THE EDINBURGH MATHEMATICAL NOTES

PUBLISHED BY

THE EDINBURGH MATHEMATICAL SOCIETY

EDITED BY D. MARTIN, M.A., B.Sc., Ph.D.

No. 41 1957

The Definition of a Tangent to a Curve

By T. M. FLETT.

1. In elementary geometry, the tangent to a curve C at a point P is defined as the limiting position of the chord PQ as Q tends to P along the curve. Further, C is said to have a continuous tangent at P if it has a tangent at every point Q in the neighbourhood of P, and if the tangent at Q tends to the tangent at P as Q tends to P along C.[1]

It would perhaps be natural to expect that a curve which possesses a continuous tangent at each of its points should be fairly well-behaved, that it should for instance be rectifiable,[2] and that it should be regular if suitably parametrized (e.g. with respect to arc length). These results, however, are false if "tangent" is interpreted in the sense above. For example, the curve

$$x = 0, \qquad y = \begin{cases} t \sin 1/t & (0 < t \leqslant 1), \\ 0 & (t = 0), \end{cases}$$

possesses a continuous tangent at every point (namely the line $x = 0$) but has infinite length. This particular curve has also an infinity of multiple points, but it is easy to construct simple curves having these properties.

[1] See e.g. Fowler, *The Elementary Differential Geometry of Plane Curves* (2nd edition, Cambridge, 1929) pp. 8, 10.

[2] This result is actually stated by Fowler (*op. cit.*, p. 12).

The reason underlying the failure of these results to hold is that the property of possessing a continuous tangent does not prevent the curve from having cusps. In view of this it seems worth while to investigate some alternative definitions of tangent which do not suffer from this defect. We shall in fact consider three types of tangent. One is defined in a similar manner to that above, save that both end-points of the chord are allowed to vary, i.e. it is defined as the limiting position of the chord Q_1Q_2 as Q_1, Q_2 tend to P along C independently of each other. The other two are directed tangents, corresponding roughly to the two types of tangent already defined, but with the added condition that the *direction* of the chords should tend to a limit. It will appear in the sequel that, with these alternative definitions, a curve with a continuous tangent is rectifiable, and is regular when parametrized with respect to arc length.

It must be emphasised that there is no correlation between the existence of (continuous) tangents and the existence of (continuous) derivatives. Thus a curve

$$x = x(t),\ y = y(t),$$

can have a continuous tangent at a point t_0 (in all the senses considered here) although the derivatives x' and y' do not exist there [e.g. $x(t) = t^{\frac{1}{3}}$, $y(t) = t^{\frac{2}{3}}$, $t_0 = 0$], and can also have continuous derivatives at a point t_0 without having a tangent there in any sense [e.g. $x(t) = t^2$, $y(t) = t \mid t \mid$, $t_0 = 0$]. Thus our arguments must not assume the existence of the derivatives x' and y'.

It is hard to believe that the results of this paper do not occur somewhere in the literature, but they are by no means well known, and this seems sufficient excuse for bringing them together here.

2. We begin with some definitions and notation, taking the case of 3 dimensions as typical. A curve C is a continuous mapping of a closed interval I, $a \leqslant t \leqslant b$, into 3-dimensional Euclidean space,

$$x = x(t),\ y = y(t),\ z = z(t),$$

where $x(t)$, $y(t)$, $z(t)$ are real-valued functions continuous in I. We refer to the point $(x(t), y(t), z(t))$ of the space as the point t of C, and we use the terms "the arc t_0t of C" and "the chord t_0t" with their obvious meanings.

For any ξ and η of I let $d(\xi, \eta)$ be the length of the chord joining the points ξ and η of C, and let $L_1(\xi, \eta)$, $L_2(\xi, \eta)$, $L_3(\xi, \eta)$ be the

https://doi.org/10.1017/S0950184300003153 Published online by Cambridge University Press

direction-cosines of the directed segment from ξ to η, i.e.

$$L_1(\xi, \eta) = \frac{x(\eta)-x(\xi)}{d(\xi, \eta)},$$

and similarly for L_2 and L_3.

We consider three types of tangent, α, β, γ, defined as follows. We say that C has an α-tangent at the point t if the direction of the chord $\xi\eta$ tends to a limit as $\xi, \eta \to t$, the points ξ, η always being in the same order along C, i.e. if there exist numbers $l_1(t)$, $l_2(t)$, $l_3(t)$, such that, as ξ and η tend to t in such a manner that $\xi < \eta$

$$L_i(\xi, \eta) \to l_i(t) \qquad (i = 1, 2, 3).$$

The α-tangent at t is then the directed line through the point t of C with direction-cosines l_1, l_2, l_3.

In the case of the β-tangent, we ask only that the chord $\xi\eta$ should tend to a limiting position. More precisely, we say that C has a β-tangent at t if there exist numbers $l_1(t)$, $l_2(t)$, $l_3(t)$, such that either

$$L_i(\xi, \eta) \to l_i(t) \qquad (i = 1, 2, 3)$$

or

$$L_i(\xi, \eta) \to -l_i(t) \qquad (i = 1, 2, 3)$$

as $\xi, \eta \to t$. The β-tangent at t is then the (non-directed) line through the point t of C with direction-cosines l_1, l_2, l_3.

The γ-tangent is again a directed tangent, but here we make one of ξ and η coincide with t. Thus we say that C has a γ-tangent at t if there exist numbers $m_1(t)$, $m_2(t)$, $m_3(t)$, such that for each i

$$L_i(\xi, t) \to m_i(t)$$

as $\xi \to t-$, and

$$L_i(t, \eta) \to m_i(t)$$

as $\eta \to t+$. The γ-tangent at t is then the directed line through the point t of C with direction-cosines m_1, m_2, m_3.[1]

In each case we can also define the tangents at t from the left and from the right by imposing the obvious restrictions on ξ and η.

It is evident that an α-tangent is also a γ-tangent, but the converse is false; a curve can have a γ-tangent although the α-tangent does not exist. It is also obvious that if C has an α-tangent at t, then it has a β-tangent there. We shall show in § 4 that the converse

[1] We could also define the direction-cosines of the γ-tangent as the limits of $L_i(\xi, \eta)$ as $\xi, \eta \to t$ in such a manner that $\xi \leqslant t \leqslant \eta$, $\xi \neq \eta$. It is easily verified that this definition of the γ-tangent is equivalent to that given above.

here is true; the existence of the β-tangent implies that of the α-tangent.

We say that an α-tangent (and similarly β- and γ-tangents) exists at each point of a set $E \subset I$ if C has an α-tangent at every point of E interior to I, and also has an α-tangent from the right at $t = a$ and an α-tangent from the left at $t = b$ whenever these points belong to E. The tangent (α, β, or γ) is continuous at a point t_0 of C if it exists at every point t of C for all t in some neighbourhood of t_0 and, for each i, $l_i(t)$ [or $m_i(t)$] is continuous at t_0.

Finally, for any ξ, η of I we denote by $s(\xi, \eta)$ the length of the arc $\xi\eta$ of C (i.e. the supremum of the lengths of inscribed polygons, the value $+\infty$ being permitted). If $s(\xi, \eta) < \infty$, the arc $\xi\eta$ is rectifiable.

3. We consider now some consequences of the existence of the α-tangent, and here our arguments are very straightforward. We have first

THEOREM 1. *If the α-tangent exists at each point of a set $E \subset I$, then it is continuous in E.*

Given $\epsilon > 0$, for each t of E we can find $\delta(t, \epsilon)$ such that

$$| L_i(\xi, \eta) - l_i(t) | < \epsilon \qquad (i = 1, 2, 3)$$

for all ξ, η of I such that $\xi < \eta$ and $| \xi - t | < \delta(t, \epsilon)$, $| \eta - t | < \delta(t, \epsilon)$. If now t_0, t are points of E such that $| t - t_0 | < \delta(t_0, \epsilon)$, and ξ, η are points of I for which $\xi < \eta$ and $| \xi - t_0 | < \delta(t_0, \epsilon)$, $| \eta - t_0 | < \delta(t_0, \epsilon)$, $| \xi - t | < \delta(t, \epsilon)$, $| \eta - t | < \delta(t, \epsilon)$, then for each i

$$| l_i(t) - l_i(t_0) | \leqslant | l_i(t) - L_i(\xi, \eta) | + | L_i(\xi, \eta) - l_i(t_0) | < 2\epsilon,$$

and this proves the theorem.

THEOREM 2. *If the α-tangent exists at each point t of a closed set $E \subset I$, then, for each i, $L_i(\xi, \eta) \to l_i(t)$ as $\xi, \eta \to t$ in such a manner that $\xi < \eta$, uniformly with respect to t in E.*

Suppose for example that L_1 does not converge uniformly to l_1 in E. Then for some $\epsilon > 0$ we can find a sequence $\{t_n\}$ of points of E and points ξ_n, η_n of I such that $\xi_n < \eta_n$, $| \xi_n - t_n | < 1/n$, $| \eta_n - t_n | < 1/n$, and

$$| L_1(\xi_n, \eta_n) - l_1(t_n) | \geqslant \epsilon. \tag{3.1}$$

The sequence $\{t_n\}$ contains a convergent subsequence, $\{t_{n_\nu}\}$ say, con-

https://doi.org/10.1017/S0950184300003153 Published online by Cambridge University Press

verging to a point t of E, and, by the continuity of l_1 in E, $l_1(t_{n_\nu}) \to l_1(t)$. But also ξ_{n_ν}, $\eta_{n_\nu} \to t$, and, by (3.1),

$$l_1(t) = \lim_{\nu \to \infty} L_1(\xi_{n_\nu}, \eta_{n_\nu}) \neq \lim_{\nu \to \infty} l(t_{n_\nu}),$$

a contradiction.

We come now to theorems concerning arc length, and here we require the following

LEMMA. *Let* $a \leqslant t_0 < t \leqslant b$, *and let* $0 < \rho < \frac{1}{3}$. *If, for each* i

$$| L_i(\xi, \eta) - L_i(t_0, t) | < \rho$$

whenever $t_0 \leqslant \xi < \eta \leqslant t$, *then*

$$1 \leqslant \frac{s(t_0, t)}{d(t_0, t)} < \frac{1}{1 - 3\rho}. \tag{3.2}$$

The first inequality in (3.2) is obvious. To prove the second, let $t_0 \leqslant \xi < \eta \leqslant t$, and for each i write $L_i = L_i(\xi, \eta)$, $L_i' = L_i(t_0, t)$. If θ is the angle between the directed segment from ξ to η and that from t_0 to t, then [since $| L_i | \leqslant 1$]

$$\cos \theta = \Sigma L_i L_i' = 1 - \Sigma L_i(L_i - L_i') \geqslant 1 - \Sigma | L_i | \, | L_i - L_i' | > 1 - 3\rho. \tag{3.3}$$

Now let $t_0 < t_1 < \ldots < t_n = t$. Then

$$\sum_{\nu=1}^{n} d(t_{\nu-1}, t_\nu) \cos \theta_\nu = d(t_0, t),$$

where θ_ν is the angle between the directed segment from $t_{\nu-1}$ to t_ν and that from t_0 to t [for, by (3.3), $| \theta_\nu | < \frac{1}{2}\pi$ for each ν]. Since $\cos \theta_\nu > 1 - 3\rho$, we thus have

$$(1 - 3\rho) \sum_{\nu=1}^{n} d(t_{\nu-1}, t_\nu) < d(t_0, t),$$

and taking the supremum of the sum on the left we obtain the required inequality.

We now have

THEOREM 3. *If C has an a-tangent at the point t_0, then as $t \to t_0$*

$$\frac{s(t_0, t)}{d(t_0, t)} \to 1.$$

We prove the result as $t \to t_0$ from the right. An exactly similar argument applies to values of t less than t_0.

Given ϵ such that $0 < \epsilon < \frac{1}{6}$, we can find δ such that for each i

$$| L_i(\xi, \eta) - l_i(t_0) | < \epsilon$$

whenever ξ, η are points of I such that $\xi < \eta$ and $|\xi - t_0| < \delta$, $|\eta - t_0| < \delta$. If now $t_0 \leqslant \xi < \eta \leqslant t < t_0 + \delta$, then for each i

$$| L_i(\xi, \eta) - L_i(t_0, t) | < 2\epsilon.$$

Hence, by the lemma,

$$1 \leqslant \frac{s(t_0, t)}{d(t_0, t)} < \frac{1}{1 - 6\epsilon},$$

and this is the required result.

THEOREM 4. *If the a-tangent exists at each point of C, then C is rectifiable.*

By Theorem 2, given ϵ such that $0 < \epsilon < \frac{1}{9}$, we can find δ such that for each i

$$| L_i(\xi, \eta) - l_i(t) | < \epsilon \tag{3.4}$$

whenever $\xi < \eta$ and $|\xi - t| < \delta$, $|\eta - t| < \delta$. Further, since l_i is continuous in I, we can find points $a = t_0 < t_1 < \ldots < t_n = b$ such that for each i

$$| l_i(t) - l_i(t_{\nu-1}) | < \epsilon \tag{3.5}$$

whenever $t_{\nu-1} \leqslant t \leqslant t_\nu$; and the points t_ν can evidently be chosen so that also $\sup_\nu (t_\nu - t_{\nu-1}) < \delta$. It follows from (3.4) and (3.5) that

$$| L_i(\xi, \eta) - L_i(t_{\nu-1}, t_\nu) | < 3\epsilon$$

whenever $t_{\nu-1} \leqslant \xi < \eta \leqslant t_\nu$, whence, by the lemma,

$$s(t_{\nu-1}, t_\nu) < \frac{1}{1 - 9\epsilon} d(t_{\nu-1}, t_\nu).$$

Since the t_ν depend only on ϵ, the result follows by addition.

Combining Theorems 3 and 4 we now have

THEOREM 5. *If the a-tangent exists at each point of C, then x, y, z are continuously differentiable functions of* $s = s(a, t)$ *in* $a < t < b$, *and* $\frac{dx}{ds}, \frac{dy}{ds}, \frac{dz}{ds}$ *are never 0 simultaneously.*

4. We consider now the β-tangent, and prove

THEOREM 6. *If C has a β-tangent at t_0, then it also has an a-tangent there.*

We have to show that in the neighbourhood of t_0 the curve C cannot have two chords pointing in approximately opposite directions.

By hypothesis, given $\epsilon > 0$ we can find a neighbourhood U of t_0 (in I) such that (i) when ξ and η belong to U the acute angle between the chord $\xi\eta$ and the β-tangent at t_0 is less than ϵ, and (ii) the arc C' of C corresponding to t in U is simple. Let now ξ belong to U, and let $V = V(\xi)$ be the cone[1] with vertex at the point ξ of C, with axis parallel to the β-tangent at t_0, and with semi-vertical angle ϵ. By (i), the arc C' of C lies in V. Moreover, the arc of C' for which $t > \xi$ lies in one half of V, say $V^+(\xi)$, while the arc of C' for which $t < \xi$ lies in the other half, $V^-(\xi)$ [for C' can only pass from one half of V to the other via the vertex, and this is impossible except at $t = \xi$, since C' is simple].

If now C' has two chords pointing in approximately opposite directions, it is easy to see that we can find ξ_1 and ξ_2 in U with $\xi_1 < \xi_2$ and such that $V^+(\xi_1)$ and $V^+(\xi_2)$ point in opposite directions. Since the point ξ_2 of C lies in $V^+(\xi_1)$, we have $V^+(\xi_2) \supset V^-(\xi_1)$, and this is a contradiction, since $V^-(\xi_1)$ and $V^-(\xi_2)$ have points in common, other than the point ξ_2 of C. This completes the proof.

From Theorem 6 we deduce immediately the

COROLLARY. *The results of Theorems* 1-5 *continue to hold if "α-tangent" is replaced throughout by "β-tangent".*

5. We consider finally the γ-tangent, and here our arguments apply only to the plane, although the results presumably continue to hold in spaces of higher dimensions. We use a geometrical form of the mean-value theorem which I do not recall having seen in print elsewhere. We include for the sake of completeness a result concerning simple curves which we do not actually require here.[2]

THEOREM 7. *If the (plane) curve $x = x(t)$, $y = y(t)$ $(a \leqslant t \leqslant b)$ has a γ-tangent at every interior point, there is at least one ξ in $a < \xi < b$ such that the line of the γ-tangent at ξ (direction being disregarded) is parallel to the chord ab. Moreover, either ξ can be so chosen that the direction of the γ-tangent at ξ is that of the directed segment from a to b, or there is at least one η in $a < \eta < b$ such that the γ-tangent at η is perpendicular to the chord ab.*

---

[1] Infinite in both directions.

[2] I have used it without giving a proof in my paper "Some remarks on schlicht functions and harmonic functions of uniformly bounded variation", *Quart. J. of Math. (Oxford 2nd Series)*, 6 (1955), 59-72.

If in addition the curve is simple, then ξ can always be chosen so that the direction of the γ-tangent at ξ is that of the directed segment from a to b.

Suppose first that the curve has a γ-tangent at t_0, and let ψ_0 be a solution of the equations

$$\cos \psi_0 = m_1(t_0), \qquad \sin \psi_0 = m_2(t_0),$$

and for given positive ϵ and δ let $W^+(t_0; \epsilon, \delta)$ be the set of points $(x(t_0) + u \cos \psi, y(t_0) + u \sin \psi)$ for which $| \psi - \psi_0 | < \epsilon$ and $0 < u < \delta$. Let also $W^-(t_0; \epsilon, \delta)$ be the corresponding set of points such that $| \psi - \psi_0 | < \epsilon$ and $-\delta < u < 0$. Then each of W^+ and W^- is a wedge-shaped region with vertex at the point t_0 of the curve and with the tangent at t_0 as axis of symmetry. It is evident that W^+ contains all points t of the curve for which t is sufficiently near t_0 and $t > t_0$ while W^- contains all those points t for which t is sufficiently near t_0 and $t < t_0$.

Consider now the proof of the theorem. We may evidently suppose that $y(a) = y(b) = 0$, and that $x(b) > x(a)$. If the curve and chord coincide, there is nothing to prove, and we may therefore suppose that y is somewhere positive or negative, say the former. Since y is continuous, it attains its supremum in (a, b) at a point ξ, and $a < \xi < b$ since $y(a) = y(b) = 0$. If now the γ-tangent at ξ is not parallel to the x-axis, we can choose ϵ so small that $W^+(\xi; \epsilon, \delta)$ and $W^-(\xi; \epsilon, \delta)$ do not meet the line $y = y(\xi)$. Then one of W^+ and W^- lies in $y > y(\xi)$, and since both contain points of the curve, we have the necessary contradiction.

Suppose next that the direction of the γ-tangent at ξ is opposite to that of the chord from a to b. Then $W^+(\xi)$ lies to the left of ξ and $W^-(\xi)$ to the right, so that $x(t) - x(\xi)$ attains positive values in $a<t<\xi$ and negative values in $\xi < t < b$. Since either $x(a) - x(\xi)$ is negative or $x(b) - x(\xi)$ is positive, $x(t) - x(\xi)$ either attains its supremum in $a \leqslant t \leqslant \xi$ at an interior point η of this interval, or attains its infimum in $\xi \leqslant t \leqslant b$ at an interior point η of this interval. The argument above now shows that the γ-tangent to the curve at η is parallel to the axis of y, i.e. is perpendicular to the chord ab.

Suppose finally that the curve is simple. Let ξ_1 be the first point (after a) at which y attains its supremum in (a, b), and let ξ_2 be the first point after ξ_1 at which y attains its infimum in (ξ_1, b). If the direction of the tangent at ξ_1 is opposite to that of the chord from a to b, $W^+(\xi_1)$ lies to the left of the point ξ_1. In order to reach

https://doi.org/10.1017/S0950184300003153 Published online by Cambridge University Press

the point b, the curve must therefore pass to the left of the point a after leaving ξ_1 [1], whence y takes negative values in (ξ_1, b) and $\xi_2 \neq b$. If now the direction of the tangent at ξ_2 is also opposite to that of the chord from a to b, $W^+(\xi_2)$ lies to the left of the point ξ_2 of C. Moreover, for small enough ϵ and δ, $W^+(\xi_2)$ does not meet the arc $\xi_1\xi_2$ of C [for there is a $\xi_3 \neq \xi_2$ on the arc $\xi_1\xi_2$ such that the open arc $\xi_3\xi_2$ lies in $W^-(\xi_2)$. Since ξ_2 is the first point after ξ_1 at which y attains its infimum in $\xi_1 \leqslant t \leqslant b$, the infimum of y in $\xi_1 \leqslant t \leqslant \xi_3$ is strictly greater than $y(\xi_2)$, and the statement follows]. But since the arc of the curve from ξ_2 to b joins points of $W^+(\xi_2)$ to b, it must cross the arc of the curve from ξ_1 to ξ_2, and this gives the necessary contradiction. Hence the direction of the tangent at either ξ_1 or ξ_2 is the same as that of the chord ab.

From Theorem 7 we deduce immediately

THEOREM 8. *If C is a plane curve, and if the γ-tangent to C exists and is continuous at t_0, then C has an α-tangent at t_0.*

If the result is false, we can find a positive ϵ and sequences $\xi_n, \eta_n, \xi_n', \eta_n' \to t_0$ such that $\xi_n < \eta_n$, $\xi_n' < \eta_n'$, and

$$| L_i(\xi_n, \eta_n) - L_i(\xi_n', \eta_n') | \geqslant \epsilon \qquad (i = 1, 2).$$

It follows now from Theorem 7 that there exist pairs of points t_n, t_n' as close to t_0 as we please such that the angle between the γ-tangents at t_n, t_n' exceeds some positive fixed number,[2] contradicting the continuity of the tangent at t_0.

We thus have the

COROLLARY. *The results of Theorems 2-5 continue to hold if the existence of the α-tangent is replaced by the existence and continuity of the γ-tangent.*

---

[1] Else it crosses itself. This, and the similar point which occurs later in the argument, seem to require something akin to the Jordan curve theorem for their disposal.

[2] Either there are points between the end-points of the chords $\xi_n\eta_n$, $\xi_n'\eta_n'$ at which the γ-tangents are parallel to the (directed) chords, or there are points at which the γ-tangents are perpendicular to each other.

DEPARTMENT OF PURE MATHEMATICS,
THE UNIVERSITY,
LIVERPOOL.

https://doi.org/10.1017/S0950184300003153 Published online by Cambridge University Press