bernoulli family, the Marquis de l'Hôpital, and was entitled *Analyse des infiniment petits, Pour l'intelligence des lignes courbes*. Central to this first work on differential geometry are the ideas developed by Huygens and his associates thirty-five or more years previously. Curiously, de l'Hôpital did not put his name to the first edition of the work, it being added in ink in many copies (Figure 1.2). The work is in fact a fairly direct translation from the original Latin of Jean Bernoulli, which came to light many years later, neither the translator nor the writer of the unsigned preface being de l'Hôpital! For an account of this ancient scandal see Truesdell (1958).

Our aim here is to give a fresh account of these ideas which remain the basis of the whole subject.

Consider as a first example the parabola in the real plane with equation $y = x^2$. An engineer wishing to cut this curve accurately out of some sheet of material has to use a cutting tool, necessarily of finite size, whose centre has to be programmed to follow some curve *offset* the right distance from the parabola to be cut. Hasty thinking might suggest that this offset is another parabola, but this is not so – compare

Copyright Museum Boerhave, Leiden

Figure 1.1

Figures 1.3 and 1.4. If one examines offsets at greater and greater distances from the original curve (on the 'inner' side) one discovers that before long these are no longer regular curves but acquire sharp points or *cusps*, where the direction of the curve reverses. Moreover these

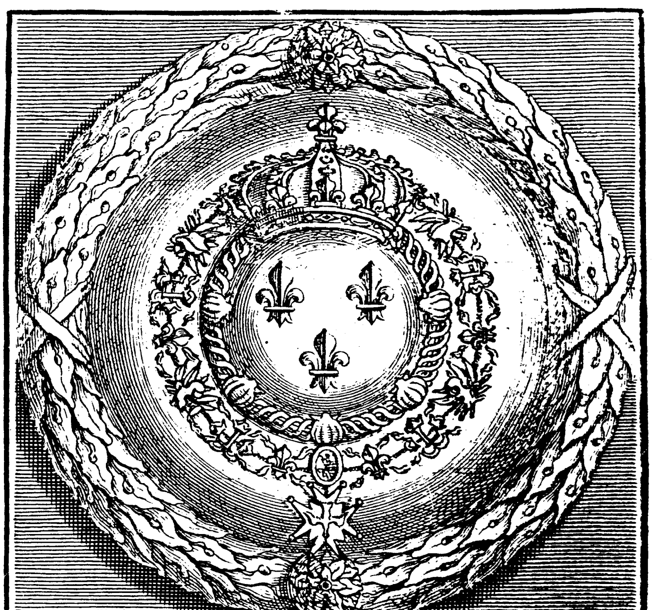
ANALYSE

DES

INFINIMENT PETITS,

Pour l'intelligence des lignes courbes.

Par le Marquis de l'Hospital



A PARIS,

DE L'IMPRIMERIE ROYALE.

M. DC. XCVI.

Figure 1.2

cusps lie along a new curve which itself sports a cusp, pointing towards the lowest point of the original parabola – see Figure 1.5.

It is a pleasant thought to think of the parabola in another way as the shoreline of a bay in which one has gone out for a swim, swimming out normally, that is at right angles, to the shore – Figure 1.6. One's first intuition probably is that, no matter how far one swims, one's starting point $*$ remains locally the nearest point of the shore. We say 'locally' here because if one goes far enough then clearly some point on the farther shore may well be nearer. But our local intuition is wrong, as Figures 1.7 and 1.8 illustrate. These display the same new cuspidal curve that we saw before, its tangents all being normal to the parabola.

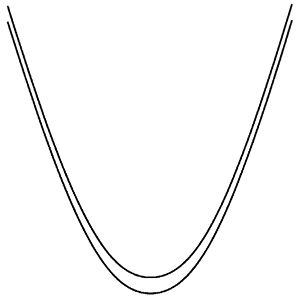


Figure 1.3

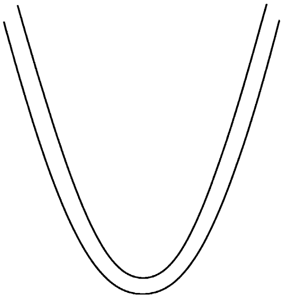


Figure 1.4

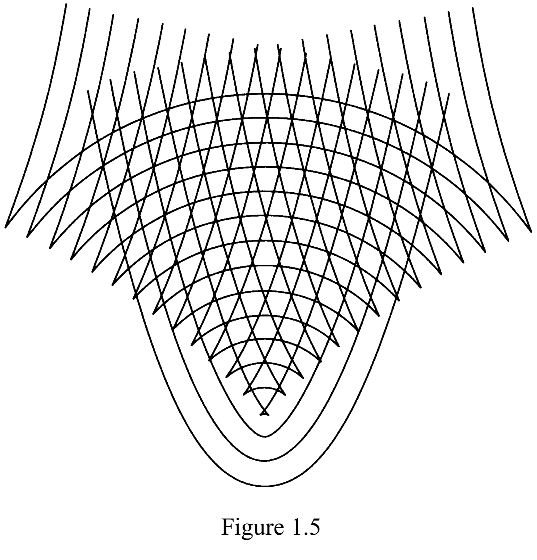


Figure 1.5

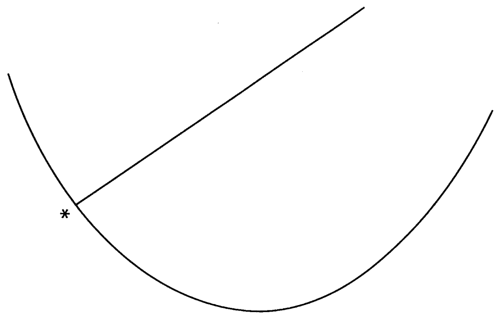


Figure 1.6

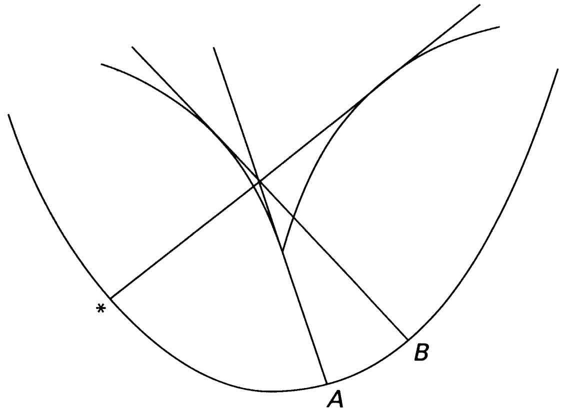


Figure 1.7

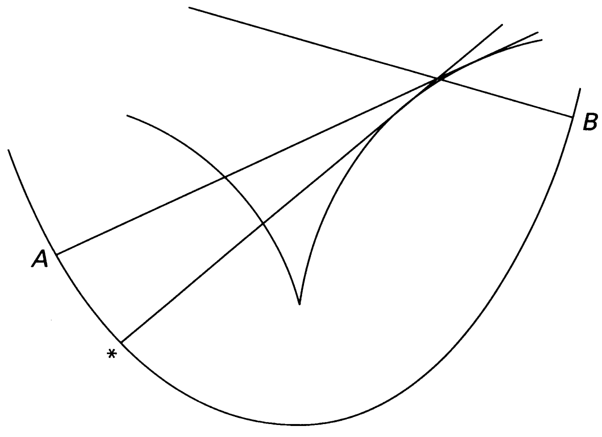


Figure 1.8

Initially one can draw only one normal to the shore from one's position $*$ in the bay, namely the path along which one has just swum, but after crossing the curve of cusps two new normals can be drawn, the three shore points $*$, A and B then being successively a local minimum at $*$, a local maximum at A and a local minimum at B, of the distance from one's position in the bay to the shoreline – Figure 1.7. As one swims on, the points A and B move round the shore in opposite directions, and as one reaches the point of tangency of the normal with the curve of cusps A comes right round to coincide with $*$. At any more distant point $*$ is a *local maximum* of distance – Figure 1.8!

The curve of cusps that falsifies both these intuitions is known as the *evolute* or *focal curve* of the original curve. In Figure 1.9 it is exhibited as the *envelope* of the family of the family of normals to the parabola. The offsets are also said to be the *parallels* or *equidistants* to the parabola.

It was Huygens who made the remarkable discovery that one can recover the original parabola from its evolute by unwinding an inextensible string laid partially along the evolute, or equivalently by rolling the tangent line to the evolute along the evolute. A bob on the string, or point of the rolling line, then describes part either of the parabola itself or, according to the position of the bob, one of the offsets to the parabola. Indeed all the offsets can be obtained in this way if one makes appropriate conventions about the unwinding process,

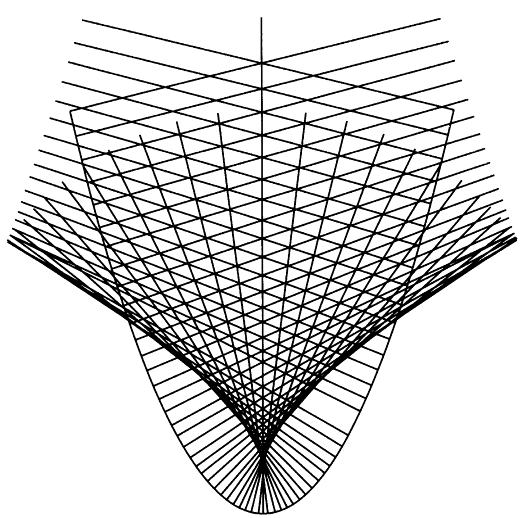
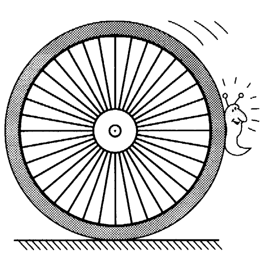


Figure 1.9

especially at a cusp of the evolute. These mutually parallel curves are known as the *involutes* or *evolvents* of the evolute.

There is nothing special about the parabola in all this. Indeed a favourite curve of Huygens, and of Wren too, is the curve which features as the solution to the following take-home problem (Figure 1.10) faced by several thousand Merseyside twelve-year olds in the Spring of 1982 (Giblin and Porteous, 1990).

The curve is the *cycloid*, consisting of a series of arches supported on a series of cusps (Figure 1.11). As we shall verify later, this curve has the remarkable property that its evolute is a congruent cycloid, whose cusps this time point away from and not towards the original curve. If we turn all this upside down (Figure 1.12) and arrange for a pendulum of suitable length to be swung from one of the jaws of the evolute cycloid one obtains the Huygens cycloidal pendulum, whose period, remarkably, turns out to be independent of the amplitude.



Arc Light

There was a young glow worm called Glim,
Who went for a ride on the rim
  Of a wheel that went round
  As it rolled on the ground.
Please draw me the arc traced by him!

Figure 1.10

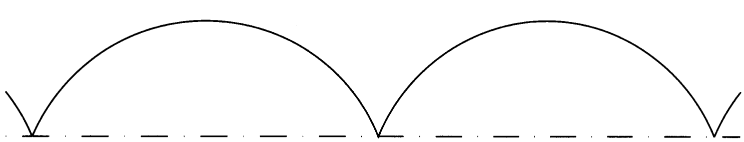


Figure 1.11

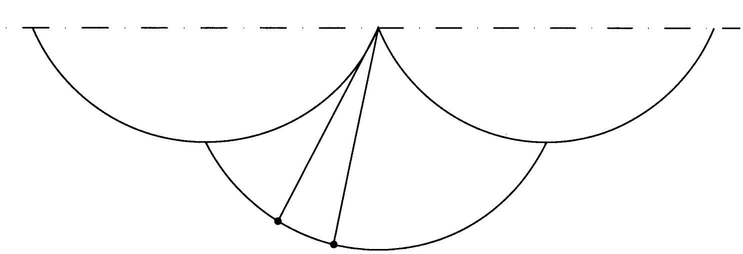


Figure 1.12

Yet a third way of regarding the evolute is as the locus of *centres of curvature* of the original curve. This is illustrated in Figure 1.13 where the circle with centre at the point of tangency of a normal to the original curve with the evolute, and passing through the base of the normal, is seen to hug the curve so closely there that it is known as the *osculating circle*, or *circle of curvature* of the curve at that point. In general, as in this example, it shares a tangent line with the original curve, but crosses the curve there. An exception to this occurs at the lowest point of the parabola, when the centre of the osculating circle lies at the cusp of the evolute and the circle lies entirely above the parabola. At this point the radius of the osculating circle, the *radius of curvature* of the curve, has a local minimum – indeed in this example an absolute minimum. In fact cusps on the evolute correspond to critical points of the radius of curvature, the cusps on the evolute pointing towards the curve at local minima and away from the curve at local maxima.

The reciprocal of the radius of curvature is known simply as the *curvature* of the curve. At a point of inflection of the curve the curvature is zero and the radius of curvature infinite, the role of osculating circle being then played by the inflectional tangent. We shall prove that the evolute of a regular plane curve does not have any points of inflection. Of course, as de l'Hôpital (or was it Jean Bernoulli?) first remarked, there is nothing to stop one swinging a pendulum from a curve with an inflection. The resulting family of non-regular involutes (see Figure 1.21) has an intimate relationship with the group of

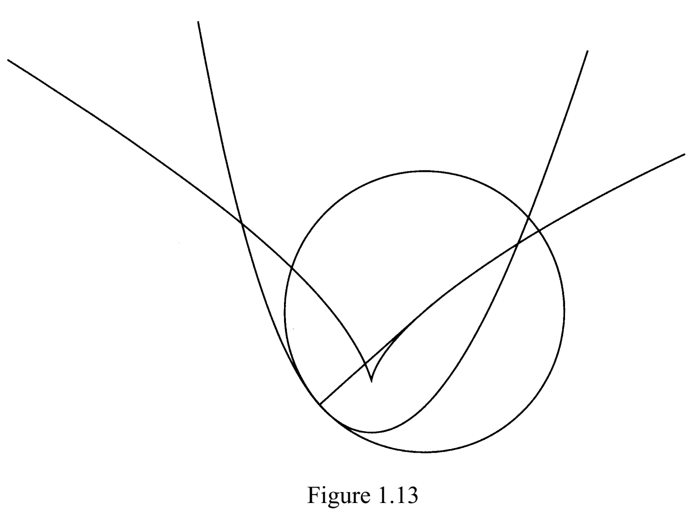


Figure 1.13

symmetries of an icosahedron – a deep and mysterious fact only recently noted by the Russian school of singularity theorists under the leadership of V.I. Arnol'd (Arnol'd, 1983, 1990b).

As we are going to be concerned in what follows with applications of the calculus to geometry we ought logically to start with reviewing the calculus. Since almost all that is required for the study of curves should already be familiar to the reader we defer this review to Chapter 4, preceded in Chapter 2 with a review of some basic frequently used facts of linear and projective geometry. For the moment it is enough to remark that the standard n-dimensional real vector space equipped with the standard Euclidean scalar product will be denoted by $\mathbb{R}^n$, the product being denoted by a dot above the line $\cdot$. The *length* of a vector $\mathbf{v} \in \mathbb{R}^n$ is $|\mathbf{v}| = \sqrt{(\mathbf{v} \cdot \mathbf{v})}$. A map $f : \mathbb{R}^n \rightarrowtail \mathbb{R}^p$ is said to be *smooth* if everywhere sufficiently many† of its derivatives exist and are continuous, the (non-standard) forked tail on the arrow indicating that the domain of definition is an open subset of $\mathbb{R}^n$ but not necessarily the whole of $\mathbb{R}^n$.

1.1 Regular plane curves and their evolutes

Curves in the plane may be presented in many different ways, for example as the zero sets of functions $\mathbb{R}^2 \rightarrowtail \mathbb{R}$, locally at least as the graphs of functions $\mathbb{R} \rightarrowtail \mathbb{R}$, or parametrically as the images of maps $\mathbb{R} \rightarrowtail \mathbb{R}^2$. For example the circle of radius 1 with centre the origin, the *unit circle*, is the zero set of the function $\mathbb{R}^2 \rightarrowtail \mathbb{R}$; $(x, y) \mapsto x^2 + y^2 - 1$, and also the image of the map $\mathbb{R} \to \mathbb{R}^2$; $\theta \mapsto (\cos\theta, \sin\theta)$. It is not globally the graph of a function from either axis to the other, but locally it is. For simplicity we begin by concentrating almost entirely on curves presented parametrically, with domains open intervals of $\mathbb{R}$. The image space will be an explicit copy of $\mathbb{R}^2$ but we occasionally will allow ourselves the luxury of choosing a fresh origin for this space, perhaps at some special point of interest of the curve, and also choosing fresh mutually orthogonal axes through this new origin. Such a change of view will, however, preserve the metric of the plane, the distance between points remaining unaltered despite the change of frame of reference.

A *smooth parametric curve* in $\mathbb{R}^2$ is a smooth map

$$\mathbf{r} : \mathbb{R} \rightarrowtail \mathbb{R}^2;\ t \mapsto \mathbf{r}(t),$$

† This usage of the word 'smooth' is slovenly but convenient. If one prefers it, take 'smooth' to mean 'infinitely differentiable', that is C^∞.

with domain an *open interval* of $\mathbb{R}$, that is an open *connected* subset of $\mathbb{R}$. It is *regular* (or *immersive*) *at t* if its first derivative $\mathbf{r}_1(t)$ is non-zero (we defy convention by using subscripts instead of ds or dots or dashes to denote differentiation with respect to the parameter). At a regular point t the vector $\mathbf{r}_1$(t), which may be regarded as the *velocity* of the curve $\mathbf{r}$ at time t, generates the *tangent vector line* to $\mathbf{r}$ at t. The *tangent line* to $\mathbf{r}$ at t is then the line

$$u \mapsto \mathbf{r}(t) + u\mathbf{r}_1(t).$$

A smooth curve may be straight! But this puts strong conditions on the higher derivatives of the curve. For suppose that the image of the curve $\mathbf{r} : t \mapsto \mathbf{r}(t)$ is the line in $\mathbb{R}^2$ with equation $ax + by = k$, or part of that line. Then, for every $t \in \mathbb{R}$, $\mathbf{c} \cdot \mathbf{r}(t) = k$, where $\mathbf{c} = (a, b)$, and for every $i \geqslant 1$ we have $\mathbf{c} \cdot \mathbf{r}_i(t) = 0$, implying that each of the derived vectors is a multiple of the first non-zero one.

It is, of course, exceptional for any of the higher derivatives $\mathbf{r}_i(t)$ of a regular smooth curve $\mathbf{r}$ at a point t to be a multiple of $\mathbf{r}_1(t)$. We say that a smooth curve $\mathbf{r}$ is *linear* at t if it is regular there and its *acceleration* $\mathbf{r}_2(t)$ is a multiple of $\mathbf{r}_1(t)$. It will be said to be A_k*-linear* at t if it is regular there and $\mathbf{r}_j(t)$ is a multiple of $\mathbf{r}_1(t)$ for $1 < j \leqslant k$, but $\mathbf{r}_{k+1}(t)$ is not a multiple of $\mathbf{r}_1(t)$. According to this definition $\mathbf{r}$ is *not* linear at an A_1-linear point, but just regular there. An A_2-linear point is an *ordinary inflection* of $\mathbf{r}$ and an A_3-linear point an *ordinary undulation* of $\mathbf{r}$.

Example 1.1 The curve $t \mapsto (t, t^3)$ (Figure 1.14) has an ordinary inflection at $t = 0$, while the curve $t \mapsto (t, t^4)$ (Figure 1.15) has an ordinary undulation at $t = 0$. □

The somewhat odd term ‘undulation’ derives from thinking of the curve $t \mapsto (t, t^4)$ as being the curve given by the value $\varepsilon = 0$ in the family of curves $t \mapsto (t, \varepsilon t^2 + t^4)$, such a curve having no inflection for $\varepsilon > 0$, but acquiring two and a consequent wiggle when ε becomes negative.

These examples are typical:

Proposition 1.2 By suitably choosing a new origin and new mutually orthogonal axes in $\mathbb{R}^2$ *the parametric equations of a smooth curve* $\mathbf{r}$ *in the neighbourhood of an ordinary inflection at* $t = 0$ *may be taken to be of the form*

$$\mathbf{r}(t) = (at + \ldots, bt^3 + \ldots), \text{ where } a \neq 0 \text{ and } b \neq 0,$$

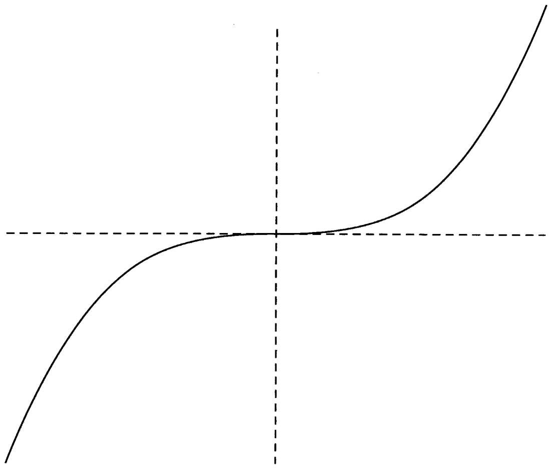


Figure 1.14

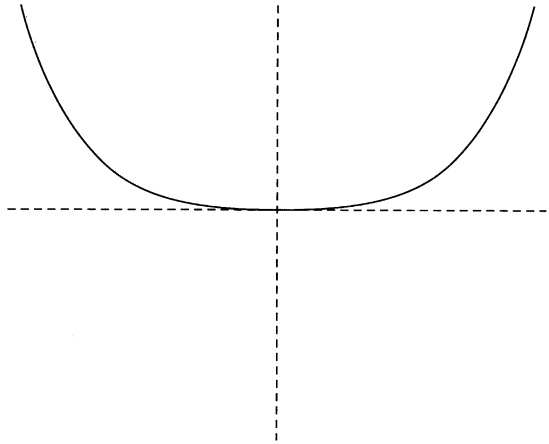


Figure 1.15

and in the neighbourhood of an ordinary undulation to be of the form

$$\mathbf{r}(t) = (at + \ldots, bt^4 + \ldots), \text{ where } a \neq 0 \text{ and } b \neq 0. \qquad \square$$

Corollary 1.3 The tangent line to a regular curve **r** *at an ordinary inflection crosses the curve, but at an ordinary undulation this is not so.* □

In the above examples the curves are actually graphs of functions. More generally consider the curve $\mathbf{r}$ given by $\mathbf{r}(t) = (t, f(t))$, where $f : \mathbb{R} \rightarrowtail \mathbb{R}; t \mapsto f(t)$ is a smooth function. Then we have

$$\mathbf{r}_1(t) = (1, f_1(t)),$$

$$\mathbf{r}_2(t) = (0, f_2(t)),$$

$$\mathbf{r}_3(t) = (0, f_3(t)),$$

$$\mathbf{r}_4(t) = (0, f_4(t)),$$

and so on. Then $\mathbf{r}_k(t)$ is a multiple of $\mathbf{r}_1(t) \Leftrightarrow f_k(t) = 0$.

Proposition 1.4 For a regular curve $\mathbf{r}$ *that is the graph of a smooth function f,* $\mathbf{r}$ *is* A_k-linear at t if and only if $\mathbf{r}_i(t) = 0$ *for* $2 \leqslant i \leqslant k$, *and* $\mathbf{r}_{k+1}(t) \neq 0$. □

Non-regular points of smooth curves also must be considered, such a point being one where the velocity of the curve is zero. Such a point is commonly called a *cusp* of the curve. In particular a smooth curve $\mathbf{r}$ is said to have an *ordinary*, or 3/2, *cusp* at t if $\mathbf{r}_1(t) = 0$ but $\mathbf{r}_2(t) \neq 0$, with $\mathbf{r}_3(t)$ linearly independent of $\mathbf{r}_2(t)$ and to have an *ordinary kink*, or $4/3$, *cusp* at t if $\mathbf{r}_1(t) = 0$ and $\mathbf{r}_2(t) = 0$, but $\mathbf{r}_3(t) \neq 0$, with $\mathbf{r}_4(t)$ linearly independent of $\mathbf{r}_3(t)$. More generally, $\mathbf{r}$ is said to have an $(n+1)/n$ *cusp* at t if $\mathbf{r}_i(t) = 0$ for $1 \leqslant i < n$ but $\mathbf{r}_n(t) \neq 0$, with $\mathbf{r}_{n+1}(t)$ linearly independent of $\mathbf{r}_n(t)$.

Example 1.5 The curve $t \mapsto (t^2, t^3)$ (Figure 1.16) has an ordinary cusp

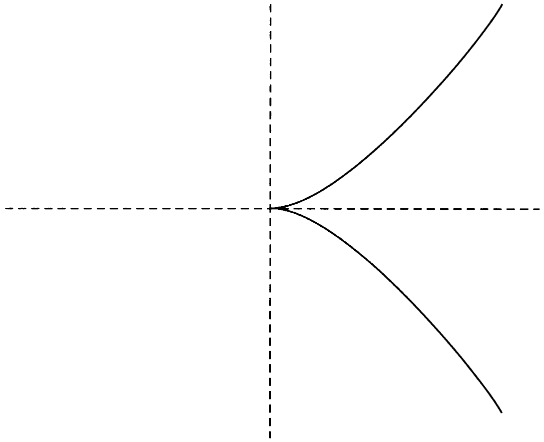


Figure 1.16

at $t = 0$, while the curve $t \mapsto (t^3, t^4)$ (Figure 1.17) has an ordinary kink at $t = 0$. □

Proposition 1.6 By suitably choosing a new origin and orthogonal axes in $\mathbb{R}^2$ *the parametric equations of a smooth curve* $\mathbf{r}$ *in the neighbourhood of an ordinary cusp at* $t = 0$ *may be taken to be of the form*

$$\mathbf{r}(t) = (at^2 + \ldots, bt^3 + \ldots), \text{ where } a \neq 0 \text{ and } b \neq 0.$$

Moreover at a cusp the vector $\mathbf{r}_2(0)$ *points 'in the opposite direction' to the cusp, lying between its two 'cheeks' (Figure 1.18).*

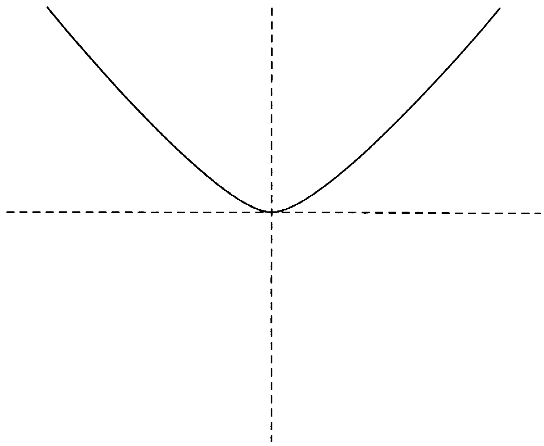


Figure 1.17

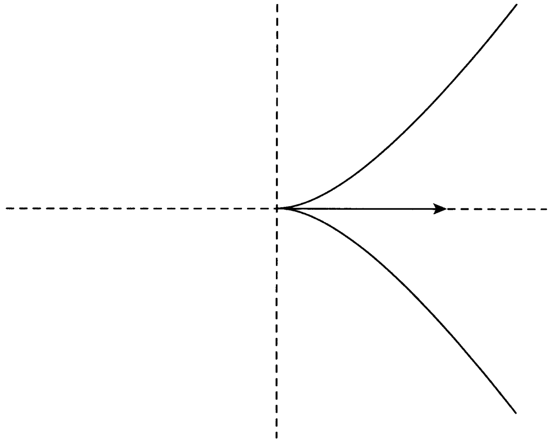


Figure 1.18

Proof Choose axes with origin at the cusp and with $\mathbf{r}_2(0) = (2a, 0)$, where $a \neq 0$. Since also $\mathbf{r}(0) = (0, 0)$ and $\mathbf{r}_1(0) = (0, 0)$ it follows from Taylor's Theorem applied to each component that $\mathbf{r}(t) = (at^2 + \ldots, bt^3 + \ldots)$. Then $\mathbf{r}_3(0) = (c, 6b)$, for some c. Since we know that $\mathbf{r}_3$ is not a multiple of $\mathbf{r}_2$, $b \neq 0$.

Suppose that $a > 0$. Then $\mathbf{r}_2(0)$ points along the x-axis in the right-hand or positive direction. Now the cusp lies entirely in the right-hand half-plane for small t. For the x-component of $\mathbf{r}(t)$ is $at^2 f(t)$, where, since $f(0) = 1$, $f(t) > 0$ for small non-zero t. Moreover, away from $t = 0$

$$\begin{aligned}\mathbf{r}_1(t) &= (2at + \ldots, 3bt^2 + \ldots) \\ &= t(2a + \ldots, 3bt + \ldots),\end{aligned}$$

which is a multiple of $(2a + \ldots, 3bt + \ldots)$, the latter tending to $(2a, 0)$ from opposite sides as t tends to 0 from either side. In particular the limit tangent direction is along the x-axis. That is $\mathbf{r}_2(0)$ points 'in the opposite direction' to the cusp, lying between its two 'cheeks'. □

We shall verify later in Example 4.20 that a smooth curve is essentially non-regular at an ordinary cusp; that is it cannot be made regular there by 'reparametrisation' of the curve.

A regular parametric curve may intersect itself; that is two or more distinct values of the parameter may have the same image point in $\mathbb{R}^2$. The common image point is said to be a *singularity* of the curve, but not a point of non-regularity of the curve.

Example 1.7 The curve $\mathbf{r} : \mathbb{R} \to \mathbb{R}^2 : t \mapsto (t^2 - 1, t(t^2 - 1))$ has a double point at the origin, for $\mathbf{r}(-1) = \mathbf{r}(1) = (0, 0)$, but $\mathbf{r}_1(t) = (2t, 3t^2 - 1)$, so that $\mathbf{r}_1(-1) = (-2, 2)$ and $\mathbf{r}_1(1) = (2, 2)$, both non-zero – Figure 1.19. □

1.2 Curvature

In studying the curvature of a regular plane curve $\mathbf{r}$ we study at each point t how closely the curve approximates there to a parametrised *circle*. Now the circle with centre $\mathbf{c}$ and radius ρ consists of all $\mathbf{r}$ of $\mathbb{R}^2$ such that $(\mathbf{r} - \mathbf{c}) \cdot (\mathbf{r} - \mathbf{c}) = \rho^2$, or equivalently such that

$$\mathbf{c} \cdot \mathbf{r} - \tfrac{1}{2}\mathbf{r} \cdot \mathbf{r} = \tfrac{1}{2}(\mathbf{c} \cdot \mathbf{c} - \rho^2),$$

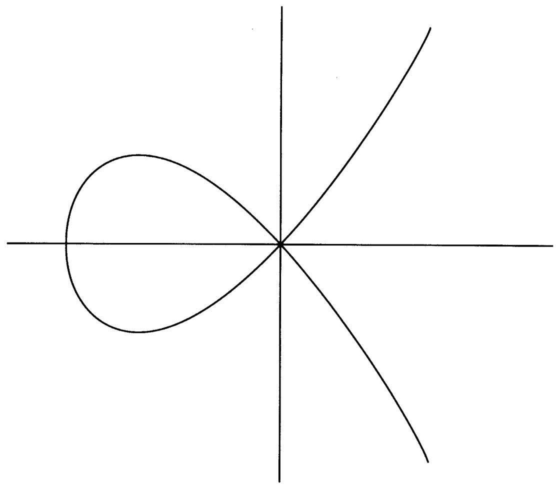


Figure 1.19

the right-hand side of this equation being constant. Accordingly for any parametrisation $t \mapsto \mathbf{r}(t)$ of this circle all the derivatives of the function

$$V(\mathbf{c}) \colon t \mapsto \mathbf{c} \cdot \mathbf{r}(t) - \tfrac{1}{2}\mathbf{r}(t) \cdot \mathbf{r}(t)$$

are everywhere zero, namely

$$V(\mathbf{c})_1 = (\mathbf{c} - \mathbf{r}) \cdot \mathbf{r}_1 = 0,$$

$$V(\mathbf{c})_2 = (\mathbf{c} - \mathbf{r}) \cdot \mathbf{r}_2 - \mathbf{r}_1 \cdot \mathbf{r}_1 = 0,$$

$$V(\mathbf{c})_3 = (\mathbf{c} - \mathbf{r}) \cdot \mathbf{r}_3 - 3\mathbf{r}_1 \cdot \mathbf{r}_2 = 0.$$

Now suppose that $\mathbf{r}$ is a regular parametric curve that is not everywhere circular. Clearly $V(\mathbf{c})_1(t) = 0$ whenever the vector $\mathbf{c} - \mathbf{r}$ is orthogonal to the tangent vector $\mathbf{r}_1(t)$, that is whenever the point $\mathbf{c}$ happens to lie on the *normal* to $\mathbf{r}$ at t, the line through $\mathbf{r}(t)$ orthogonal to the tangent line there. It may be that $\mathbf{r}$ is linear at t, that is that $\mathbf{r}_2(t)$ is linearly dependent on $\mathbf{r}_1(t)$. When this is not so, as will generally be the case, there will be a unique point $\mathbf{c} \neq \mathbf{r}(t)$ on the normal line such that also $V(\mathbf{c})_2(t) = 0$.

This point, which we denote by $\mathbf{e}(t)$, is called the *centre of curvature* or *focal point* of $\mathbf{r}$ at t, the curve $\mathbf{e} : t \mapsto \mathbf{e}(t)$ being called the *evolute* or *focal curve* of $\mathbf{r}$. The distance $\rho(t)$ of $\mathbf{e}(t)$ from $\mathbf{r}(t)$ is called the *radius of curvature* of $\mathbf{r}$ at t and its reciprocal $\kappa(t) = 1/\rho(t)$ the *curvature* of $\mathbf{r}$ at t.

Example 1.8 Let $\mathbf{r}$ be the parabola $t \mapsto (t,\ t^2)$. Then we have

$$\mathbf{r}(t) = (t,\ t^2),$$
$$\mathbf{r}_1(t) = (1,\ 2t),$$
$$\mathbf{r}_2(t) = (0,\ 2).$$

So the equation for $\mathbf{e}(t)$ becomes

$$\begin{bmatrix} 1 & 2t \\ 0 & 2 \end{bmatrix} \mathbf{e}(t) = \begin{bmatrix} (t,\ t^2)\cdot(1,\ 2t) \\ (t,\ t^2)\cdot(0,\ 2) + (1,\ 2t)\cdot(1,\ 2t) \end{bmatrix};$$

that is

$$\mathbf{e}(t) = \frac{1}{2}\begin{bmatrix} 2 & -2t \\ 0 & 1 \end{bmatrix}\begin{bmatrix} t + 2t^3 \\ 2t^2 + 1 + 4t^2 \end{bmatrix}$$
$$= \frac{1}{2}\begin{bmatrix} -8t^3 \\ 1 + 6t^2 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix} + \begin{bmatrix} -4t^3 \\ 3t^2 \end{bmatrix}.$$

Now the curve $t \mapsto (-4t^3,\ 3t^2)$ clearly has an ordinary cusp at $t = 0$, at the origin in $\mathbb{R}^2$. So the curve $\mathbf{e}$ has an ordinary cusp at $t = 0$, at $(0,\ \frac{1}{2})$ in $\mathbb{R}^2$ (Figure 1.9). □

The curvature of a regular curve $\mathbf{r}$ may be defined directly by assigning to each t either of the unit normal vectors $\mathbf{n}(t)$ to $\mathbf{r}$ at t. The choice does not matter, except that it should be made continuously along the curve. For definiteness we shall generally tacitly choose $\mathbf{n}$ so that one turns through $+\frac{1}{2}\pi$ in turning from $\mathbf{r}_1(t)$ to $\mathbf{n}(t)$. (But this is not so straightforward for a curve with non-regular points. See Exercise 1.27 – the cardioid (Figure 1.25).)

In this way we have associated to the regular curve $\mathbf{r}$ a smooth circular curve $\mathbf{n}$, the image of $\mathbf{n}$ being a subset of the unit circle, the circle with centre 0 and radius 1. As one travels along the curve $\mathbf{r}$ in time the unit vector $\mathbf{n}(t)$ swings to and fro, like a pointer on a dial (Figure 1.20).

From the definition of $\mathbf{n}$ it follows that $\mathbf{n}\cdot\mathbf{r}_1 = 0$ everywhere, the vectors $\mathbf{n}(t)$ and $\mathbf{r}_1(t)$ being linearly independent for each t, any vector orthogonal to each necessarily being the zero vector, a remark that will be relevant again and again in what follows. Moreover, since $\mathbf{n}\cdot\mathbf{n} = 1$ everywhere it follows that $\mathbf{n}\cdot\mathbf{n}_1 = 0$ (cancelling a 2). So, for all t, the vectors $\mathbf{n}_1(t)$ and $\mathbf{r}_1(t)$ are linearly dependent, both being orthogonal to $\mathbf{n}(t)$. With $\mathbf{r}_1(t) \neq 0$ it follows that $\mathbf{n}_1(t)$ is a (possibly zero) multiple of $\mathbf{r}_1(t)$.

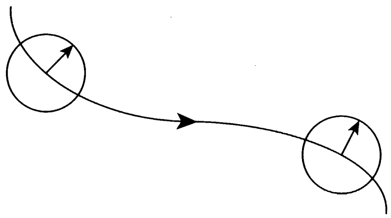


Figure 1.20

For the circle with centre $\mathbf{c}$ and radius ρ we have $\mathbf{r} + \rho\mathbf{n} = \mathbf{c}$, from which it follows that $\mathbf{r}_1 + \rho\mathbf{n}_1 = 0$, that is that $\kappa\mathbf{r}_1 + \mathbf{n}_1 = 0$, where $\kappa = \rho^{-1}$. This suggests that for a regular curve $\mathbf{r}$ the *curvature* of $\mathbf{r}$ at t, $\kappa(t)$, should be defined by the equation

$$\mathbf{n}_1(t) = -\kappa(t)\mathbf{r}_1(t),$$

the actual sign of the curvature depending on the choice of normal. That is up to sign the curvature of a curve is the ratio of the velocity of the end of the normal vector on the dial to one's velocity along the curve.

Suppose that κ is so defined. Then we have the following proposition:

Proposition 1.9 A regular curve $\mathbf{r}$ *is linear at a point* t *if and only if* $\kappa = 0$ *at* t. *Equivalently,* $\kappa \neq 0$ *at* t *if and only if* $\mathbf{r}$ *is not linear at* t (*is* A_1*-linear at* t).

Proof Since $\mathbf{r}_1(t) \neq 0$ it is clear that $\kappa = 0$ if and only if $\mathbf{n}_1 = 0$. Now $\mathbf{n} \cdot \mathbf{r}_1 = 0$. So $\mathbf{n} \cdot \mathbf{r}_2 + \mathbf{n}_1 \cdot \mathbf{r}_1 = 0$, from which it follows that if $\mathbf{n}_1 = 0$ then $\mathbf{n} \cdot \mathbf{r}_2 = 0$ and therefore that $\mathbf{r}$ is linear at t. Conversely if $\mathbf{r}$ is linear at t then $\mathbf{n}_1 \cdot \mathbf{r}_1 = 0$. But also $\mathbf{n}_1 \cdot \mathbf{n} = 0$. Thus $\mathbf{n}_1 = 0$. □

One can extend the last proposition.

Proposition 1.10 A regular curve $\mathbf{r}$ *is* A_k*-linear* ($k \geqslant 2$) *at a point* t *if and only if* $\kappa_i = 0$ *at* t, *for* $0 \leqslant i \leqslant k-2$, *but* $\kappa_{k-1}(t) \neq 0$. *In particular* $\mathbf{r}$ *has an ordinary inflection* (*is* A_2*-linear*) *at* t *if and only if* $\kappa = 0$ *but* $\kappa_1 \neq 0$ *at* t, *and has an ordinary undulation* (*is* A_3*-linear*) *at* t *if and only if* $\kappa = \kappa_1 = 0$ *but* $\kappa_2 \neq 0$ *at* t. □

Note that the curvature of a regular curve changes sign as one passes through an ordinary inflection of the curve.

To relate the two definitions of curvature suppose that at t $\kappa(t) \neq 0$, with $\rho(t) = 1/\kappa(t)$. Then $\mathbf{r}_1(t) + \rho(t)\mathbf{n}_1(t) = 0$. By an earlier remark this is equivalent to the pair of equations

$$(\mathbf{r}_1(t) + \rho(t)\mathbf{n}_1(t)) \cdot \mathbf{r}_1(t) = 0 \text{ and } (\mathbf{r}_1(t) + \rho(t)\mathbf{n}_1(t)) \cdot \mathbf{n}(t) = 0,$$

the second of these being true for all t since $\mathbf{r}_1 \cdot \mathbf{n} = 0$ and $\mathbf{n}_1 \cdot \mathbf{n} = 0$. As to the first, since $\mathbf{n} \cdot \mathbf{r}_1 = 0$, $\mathbf{n}_1 \cdot \mathbf{r}_1 = -\mathbf{n} \cdot \mathbf{r}_2$ as before, so that the equation takes the form

$$\rho(t)\mathbf{n}(t) \cdot \mathbf{r}_2(t) - \mathbf{r}_1(t) \cdot \mathbf{r}_1(t) = 0,$$

in accordance with our earlier definition, with $\rho(t)\mathbf{n}(t) = \mathbf{e}(t) - \mathbf{r}(t)$. Of course ρ may now take either sign, the actual sign of ρ at any point t depending on the choice of the unit normal vector $\mathbf{n}(t)$.

We shall say that a smooth curve $\mathbf{r}$ is *circular* at a point t if it is regular and not linear at t and if also not only $V(\mathbf{c})_1(t) = V(\mathbf{c})_2(t) = 0$ but also $V_3(\mathbf{e})(t) = 0$, where $V(\mathbf{c}) = \mathbf{c} \cdot \mathbf{r} - \frac{1}{2}\mathbf{r} \cdot \mathbf{r}$ and $V_k(\mathbf{e})$, for any positive integer k, is a convenient shorthand notation for $V(\mathbf{c})_k$ with $\mathbf{c}$ *after* the differentiation put equal to $\mathbf{e}$. It has an *ordinary vertex* at t if also $V_4(\mathbf{e})(t) \neq 0$.

Several equations relate the derivatives of a regular curve $\mathbf{r}$ to the derivatives of $\mathbf{e}$, of ρ and of $\mathbf{n}$. Apart from $\mathbf{n} \cdot \mathbf{r}_1 = 0$ and $\mathbf{n} \cdot \mathbf{n}_1 = 0$ we have

$$(\mathbf{e} - \mathbf{r}) \cdot \mathbf{r}_1 = 0 \tag{1}$$

$$(\mathbf{e} - \mathbf{r}) \cdot \mathbf{r}_2 = \mathbf{r}_1 \cdot \mathbf{r}_1$$

defining $\mathbf{e}$,

$$\mathbf{e}_1 \cdot \mathbf{r}_1 = 0 \tag{3}$$

obtained by differentiating (1) and using (2),

$$\mathbf{e}_1 \cdot \mathbf{r}_2 + V_3(\mathbf{e}) = 0, \tag{4}$$

obtained by differentiating (2),

$$\mathbf{r}_1 + \rho\mathbf{n}_1 = 0 \tag{5}$$

and

$$\mathbf{e}_1 = \rho_1\mathbf{n} \tag{6}$$

obtained by differentiating the equation $\mathbf{e} = \mathbf{r} + \rho\mathbf{n}$ and using (5).

We employ these in the proof of the following proposition listing some elementary properties of the evolute of a regular plane curve.

Proposition 1.11 Let $\mathbf{r}$ *be a regular curve in the plane with evolute* $\mathbf{e}$. *Then*

(*a*) *for each* t *at which the evolute* $\mathbf{e}$ *is regular the tangent line to* $\mathbf{e}$ *at* t *coincides with the normal line to* $\mathbf{r}$ *at* t;
(*b*) *the curve* $\mathbf{e}$ *has no linear points – in particular no ordinary point of inflection or undulation*;
(*c*) *if* $\mathbf{e}_1(t) = 0$ *but* $\mathbf{e}_2(t) \neq 0$ *then* $\mathbf{e}$ *has an ordinary cusp at* t;
(*d*) *the curve* $\mathbf{e}$ *is regular at* t *if and only if the curve* $\mathbf{r}$ *is non-circular at* t;
(*e*) *the curve* $\mathbf{e}$ *has an ordinary cusp at* t *if and only if the curve* $\mathbf{r}$ *has an ordinary vertex at* t;
(*f*) *the curve* $\mathbf{r}$ *has an ordinary vertex at* t *if and only if the radius of curvature* ρ *has an ordinary critical point at* t, *the cusp on* $\mathbf{e}$ *pointing towards or away from the vertex according as (the absolute value of)* ρ *has a local minimum or maximum at* t.

Proof By equation (3) in the preamble to this proposition $\mathbf{e}_1 \cdot \mathbf{r}_1 = 0$. Thus for each regular point t of $\mathbf{e}$ not only does $\mathbf{e}(t)$ lie on the normal line to $\mathbf{r}$ at t but also the tangent to $\mathbf{e}$ is normal to the tangent to $\mathbf{r}$; that is the tangent line to $\mathbf{e}$ at t coincides with the normal to $\mathbf{r}$ at t, which is assertion (*a*). That is the solution set of the equation for the normal to $\mathbf{r}$ at t, when put in the form

$$\text{‘}\mathbf{c} = \text{particular solution} + \text{kernel’},$$

is

$$\mathbf{c} = \mathbf{e}(t) + \lambda \mathbf{e}_1(t), \text{ for all } \lambda \in \mathbb{R}.$$

On differentiating the equation $\mathbf{e}_1 \cdot \mathbf{r}_1 = 0$ we get

$$\mathbf{e}_2 \cdot \mathbf{r}_1 + \mathbf{e}_1 \cdot \mathbf{r}_2 = 0.$$

Now, for $\mathbf{e}(t)$ to be defined, the vector $\mathbf{r}_2(t)$ is linearly independent of $\mathbf{r}_1(t)$, so that if $\mathbf{e}_1(t) \neq 0$ then $\mathbf{e}_1(t) \cdot \mathbf{r}_2(t) \neq 0$, from which it follows that $\mathbf{e}_2(t) \cdot \mathbf{r}_1(t) \neq 0$. But $\mathbf{e}_1(t) \cdot \mathbf{r}_1(t) = 0$. It follows that $\mathbf{e}_2(t)$ is linearly independent of $\mathbf{e}_1(t)$, which is assertion (*b*).

On differentiating the same equation a second time we get

$$\mathbf{e}_3 \cdot \mathbf{r}_1 + 2\mathbf{e}_2 \cdot \mathbf{r}_2 + \mathbf{e}_1 \cdot \mathbf{r}_3 = 0.$$

So if $\mathbf{e}_1(t) = 0$ but $\mathbf{e}_2(t) \neq 0$ then, since $\mathbf{e}_2(t) \cdot \mathbf{r}_1(t) = 0$, we must have $\mathbf{e}_2(t) \cdot \mathbf{r}_2(t) \neq 0$ and so also $\mathbf{e}_3(t) \cdot \mathbf{r}_1(t) \neq 0$. But then $\mathbf{e}_3(t)$ is linearly independent of $\mathbf{e}_2(t)$, which is assertion (*c*).

Next consider equation (4) of the preamble, namely

$$\mathbf{e}_1 \cdot \mathbf{r}_2 + V_3(\mathbf{e}) = 0.$$

Now at a point t where $\mathbf{r}$ is not circular $V_3(\mathbf{e})(t) \neq 0$, implying that $\mathbf{e}_1(t) \cdot \mathbf{r}_2(t) \neq 0$ and hence that $\mathbf{e}_1(t) \neq 0$. Conversely, if $\mathbf{e}_1(t) \neq 0$ then, since $\mathbf{e}_1(t) \cdot \mathbf{r}_1(t) = 0$, $\mathbf{e}_1(t) \cdot \mathbf{r}_2(t) \neq 0$, so that $V_3(\mathbf{e})(t) \neq 0$. This establishes assertion (d).

Differentiating (4) (where $V_3(\mathbf{e}) = (\mathbf{e} - \mathbf{r}) \cdot \mathbf{r}_3 - 3\mathbf{r}_1 \cdot \mathbf{r}_2$) gives

$$\mathbf{e}_2 \cdot \mathbf{r}_2 + 2\mathbf{e}_1 \cdot \mathbf{r}_3 + V_4(\mathbf{e}) = 0.$$

Now at an ordinary vertex $V_3(\mathbf{e}) = 0$ but $V_4(\mathbf{e}) \neq 0$, implying that at such a point $\mathbf{e}_1 \cdot \mathbf{r}_2 = 0$ but $\mathbf{e}_2 \cdot \mathbf{r}_2 + 2\mathbf{e}_1 \cdot \mathbf{r}_3 \neq 0$. Since also, by (3), $\mathbf{e}_1 \cdot \mathbf{r}_1 = 0$ it follows that $\mathbf{e}_1 = 0$ there. But then $\mathbf{e}_2 \cdot \mathbf{r}_2 \neq 0$, implying that $\mathbf{e}_2 \neq 0$ there and so, by assertion (c), that $\mathbf{e}$ has an ordinary cusp.

Conversely, at a point where $\mathbf{e}$ has an ordinary cusp $\mathbf{e}_1 = 0$, implying that $V_3(\mathbf{e}) = 0$ but $\mathbf{e}_2 \neq 0$, implying that $\mathbf{e}_2 \cdot \mathbf{r}_2 \neq 0$ and therefore that $V_4(\mathbf{e}) \neq 0$, so that $\mathbf{r}$ has an ordinary vertex there. Thus ordinary vertices of $\mathbf{r}$ and ordinary cusps of $\mathbf{e}$ correspond. This is assertion (e).

To prove the correspondence of each of these with ordinary critical points of ρ we note first that we may at any particular point assume that ρ is positive, by choosing the circular curve $\mathbf{n}$ appropriately near that point. Now consider equation (6) of the preamble, namely $\mathbf{e}_1 = \rho_1 \mathbf{n}$. Differentiating this we obtain $\mathbf{e}_2 = \rho_2 \mathbf{n} + \rho_1 \mathbf{n}_1$. Clearly $\mathbf{e}_1 = 0$ if and only if $\rho_1 = 0$, with $\mathbf{e}_2 \neq 0$ also if and only if $\rho_2 \neq 0$. Finally, when $\rho_1 = 0$, it follows by Proposition 1.6 from the equation $\mathbf{e}_2 = \rho_2 \mathbf{n}$ that the cusp on $\mathbf{e}$ points towards or away from the vertex of $\mathbf{r}$ according as the radius of curvature ρ has a local minimum or maximum there. This completes the proof of (f). □

Clearly the ordinary critical points of ρ are also the critical points of ρ^2. In fact earlier on when we went for a swim in the bay we were concerned with knowing at all times the points of the shore-line that were nearest to us. We develop this line of thought in Exercise 1.22.

Equally clearly the ordinary critical points of ρ are also the ordinary critical points of κ, provided that $\kappa \neq 0$.

The following proposition complements the one we have just proved.

Proposition 1.12 An ordinary undulation of $\mathbf{r}$ *is a point where the curvature has an ordinary critical point and where also the curvature is zero.* □