Local corner cutting and the smoothness of the limiting curve

Carl de Boor

submitted dec88, revised apr89, published CAGD; 7; 1990; 389-397

Abstract. Stimulated by recent work by Gregory and Qu, it is shown that the limit of local corner cutting is a continuously differentiable curve in case the corners of the iterates become increasingly flat.

It was proved in [de Boor '87] that corner cutting of any kind converges to a Lipschitzcontinuous curve, but the question of how one might guarantee that the limiting curve be smoother than that was not considered there. Recently, Gregory and Qu [Gregory, Qu '88] took up this question and established sufficient conditions for a certain systematic and local corner cutting scheme to give a limiting curve in C^1 . Since [Gregory, Qu '88] use the same parametrization of the successive broken lines that made the argument in [de Boor '87] so simple, I became intrigued and took a look at what one might say in greater generality. Specifically, I looked for conditions under which continuous differentiability of the limiting curve could be inferred from the fact that the corners of the broken lines flatten out eventually.

It is the purpose of this note to prove that the limit of any 'local' corner cutting scheme is in C^1 provided the corners of the broken lines become increasingly flatter. A simple example is given to show that this condition is not necessary, while another example shows that, without 'localness', the condition is not sufficient, in general. Finally, as an application, the nice argument in [Gregory, Qu '88] is redone.

1. Cutting corners

In this section, we recall the setup of [de Boor '87].

We deal with a sequence $(b_n)_{n=0}^{\infty}$ of broken lines in which, for n > 0, b_n is obtained from b_{n-1} by a 'cut', i.e., by replacing a curve segment by the subtended secant to the curve. This means that all the vertices of b_n lie on b_{n-1} , i.e., b_n can be thought of having been obtained from b_{n-1} by interpolation. This observation is used in [de Boor '87] to prove that, no matter just how the cutting was done to generate the sequence (b_n) from an initial broken line b_0 with finitely many vertices, $b_{\infty} := \lim_{n \to \infty} b_n$ exists as a Lipschitzcontinuous curve which is approached uniformly by b_n , i.e., $\lim_{n\to\infty} \text{dist}(b_n, b_{\infty}) = 0$.

The argument in [de Boor '87] was based on parametrizing the curves appropriately. If (v_i) is the sequence of vertices of b_n and (t_i) is a corresponding arbitrary increasing sequence of numbers, then b_n can be parametrized by

(1.1)
$$b_n(t) := v_{i-1} \frac{t_i - t}{t_i - t_{i-1}} + v_i \frac{t - t_{i-1}}{t_i - t_{i-1}}, \quad t_{i-1} \le t \le t_i, \quad \text{all } i.$$

Since b_n is obtained from b_{n-1} by interpolation, it is natural to choose the sequence (t_i) in dependence on the parametrization of b_{n-1} , i.e., so that $b_n(t_i) = v_i = b_{n-1}(t_i)$ for all *i*. With this,

$$b_n = P_n b_{n-1},$$

where P_n is broken line interpolation at the points (t_i) . Therefore, ultimately, $b_n = P_n \cdots P_1 b_0$, with $P_n \cdots P_1$ a *linear* map. Hence, although the process of generating the sequence (b_n) is nonlinear (in that it is quite arbitrary), once we have decided on how to cut, we can think of each b_n as a linear function of b_0 . In particular, writing b_0 in any one of many reasonable ways as a sum

$$b_0 = \sum_i w_i \varphi_i$$

of scalar-valued functions φ_i with vector coefficients $w_i \in \mathbb{R}^d$, we have

$$b_n = \sum_i w_i P_n \cdots P_1 \varphi_i,$$

and questions of convergence or of smoothness of the limit can be settled by settling them for the (presumably simpler) sequences $(P_n \cdots P_1 \varphi)$, with φ any one of the φ_i . A particularly simple choice for the φ_i are the truncated powers $(\cdot - \tau_i)_+$ (in addition to the constant function), and this leads to the conclusion that the nature of the limiting curve can be understood if one understands what the particular corner cutting process does to the **standard corner**, i.e., the broken line with vertices (0,0), (1,0), (2,1).

2. Examples

We are now ready to consider the smoothness of the limiting curve, having understood that it is sufficient to consider the case that b_0 is a piecewise linear (real-valued) function on some interval. Implicit in this statement is the claim that the corners of the *curve* b_n become flat if and only if the corners of its component *functions* become flat. While the increasing flatness of the component functions does indeed imply the increasing flatness of the curve so parametrized, the converse does not hold for general corner cutting since it is possible to obtain a nonregular parametrization thereby. (Take, e.g., b_0 to have the vertices (-1,0),(1,0),(1,1),(0,0),(2,0), use arclength parametrization for b_0 , and obtain b_1 by cutting out the loop, e.g., by cutting across the parameter interval [1,5]. Then b_1 has no corners, while b'_1 has jumps.) But in 'local' corner cutting to be discussed, such examples are not possible, as is argued at the end of the next section.

With b_0 a piecewise linear *function*, each b_n is of the same nature, and its derivative, $d_n := b'_n$, is a step function. In the notation adopted in the preceding section,

(2.1)
$$d_n = \sum_i (\cdot - t_i)^0_+ \operatorname{jump}_{t_i} d_n,$$

with $\operatorname{jump}_t d := d(t+) - d(t-)$ the difference between the limit from the right and the limit from the left at t. We can take the absolutely largest jump, i.e., the number $\|\operatorname{jump}_{()}d_n\|_{\infty}$, as a measure of the extent to which b_n fails to be in C^1 .

It will be useful to visualize the process by which d_n is obtained from d_{n-1} . Suppose that b_n is obtained from b_{n-1} by replacing b_{n-1} on [s, t] by the linear interpolant to b_{n-1} at s and t. Then

$$\int_{s}^{t} (d_{n} - d_{n-1})(x) dx = (b_{n} - b_{n-1})(t) - (b_{n} - b_{n-1})(s) = 0.$$

Also, d_n is a constant (viz., the difference quotient $[s,t]b_{n-1}$) on [s,t]. If now b_{n-1} has just one vertex in [s,t], then d_{n-1} has just two steps there, hence $d_n - d_{n-1}$ has just two steps there and, as $\int_s^t (d_n - d_{n-1}) = 0$, the two rectangles which make up this integral must balance; see (2.2)Figure.



(2.2) Figure The change in the derivative as the result of a corner cut. The two areas are of equal size.

More precisely, let J be the jump in d_{n-1} at that sole vertex in [s, t], let l and r be the parametric distances of the corner from its left and right neighbor, and assume that the two new vertices (which replace the vertex being cut off) occur at parametric distances $l\beta$ and $r\alpha$, respectively (with $\alpha, \beta \in [0, 1]$). Then the two new jumps are of size $h_l J$ and $h_r J$, with $h_l + h_r = 1$ and $l\beta h_l = r\alpha h_r$, or,

(2.3)
$$h_l = \frac{r\alpha}{l\beta + r\alpha} = 1 - h_r.$$



(2.4) Figure These uniformly nonsmooth broken lines converge to a smooth limit.

Here is an example to show that the maximum jump need not go to zero for the limit function to be C^1 . As shown in (2.4)Figure, the limiting function is smooth (it is the zero function), while the maximum jump in the first derivative stays above 1 (in absolute value). Note that, in this example, we have done two cuts simultaneously, i.e., in terms of the setup adopted earlier, we are showing only every other iterate. Note also that, strictly speaking, each of our cuts involves two corners. It will be important later on to know that such an example (of a limiting curve being smooth even though the absolutely largest jump in the derivative of the iterates is bounded away from zero) can also be given when each cut involves only exactly one corner. Such an example can be supplied by applying, e.g., Chaikin's algorithm to the initial broken line in the above example, except that the points on the two segments flanking the middle segment are chosen closer and closer to the farther endpoint.

We also illustrate the fact that having the maximum jump in the first derivative go to zero is, in general, no guarantee that the limiting curve is C^1 . The essence of such an example is the observation that the limiting d_{∞} will have a jump at any point at which sufficiently many jumps (of the same sign) of the iterates d_n accumulate. Individual jumps may well become arbitrarily small, yet their combined strength may force a jump in d_{∞} . Since corner cutting does not increase the sum of the absolute jumps, this requires **migration** of jumps, and this can already be accomplished by cuts involving just one corner, as long as such a cut is allowed to start at a neighboring corner. For (see (2.2)Figure) in that case, part of the jump at the corner being cut away is transferred to the jump at the neighboring corner, thus increasing it if it is in the same direction. A subsequent corner cut can move most of that jump further.



(2.5) Figure By a sequence of edge cuts (as the one illustrated on the left), all vertices (o) of the original broken line are shifted by the same amount to the NNW (*). After a succession of such cuts, a broken line (•) with the same angles but vertices much more clustered is obtained.
(Only the left half of a symmetric situation is shown.)

Without going into all the details, here is a specific example whose final form has benefitted much from comments by Hartmut Prautzsch [Prautzsch '89]. As he suggests, only two kinds of cuts need to be used, a symmetric cut across a corner which therefore replaces that corner's jump by two jumps of half the size, and an **edge cut**, i.e., a cut across two corners which results in an edge parallel to the edge cut away, hence leaves the jumps at the edge's endpoints unchanged. Note that such an edge cut can be effected by two cuts across just one corner but including a neighboring corner as an endpoint (see the detail in the left part of (2.5)Figure). The resulting (temporary) increase in vertex angle cannot exceed the sum of two neighboring vertex angles.

(2.5)Figure shows the left half of the symmetric starting configuration (o). A sequence of edge cuts starting with the third edge from the left (with a symmetric sequence of edge cuts on the right, but not shown here) shifts all but the first two vertices shown the same amount to the left, to the locations marked (*). A next move would shift all but the first three of these vertices the same amount to the left, then all but the first four vertices shown the same amount to the left, etc, always maintaining symmetry. The new configuration obtained in the end is also shown, with the new vertices marked (•). It has exactly the same angles as the original configuration, but the corners are now much more closely clustered near the second (and second last) vertex.

Repetition of this process produces a sequence of broken lines converging to a broken line with a corner at the second (and second last) vertex, while maintaining the original angles (and not introducing any new or bigger angles). It is now a simple matter also to make all the corners flatter: Start off the beginning of each repetition with a symmetric cut across each active corner, i.e., each vertex other than the second and second last, (say from position 2/3 on the left segment to position 1/3 on the right). This will double the number of vertices and halve all angles. Since the process preserves angles, the resulting sequence of broken lines has all its vertex angles go to 0 uniformly, yet converge to a curve which is not C^1 .

Nevertheless, if the corner cutting is local, then having the absolutely largest jump in the first derivative (or, equivalently, the largest vertex angle) go to zero does imply that the limiting function is in C^1 . This is the content of the next section.

3. Local corner cutting

We say that the corner cutting is **local** in case any cut involves exactly one corner. This means that the cut endpoints must lie in the interior of the two segments which form the corner being cut. Schemes that cut all corners simultaneously fall into this category as long as the cuts of neighboring corners do not share an endpoint. For we can then think of them as having been carried out one cut at a time. In particular, the corner cutting scheme considered in [Gregory, Qu '88] is local in this sense, as are the schemes considered in [de Rham '47] so many years ago. It follows that every segment, of the original broken line as well as of any subsequently generated broken line, is tangent to the limiting curve, hence the situation depicted in (2.5)Figure could not have been generated by *local* corner cutting. On the other hand, we can obtain any Lipschitz-continuous curve (with finitely many regions of concavity/convexity) by local corner cutting, by starting off with any sufficiently articulated broken line whose segments are tangent to the target curve, and then using only cuts whose secant touches the curve.

(3.1)**Theorem.** The limiting curve produced by a local corner cutting scheme is C^1 in case the maximum jump in the first derivative of the iterates b_n goes to zero as $n \to \infty$. The converse holds in case b_0 is convex and for arbitrary corner cutting.

Proof: It is sufficient to consider the special case that b_0 is the 'standard corner'. Then b_0 is convex, hence so are all the iterates, with b_n growing uniformly, and pointwise monotonely, toward the limit function b_{∞} , which is also convex. Let b be one of the iterates. Then d := b' is a monotone increasing step function. Recall from (2.1) that

$$d = \sum_{j} (\cdot - t_j)^0_+ \operatorname{jump}_{t_j} d,$$

with $0 < t_1 < \ldots < t_m < 2$ its breakpoints. We consider also the two step functions d^+ and d^- , given by the rule

$$d^{\pm} = \sum_{j=1}^{m} (\cdot - t_{j \mp 1})^{0}_{+} \operatorname{jump}_{t_{j}} d_{j}$$

with $t_0 := 0$ and (for the sake of neatness) $t_{m+1} := 2$. Then $||d^+ - d^-||_{\infty} \le 2||\text{jump}_{()}d||_{\infty}$, while $d^- \le d \le d^+$ pointwise, since d is monotone increasing.



(3.2) Figure Local corner cutting contracts the 'envelope' formed around the derivative d by the step functions d^- and d^+ .

Further, if b^* is obtained from b by cutting off exactly one corner, and d^* is, correspondingly, the derivative of b^* , then (see (3.2)Figure)

(3.3)
$$d^{-} \le (d^{*})^{-} \le (d^{*})^{+} \le d^{+}.$$

This implies that the first derivative of all subsequent iterates lies between d^- and d^+ . Hence, if $\|\operatorname{jump}_{()}d_n\|_{\infty} \longrightarrow 0$ as $n \to \infty$, then (d_n) is a Cauchy sequence in the complete normed linear space of all bounded functions on [0, 2] (with the max-norm). Consequently, $d_n = b'_n$ converges uniformly to some bounded function d_{∞} . Now consider the modulus of continuity ω_{∞} of this limiting function. For any h > 0 and any n,

$$\omega_{\infty}(h) = \sup_{0 < t-s < h} (d_{\infty}(t) - d_{\infty}(s)) \le \sup_{0 < t-s < h} (d_n^+(t) - d_n^-(s)) =: \overline{\omega}_n(h).$$

Note that $\overline{\omega}_n$ is a nondecreasing step function, with $0 \leq \overline{\omega}_n(0+) = \sup_j (d_n(t_{j+2}+) - d_n(t_{j-1}-)) \leq 3 \|\operatorname{jump}_{()} d_n\|_{\infty} \longrightarrow 0$ as $n \to \infty$. Consequently, for every $\varepsilon > 0$, we can find n and $\delta > 0$ so that, for all $h < \delta$, $\overline{\omega}_n(h) < \varepsilon$. This proves that $\omega_{\infty}(0+) = 0$, and so establishes that d_{∞} is continuous.

We now know that b_n converges uniformly to some Lipschitz-continuous function b_{∞} , while $d_n := b'_n$ converges uniformly to some continuous function d_{∞} . It is a standard result that, therefore, $b'_{\infty} = d_{\infty}$, i.e., the limiting function b_{∞} is in C^1 .

For the converse, assume that the sequence (b_n) of convex broken lines, all defined on the interval [0,2] say, converges uniformly to some function b. Assume further that b is continuously differentiable at the interior point p. This means that, for some modulus of continuity ω (i.e., some positive function ω on $(0, \infty)$ with $\omega(0+) = 0$),

$$[t,p]b := (b(t) - b(p))/(t-p) = b'(p) + O(\omega(|t-p|)).$$

Now consider $J := \operatorname{jump}_p b_n$ for some n, and let $\varepsilon := \|b - b_n\|_{\infty}$. Then, for any small positive h,

$$0 \le J \le [p+h,p]b_n - [p,p-h]b_n \le [p+h,p]b - [p,p-h]b + 4\varepsilon/h = O(\omega(h)) + 4\varepsilon/h.$$

Since $\omega(0+) = 0$ and $\varepsilon = ||b - b_n|| \to 0$ as $n \to \infty$, this implies that J must be small when n is large.

Remarks (i) The theorem fails already if the concept of a 'local' cut is relaxed to permit inclusion of the segment endpoint(s), as the earlier example illustrates (and [de Boor '88] ignorantly contradicts).

(ii) In local corner cutting as defined here, each original segment and all segments subsequently generated are tangent to the limiting curve. In particular, whatever the parameter values τ_i assigned to the vertices of the initial broken line b_0 , there exists a point σ_i between τ_i and τ_{i+1} so that, for all n, $b_n(\sigma_i) = b_0(\sigma_i)$. For each n, the curve segment $b_n([\sigma_{i-1}, \sigma_i])$ turns monotonely through the same total angle, with the maximum angle (weakly) decreasing as n increases. This implies that the component functions are convex on the parameter inverval $[\sigma_{i-1}, \sigma_i]$, hence converge uniformly and monotonely there. Further, arclength s_n of b_n as a function of the parameter t is a strictly monotone broken line on that parameter interval (its slopes being bounded above by 1 and below by $||b_0(\sigma_{i-1}) - b_0(\sigma_i)||/|\sigma_{i-1} - \sigma_i|$).

Consider now the angle α at some vertex $v = b_n(\tau)$ with $\tau \in [\sigma_{i-1}, \sigma_i]$. Then, using a dot to denote differentiation with respect to arclength,

$$\|\operatorname{jump}_{\tau} b'_n / s'_n\|^2 = \|\dot{b}_n(\tau+) - \dot{b}_n(\tau-)\|^2 = 2(1 - \cos(\alpha)).$$

But $s'_n = ||b'_n||$ is bounded above and below. Hence $||\text{jump}_{\tau}b'_n||$ can be bounded in terms of $\sqrt{1 - \cos(\alpha)}$. This proves that the jumps in the slope of the component functions must go to zero if the corners of b_n flatten as $n \to \infty$, hence shows that b_{∞} is C^1 in that case.

4. The Gregory-Qu result

As an application of (3.1)Theorem, we now consider the Gregory-Qu scheme, in which b_n is obtained from b_{n-1} by a simultaneous, non-interfering, cutting of all corners; hence the scheme is local in the sense defined earlier. In fact, [Gregory, Qu '88] assumes that the new vertices generated are in the interior of old segments (i.e., that all α and β are positive, in the notation used there and introduced below). But if we allow also trivial cuts, i.e., cuts that begin and end at the same vertex ($\alpha = 0$ or $\beta = 0$), then this scheme models any local corner cutting.



(4.1) Figure The change in the derivative due to one step of the Gregory-Qu process.

To prove that the limit is in C^1 , it is therefore sufficient to prove that the jumps in the first derivative go to zero. For this, we discuss the scheme in terms of the step function which is the first derivative of the broken line in question. A look at (4.1)Figure might be helpful.

The single jump of height J, with left and right segments of length l, r, spawns two jumps, a left one of height Jh_l , with $h_l := r\alpha_r/(l\beta_l + r\alpha_r)$ and with segments $l_l := l(1 - \alpha_l - \beta_l)$ and $r_l := l\beta_l + r\alpha_r$, and a right one of height Jh_r , with $h_r := 1 - h_l$ and with segments $l_r := r_l$ and $r_r := r(1 - \alpha_r - \beta_r)$. Since

$$h_l = \frac{1}{1 + (\beta_l / \alpha_r) l/r} = 1 - h_r = 1 - \frac{1}{1 + (\alpha_r / \beta_l) r/l}$$

we can be assured that h_l and h_r are uniformly smaller than 1 (hence the limiting curve is in C^1) provided we can show that the local mesh ratio l/r is bounded away from 0 and ∞ . For this, consider the local mesh ratios l_l/r_l and l_r/r_r spawned by the cutting of this corner. We find

$$l_l/r_l = \frac{l(1 - \alpha_l - \beta_l)}{l\beta_l + r\alpha_r} = L(l/r),$$

with

$$L(t) := \frac{t(1 - \alpha_l - \beta_l)}{t\beta_l + \alpha_r}.$$

This function is increasing on $[0, \infty)$, starting at 0 with a value of 0 and a slope of $L'(0) = (1 - \alpha_l - \beta_l)/\alpha_r$ at 0 and taking the limiting value $L(\infty) = (1 - \alpha_l - \beta_l)/\beta_l$. Consequently, L maps the interval $[0, \infty)$ into the interval $[0, L(\infty)]$. Further, if L'(0) > 1, then L has an attracting fixed point in that interval, viz. the point $F := (1 - \alpha_l - \beta_l - \alpha_r)/\beta_l = (L'(0) - 1)\alpha_r/\beta_l$. This means that L maps any interval [1/M, M] containing F (and contained in $[0, \infty]$) into itself.

By symmetry, $r_r/l_r = R(r/l)$, with

$$R(t) := \frac{t(1 - \alpha_r - \beta_r)}{\beta_l + t\alpha_r}$$



(4.2) Figure The function L contracts around the point F.

a function which maps $[0, \infty)$ into the interval $[0, R(\infty)]$, and which has the fixed point $G := (1 - \alpha_r - \beta_r - \beta_l)/\alpha_r = (R'(0) - 1)\beta_l/\alpha_r$ in case $R'(0) = (1 - \alpha_r - \beta_r)/\beta_l$ is greater than one. This means that R maps any interval [1/M, M] containing G (and contained in $[0, \infty)$) into itself.

We conclude that the local meshratios are bounded away from 0 and ∞ provided the fixed points F and G are eventually bounded away from 0 and ∞ . As [Gregory, Qu '88] point out, this can be guaranteed by having α and β eventually bounded away from zero, i.e. having both $\underline{\alpha} := \liminf \alpha$ and $\underline{\beta} := \liminf \beta$ be positive, and having L'(0), R'(0) > 1 when formed with $\alpha_l, \alpha_r = \overline{\alpha} := \limsup \alpha$ and $\beta_l, \beta_r = \overline{\beta} := \limsup \beta$. This amounts to the conditions

$$0 < \underline{\alpha}, \beta$$
 and $\overline{\alpha}, \overline{\beta} < 1 - \overline{\alpha} - \overline{\beta}.$

When these conditions are violated, we cannot be certain that the local meshratios stay away from 0 or ∞ , hence the reduction factors h_l or h_r may come close to 1. This does not, of itself, imply that the limiting curve has corners. But the above discussion is sufficient to show that the limiting curve has corners if L'(0) or R'(0) are uniformly below 1.

We discuss this only for the case of constant α and constant β . Assume, for example, that α and β are such that

$$R'(0) = \frac{1 - \alpha - \beta}{\beta} < 1.$$

Then, starting with the 'standard corner', the cutting process generates a sequence of vertices proceeding to the right with associated local mesh ratios r/l equal to

$$R(1), R^{2}(1) = R(R(1)), R^{3}(1), \dots$$

which decay geometrically to zero. In fact, $R^n(1) \sim (R'(0))^n$ as $n \to \infty$. The corresponding reduction factors therefore satisfy

(4.3)
$$h_r^{(n)} = \frac{1}{1 + (\alpha/\beta)R^n(1)} = 1 - (\alpha/\beta)R^n(1) + O((R^n(1))^2) \sim 1 - (\alpha/\beta)(R'(0))^n.$$

We want to show that the corresponding sequence of jumps is bounded away from zero. Since this is a decreasing sequence, it is sufficient to show that its limit, the infinite product

$$\prod_{n=1}^{\infty} h_r^{(n)},$$

is positive. This is the same as proving that the infinite series

$$\sum_{n=1}^{\infty} \ln h_r^{(n)}$$

is finite. From (4.3), $\ln h_r^{(n)} \sim -(\alpha/\beta)(R'(0))^n$, i.e., the terms of the sum behave like that of a convergent geometric series, hence the series converges.

The foregoing analysis also explains the **fractal** nature of the resulting curves (when using constant α and β). For it shows that the height of a particular jump or the meshratio at a particular breakpoint of the *n*th iterate is the result of two **contending** fixed point iterations, *L* and *R*, with the influence of each entirely determined by the particular sequence of right and left turns taken to reach the breakpoint in question from the original breakpoint. In particular, we expect any collection of jumps sharing the first few of these turns to look like any other collection of jumps sharing the first few of these turns.

It would be interesting to explore further special situations when the fixed points coincide. For example, both fixed points (for L and R) are 1 exactly when $\alpha + \beta = 1/2$. In this case, all the mesh ratios are the same, hence the left factors h_l are all the same as are all the right factors h_r . The requirement that $h_l = h_r$ is satisfied exactly when $\alpha = \beta$. Thus the imposition of both requirements leads to $\alpha = 1/4 = \beta$ which is Chaikin's algorithm.

References

de Boor, C. (1987), Cutting corners always works, Computer Aided Geometric Design 4, 125–131.

de Boor, C. (1988), Local corner cutting and the smoothness of the limiting curve, CMS Technical Summary Report #89-18.

Gregory, John A. and Qu, R. (1988), Non-uniform corner cutting, ms.

Prautzsch, H. (1989), letter of Mar.3.

de Rham, G. (1947), Un peu de mathematique à propos d'une courbe plane, Elem. Math. 2, 73–76; 89–97. (*Collected Works*, 678–689).