Binary Subdivision Schemes for Functions over Irregular Knot Sequences

J. Warren

Abstract. For a wide class of stationary subdivision methods, we derive necessary and sufficient conditions for these schemes to produce C^k continuous limit curves. These stationary schemes include those arising from midpoint subdivision of irregularly-spaced knot sequences. We also describe a matrix method for computing various derivative schemes associated with such stationary schemes.

§1. Introduction

Subdivision is a powerful tool for creating and representing complex curved shapes. Consider the process due to Chaikin [2] depicted in figure 1. A polygon is transformed into a new polygon with twice as many segments. For this particular transformation, the vertices of the new polygon are placed $\frac{1}{4}$ and $\frac{3}{4}$ of the way between the old vertices. Applying this process repeatedly yields a polygon with a great number of segments that closely approximate a smooth curve. What is this smooth curve? Riesenfeld [10] shows that the curve is a uniform quadratic B-spline whose control points are the original polygon.

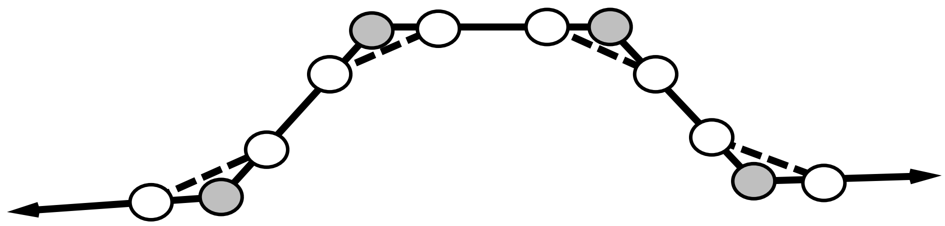

Fig. 1. A subdivision method.

Mathematical Methods in CAGD III
M. Dæhlen, T. Lyche, and L. L. Schumaker (eds.), pp. 1–3.

ISBN xxx.

Implementing this subdivision process is simple. Two copies of each vertex are made. These vertices are then positioned by taking fixed affine combinations of nearby vertices. The key to this process is the choice of the affine combinations. Some combinations lead to fractal curves. Some combinations lead to very smooth curves.

An extensive theory of univariate subdivision has been developed that relates the combinations used to the smoothness of the corresponding limit curve. [1,5] give a comprehensive survey of this theory. However, this theory was developed with the assumption that the parameterization underlying the subdivision method is uniform, that is the ith control point is associated with a parameter value proportional to i. In the parametric case, this is not a substantial restriction. However, in the functional case, this is a non-trivial restriction. In the multi-variate case, this restriction is even worse. The uniform approach cannot deal with the irregular triangulations that often arise during the modeling of complicated shapes.

Our approach is to develop a theory of univariate subdivision for irregularly spaced knot sequences in the functional setting. The corresponding theory for the parametric case can then be derived easily. A theory for the irregular case is also useful in applications such as finite element analysis that are intrinsically functional. Eventually, we hope that this theory will be a special case of a general multi-variate theory of subdivision over irregular triangulations.

§2. Preliminaries

§2.1 Knot Sequences

Our approach in defining subdivision is similar to that of B-splines. A subdivision method is driven by a sequence of scalar values called *knots*. We require that there exist an $\epsilon > 0$ such that the initial knot sequence $n = n^0$ satisfies $n_{i+1} - n_i > \epsilon$ for all i. This restriction ensures that the knots appear in ascending order and fill the parameter line. Subsequent knot sequences n^j are defined by midpoint insertion.

$$
\begin{aligned}
n_{2i}^{j+1} &= n_i^j, \\
n_{2i+1}^{j+1} &= \frac{n_i^j + n_{i+1}^j}{2}.
\end{aligned}
$$

§2.2 Subdivision Schemes

A subdivision scheme is a map from knot sequences n to a collection of subdivision rules $S[n]$. Given the sequence of knot vectors n^j, the subdivision rules $S[n^j]$ may be viewed as a matrix that maps the jth set of control points p^j into a new set of control point p^{j+1}.

$$p^{j+1} = S[n^j]p^j. \tag{1}$$

To simplify our indexing notation, we treat the p^j as an infinite set of control points. Likewise, $S[n^j]$ is a bi-infinite matrix. We restrict our attention to subdivision schemes that satisfy four important properties

Compact support: There exists nonnegative a and b such that, for all k, the kth column of $S[n]$ is zero except from row $2k-a$ to $2k+b$.

Affine invariance: For any scaled and translated knot sequence $\alpha n+\beta$ with $\alpha > 0$,

$$S[n] = S[\alpha n + \beta].$$

Index invariance: If $r(n)$ is the vector whose ith entry is n_{i+1}, then

$$S[r(n)]_{i,j} = S[n]_{i+2,j+1} \quad \forall\ j.$$

Local definition: The kth column of $S[n]$ depends only on the knots n_i where $|k-i|$ is bounded independent of k.

Compact support ensures that the infinite sum in equation 1 is well-defined. The number $a+b+1$ is the *column height* of S. The 2-1 slant of $S[n]$ doubles the number of control points during each step of subdivision. If the knot sequence n is regular (i.e $n_i = i$), then affine invariance ensures the resulting subdivision scheme is *stationary*, that is $S[n] = S[\frac{n}{2^j}] = S[n^j]$ for all j. For regular knot sequences n, the index shift $r(n)$ is also an affine transformation of n and $S[n] = S[r(n)]$. In this case, index invariance forces the subdivision scheme to be *uniform* ($S[n]_{i,j} = S[n]_{i+2,j+1}$). Local definition is critical in showing that midpoint subdivision of irregular knot sequences leads to stationary subdivision schemes.

§2.3 Parameterization of Subdivision Methods

In standard subdivision, the control points p^j are associated with a piecewise linear function that interpolates the value p_i^j at the parameter value $\frac{i}{2^j}$. In our analysis, the parameter values $\frac{i}{2^j}$ are replaced by the knot vector n^j. Given n^j and a vector of control points p^j, we let $L[n^j,p^j](t)$ denote the piecewise linear function interpolating the points (n_i^j, p_i^j) for all i (see figure 2). $L[n^j,p^j](t)$ can be viewed as a function in t since the n_i^j are indexed in ascending value. The knot vectors n^j provide *parameter* values for the control points p^j.

The natural object to consider here is the limit of the functions $L[n^j,p^j](t)$ as j goes to infinity. Given an initial set of control points $p = p^0$ and an initial knot vector $n = n^0$, we define the limit function associated with the process of equation 1

$$F[n,p](t) = \lim_{j\to\infty} L[n^j,p^j](t).$$

Here, the limit (if it exists) is taken *point-wise*, that is individually for each distinct t. By construction, the limit operator F is linear in p. Specifically,

$$F[n,\alpha p](t) = \alpha F[n,p](t),$$
$$F[n,p+q](t) = F[n,p](t) + F[n,q](t).$$

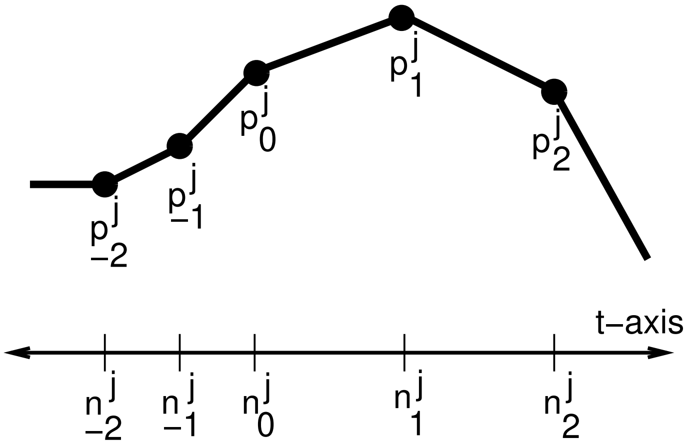

Fig. 2. Parameterization of a subdivision method.

Scaling the knot vector n is also equivalent to scaling the parameter t,

$$F[\alpha n, p](t) = F[n, p](\frac{t}{\alpha}).$$

§2.4 Basis Functions

Let e^i be the vector whose ith entry is one with the remaining entries being zero. Given the knot sequence n^j, we associate the function $F[n^j, e^i](t)$ with each control point p^j_i. By the linearity of F, these functions are basis functions since

$$F[n^j, p^j](t) = \sum_i p^j_i F[n^j, e^i](t).$$

Just as in the case of B-splines, the knots of n^j determines the support of $F[n^j, e^i](t)$. The column height, $a + b + 1$, of $S[n^j]$ determines the width of the support of $F[n^j, e^i](t)$.

Theorem 1. *$F[n^j, e^i](t) = 0$ for $t \notin [n^j_{i-a}, n^j_{i+b}]$.*

Proof: Without loss of generality, we show that $F[n^0, e^0]$ is zero for $t < n^0_{-a}$ and $t > n^0_b$. We keep track of the range of indices of non-zero coefficients during subdivision. After one round of subdivision, the non-zero coefficients range from p^1_{-a} to p^1_b. After k rounds of subdivision, the non-zero coefficients range from $p^k_{-a(2^k-1)}$ to $p^k_{b(2^k-1)}$. The limit of the parameter values associated with this range is

$$\lim_{k\to\infty} n^k_{-a(2^k-1)} = \lim_{k\to\infty} n^k_{-a2^k} = n^0_{-a},$$
$$\lim_{k\to\infty} n^k_{b(2^k-1)} = \lim_{k\to\infty} n^k_{b2^k} = n^0_b.$$

■

§2.5 Example: Interpolating Subdivision

As a running example, we focus on a general interpolatory scheme. Let r be a non-negative integer. Given a knot sequence n, $S[n]$ is defined a row at a time.

$$S[n]_{2i} = e^i,$$
$$S[n]_{2i+1} = m^i$$

where the $m^i_k = 0$ for $k < i - r$ and $k > i + r + 1$. The nonzero entries of m^i are defined by

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ n_{i-r} & n_{i-r+1} & \cdots & n_{i+r+1} \\ \cdots & \cdots & \cdots & \cdots \\ (n_{i-r})^{2r+1} & (n_{i-r+1})^{2r+1} & \cdots & (n_{i+r+1})^{2r+1} \end{pmatrix} \begin{pmatrix} m^i_{i-r} \\ m^i_{i-r+1} \\ \cdots \\ m^i_{i+r+1} \end{pmatrix} = \begin{pmatrix} 1 \\ w \\ \cdots \\ w^{2r+1} \end{pmatrix}$$

where $w = \frac{n_i + n_{i+1}}{2}$. The masks m^i can be thought of in the following way. Consider the degree $2r+1$ polynomial $f(t)$ that interpolates the value f_k at knot n_k for $i-r \leq k \leq i+r+1$. The nonzero entries of m^i are the combinations of the values f_k necessary to reproduce the value $f(\frac{n_i+n_{i+1}}{2})$ independent of the f_k chosen. The masks m^i are chosen to force the subdivision scheme to have polynomial precision of degree $2r+1$.

By the construction of $S[n]$, this subdivision scheme is interpolating. p^{j+1}_{2i} and p^j_i agree and share the same parameter value $n^{j+1}_{2i} = n^j_i$. For the case of regularly spaced n, this subdivision scheme has been heavily analyzed. If $r = 0$, then the non-zero portion of $m^i = (\frac{1}{2}, \frac{1}{2})$. This subdivision scheme is simply piecewise linear interpolation. For $r = 1$, the non-zero portion of the mask m^i is

$$(\frac{-1}{16}, \frac{9}{16}, \frac{9}{16}, \frac{-1}{16}).$$

This subdivision scheme is the four-point rule of [3]. This method converges to a C^1 continuous function.

For irregularly spaced n, this scheme has not been studied. Of course for $r = 0$, the scheme produces piecewise linear interpolation. For $r = 1$, this scheme produces an irregular four point rule that depends on the local knot spacing. Figure 3 show an example of the method. The purpose of this paper is to develop tools for analyzing the smoothness of schemes such as this one.

§2.6 Reduction to the Stationary Case

The key to analyzing the smoothness of subdivision schemes that are compact, affinely invariant, index invariant and locally defined is reducing the scheme locally to an equivalent stationary scheme. Without loss of generality, we assume that $n^0_0 = 0$ and focus our analysis at the origin. Other case can be handled through translation and re-indexing. During subdivision, larger and larger clusters of knots on either side of $n^j_0 = 0$ are regularly spaced. In particular, the knots from $n^j_{-2^j}$ to n^j_0 and from n^j_0 to $n^j_{2^j}$ are regularly spaced.

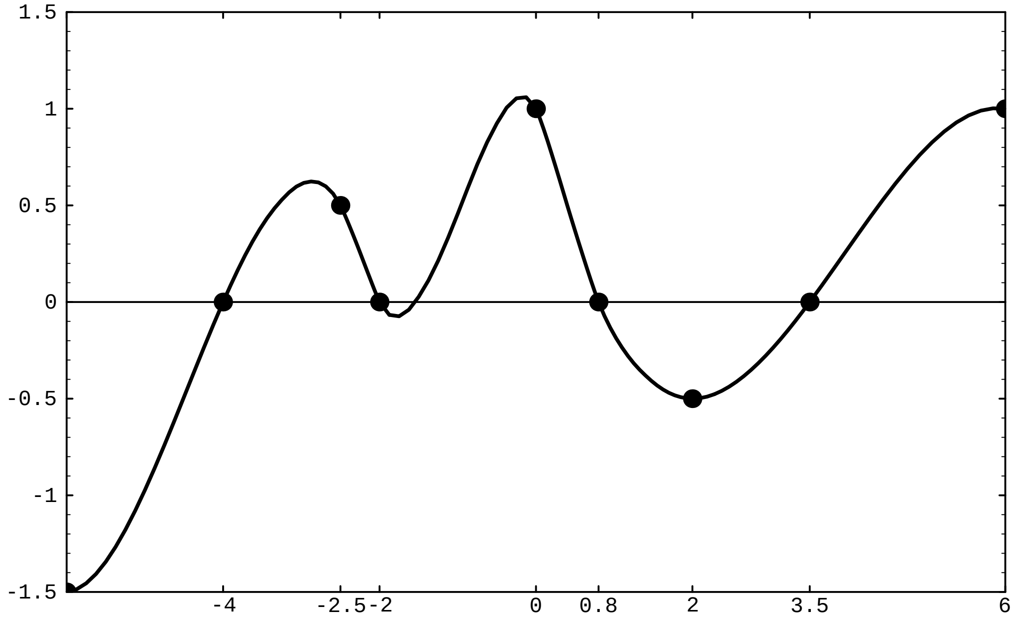

Fig. 3. An example of the irregular four point method .

Theorem 2. *Let $n_0^0 = 0$. There exists an integer j and knot sequence $\hat{n}$ of the form*

$$\begin{aligned} \hat{n}_i &= (n_{-1}^j)i \quad (\forall \ \ i \leq 0), \\ \hat{n}_i &= (n_1^j)i \quad (\forall \ \ i \geq 0), \end{aligned} \tag{2}$$

such that

$$F[n^j, p^j](t) = F[\hat{n}, p^j](t)$$

for all $t \in (n_{-1}^j, n_1^j)$.

Proof: Track those control points produced after k subsequent subdivisions using n^j and $\hat{n}$ as initial knot sequences. By theorem 1, only those control points whose lower indices are in the range $-b+1-2^k$ and 2^k+a-1 affect the limit functions in the parameter range (n_{-1}^j, n_1^j).

We next show that if those control point in this range agree after k subdivision, then the control points with indices in the range $-2^{k+1}-b+1$ and $2^{k+1}+a-1$ agree after $k+1$ subdivisions. By construction, n_i^{j+k} and $\frac{\hat{n}_i}{2^k}$ agree for $-2^{j+k} \leq i \leq 2^{j+k}$. If j is chosen large enough, then the subdivision rules generated by n^{j+k} and $\frac{\hat{n}}{2^k}$ must agree for all control points whose indices are in the range $-2^{k+1}-b+1$ and $2^{k+1}+a-1$ (independent of k). This fact is a consequence of the local definition of the subdivision scheme.

So by induction, the two scheme agree on all control points assigned to the parameter range (n_{-1}^j, n_1^j) for all subdivisions. Therefore, the two associated limit functions agree. ∎

This theorem allows us to focus our analysis on the knot sequence in equation 2. Henceforth, we take the initial knot sequence $n = n^0$ to be of this form. Applying midpoint subdivision to this sequence is equivalent to division by two,

$$n^j = \frac{n}{2^j}.$$

Since our subdivision scheme is affinely invariant,

$$S[n^j] = S[\frac{n}{2^j}] = S[n].$$

Therefore, in a neighborhood of the origin, we may focus on analyzing the stationary subdivision scheme of the form

$$p^{j+1} = S[n]p^j.$$

Note that $S[n]$ depends only on the relative spacing on the knots on either side of the origin. To simplify, we refer to the matrix $S[n]$ as S.

§3. Basic Properties of Stationary Subdivision Schemes

§3.1 Local Analysis

Our approach to analyzing the smoothness of $F[n,p](t)$ is to express this function locally as a linear combination of functions $F[n,x_i](t)$ where the x_i are eigenvectors of S. At first glance, determining these eigenvectors seems daunting because S is an infinite matrix. However, in practice, all of the interesting spectral properties of S are captured by a finite submatrix of S.

If the column height of S is $a+b+1$, then let the *bar* operator, $\bar{p}$, select the entries $p_{-b}, ..., p_a$ from the infinite vector p. These are exactly the entries of p that affect the limit function near the origin. The bar operator applied to the matrix S yields $\bar{S}$, the $a+b+1$ by $a+b+1$ matrix with entries S_{ij} where $-b \geq i, j, \geq a$.

Consider the regular, four point subdivision scheme of section 2.5. Since $r=1$, a and b are 3. So $\bar{S}$ is the $7x7$ matrix

$$\begin{pmatrix} -\frac{1}{16} & \frac{9}{16} & \frac{9}{16} & -\frac{1}{16} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{16} & \frac{9}{16} & \frac{9}{16} & -\frac{1}{16} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{16} & \frac{9}{16} & \frac{9}{16} & -\frac{1}{16} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{16} & \frac{9}{16} & \frac{9}{16} & -\frac{1}{16} \end{pmatrix}. \tag{3}$$

The eigenvalues of this matrix are

$$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{-1}{16}, \frac{-1}{16}.$$

By construction, $\bar{S}$ contains exactly the nonzero portion of the main diagonal of S. Thus, these values are exactly the nonzero eigenvalues of S. These eigenvectors of $\bar{S}$ can be extended to form eigenvectors of S.

Theorem 3. *Let $\lambda \neq 0$ be an eigenvalue of $\bar{S}$ with associated eigenvector y. Then S has eigenvalue λ with a unique associated eigenvector x such that $\bar{x} = y$.*

The proof of this theorem is simple and left to the reader. The eigenvectors of $\bar{S}$ with eigenvalue zero have no effect on the final limit curve at the origin since after one round of subdivision the control points are mapped to zero. For the sake of simplicity, we assume that $\bar{S}$ has no zero eigenvalues.

If $\bar{S}$ does not have a full set of eigenvectors, then $\bar{S}$ is *defective.* A non-defective $\bar{S}$ has a full set of linearly independent eigenvectors $\bar{x}_0$, ..., $\bar{x}_{a+b}$. The extension of these eigenvectors, x_0, ..., x_{a+b}, can be used in the following manner.

Theorem 4. *Let $\bar{S}$ be a non-defective matrix with eigenvectors $\bar{x}_0$, ..., $\bar{x}_{a+b}$. If*

$$\bar{p} = \sum_{i=0}^{a+b} c_i \bar{x}_i,$$

then, for all $t \in (t_{-1}^0, t_1^0)$,

$$F[n,p](t) = \sum_{i=0}^{a+b} c_i F[n, x_i](t).$$

Proof: The vector $p - \sum_{i=0}^{a+b} c_i x_i$ is zero for entries $-b$ to a. By theorem 1,

$$F[n, p - \sum_{i=0}^{a+b} c_i x_i](t) = 0$$

for all $t \in (n_{-1}^0, n_1^0)$. The theorem follows by the linearity of the limit operator F. ■

For defective $\bar{S}$, *generalized* eigenvectors can be used in place of eigenvectors. Each of the generalized eigenvectors, $\bar{x}_0$, ..., $\bar{x}_{a+b}$, is either an eigenvector of $\bar{S}$ or satisfies

$$\bar{S}\bar{x}_i = \lambda_i \bar{x}_i + \bar{x}_{i-1}.$$

These generalized eigenvectors can be extend to a set of infinite vectors, x_0, ..., x_{a+b}, satisfying

$$S x_i = \lambda_i x_i + x_{i-1}. \tag{4}$$

The proof of this fact is exactly the same as the proof of theorem 3. Since the generalized eigenvectors are linearly independent, theorem 4 also holds for these vectors. For subsequent theorems, we assume that $\bar{S}$ is non-defective. Where appropriate, we state the variant of the theorem that holds for defective $\bar{S}$ using generalized eigenvectors.

§3.2 A Spectral Recurrence

The main message of this paper is that the smoothness properties of a stationary subdivision scheme are tied to the spectral properties of its subdivision matrix S. In particular, the limit function associated with an eigenvector of a stationary subdivision scheme satisfies a fundamental relation.

Theorem 5. *Let x be an eigenvector of S satisfying $Sx = \lambda x$. Then,*

$$\lambda F[n, x](t) = F[n, x](\frac{t}{2}). \tag{5}$$

Proof: The proof consist of simply recalling the definition of F.

$$\begin{aligned} \lambda F[n, x](t) &= F[n, \lambda x](t), \\ &= F[n, Sx](t), \\ &= \lim_{j \to \infty} L[n^j, S^j(Sx)](t), \\ &= \lim_{j \to \infty} L[2n^{j+1}, S^{j+1}x](t), \\ &= \lim_{j \to \infty} L[n^{j+1}, S^{j+1}x](\frac{t}{2}), \\ &= F[n, x](\frac{t}{2}). \end{aligned}$$

■

If x_i is a generalized eigenvector as in equation 4, then

$$\lambda_i F[n, x_i](t) + F[n, x_{i-1}](t) = F[n, x_i](\frac{t}{2}). \tag{6}$$

Again, the proof is exactly as above.

§3.3 Properties of the Recurrence

The recurrence of theorem 5 is a powerful tool for analyzing stationary subdivision schemes. The following lemma illustrates several properties of such recurrences.

Lemma 1. *Let $g(t)$ be a function satisfying*

$$\lambda g(t) = g(\frac{t}{2}). \tag{7}$$

(1) *If $g(t)$ is bounded on the interval $(\frac{1}{2}, 1]$ and $|\lambda| < \frac{1}{2^k}$, then the limit as $t \to 0$ of $\frac{g(t)}{t^k}$ is zero.*
(2) *If $\lambda = 1$ and $g(t)$ is continuous at the origin, then $g(t)$ is the constant function.*

(3) *If $g(d) \neq 0$ for $d \neq 0$ and $|\lambda| > 1$ then the limit as $t \to 0$ of $g(t)$ diverges.*

Proof: For part one, let

$$v = \max_{\frac{1}{2} < t \leq 1} |g(t)|.$$

If $\frac{1}{2^{i+1}} < t \leq \frac{1}{2^i}$, then iterating equation 7 yields

$$\begin{aligned} |g(t)| &= |\lambda^i g(2^i t)|, \\ &\leq |\lambda|^i v. \end{aligned}$$

Dividing the lefthand side of this relation by t^k and the righthand side by $\frac{1}{2^{k(i+1)}}$ respectively yields

$$\begin{aligned} |\frac{g(t)}{t^k}| &\leq 2^{k(i+1)} |\lambda|^i v, \\ &\leq (2^k |\lambda|)^i 2^k v. \end{aligned}$$

Since $|\lambda| < \frac{1}{2^k}$, the limit as $i \to \infty$ and therefore as $t \to 0$ of $\frac{g(t)}{t^k}$ must be zero.

For part two, we observe that if there exists $d \neq e$ such that

$$g(d) - g(e) = c \neq 0,$$

then by equation 7

$$g(\frac{d}{2^i}) - g(\frac{e}{2^i}) = c \neq 0. \tag{8}$$

As i goes to infinity, $\frac{d}{2^i}$ and $\frac{e}{2^i}$ approach zero. Equation 8 contradicts that fact that $g(t)$ is continuous at the origin. Therefore, $g(d) = g(e)$ for all d and e.

For part three, we note that by hypothesis there exists $d \neq 0$ such that $g(d) \neq 0$. Since $g(\frac{d}{2^i}) = \lambda^i g(d)$ and $|\lambda| > 1$, $g(t)$ must diverge as $i \to \infty$ and $t = \frac{d}{2^i} \to 0$. ■

§4. Necessary Conditions for C^k Continuity

By theorem 4, we can restrict our smoothness analysis to those limit functions $F[n, x](t)$ where x is an eigenvector of S with eigenvalue λ. If $F[n, x](t)$ is a C^k continuous function, then x and λ must have special properties. Let $F^{(i)}[n, x](t)$ denote the ith derivative of $F[n, x](t)$.

Theorem 6. *Let $Sx = \lambda x$ with $|\lambda| \geq \frac{1}{2^k}$. If $F[n, x](t)$ is C^k continuous and not identically zero, then there exist $0 \leq i \leq k$ such that*

- $\lambda = \frac{1}{2^i}$,

- $F[n,x](t) = c_i t^i$ *for* $c_i \neq 0$.

Proof: We first show that $F^{(k)}[n,x](t)$ is a constant function. Take the kth derivative of equation 5.

$$(2^k\lambda)F^{(k)}[n,x](t) = F^{(k)}[n,x](\frac{t}{2}).$$

If $F^{(k)}[n,x](t)$ is not a constant function, then $F^{(k)}[n,x](t)$ must diverge as $t \to 0$ by either part two of lemma 1 ($2^k\lambda = 1$) or part three of lemma 1 ($|2^k\lambda| > 1$). However, this contradict the continuity of $F^{(k)}[n,x](t)$.

Since $F^{(k)}[n,x](t)$ is a constant function, $F[n,x](t)$ is a polynomial function of degree k. Since $F[n,x](t)$ is not identically zero, choose the minimal i such that $F^{(i)}[n,x](0) \neq 0$ and take the ith derivative of equation 5.

$$\lambda F^{(i)}[n,x](t) = \frac{1}{2^i}F^{(i)}[n,x](\frac{t}{2}).$$

For lefthand and righthand side of this equation to agree at $t = 0$, λ must be $\frac{1}{2^i}$. Since $F^{(i)}[n,x](t)$ is continuous, by part two of lemma 1, $F^{(i)}[n,x](t)$ must be the constant function. Since the lower order derivatives are zero at the origin, $F[n,x](t)$ is multiple of t^i. ■

If a generalized eigenvector (equation 4) produces a C^k limit curve, then its eigenvalue must have modulus less than $\frac{1}{2^k}$. Consider the continuous (C^0) case. If $\lambda_0 = 1$, then $F[n,x_0](t)$ is the constant function. If x_1 were a generalized eigenvector satisfying

$$Sx_1 = \lambda_1 x_1 + x_0,$$

then $\lambda_1 = \lambda_0 = 1$. In terms of limit functions,

$$\begin{aligned} \lambda_1 F[n,x_1](t) + F[n,x_0](t) &= F[n,x_1](\frac{t}{2}), \\ F[n,x_1](t) + 1 &= F[n,x_1](\frac{t}{2}). \end{aligned}$$

As $t \to 0$, $F[n,x_1](t)$ must diverge.

The four point rule is known to yield C^1 continuous functions. This theorem states that the only possible eigenvalues with modulus greater or equal to $\frac{1}{2}$ are 1 and $\frac{1}{2}$. Also the eigenvectors x_0 and x_1 must produce the limit functions 1 and t up to a constant multiple. Using matrix 3, one can verify that the kth entries of x_0 and x_1 are 1 and k respectively. The eigenvector $\frac{1}{4}$ has multiplicity two, but has only a single associated eigenvector. Therefore, this scheme is strictly a C^1 continuous.

§5. Sufficient Conditions for C^k Continuity

For regularly spaced knot sequences, the subdivision schemes described here generate subdivision matrices whose columns are just shifts of each other. Powerful tools such as Laurent polynomials are available for analyzing such schemes. For knot sequence whose knots are regularly-spaced away from the origin, all but a finite number of columns of the subdivision matrix must agree with the regular case (due to the local definition of the subdivision rules). If the regular case produces C^k continuous curves, then this irregular case also produces a C^k curve except possibly at $t = 0$.

By theorem 4, we need only analyze the smoothness of the functions $F[n,x](t)$ where x is an eigenvalue of S. By theorem 6, an eigenvector x with eigenvalue $|\lambda| \geq \frac{1}{2^k}$ must reproduce a polynomial if the scheme is to be C^k continuous. This reproduction is usually ensured during construction of S (e.g. interpolatory subdivision). Only if $|\lambda| < \frac{1}{2^k}$ does the smoothness of $F[n,x](t)$ need to be verified.

Theorem 7. *Let $Sx = \lambda x$ with $|\lambda| < \frac{1}{2^k}$. If $F[n,x](t)$ is C^k continuous everywhere except at $t = 0$, then $F[n,x](t)$ is C^k continuous everywhere.*

Proof: We show that $F^{(i)}[n,x](t)$ exists and is continuous at $t = 0$ for $0 \leq i \leq k$. The proof proceeds by induction on i. For the base case $i = 0$, we note that $F[n,x](0)$ must be zero by theorem 5 since $|\lambda| < 1$. The limit as t goes to zero of $F[n,x](t)$ is also zero by part one of lemma 1. Therefore, $F[n,x](t)$ is continuous at $t = 0$.

We proceed with the induction step. First, we show that $F^{(i)}[n,x](0)$ exists and is zero at $t = 0$. By definition,

$$\begin{aligned} F^{(i)}[n,x](0) &= \lim_{t \to 0} \frac{F^{(i-1)}[n,x](t) - F^{(i-1)}[n,x](0)}{t}, \\ &= \lim_{t \to 0} \frac{F^{(i-1)}[n,x](t)}{t}, \end{aligned}$$

since $F^{(i-1)}[n,x](0)$ is zero by the inductive hypothesis. Taking the $i-1$st derivative of equation 5 yields

$$\lambda 2^{i-1} F^{(i-1)}[n,x](t) = F^{(i-1)}[n,x](\frac{t}{2}).$$

The $(i-1)$st derivatives exists by our inductive hypothesis. Now, if $|\lambda| < \frac{1}{2^k}$ and $i \leq k$, then $|\lambda| 2^{i-1} < \frac{1}{2}$. Therefore, by part one of lemma 1, the limit as t goes to zero of $\frac{F^{(i-1)}[n,x](t)}{t}$ must also be zero.

Given that the ith derivative of $F[n,x](t)$ exists at $t = 0$, we can take the ith derivative of equation 5. Applying part one of lemma 1 shows that $F^{(i)}[n,x](t)$ is continuous at $t = 0$. This step completes the induction and the proof. ∎

This theorem also holds for generalized eigenvectors (equation 4) whose eigenvalues have moduli less than $\frac{1}{2^k}$. The proof involves modifying part one of lemma 1 to use equation 4.

In theorem 6, eigenvectors produced the monomials as limit functions. In thoerem 7, eigenvectors produce limit functions whose first k derivatives are zero at the origin. Together, these two theorems yield necessary conditions for a subdivision scheme to produce C^k continuous limit functions. If S necessarily produces C^k limit curves and there exists p such that the ith derivative of $F[n,p](0) \neq 0$ for all $0 \leq i \leq k$, then there must exist eigenvectors of S that reproduce each monomial t^i up to degree k. So, the subdivision scheme defined by S can reproduce any polynomial up to degree k. The spectrum of S includes 1, $\frac{1}{2}$, ..., $\frac{1}{2^k}$. If the basis functions associated with the scheme are linearly independent, then these eigenvalues must also be unique.

Under these assumptions, we can summarize the necessary conditions for a C^k continuous scheme. Indexing the eigenvalues of S in order of decreasing moduli, $|\lambda_i| \geq |\lambda_{i+1}|$,

(1) $F[n,x_i](t) = c_i t^i$ for all $0 \leq i \leq k$,
(2) $\lambda_i = \frac{1}{2^i}$ for all $0 \leq i \leq k$,
(3) $|\lambda_i| < \frac{1}{2^k}$ for all $i > k$.

If a full span of derivatives at the origin does not exist, then the spectrum may be missing certain powers of two. If the scheme has linearly dependent basis functions, several powers of two may be repeated. This result is a generalization of a similar result for uniform subdivision in [1].

Since the regular four point rule produces C^1 functions, we can now use theorem 7 to show that the irregular four point rule also produces C^1 functions. For $c > 0$, consider the knot sequence n

$$..., -3, -2, -1, 0, c, 2c, 3c, ...$$

where $n_i = i$ is $i \leq 0$ and $n_i = ci$ if $i \geq 0$. Then the matrix $\bar{S}$ is

$$\begin{pmatrix} -\frac{1}{16} & \frac{9}{16} & \frac{9}{16} & -\frac{1}{16} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{-(1+2\,c)}{16\,(2+c)} & \frac{3\,(1+2\,c)}{8\,(1+c)} & \frac{3}{8} + \frac{3}{16\,c} & \frac{-3}{8\,c\,(2+3\,c+c^2)} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{-3\,c^3}{8\,(1+3\,c+2\,c^2)} & \frac{3\,(2+c)}{16} & \frac{3\,(2+c)}{8\,(1+c)} & \frac{-(2+c)}{16\,(1+2\,c)} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{16} & \frac{9}{16} & \frac{9}{16} & -\frac{1}{16} \end{pmatrix}.$$

The matrix, parameterized by c, has eigenvalues that are independent of c.

$$1, \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{-1}{16}, \frac{-1}{16}.$$

The leading eigenvalues are 1 and $\frac{1}{2}$ while the remaining eigenvalues have moduli less that $\frac{1}{2}$. Since two leading eigenvectors reproduce 1 and t (by construction), this scheme produces C^1 limit curves.

For interpolatory schemes with higher order polynomial precision, this analysis becomes more difficult. For quintic precision ($r = 2$), Mathematica can be used to derive the characteristic polynomial of $\bar{S}$. This polynomial and its roots vary as a function of c. By representing the polynomial in various bases, one can show that for $c > 0$, the spectrum of $\bar{S}$ satisfies the C^2 conditions. Unfortunately, this type of analysis fails for $r > 2$. We conjecture that such interpolatory schemes with order $2r + 1$ polynomial precision are C^r continuous.

§6. Convergence to a Continuous Scheme

Theorem 7 relies on knowing the subdivision scheme is C^k continuous away from the origin. In this section and the next, we give direct conditions on S for the scheme to be C^k continuous everywhere. This technique is a direct generalization of the method of [4]. We first establish conditions for convergence to a continuous function.

For function $g(t)$, infinite vector p, and bi-infinite matrix S, define the norm

$$||f(t)|| = \max_t |f(t)|,$$

$$||p|| = \max_i |p_i|,$$

$$||S|| = \max_i (\sum_k |S_{ik}|),$$

where S_i is the ith row of S.

Point-wise convergence is not always sufficient to establish differentiability of the limit functions. A stronger type of convergence is often required. A sequence of functions $g_i(t)$ *uniformly* converges to a limit function $g(t)$ if for all $\epsilon > 0$ there exists n such that for all $i \geq n$

$$||g_i(t) - g(t)|| < \epsilon.$$

We now proceed with the convergence analysis. Let Δ_u denote the bi-infinite matrix whose (i, i)th entry is -1 and whose $(i, i+1)$th entry is 1 for all i. The remaining entries of Δ_u are zero. A finite portion of Δ_u is

$$\begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

Applying Δ_u to a vector p returns a vector whose ith entry is $p_{i+1} - p_i$. As we shall see, proving uniform convergence of the sequence of linear functions $L[n^j, S^j p](t)$ is equivalent to showing that the sequence $\Delta_u (S)^j p$ uniformly converges to zero.

To analyze this sequence, we build a matrix D_u satisfying

$$D_u \Delta_u = \Delta_u S.$$

If Δ_u^* is the matrix whose ijth entry is one if $i > j$ and zero otherwise, then Δ_u^* is both a left and right inverse of Δ_u. So, D_u can be written as

$$D_u = \Delta_u S \Delta_u^*.$$

If S has column height $a + b + 1$, then $\Delta_u S$ has column height $a + b + 2$. If S yields a continuous scheme, then by theorem 5 and lemma 1, S had an eigenvector $\underline{1} = (\ldots, 1, 1, 1, \ldots)$. Since

$$(\Delta_u S)\underline{1} = \Delta_u \underline{1} = \underline{0},$$

the rows sums of $\Delta_u S$ must be zero. Therefore, post-multiplying by Δ_u^* trims one element off the non-zero portion of each row of $\Delta_u S$. The resulting matrix D_u has column height $a + b$. Therefore, D_u defines a compactly supported subdivision scheme.

By construction, D_u satisfies

$$(D_u)^j \Delta_u = \Delta_u (S)^j.$$

The next theorem follows from this observation.

Theorem 8. *Let $||p||$ be bounded. The following three properties are equivalent.*

- *The sequence $L[n^j, S^j p](t)$ uniformly converges to a continuous function as $j \to \infty$.*
- *The sequence $L[n^j, \Delta_u(S^j p)](t)$ uniformly converges to zero as $j \to \infty$.*
- *There exists $k \geq 1$ such that $||(D_u)^k|| < 1$.*

Proof: See theorem 3.1 in [4]. ■

§7. Convergence to Smooth Schemes

§7.1 Linear Parameterizations

We previously showed that any stationary subdivision scheme that is at least C^1 continuous and that can produce functions with non-zero first derivatives at the origin must have linear precision. In particular, the eigenvector x_1 associated with eigenvalue $\frac{1}{2}$ must produce a constant multiple of the function t. Let l denote the multiple of x_1 such that

$$F[n, l](t) = t.$$

To simplify the following analysis, we restrict our attention to a fairly specific class of subdivision schemes, those in which the function t has a locally unique representation l in terms of coefficients that form a monotonically increasing sequence. For example, this restriction rules out schemes that produce linearly dependent basis functions. (However, a similar, but more complex theory is possible if t has multiple representations.)

Our functional scheme can also be expressed as the parametric curve

$$(t, F[n,p](t)) = (F[n,l](t), F[n,p](t)).$$

Applying the subdivision matrix S repeatedly to the initial parametric control polygon $(l,p)^T$ produces the appropriate functional limit curve. This observation suggests that the spacing given by $S^j l$ might be a more appropriate *linear* parameterization for the piecewise linear function defined by p^j. After j subdivisions, this linear parameterization satisfies

$$S^j l = \frac{l}{2^j} = l^j.$$

Note that the knot sequence n^j is still used to define the subdivision matrices $S[n^j]$. Only the parameter values associated with the p^j are changed.

The following example highlights the main reason for using this parameterization. Consider cubic B-splines over the initial knot sequence n

$$..., -9, -6, -3, 0, 6, 12, 18, ...$$

Applying midpoint subdivision to this knot sequence leads to a stationary subdivision scheme with matrix $\bar{S}$

$$\begin{pmatrix} \frac{1}{8} & \frac{25}{32} & \frac{3}{32} & 0 & 0 \\ 0 & \frac{5}{8} & \frac{3}{8} & 0 & 0 \\ 0 & \frac{5}{24} & \frac{29}{40} & \frac{1}{15} & 0 \\ 0 & 0 & \frac{3}{5} & \frac{2}{5} & 0 \\ 0 & 0 & \frac{3}{20} & \frac{29}{40} & \frac{1}{8} \end{pmatrix}.$$

The derivative of $F[n,p](t)$ is a C^2 function. The derivatives of $L[n^j, S^j p](t)$ (defined using the knot parameterization) do not converge to a C^1 function. The eigenvector l for this scheme is

$$..., -9, -6, -3, 1, 6, 12, 18, ...$$

The derivatives of $L[l^j, S^j p](t)$ (defined using the linear parameterization) do converge to a C^1 function. In fact this function is the derivative of $F[l,p](t)$. In summary, only the linear parameterization ensures that the limit of the derivatives of control polygons is the derivative of the final limit curve. Section 7.3 proves this result.

Parameterizing by either n or l produce the same results as long as the convergence is uniform.

Theorem 9. *Let $||p||$ be bounded. If $L[n^j, S^j p](t)$ uniformly converges to a continuous function as $j \to \infty$, then*

$$F[n,p](t) = F[l,p](t).$$

Proof: We first show that the difference between n_i and l_i is bounded for all i. If n is regularly spaced, then $S[n]$ must be uniform due to affine and index invariance. Since the subdivision rules are locally defined and n has a single non-uniformity, only a finite number of columns of $S[n]$ can deviate from the uniform case. For the uniform case, [1] and [6] show that the polynomial t can be reproduced by a vector that is a sample of a linear function over a regular knot sequence. Since this representation is unique (as assumed at the start of the section), l_i must agree with the uniform analysis for large $|i|$. Therefore, there must exist k such $n_i = l_i + c_0$ for all $i \le -k$ and $n_i = l_i + c_1$ for all $i \ge k$. So, the difference $||n - l||$ must be bounded.

The ith entry of p is plotted at either n_i or l_i depending on the parameterization chosen. The maximum deviation between the two functions $L[n,p](t)$ and $L[l,p](t)$ at $t = l_i$ is bounded by the number of knots in n between n_i and l_i and the maximum change in $L[n,p](t)$ between a pair of knots. If d is the maximum number of knots between l_i and n_i over all i (d exists since $||n - l||$ is bounded), then this difference is bounded by $d||\Delta_u p||$. Repeating this argument after j subdivisions,

$$\begin{aligned} ||L[n^j, S^j p](t) - L[l^j, S^j p](t)|| &\le d||\Delta_u S^j p||, \\ &\le d|(D_u)^j \Delta_u p||, \\ &\le d||(D_u)^j|| \cdot ||\Delta_u p||. \end{aligned}$$

Since $||p||$ is bounded, $||\Delta_u p||$ must be bounded. By theorem 8, $||(D_u)^j||$ converges to zero as $j \to \infty$. Therefore, the difference between these two function converges to zero and the two functions uniformly converge to the same limit. ■

§7.2 Non-uniform Differencing Operator

Our ultimate goal is to build a subdivision matrix D whose limit functions are the derivatives of limit functions produced by S. To this end, we use an approach similar to that of Gregory and Qu in [7,8]. The key is to construct a differencing operator Δ that behaves like a discrete derivative. Δ maps a set of control points for S to a new set of control points for D. S, D, and Δ satisfy a commutative relation. Differencing a set of subdivided control points, Sp, should produce the same results as subdividing (using D) a set of differenced control points, Δp. In terms of Δ, S, and D,

$$D\Delta = (2\Delta)S. \tag{9}$$

The cause of the extra factor of 2 will become clear momentarily.

If Δ is to act as a derivative, Δ should annihilate the vector $\underline{1}$ since the derivative of the function one is zero. Applied before any subdivision, Δ should map l to the vector $\underline{1}$ since the derivative of t is one. Thus, we want

$$\Delta\underline{1} = \underline{0},$$
$$\Delta l = \underline{1}.$$

If we let Γ be the diagonal matrix whose (i,i)th entry is $1/(l_{i+1} - l_i)$, then Δ can be expressed as

$$\Delta = \Gamma\Delta_u.$$

Δ_u annihilates the vector $\underline{1}$ and Γ scales $\Delta_u l$ to yield $\underline{1}$.

Note that Δ behaves correctly only if appled before subdivision. After j rounds of subdivision, the function t is represented by the control points $\frac{l}{2^j}$. Therefore, the correct differencing operator in this case is $2^j\Delta$. This observation explains the factor of two in equation 9.

Given Δ, we can explicitly construct D. Let

$$\Delta^* = \Delta_u^*\Gamma^{-1}.$$

Δ^* is a left and right inverse of Δ. Now, the matrix D satisfies

$$D = 2\Delta S\Delta^*.$$

If S has column height $a+b+1$, then D has column height $a+b$ as argued in the previous section.

§7.3 Derivative Schemes

D defines a stationary subdivision scheme with associated limit functions

$$\lim_{j\to\infty} L[l^j, D^j q](t).$$

It remains to show that these functions are derivatives of those produced by S.

Theorem 10. *Let the sequence $L[l^j, D^j q](t)$ uniformly converge to a continuous function for all bounded $||q||$. Then, for all bounded $||p||$,*

$$\lim_{j\to\infty} L[l^j, D^j(\Delta p)](t)$$

is the derivative of $F[l,p](t)$ with respect to t.

Proof: Recall that by definition $F[l,p](t) = g(t)$ is the limit of the functions

$$g_j(t) = L[l^j, S^j p](t).$$

The derivatives, $g_j'(t)$, of the $g_j(t)$ are piecewise constant functions over the knot sequence l^j with piecewise values $(2^j\Delta)S^jp$. By the construction of D,

$$(2^j\Delta)S^j = D^j\Delta.$$

So, the $g_j'(t)$ have piecewise values $D^j(\Delta p)$.

By hypothesis, the sequence

$$h_j(t) = L[l^j, D^j(\Delta p)](t)$$

uniformly converges to a continuous function, call it $h(t)$. By theorem 8, the difference between the function $g_j'(t)$ and $h_j(t)$ must uniformly converge to zero. Therefore, the sequence $g_j'(t)$ uniformly converges to $h(t)$. Figure 4 illustrates this process.

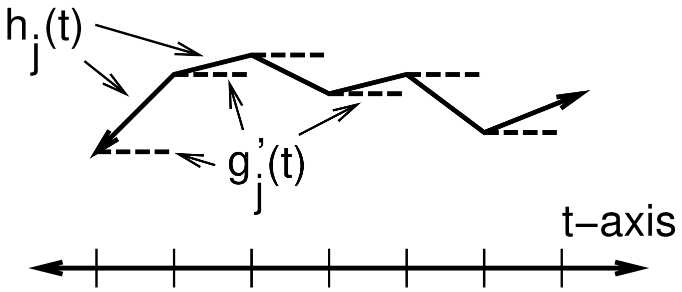

Fig. 4. Convergence of $g_j'(t)$ and $h_j(t)$.

We next prove that $\int_0^t g_j'(s)ds$ is point-wise convergent to $\int_0^t h(s)ds$ for any t. Fix t. By the uniform convergence of the $g_j'(t)$ to $h(t)$, for all $\epsilon > 0$, there exist an n such that for all $j \geq n$

$$|g_j'(t) - h(t)| < \epsilon.$$

In terms of the integrals, for all $j \geq n$,

$$\begin{aligned} |\int_0^t g_j'(s)ds - \int_0^t h(s)ds| &\leq \int_0^t |g_j'(s) - h(s)|ds, \\ &\leq \epsilon t. \end{aligned}$$

Therefore, $g_j(t) - g_j(0)$ converges to $\int_0^t h(s)ds$ for any t. However, $g_j(t) - g_j(0)$ also converges to $g(t) - g(0)$. So,

$$g(t) = \int_0^t h(s)ds + g(0)$$

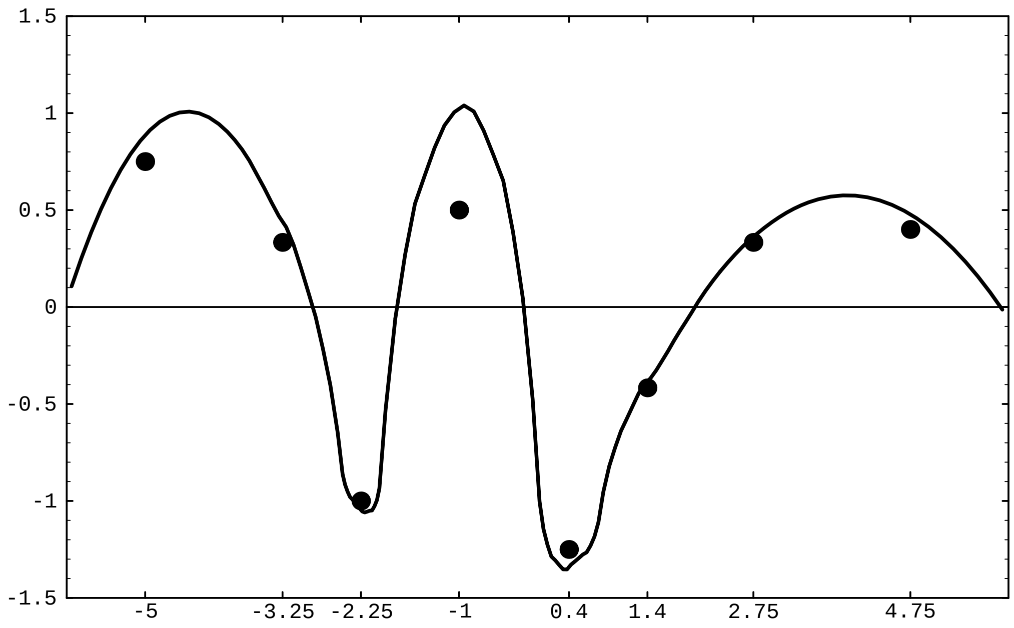

Fig. 5. The derivative of an irregular scheme.

for any t. Thus, $h(t)$ is the derivative of $g(t)$. ∎

This theorem is a slight variant of theorem IV on page 599 of [11]. Figure 5 shows the derivative curve associated with the irregular four point curve of figure 3. The control points for this curve were computed from the original control points by applying Δ.

The previous analysis can be repeated with D in place of S. Since D is a stationary subdivision scheme, theorem 8 can determine whether the new scheme is uniformly convergent to a continuous function. If D satisfies the necessary conditions for a C^1 scheme, then a new linear parameterization for its limit curves can be computed. If $F[n, x_2](t) = \frac{1}{2}t^2$ for S, then Δx_2 is the correct linear parameterization for D. A new non-uniform difference operator can be defined with this parameterization. Finally, theorem 10 can be applied to build a derivative scheme for D.

§8. Parametric Analysis for the Stationary Case

Given a control polyon $(p_1, p_2)^T$, we can apply the stationary subdivision matrix S to each component separately and build a sequence of polygons $(S^j p_1, S^j p_2)^T$. When does a limit polygon exist and when is it smooth?

The spectral properties of S can be used to give conditions for smoothness. Given subdivision matrix S, we can compute in descending order its eigenvalues $1, \gamma, ...$ and their associated eigenvectors $x_0, x_1, ...,$. If $0 < \gamma < 1$, then the eigenvector $l = x_1$ provides the natural parameterization for this subdivision scheme. Note that γ need not be $\frac{1}{2}$. $\frac{1}{2}$ arose in the preceding functional analysis due to our use of binary subdivision. The only restriction placed on l is that

$$l_{i+1} - l_i \geq \epsilon \geq 0$$

for all i. This restriction guarantees that the parameterization is $1-1$ and fills the parameter line. (Here the ordering of the control points determines the ordering of the rows and columns of S and thus, the ordering of the entries of l.)

Given a parameterization for our control polygon, we now can apply the functional analysis of the previous section to each component of our parametric curve (with $\frac{1}{2}$ replaced by γ). If each component is a C^k function, then the parametric curve

$$(F[l, p_1](t), F[l, p_2](t))$$

is a C^k curve for those t such that the derivative of $F[l, p_1](t)$ or $F[l, p_2](t)$ is non-zero.

§9. Conclusions

The true purpose of this paper is to lay the groundwork for a theory of functional subdivision over irregular triangulations. The author believes that such a theory is necessary to fully understand the problem of building parametric subdivision surfaces of arbitrary topological type. Subsequent papers will address this topic.

For a subdivision matrix S to produce a C^k continuous scheme, S must have $k+1$ eigenvectors that reproduce 1, t, ..., t^k. Instead of specifying a subdivision scheme as a map from knots to matrices, another approach would be to specify a subdivision scheme as a map from knot sequences to eigenvectors. In the case of B-splines and interpolatory subdivision, this map is much simpler than the map from knots to the subdivision matrix. The subdivision matrix S could then be derived from the eigenvectors using linear algebra (while attempting to minimize the support of the subdivision masks).

Finally, better tools for analyzing the eigenstructure of parametrized matrices are needed. In the irregular case, the subdivision matrix $\bar{S}$ is a function of the underlying knots. Tools for analyzing the spectral properties of such matrices would be very useful.

References

1. Cavaretta, A., W. Dahmen, and C. Micchelli, Stationary Subdivision, volume 453 of *Memoirs of the AMS*. AMS, 1991.
2. Chaikin, G., An algorithm for high speed curve generation, Computer Graphics and Image Processing, **3**, (1974), 346–349.
3. Dyn, N., J. Gregory, and D. Levin, A 4-point interpolatory subdivision scheme for curve design, Computer Aided Geometric Design, **4**, (1987), 257–268.

4. Dyn, N., J. Gregory, and D. Levin, Analysis of uniform binary subdivision schemes for curve design, Constructive Approximation, **7**, (1991), 127–147.
5. Dyn, N., Subdivision schemes in computer aided geometric design, in *Advances in Numerical Analysis II*, W. Light (ed.), Oxford University Press, 1992, 36–104.
6. Dyn, N., J. Gregory, and D. Levin, Piecewise uniform subdivision schemes, in preparation.
7. Gregory, J.A. and Qu, R.B, A subdivision algorithm for non-uniform B-splines, in *Approximation Theory, Spline Functions and Applications*, S. Singh (ed.), NATO ASI Series C, 356, 1992, 423-436.
8. Gregory, J.A. and Qu, R.B, Non-uniform corner cutting, to appear in Computer Aided Geometric Design.
9. Michelli, C., and H. Prautzsch, Uniform refinement of curves, Linear Algebra and Applications, **114**, (1989), 841–870.
10. Riesenfeld, R., On Chaikin's Algorithm, Computer Graphics and Image Processing, **4**, (1975), 304–310.
11. Taylor, A., *Advanced Calculus*, Blaisdell Publishing, 1955.

Acknowledgements. The author would like to thank Nira Dyn, David Levin, Ron Goldman, Tom Lyche, Tony DeRose, and Tom Duchamp for their discussion of the various issues raised in this paper.

Joe Warren
Department of Computer Science
Rice University
Houston, TX 77251-1892 USA