Fast Approximation Methods for Fitting Surfaces to Unorganized Point Clouds

Karl-Heinz Brakhage
Fast Approximation Methods
for Fitting Surfaces
to Unorganized Point Clouds

Karl-Heinz Brakhage
Institut für Geometrie und Praktische Mathematik, RWTH Aachen,
Templergraben 55, 52056 Aachen, Germany
brakhage@igpm.rwth-aachen.de

Abstract
We present and analyze a novel fast method for scattered data approximation with
curves / surfaces which have a representation as a linear combination of smooth basis
functions associated with the control points. Our technique can be applied to stan-
dard Bézier and B-spline curves / surfaces as well as for subdivision schemes. The
approach can be formulated in such way that for the iteration we have a standard least
squares problem in each step. A regularization term that expresses the fairness of the
intermediate and / or final result can be added. Adaptivity is easily integrated in our
concept. Furthermore our approach is well suited for reparameterization occurring in
grid generation.

Keywords: Splines, Multivariate Approximation, Fairing, Numerical Analysis, Numerical
Linear Algebra

Introduction
The problem of computing a smooth curve or surface representation of a target shape
given by a set of unorganized data points has many applications in engineering and CAD.
For a long period surface fitting methods use the distance between a point on the fitting
surface and the corresponding foot point on the target surface or vice versa for minimizing
the objective function. These methods are called point distance minimization (PDM).
The parameters for the projected points have to be adjusted in an outer loop. Later
Pottmann et al. ([6, 7]) introduced an approach based on the minimization of a quadratic
approximant of the squared distance function. Their aim was to avoid the parameterization
problem and to construct algorithms of second order convergence. Methods based on
this idea are called SDM, standing for squared distance minimization. Unfortunately in
general the second order Taylor approximant does not lead to symmetric positive definite
system matrices and for this reason the existence and uniqueness of the minimum cannot
be guaranteed. Thus the second order Taylor approximation was modified to ensure
positive definiteness. But this modification destroys the second order and thus the claimed
quadratic convergence of those methods. Nevertheless such approaches need less iterations.
On the other side the main drawback is a large computational overhead (see [3]), e.g. due to
principal curvature computations. Furthermore the parameterization is not really totally
avoided.
We use a different approach. This sidesteps the curvature computation. Furthermore our method can still be written in the form of a standard least squares problem which is not the case for SDM. Thus we can use the normal equation and iterative solvers for them as well as orthogonal transformations that have a better condition number than the normal equations. To reduce the number of parameter corrections (the outer loop) we minimize a combination of the PDM and the distances of the data points to the linear approximation of the target surface at the projection point. Note that the later one coincides with the squared distance function for points on the surface. We will demonstrate the idea of our method for the approximation with Beziers. It turned out that our method has super linear convergence with only a small computational overhead for surface normals.

The rest of this paper is organized as follows. First we give some basic notations and properties of Bézier and B-spline surfaces. Next we analyze the behavior of the standard PDM iteration and that of our new one. Details on the used approximation strategies and algorithms for splines will be given. Finally a summary and an evaluation regarding computation time and accuracy of the algorithms usually applied for these purposes will be given.

**Basics on Splines and Subdivision**

Bézier curves of order \( n \) are given by \( n + 1 \) control or Bézier points \( p_i \in \mathbb{R}^m \), \( i = 0 \ldots n \) (here \( m = 2, 3 \)), that build the so-called control polygon, for \( u \in [0,1] \) by

\[
 x(u) = \sum_{i=0}^{n} B^n_i(u) p_i = \sum_{i=0}^{n} \binom{n}{i} u^i (1-u)^{n-i} p_i .
\]

The \( B^n_i(u) \) are called Bernstein polynomials. Derivatives of Bézier curves of order \( n \) are Bézier curves of order \( n - 1 \).

To keep the technical descriptions as simple as possible, we will not touch the additional constraints coming from continuity conditions. Let us only state here that for B-splines of order four or higher the first and second derivative are continuous. With the exception of extraordinary vertices the same is true for Catmull Clark and Loop subdivision.

One main advantage for interpolation and approximation with curves like (1) is the linearity in the control points \( p_i \). B-spline curves have these property, too. They are given by

\[
 x(u) = \sum_{i=0}^{n} N_{i,p,U}(u) p_i
\]

where \( N_{i,p,U}(u) \) is the \( i \)-th normalized B-spline function of order \( p \) (degree \( p - 1 \)) corresponding to the generally non-uniform knot vector \( U = (u_0, u_1, \ldots, u_{n+p}) \). We usually assume that \( U \) is clamped and the internal knots are single knots, i.e., \( u_0 = \ldots = u_{p-1} < u_p < \ldots < u_n < u_{n+1} = \ldots = u_{n+p} \). For the sake of simplicity we write \( n_{i,p} \) instead of \( n_{i,p,U} \). Furthermore, without loss of generality in this paper we assume \( U \) to be scaled in such a way that \( u_0 = 0 \) and \( u_{n+p} = 1 \). Surfaces are represented by tensor products of the form

\[
 x(u,v) = \sum_{i=0}^{n} \sum_{j=0}^{m} N_i(u) N_j(v) p_{ij} \text{ or } x(u,v) = \sum_{i=0}^{n} \sum_{j=0}^{m} B^n_i(u) B^m_j(v) p_{ij}
\]
for B-spline and Bézier surfaces, respectively. Again these representations are linear in the control points. The same applies for stationary subdivision curves and surfaces. The algorithms presented below can be used for all these curve and surface classes. We rewrite all the above representations in the form

\[ x(u,v) = \sum_{j=1}^{N} B_j(u,v) p_j \quad (4) \]

with general basis functions \( B_j(u,v) \). In [10] and [11] it was shown that for Loop and Catmull Clark subdivision the corresponding \( B_j(u,v) \) can be evaluated at arbitrary points.

Further information on Bézier- and B-splines can be found in standard literature on CAGD (Computer Aided Geometric Design), e.g. [4]. For more details on subdivision surfaces see [5].

**Approximation of Curves and Surfaces**

To be more precise with our least squares formulation we have to introduce some technical notations. For an optimal approximation of a given surface \( y(s,t), (s,t) \in [s_{\min}, s_{\max}] \times [t_{\min}, t_{\max}] =: D \subset \mathbb{R}^2 \) by a parametric surface we have to determine the control points \( p_j \) associated to (4) in such a way that

\[ \max_{(s,t) \in D} \min_{(u,v) \in [0,1]^2} \| y(s,t) - x(u,v) \|_2 \quad (5) \]

is minimized. Since in practice these problems are too complex to solve we switch to discrete approximation problems. For Béziers the whole domain \( D \) is normally a-priori splitted into a couple of subdomains. In each (sub-)domain a set of approximation points \( y_i = y(s_i, t_i) \) is chosen. In this paper we assume that an initial simple mesh corresponding to the correct topology of the target surface and allowing the computation of the \((u_i, v_i)\) is already given. We further use the error estimator

\[ \delta = \max_i \| y_i - x(u_i, v_i) \|_2. \quad (6) \]

If the error is too large, we (recursively) subdivide the parameter domain. For subdivision surfaces this is a normal subdivision step. For B-splines we use knot insertion and for Béziers we split each subdomain into four equal parts. The corresponding parameter values \((u_i, v_i)\) are recomputed for the new domains. This step is not necessary for B-splines.

The basic approximation principle, explained for surfaces, is as follows. Let \( y_i, i \in \{1, \ldots, M\} =: I \) be given data points or samples on a given target surface. We want to compute a good approximating parametric fitting surface with a representation of the form (4). For the B-spline case we further assume that the knot vectors \( U \) and \( V \) are already determined with the constraints of the previous section. In a first step we have to find (at least approximately) the nearest points \( x_i = x(u_i, v_i) \approx y_i \) on the fitting surface. A good estimation of the parameter values \((u_i, v_i)\) might be a difficult task. Furthermore it is well known that the parameterization problem is a fundamental one for the whole approximation process and the final result. Therefore, parameter correction procedures
have to be used to improve the quality of the final approximation. Unfortunately the decoupling of the overall fitting procedure into the two independent steps parameter correction and solving a linear least squares problem with fixed parameters converges very slowly. On the other hand the overhead due to the curvature computation and the set up of a more complex SDM error function leads to computational inefficiency of SDM. A comparison in [3] shows that the time to attain comparable results used by SDM on iterative optimization is about 30% to 50% more than PDM. The goal of our development is the avoidance of the curvature computation in unity with the construction of a cheap error function that accelerates the parameter correction.

Figure 1: Recursive approximation with quadratic Bézier curves. Upper plot: Initial approximation and parameter correction. Lower plot: First subdivision step.
To get a better understanding of the key ingredients of our methods that lead to convergence acceleration, we first analyze the behavior with planar curves. Figure 1 shows the approximation of one quarter of an ellipse with quadratic Béziers. We force interpolation at the beginning and the end of the interval. Thus $p_0 = y(t_{\text{min}})$, $p_2 = y(t_{\text{max}})$ and the only free control point is $p_1$. Only one level of subdivision is shown (lower plot). This simple example already shows the power of spline approximation and the influence of parameter corrections. For quadratic splines we expect the convergence rate $O(h^3)$. Of course this can only be achieved if the underlying curve is in $C^3[t_{\text{min}}, t_{\text{max}}]$. Thus for every recursion the error will be (asymptotically) reduced by a factor of 8.

We chose $\Delta t = (t_{\text{max}} - t_{\text{min}})/8$ and started with $y_i = y(t_{\text{min}} + i \Delta t)$ for $i \in \{1, 2, \ldots, 7\}$. As initial guess for the $u_i$ we used $u_i = i/8$. In the upper row we see the initial approximation with different (scaled) error vectors. Those without parameter correction are not orthogonal to the target curve. Even more, without parameter correction the error vectors become more and more tangential to the curves in higher recursion levels. For a measurement of the real curve to curve distance at these points the error vectors should be orthogonal to them. This can be achieved with parameter corrections. The error vectors after parameter correction are shown too. Finally the result after some iterations is shown. To make the changes of the parameter values visible we projected onto the target curve. All error vectors are scaled with a factor of ten. In the lower part the result after one subdivision is shown. Here the error vectors are scaled with 80 to make the effect of error reduction visible. Since the error vectors scaled in such a way are shorter for the subdivided curve the errors have reduced by factors larger than 8.

![Figure 2](image-url)

**Figure 2:** Approximation with cubic Bézier curves – PDM 100 iterations and our method 3 iterations. The error vectors are scaled with a factor of 100.

Next we want to study the influence of good parameter corrections. Again we start with an illustrative example. In Figure 2 we have used a cubic Bézier with interpolation at the boundaries for the approximation of a trochoid. Only one approximation interval is shown with error estimates. Again as reference an approximation with desired relative error of $10^{-6}$ is shown. Very few control points lead to a good approximation of the trochoid curve. For cubic splines we expect the convergence rate $O(h^4)$. Here the curve for approximation has to be in $C^4[t_{\text{min}}, t_{\text{max}}]$. Furthermore we see that we need much less iterations with an optimized parameter correction strategy.

For a better understanding of fruitful concepts for parameter correction we take a look on the related analysis. Pottmann et al. ([6, 7]) introduced a concept of using
local quadratic approximants of the squared distance function to a surface for approximation strategies. At each regular point \( y(s,t) \) of a smooth target surface we have a local right-handed Cartesian system whose first two vectors \( t_k(s,t) \), \( k = 1, 2 \) are the principal curvature directions belonging to the signed principal curvature \( \kappa_k(s,t) \). The third one \( n(s,t) \) is orthogonal to them. Furthermore these vectors are normed. The thereby defined frame \( \Sigma \) is called principle frame. If we define \( \rho_k(s,t) = 1/\kappa_k(s,t) \) we can write the second order taylor approximation for a point \( x_i \) with coordinates \((0,0,d)\) in \( \Sigma \) (i.e. \( y \) is the closest point to \( x_i \) on the target surface) as

\[
F_d(x) = \frac{d}{d - \rho_1} \left( t_1^T (x - y) \right)^2 + \frac{d}{d - \rho_2} \left( t_2^T (x - y) \right)^2 + (n^T (x - y))^2. \tag{7}
\]

Unfortunately there are points \( x_i \) for which \( F_d \) is not a positive definite quadratic form. This leads to a system matrix that is not symmetric positive definite such that we can not guarantee the existence and uniqueness of the minimum. For this reason (7) is modified to

\[
F_d^+(x) = \frac{|d|}{|d| + |\rho_1|} \left( t_1^T (x - y) \right)^2 + \frac{|d|}{|d| + |\rho_2|} \left( t_2^T (x - y) \right)^2 + (n^T (x - y))^2. \tag{8}
\]

Now we end up with a positive definite system matrix but the second order of our approximant is destroyed and thus the claimed quadratic convergence of that method. Nevertheless using (8) for minimization needs much less iterations. For this reason we use a scaled combination of the standard minimization of \((x - y)^2\) and \((n^T (x - y))^2\). The later one coincides with \( F_d^+(x) \) if \( d = 0 \), which is the case if \( x_i \) lies on the target surface. From this we conclude that it is a good choice to make the scaling dependent of \( d \).

\[
F_d^{new}(x) = (x - y)^2 + \lambda(d) \left( n^T (x - y) \right)^2. \tag{9}
\]

For using (9) in our minimization concept we only need a normal \( n_i \) for each sample point \( y_i \). Since the normal of the fitting surface \( x(u,v) \) approximates the normal of the target surface, we can even use the normal of \( x(u,v) \). The high cost for the curvature computation is avoided, too.

For the sample points \( y_i, i = 1, 2, \ldots M \) with associated parameter values \((u_i,v_i)\) of the fitting surface we can use the following setup. According to (4) the \( x_i \) are given as \( x_i = x(u_i,v_i) \). We collect the control points \( p_j \) in a vector (of 3d vectors) \( p \), the \( y_i \) in a vector \( y \) and all coefficients \( a_{ij} := B_j(u_i,v_i) \) in a matrix \( A \in \mathbb{R}^{M \times N} \). With this notation PDM is simply

\[
p^* = \arg\min_p \| A p - y \|_2. \tag{10}
\]

For the solution of (10) we can use orthogonal transformations or the normal equation

\[
A^T A p^* = A^T y. \tag{11}
\]

Using \( x_i = \sum_j a_{ij} p_j \) we can write the distance to the tangent plane as

\[
\| n_i^T (x_i - y_i) \|_2 = \| n_i^T \left( \sum_j a_{ij} p_j - y_i \right) \|_2 = \left\| \sum_j a_{ij} n_i^T p_j - n_i^T y_i \right\|_2. \tag{12}
\]
Notice that (12) leads to a minimization over the control points \( p_j \). To use matrix notations we have to separate the \( x \), \( y \) and \( z \) components. We define \( N_x = \text{diag}\{ n_{i,x} \} \), \( N_y = \text{diag}\{ n_{i,y} \} \), \( N_z = \text{diag}\{ n_{i,z} \} \) and collect the terms \( n_y^T y_i \) in a vector \( d \). Now we can write this part as

\[
\| N_x A p_x + N_y A p_y + N_z A p_z - d \|_2 \rightarrow \min. \tag{13}
\]

For our final minimization problem according to (9) we use (10) and a scaled portion of (13). Here we have to split \( p \) into its \( x \)-components \( p_x \), \( y \)-components \( p_y \) and \( z \)-components \( p_z \). The same way we split \( y \) into \( y_x \), \( y_y \) and \( y_z \). With

\[
\hat{A} = \begin{pmatrix}
A & 0 & 0 \\
0 & A & 0 \\
\lambda N_x A & \lambda N_y A & \lambda N_z A
\end{pmatrix}, \quad \hat{p} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} \quad \text{and} \quad \hat{y} = \begin{pmatrix} y_x \\ y_y \\ y_z \\ \lambda d \end{pmatrix} \tag{14}
\]

our minimization problem now reads

\[
\hat{p}^* = \arg \min_{\hat{p}} \| \hat{A} \hat{p} - \hat{y} \|_2. \tag{15}
\]

Thus again we can use orthogonal transformations, the normal equations or iterative methods. Especially for recursive approximation of surfaces the iterative methods are much faster. The parameter \( \lambda \geq 0 \) is chosen depending on the error estimator. For small errors we use a large \( \lambda \).

As iterative solver we normally use a modification of the Conjugate Gradient method for linear Least Squares. CGNR (see [9]) can directly be applied to (15). But we have developed a more efficient variant for this special case. The only costly part of CG applied to \( A x = b \) is the matrix-vector product of the system matrix \( A \) with an intermediate direction in each step. The normal equation for (15) is

\[
\hat{A}^T \hat{A} \hat{p}^* = \hat{A}^T \hat{y}. \tag{16}
\]

For the standard CG method applied to (11) we have to compute one product of the form \( z_k = A^T A d_k \) in each step. \( d_k \in \mathbb{R}^3N \) is a \( N \)-vector of 3d-vectors, \( A \in \mathbb{R}^{M \times N} \). Thus \( A d_k \in \mathbb{R}^M \) and \( z_k \in \mathbb{R}^3N \). As in CGNR we avoid the computation of \( \hat{A}^T \hat{A} \). We rewrite \( \hat{A}^T \hat{A} \) as

\[
\hat{A}^T \hat{A} = \begin{pmatrix}
A^T A + A^T \lambda^2 N_x^2 A & A^T \lambda^2 N_x N_y A & A^T \lambda^2 N_x N_z A \\
A^T \lambda^2 N_x N_y A & A^T A + A^T \lambda^2 N_y^2 A & A^T \lambda^2 N_y N_z A \\
A^T \lambda^2 N_x N_z A & A^T \lambda^2 N_y N_z A & A^T A + A^T \lambda^2 N_z^2 A
\end{pmatrix}
\]

\[
=: A^T \otimes \left( I + \lambda^2 N_x^2 \right) \left( I + \lambda^2 N_y^2 \right) \left( I + \lambda^2 N_z^2 \right) \otimes A =: A^T \otimes B \otimes A \tag{17}
\]

\( B \) is a \( 3 \times 3 \) block matrix. Each of the 9 blocks is a diagonal matrix. The multiplication \( \otimes \) is implicitly defined in (17) as follows. \( A^T \otimes B \) means that each block of \( B \) is multiplied with \( A^T \) from the left. \( B \otimes A \) means that each block of \( B \) is multiplied with \( A \) from the right. Let \( d_k \in \mathbb{R}^3N \). The first product is the same as above: \( \hat{z}_k = A d_k \). Next we apply \( B \): \( \bar{z}_k = B \otimes \hat{z}_k \). \( \otimes \) describes an update of the elements in \( \hat{z}_k \). Note that due to
the structure of $B$ only three reals are added to each component in $\hat{z}_k$. The last step is a normal multiplication: $z_k = A^T \bar{z}_k \in \mathbb{R}^3$. Let us summarize these steps:

$$\hat{z}_k = A d_k, \quad \bar{z}_k = B \otimes \hat{z}_k, \quad z_k = A^T \bar{z}_k$$

(18)

For the right hand side we use

$$\hat{A}^T \hat{y} = \begin{pmatrix} A^T y_x + A^T \lambda^2 N_x d \\ A^T y_y + A^T \lambda^2 N_y d \\ A^T y_z + A^T \lambda^2 N_z d \end{pmatrix} = A^T \otimes \begin{pmatrix} y_x + \lambda^2 N_x d \\ y_y + \lambda^2 N_y d \\ y_z + \lambda^2 N_z d \end{pmatrix} =: A^T \otimes \hat{b}. \quad (19)$$

Similar to $z_k = B \otimes \hat{z}_k$ above $\hat{b}$ is an update of $y$. Here only one real is added to each component. We rewrite this as

$$b = A^T \left( y \oplus \begin{pmatrix} N_x \\ N_y \\ N_z \end{pmatrix} \otimes \lambda^2 d \right). \quad (20)$$

With (20) (has to be computed only once) and (18) we have described an efficient form for the costly parts of the CG method. The rest of the computations in CG are only two inner products and three vector additions in each step. Note that we can make use of the sparsity of the matrix $A$ in (18).

Figure 3: Convergence studies: 30 iterations with PDM (left) versus 3 iterations of our new method (right).

To analyze the coupling effect of the parameter correction with the solution of the linear system of equations we again use a simple curve approximation. Figure 3 shows the approximation of one quarter of an ellipse with cubic Béziers. We have chosen $a = 25$ and $b = 17$ for the principal axis of the ellipse. Again we force interpolation at the beginning and the end of the interval – here including the tangent – and the same initial distribution of the sample points as in Figure 3. The left plot shows the result after 30 iterations with PDM. The error vectors are scaled with a factor of 1000. This is compared with the result of our method after only three iterations. The main reason for the improvement in our method can not be seen from these plots. In Figure 4 we have plotted the signed distance of the parameter values in $k$-th iteration to those of the best approximation. Due to the coupling in our method we reach the optimal parameter values very quickly. We show
Figure 4: Differences of the parameter values from their optimal place: Upper plot PDM, lower plot our method.
the initial values and 8 iterations for our method and about fifty iterations of PDM. Here PDM needs more than a million iterations to get the optimal parameter values with a precision that leads to a solution that coincides with the optimal solution up to 10 digits.

The table below shows the error estimator according to (6) for the iteration. The errors are shown as multiple of $10^{-3}$. Our method shows a quadratic convergence. The number of significant digits is doubled in each step. Due to influence of the machine precision this stops at iteration 5. With iteration 6 we have reached the minimum up to 10 significant digits. PDM has 9 significant digits after one million iterations.

<table>
<thead>
<tr>
<th>#itera</th>
<th>PDM method</th>
<th>our method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>max error [10^{-3}]</td>
<td>max error [10^{-3}]</td>
</tr>
<tr>
<td>1</td>
<td>51.0291992263</td>
<td>1.2156378925</td>
</tr>
<tr>
<td>2</td>
<td>45.5340955079</td>
<td>1.1050648499</td>
</tr>
<tr>
<td>3</td>
<td>40.8463545168</td>
<td>1.0932771389</td>
</tr>
<tr>
<td>4</td>
<td>36.6556730174</td>
<td>1.0932474969</td>
</tr>
<tr>
<td>5</td>
<td>32.8881440210</td>
<td>1.0932475456</td>
</tr>
<tr>
<td>6</td>
<td>29.4980591588</td>
<td>1.0932475452</td>
</tr>
<tr>
<td>10^4</td>
<td>2.95707434807</td>
<td>1.0932475454</td>
</tr>
<tr>
<td>10^5</td>
<td>1.1713527195</td>
<td>1.0932475453</td>
</tr>
<tr>
<td>10^6</td>
<td>1.0932479122</td>
<td>1.0932475453</td>
</tr>
</tbody>
</table>

Conclusion and Future Research

We have presented and analyzed a novel and fast approximation approach for curves and surfaces. It can be used for the construction of smooth surfaces from point clouds as well as for the reparameterization of given surfaces. The methods are not restricted to be applied to Bézier or B-spline surfaces. They can be used for subdivision surfaces as well. In particular, the focus has been on an analytical understanding of the coupling of parameter correction with the linear Least squares approximation step and the transformation of the mathematical model to a form that can be used efficiently for fast solvers. In comparison to the standard PDM approach with decoupling of the linear solver and the parameter correction our method has a tremendous speed up. We need much fewer iterations without the drawback of the computational overhead of SDM.

Encapsulated Postscript (EPS), PDF and Scalable Vector Graphics (SVG) files have the entities quadratic an cubic Bézier curves. EPS and PDF files are for use in printed publications. SVGs can be used for websites. The major advantage of all three formats is the very small size of the files and the possibility to zoom without grid pattern effects. All the curves shown in this paper are piecewise quadratic or cubic Béziers. The reference curves are approximated with our approach by cubic Béziers to a final relative error of $10^{-6}$.

There is still a large amount of work left for future research. We want to improve the stop criteria for the overall iteration and implement more flexible subdivision of the parameter domain. Up to know we only use uniform fragmentation for the surface case. For very large problems we want to use preconditioning of the CG method. Furthermore, we will test our algorithms with other surface classes. The integration of this methodology into
our algorithms on grid generation and reparameterization, for instance given in [1, 2, 8], will lead to a significant reduction of computation in those packages, too.

References


