Curvature-based energy for simulation and variational modeling

Denis Zorin
New York University
719 Broadway, 12th floor
New York, New York, 10012
dzorin@mrl.nyu.edu

Abstract

Curvature-based energy and forces are used in a broad variety of contexts, ranging from modeling of thin plates and shells to surface fairing and variational surface design.

The approaches to discretization preferred in different areas often have little in common: engineering shell analysis is dominated by finite elements, while spring-particle models are often preferred for animation and qualitative simulation due to their simplicity and low computational cost. Both types of approaches have found applications in geometric modeling. While there is a well-established theory for finite element methods, alternative discretizations are less well understood: many questions about mesh dependence, convergence and accuracy remain unanswered.

We discuss the general principles for defining curvature-based energy on discrete surfaces based on geometric invariance and convergence considerations. We show how these principles can be used to understand the behavior of some commonly used discretizations, to establish relations between some well-known discrete geometry and finite element formulations and to derive new simple and efficient discretizations.

1. Introduction

In physics, mechanical engineering and biology curvature-based functionals are used to model the behavior of bending surfaces: car bodies, roofs of buildings, polymer vesicles and biological membranes. In geometric modeling these functionals are used for surface fairing, feature detection and defining variational surfaces; in computer animation, curvature-based functionals are used for simulating motion of deformable objects, most commonly, cloth animation.

While in most cases the continous energy functionals are very similar, their discretizations may differ greatly. The most mathematically consistent approach to discretization is to use a smooth parametric surface representation with sufficient approximation power. In this case curvature-based discrete energy can be computed by analytic or numerical integration of standard expressions for curvature of parametric surfaces. However, this aproach is often

impractical, due to relative inefficiency and complexity of basis functions used in high-order constructions. Commonly used approaches in different application areas range from nonconforming finite elements in engineering applications to bending springs in animation. Finite elements typically rest on more rigorous mathematical foundation and emphasize convergence to the solution of the continuum problem. Discretizations applied in other areas, geometric modeling and animation in particular, often are based on qualitative considerations and are designed for efficiency and simplicity.

In this paper we present a brief overview of curvature-based energy functionals and their discretizations, and discuss a set of natural requirements that can be used to evaluate different discretizations. Several simple general principles result in surprisingly specific conditions, especially for discretizations using a small number of degrees of freedom per term. Our focus is on the relationship between the solution of the continuum energy-minimization problems and discrete problems. In particular, we examine convergence of discrete solutions to the continuum solutions. We discuss how simple convergence conditions can be used to understand the behavior of an energy functional, as well as establish requirements for meshes used to represent deformable surfaces.

2. Continuous curvature-based functionals

The simplest form of the curvature-based energy is the integral of the mean curvature:

$$k \int H^2 dS \tag{1}$$

where the integral is computed over the surface and k is a constant.

The relationship between curvature and bending energy first appeared in an explicit form in Euler's elastica, curves minimizing the integral of squared curvature. The exact

nature of the relationship between the bending forces in thin plates and their curvature has initially emerged in the work of Sophie Germain [7]; a more complete theory was developed by Kirchhoff [8]. Initially the focus was on small deformations, in which case the curvature could be approximated by derivatives of displacements; Love [22, 21] has extended this theory to deformations of thin shells of arbitrary undeformed shape. The case of large deformations of plates was considered by von Kármán [18]; Complete invariant equations for thin shells are often referred to as Koiter's model [20]. In all cases a thin plate or a shell is modelled by a two-dimensional surface, with shear effects in the volume of the shell ignored. Taking these effects into account leads to a different set of models, with additional energy terms which we do not consider here.

Among the models that we consider Koiter's model has the most general form. The model is derived under several assumptions: shell thickness h is much smaller than the minimal radius of curvature; the normal lines to the middle surface in the undeformed state are deformed into normal lines to the middle surface in the deformed state (the Kirchhoff-Love assumption); components of stress in the normal direction can be assumed to be zero. In addition, small strains are assumed and linear elasticity constitutive law is used.

Under these assumptions, the total energy of the deformed surface can be separated into two parts: the membrane energy, which we do not consider, and the bending energy. The expression for the bending energy can be written in a compact form using the *shape operator*. Recall that the the shape operator $\Lambda(p)$ at a point p of a smooth surface is defined by its action on tangent vectors: for a tangent vector t, $\Lambda t = -\partial_t n$, where ∂_t is the derivative in the direction of t and n is the unit normal to the surface. The eigenvalues of this operator are principal curvatures, and its eigenvectors are principal curvature directions. If S denotes the gradient of surface deformation at a point, then Koiter's bending energy can be written in the form

$$\frac{D}{2}\left(\nu(\mathrm{Tr}\left(S^T\Lambda S - \Lambda_0\right))^2 + (1-\nu)\mathrm{Tr}\left((S^T\Lambda S - \Lambda_0)^2\right)\right) \tag{2}$$

where ν is the Poisson ratio, Λ and Λ_0 are the shape operators for the deformed and undeformed surfaces respectively. The flexural rigidity D is related to the Young modulus and is given by $Eh^3/24(1-\nu^2)$. This expression can be given more intuitive form by observing that for small strains the difference between $S^T\Lambda S$ and Λ are of higher order and can be neglected, yielding

$$\frac{D}{2}\left(\nu(\mathrm{Tr}\left(\Delta\Lambda\right))^2 + (1-\nu)\mathrm{Tr}\left((\Delta\Lambda)^2\right)\right) \tag{3}$$

where $\Delta\Lambda = \Lambda - \Lambda_0$. To make the geometric meaning of this expression clear it is useful to write it in terms of familiar invariants: mean curvature and Gaussian curvature. As two surfaces participate in the expression, an additional needed geometric invariant is the rotation β of the principal curvature directions under deformation. One can show that (3) can be written as

$$D\left((1+\nu)\Delta H^2 + (1-\nu)(\Delta A^2 + 4AA_0\sin^2\beta)\right) \tag{4}$$

where $A = \sqrt{H^2 - K}$ is the half-difference of the principal curvatures. This expression clarifies the relationship between Koiter's formulation and the simplest energy (1): if we assume that the undeformed state is flat, then $A_0 = H_0 = 0$ and the expression reduces to $D\int(2H^2 + (1-\nu)K)$. The second term remains constant on surfaces with fixed boundaries and fixed normals on the boundary. This term can also be neglected if the deformations are close to isometric. This leaves only the term of the form (1). We also observe that in (3) for flat undeformed configurations the first term $\int(\mathrm{Tr}\Lambda)^2$ is proportional to the integral of the squared mean curvature, and the second term $\int\mathrm{Tr}\Lambda^2 = \int\kappa_1^2 + \kappa_2^2$ is often called the total curvature functional in geometric modeling literature.

In mathematical literature the functional (1) is known as the Willmore energy as the investigation of surfaces minimizing this energy started with the work of Willmore [37] ([38] is a more accessible reference). The Euler-Lagrange equation for this functional has the form

$$\Delta_M H + 2H(H^2 - K) = 0 \tag{5}$$

where Δ_M denotes the Laplace-Beltrami operator. This equation can be used to define the Willmore flow, which was demonstrated to be useful for surface restoration [4].

A variation of this functional, possibly with a nonzero undeformed mean curvature, is known in membrane physics literature as Canham-Helfrich energy [3, 12]:

$$\int(H - H_0)^2 dS \tag{6}$$

We note that this is not a special case of Koiter's bending energy, unless $\kappa_1 = \kappa_2$ for both deformed and undeformed configurations, as it is for spheres. The reason for this is that membranes modeled by this equation are liquid. For liquid membranes, one can neglect the work required to transform the surface into a configuration with the same principal curvatures but with principal curvature directions aligned with a material coordinate system a different way.

Functionals used in geometric modeling. Functional (1) is the most common foundation for various discrete functionals used to construct variational surfaces. In many cases a linearized form of the functional is used, replacing curvatures with second partial derivatives of the parameterization, or with Laplace-Beltrami operators computed with respect to a fixed approximate surface shape. Discrete functionals are often derived not by formally discretizing (1) but by using simplified expressions capturing its properties qualitatively.

Another important family of functionals are based on minimizing curvature variation; we do not consider this types of functionals in this paper, but similar considerations would apply to their discretization.

3. Desirable discrete energy properties

To simplify our presentation, we focus on the functional (1), but similar considerations apply for the more general form (3). We also use the curve energy $\int \kappa^2 ds$ to provide simple examples. Before we discuss several discretizations of this functional, we formulate a number of conditions that the discrete functional should satisfy.

We use $\mathbf{p}$ to denote the vector of the vertices of the surface or curve and $E(\mathbf{p}) \geq 0$ to denote the discrete energy function. We will also consider discretizations using additional angular degrees of freedom; the conditions we discuss also apply in these cases.

We assume that the mesh is split into nonoverlapping domains D_i with areas A_i. For each domain we use vertices $\mathbf{p}_i$ of a submesh N_i contained in a fixed-size neighborhood of D_i to compute a local curvature estimate $H_i(\mathbf{p}_i)$. For example, D_i can be a triangle or a Voronoi area of a vertex or an edge, and N_i can be all vertices connected by an edge to a vertex in D_i.

The complete curvature functional is defined by

$$E(\mathbf{p}) = \sum_i A_i (H_i(\mathbf{p}_i))^2 \qquad (7)$$

Similar definitions apply in the curve case which we use in examples.

Geometric invariance conditions. This group of conditions ensures that the geometric behavior of the discrete energy function matches the essential features of the continuous functional.

G1. *For planar (collinear for curves) configurations of control points the energy vanishes: $E(\mathbf{p}^{lin}) = 0$. This requirement ensures that planar surfaces and straight lines have zero curvature.*

G2. *The energy is invariant with respect to rigid transformations: $E(\mathbf{p}) = E(T\mathbf{p})$, where $T\mathbf{p}$ is defined by applying a rigid transform to each control point.*

G3. *For surfaces, the energy is invariant with respect to uniform scale: $E(s\mathbf{p}) = E(\mathbf{p})$; for curves, the energy scales as $1/s$: $E(s\mathbf{p}) = (1/s)E(\mathbf{p})$.*

We note that functions based on angles between normals and discrete geometric measures of curvature satisfy requirements G1 and G2; for careful choice of scaling coefficients G3 is also satisfied for any reasonable functional.

Convergence conditions. Convergence conditions are formulated for sequences of meshes with triangle size, as measured by the radius of the circumscribed circle, going to zero. Additional constraints are typically needed to ensure convergence; a weak constraint may be that the triangle aspect ratio stays bounded; a stronger constraint may require a sequence of regularly subdivided meshes. We call meshes belonging to a specified class *admissible*. Our discussion here is informal; more details will be provided in [42].

In general, for a complete theory of convergence one needs to introduce suitable norms and use functional analysis. Fortunately, if we are interested only in the *necessary*

conditions for convergence, it is sufficient to look at the convergence of the total energy.

We assume that we have a well-posed continuous problem, and a sequence of discrete problems with fixed boundary values for the discrete problems obtained by sampling the fixed boundary values of the continuum problem. For any reasonable definition of convergence of surface approximations, we would expect the following conditions to hold:

C. *For any sequence of solutions of discrete problems for admissible meshes and mesh triangle size going to zero, the energy of the solution of the discrete problem converges to the energy of the solution of the continuous problem.*

This condition results in very specific constraints on the *linearized* version of the discrete energy (7), which we discuss next.

Linearized energy. To convert the general conditions on the energy to specific constructive requirements, we consider a very specific case: small deformations of planar surfaces (plates). Any "good" general discrete energy functional should also work in this case. Let $\mathbf{u}$ be the vector of displacements at each vertex; for small displacements, $\mathbf{u}$ we can replace $E(\mathbf{p})$ by its quadratic approximation: $E(\mathbf{p}_0 + \mathbf{u}) = E(\mathbf{p}_0) + \nabla E^T \mathbf{u} + (1/2)\mathbf{u}^T \nabla^2 E \mathbf{u}$, where $\nabla^2 E$ is the Hessian of E at $\mathbf{p}_0$. As the planar configuration $\mathbf{p}_0$ has zero energy, and this is the minimal possible energy, $E(\mathbf{p}_0) = 0$ and $\nabla E(\mathbf{p}_0) = 0$, which leaves only the quadratic term. Using (7), we obtain $E^{lin}(\mathbf{u}) = (1/2)\sum_i A_i \mathbf{u}_i^T \nabla^2 (H_i^2) \mathbf{u}_i$. As $\nabla^2(H_i^2) = 2H_i \nabla^2 H_i + 2\nabla H_i \nabla H_i^T$ and the energy is zero for the planar configuration the expression for the linearized energy is

$$E^{lin}(\mathbf{u}) = \frac{1}{2} \sum_i A_i (\nabla H_i(\mathbf{p}_i)^T \mathbf{u})^2 \qquad (8)$$

The conditions on the linearized energy are the conditions on the gradient of the mean curvature estimate for planar configurations. Geometric invariance and scaling conditions G1–G3 for $E(\mathbf{p})$ should also hold for $E^{lin}(\mathbf{u})$, as it can be easily shown by differentiating these conditions; furthermore these conditions have to hold for each term in the expression for energy. One immediate consequence of this observation is that only the components of displacements perpendicular to the plane of the initial configuration can have influence on $\nabla H_i(\mathbf{u}_i)$: for any in-plane deformation the initial configuration remains planar, and the curvature remains zero: $\nabla H_i^T \mathbf{u}_i = 0$ for any $\mathbf{u}_i$ perpendicular to the plane normal. This allows us to simplify the expression for the linearized energy further:

$$E^{lin}(\mathbf{u}) = \frac{1}{2} \sum_i A_i ((\nabla H_i^n)^T (\mathbf{p}_i) u)^2 \qquad (9)$$

where $u_i = (\mathbf{n} \cdot \mathbf{u}_i)$ are the normal displacements and ∇H_i^n are the corresponding components of the gradient.

It is instructive to consider the constraints imposed by geometric conditions only on $\nabla(H_i^n)^T u$ for the minimal number of degrees of freedom, which in the case of curves

is 3. In this case $\nabla(H_i^n)^T u$ can be written as $\alpha_i u_{i-1} + \beta_i u_i + \gamma_i u_{i+1}$. As the energy remains zero for any planar configuration, if u_i are sampled from a line it should remain zero; this leads to the conditions of orthogonality of the gradient to constants and linear functions: $\alpha_i + \beta_i + \gamma_i = 0$ and $l_{i-1}\beta_i + (l_i + l_{i-1})\gamma_i = 0$, where l_i is the distance between $\mathbf{p}_i$ and $\mathbf{p}_{i+1}$. This immediately yields

$$\nabla H_i^n = k_i(1/l_{i-1}, -(1/l_{i-1} + 1/l_i), 1/l_i) \quad (10)$$

i.e. the second divided difference up to a scale factor; by scale invariance, the factor k_i depends only on the ratio l_i/l_{i-1}. We observe that invariance considerations define the linearized energy up to a scale factor for a small number of degrees of freedom. We note that the choice of k_i is not arbitrary: incorrect choices may yield incorrect ratios of energy terms for different i. Only the convergence condition C1 allows one to find the correct factor. One can easily check that the expression (10) is proportional to the gradient of $2\theta_i/(l_i + l_{i-1})$, where θ_i is the angle between normals to two adjacent line segments $(\mathbf{p}_{i-1}, \mathbf{p}_i)$ and $(\mathbf{p}_i, \mathbf{p}_{i+1})$, i.e. all geometrically invariant expressions for the curvature coincide with this one up to a non-dimensional scaling constant.

Convergence conditions for the linearized energy. To convert the general condition C to specific conditions for the gradient of the discrete curvature functional, we further specialize the situation that we consider, by considering solutions of the continuum problem that can be computed explicitly.

One can show that the energy of the solution of the linearized discrete problem should converge to the solution of the linearized continuum problem. The linearized form of the energy (1) is simply

$$\int (\partial_{xx} u + \partial_{yy} u)^2 dS \quad (11)$$

where (x, y) is a coordinate system in the plane, and $\int (u'')^2 dl$ for the curve case.

The Euler-Lagrange equation for this functional is the biharmonic equation. It is known that under certain assumptions on the boundary this problem has a unique solution for Dirichlet and Neumann boundary conditions specified. One can easily check that it is a solution, if we specify the boundary conditions by evaluating a second-degree polynomial $u = q(x, y)$ on the boundary of the domain. Second-degree polynomials are special solutions, as the value of the integrand in (11) remains constant.

There are two explicit and easy-to-check conditions that follow from the convergence requirement. The crucial observation, on which these conditions are based, is related to resampling a quadric using the same sampling pattern but possibly scaled and translated. We focus on the one-dimensional case, the two-dimensional case is similar. Consider a local curvature estimate $\nabla H^T(\mathbf{p}^0)u$, associated with vertices $p_l \ldots p_m$, $\mathbf{p}_k = (x_k, u_k)$, $\mathbf{p}_k^0 = (x_k, 0)$. Suppose u is a quadric, i.e. $u_k = a x_k^2$ (we do not include linear and constant terms as these make no contribution). Suppose we

sample the same quadric at points $x'_k = h x_k + t$. It is easy to verify that the corresponding *local curvature energy term scales exactly as h as the sample point spacing is scaled by h.*

We are interested in sampling pattern x which are *tilable*. We consider curves defined by normal perturbation of an initial linear state. A sequence of sample points is a tiling if the linearized energy of the complete curve is a sum of energy terms defined by subsequences obtained by translation and scale from the original pattern x. We call a sampling pattern tilable if one can have an arbitrary long tilings with respect to this pattern.

For example, for three-point curvature approximation described above the simplest family of tilable patterns are sequences of points $x_1, \ldots x_n$ for which the distances between $x_2 - x_1$ and $x_n - x_{n-1}$ are equal: one can link copies of one of these patterns together overlapping the last two points of each copy with the first two points of the next.

We call a vertex of a mesh or a polygonal line *interior*, if for any triangle (segment) added to the boundary of the mesh, its energy is not affected by this vertex; for curves and three-point curvature estimates, this excludes two points on each side.

Exact scaling of the linearized energy for quadrics has two remarkable consequences for tilable sampling patterns:

C1. *If a mesh or a polygonal line is sampled from a quadric, and the sampling pattern can be tiled, the discrete energy should reproduce the continuous energy of quadrics exactly;*

C2. *If the discrete energy density is constant for a polygonal line and its sampling pattern can be tiled, the total energy should be minimal with respect to variation of all interior vertices.*

The reason for the first requirement can be understood as follows. Suppose for a tilable sampling pattern and for a quadric, the energy estimate differs from the correct value by a factor $a \neq 1$. For a given line segment we can construct a series of approximations of the quadric by shrinking the original pattern by a factor of 2^n, and tiling 2^n copies of the original pattern. The energy error for each tile will scale as the tile size, i.e. 2^{-n}; as the number of tiles grows as 2^n, the total error does not decrease.

The line of reasoning for the second requirement is practically the same: if we assume the first requirement satisfied, we know that we can at least reproduce the energy of the quadric exactly. However, nothing guarantees that we cannot obtain a lower energy aE, for some $a < 1$, and if we can, we can obtain the same error for an arbitrarily fine mesh.

We note that the notion of tilable patterns clarifies how one can obtain convergence even for energy functionals satisfy the requirements only for a limited class of patterns (e.g. with regular connectivity): it is necessary to restrict allowable refinement procedures to those producing only the "good" patterns almost everywhere.

In all cases that we have considered, any pattern for which conditions C1 or C2 are violated can be used to construct a tilable pattern containing it, which also violates

IEEE
COMPUTER
SOCIETY

C1 or C2. As a result it makes sense to enforce these conditions for any sampling pattern, or to exclude "bad" patterns by enforcing special refinement rules.

These conditions closely parallel the famous patch test widely used in the finite element community. However, it is based primarily on geometric considerations, and quite general assumptions on the way the discrete energy is computed without introducing a basis or explicit high-order approximations of the continuous surface. One important difference is the need to use tilable sampling patterns, as we do not assume that all degrees of freedom used for an area A_i correspond to points inside the area.

One-dimensional examples. Applying condition $C1$ to the three-point linearized energy (10), we immediately obtain the value for the constant; not surprisingly, the resulting expression is the second finite difference, scaled by the length of the segment associated with these three points. This length, as well as the point at which we compute the energy density, can be chosen in different ways. Consider two possible choices; for simplicity, we assume that points are equispaced in the undeformed configuration. In the first case, we compute the total energy as

$$\sum_i \frac{1}{h^3}(2u_i - u_{i-1} - u_{i+1})^2 \qquad (12)$$

associating an interval of length h with each term. In the second case, we use a sparser sampling, with one term for each interval of length $2h$:

$$\sum_i \frac{2}{h^3}(2u_{2i} - u_{2i-1} - u_{2i+1})^2 \qquad (13)$$

Both approaches satisfy condition C1. To check condition C2, we need to verify that the gradient of the energy evaluated for a quadric is zero.

We observe that it is equivalent to verifying that for any displacement of the interior nodes from the planar configuration, the total discretized curvature remains zero, as the total (linearized) curvature in the first case is computed as

$$\sum_{i=1}^{n-1} \frac{2}{h}(2u_i - u_{i-1} - u_{i+1}) = \frac{2}{h}(u_1 - u_0 + u_{n-1} - u_n) = 0 \qquad (14)$$

assuming that two points on each side are fixed.

In the second case, it is easy to see that C2 is not satisfied, and the energy can be decreased below the energy of the quadric, e.g. by setting $u_{2i} = (u_{2i-1} - u_{2i+1})/2$ for interior points. The discrete energy in this case does not even have a unique minimum (Figure 1): one can make it zero for arbitrary odd number of points, by placing every even-numbered point precisely at the midpoint of its neighbors. This functional can be used to estimate the energy of a curve, but cannot be used to compute the minimal curve.

We can easily construct artificial examples for which the condition C2 is violated in a more subtle way, and a unique solution still exists e.g. by assigning different areas to sequential points (Figure 1, curve b).
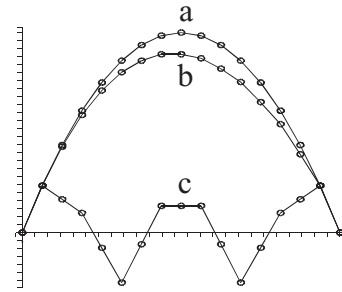

Figure 1. Results of optimizing different linearized squared curvature functionals in 1d; a) The functional satisfies both conditions C1 and C2; a quadric is exactly reproduced from boundary values. b) Condition C2 is not satisfied; the area allocated to each local energy term oscillates from one point to the next, with increasing amplitude from left to right. c) Condition C2 is not satisfied, functional (13) is used, i.e. there are half as many terms. The solution is not unique in this case.

We also observe that it is possible to satisfy condition C2 without satisfying C1: one can rescale the energy arbitrarily.

Although the presentation above was restricted to vertices as degrees of freedom, one can easily see that any local geometric degrees of freedom (surface normals, mean curvature etc) could be used, and the reasoning would still hold with only minor adjustments. For example, if normals are added to the mesh as degrees of freedom, one needs to take into account invariance of normal with respect to scale, and for linear analysis observe that for small displacements the projection of the deformed normal to the original plane is the gradient of the displacement function.

4. Discrete energy functionals

Having established the basic requirements for the discrete energy functionals in the previous section, we now apply them to evaluate existing functionals for meshes. The brief overview in this section is not by any means intended as comprehensive: we focus on what we consider representative examples, which demonstrate the relationships between different approaches. The situation for two dimensions is much more complex than what we have observed for curves: for the minimal number of points, it is impossible to satisfy C1 and C2 for general meshes.

An overview of discretization approaches. In *engineering* applications, the principal tool for discretizing thin shell problems is the finite element method, which reflects the fact that accuracy is a primary concern. The number of shell finite elements is quite large (see e.g. a survey [6] of just one family of such elements). The two main groups of finite elements for plates and thin shells are conforming elements

with C^1 continuous basis functions and reproducing polynomials of degree two, and nonconforming, which are not C^1-continuous.

Most finite element techniques share two distinctive features. First, it is common to linearize the problem first, and construct methods for solving the linear problem, rather than discretize the nonlinear formulation and then linearize the discrete formulation. Second, all degrees of freedom associated with an area of the mesh are typically assigned to nodes in this mesh (some notable exceptions include [25] and [26, 27] which we discuss below). Most of such elements use rotational degrees of freedom.

The advantage of linearizing the continuous problem is that that for many engineering problems deformations are small, and the linearized form is sufficient. In addition, this makes it possible to ensure that the method satisfies convergence conditions for small deformations by construction. For conforming elements, convergence is assured by the general FEM theory. and most practical nonconforming elements are shown to satisfy the patch test, [17] which is analogous to our conditions C1 and C2 (whether various forms of the patch test are sufficient or necessary for convergence is discussed in a number of books and papers: [41, 31, 16, 32, 34]).

As a downside, geometric invariance properties for large deformations are often not guaranteed by construction, and in a few cases the transition from linear to nonlinear problem results in the loss of invariance e.g. with respect to the change of node numbering.

In contrast, high order finite elements are rarely used in *geometric modeling*, possibly because they usually provide visually poor surface approximation (e.g. nonconforming elements can be even discontinuous). Most early work on fairing applied functionals directly to spline patches. The geometric functional (1) is often replaced by its linearization or linearization of the equivalent total curvature functional $\int 2H^2 - K dS$ (e.g. [35, 10]). Variational modeling for meshes based on curvature-based functionals was explored by many authors; some examples are [36, 19, 29, 40, 30].

Clarenz et al. [4] recently demonstrated that linear finite elements can be successfully applied to curvature-based functionals if additional degrees of freedom are introduced. The most commonly used approximation to the integral mean curvature for meshes was described in [15, 28]; variations of this approach were systematically considered in [23] and it was extended to shape operators in [5]. For this type of approaches, the situation is quite different from finite elements: the nonlinear continuous functional is discretized directly, without linearizing it first; the geometric invariance conditions are satisfied automatically, however, convergence is not guaranteed and indeed can be shown to hold only for a restricted class of meshes [39]. Yet another set of techniques were developed for *curvature estimation*: most of these are based on local polynomial fits (see [9] for a survey). One can use these estimates to obtain a discrete energy approximation. Geometric invariance is guaranteed by construction, and the convergence condition C1 is usually satisfied, however, nothing in these construc-

tions ensures that condition C2 is satisfied.

In *computer animation*, a functional similar to Koiter's model (although using second fundamental forms instead of shape operators) was described in [33]. However, the primary focus in most of animation literature was on efficient cloth simulation, and bending is accounted for by using springs with empirically determined parameters. Another common way to account for bending of surfaces represented by meshes is to use forces proportional to the angles between normals of adjacent faces (e.g. [1, 11, 2]). In [11] the connection between this approach and the definition of discrete mean curvature was observed.

While radically different approaches can be used to derive curvature energies, if only few degrees of freedom are used for every local energy term, the results may be quite similar especially if we consider linearized form of the energy. To illustrate the relationship between different approaches we discuss several representative examples.

Discrete shell energy. The most basic type of energy which can be regarded as a curvature approximation uses local terms associated with edges. Specifically, the mean curvature associated with an edge of length l_e and the angle between normals of faces adjacent to the edge equal to θ is taken to be

$$H_{DS}(e) = l_e \theta / A \qquad (15)$$

The corresponding energy term is $(H(e))^2 A = \theta^2 l_e^2 / A$, where A is the area associated with the edge which can be chosen in different ways (e.g. the Voronoi cell area). This formula is derived based on the observation that the mean curvature normal (that is, Hn, where n is the unit surface normal) can be defined as the gradient of the area of the surface with respect to point position.

This simple formula closely resembles the one-dimensional formula for curvature considered in the previous section. Similar discretizations are widely used in computer animation. While this formula satisfies geometric invariance conditions, it cannot possibly satisfy the convergence condition C1: as one cannot recover the second derivatives of the quadric uniquely from just four points (vertices of two adjacent triangles) which we need to compute this energy. Remarkably, condition C2 can be shown to be satisfied.

Averaged energy. While local mean curvature approximation (15) does not converge pointwise, it does converge when averaged over sufficiently large areas, under certain assumptions on mesh refinement. One can extend the idea of the discrete shell energy to triangle-based and vertex-based local energy terms by summing the discrete mean curvatures or curvature normals first, and then squaring to obtain a local energy term. For example for a triangle we can use

$$H_{DS}(T) = 1/2A(l_1\theta_1 + l_2\theta_2 + l_3\theta_3) \qquad (16)$$

where the factor $1/2$ reflects the fact that each edge is shared by two triangles, and A is the area of the triangle. This corresponds to computing the gradient for a larger area, and may improve convergence in some cases. Unfortunately, for
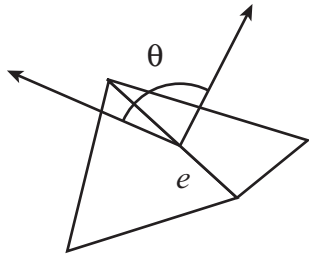
Figure 2. Discrete shell energy local term depends only on the angle between triangle normals θ and the length of edge e.

Figure 3. Averaged curvature for a triangle depends on the angles θ_i between the triangle normal and the normals to the adjacent triangles. The linearization of this energy is similar to the rotation-free finite element construction in [26].

Figure 4. Variables used for the midpoint normal energy. The extra variables φ_i measure deviation of the actual (unknown) surface normal n_{mid} from the average of normals $n_{avg} = (n + n_i)/\|n + n_i\|$ (variables for only one edge out of three are shown).

any fixed-size area, one cannot achieve convergence for a general class of meshes: one needs to restrict consideration to regular grids [39, 14].

This type of triangle-based discretizations are closely related to a special case of rotation-free elements described in [27], although these elements are derived using different considerations.

When averaging is applied to the mean curvature normals for each edge surrounding a vertex v, we get the expression

$$H_{DS}(V)n = \frac{1}{2A} \sum_i l_i \theta_i n_i \qquad (17)$$

where l_i is the length of the i-th edge with endpoint v and A is the area associated with the vertex (see [23]). This functional is equivalent, up to second-order terms in θ_i to the well-known cotangent formula

$$H_{DS}(V)n = \frac{1}{2A}(\sum_i e_i(\cot \alpha_i + \cot \beta_i)) \qquad (18)$$

where α_i and β_i are the two triangle angles adjacent to the i-th edge at its endpoint $v_i \neq v$ and e_i is the vector from v to v_i.

The vertex-based formula, which appeared in [15, 28] is widely used, but also lacks convergence for general meshes. For a proper choice of areas associated with each vertex, it satisfies condition C2, as all energies based on summing of terms (15) over areas: for all such energies, constant discrete curvature solutions are stationary.

Quadric fit energy. An alternative approach would be to use a quadric or higher-order polynomial fit for a part of the mesh and compute the energy of the quadric, as it is done e.g. in [36]. If we use at least six vertices, we can guarantee that condition C1 is satisfied under mild assumptions. However, for a minimal set of samples (Figure 3) using exactly six degrees of freedom, in general one cannot simultaneously satisfy condition C2, which was possible in the linear case for three-point local energy terms.

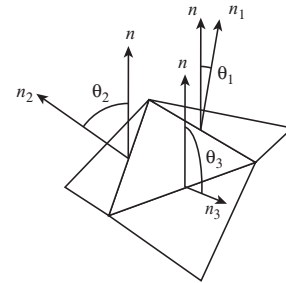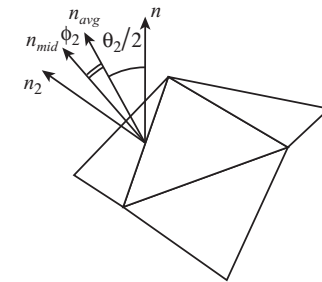The examples of [39] and [14] indicate that it may be impossible to design efficient and general local energy terms using positional degrees of freedom only, without either excluding certain local mesh configurations or treating them as special cases.

Midpoint normal energy. The observations about curvature approximations based on positions only leads us to considering additional degrees of freedom.

One way to introduce additional degrees of freedom to a mesh approximating a smooth surface is to assign independent unit normals to mesh points. Just as mesh vertices are regarded as samples of a continuous surface, one can regard normals as samples of surface normals; each normal represents up to two rotational degrees of freedom. However, we can choose normals for which one degree of freedom is fixed up to higher-order terms, by using normals assigned to edge midpoints.

Consider the plane spanned by the normal and the edge; then if we approximate the intersection of the surface with

this plane by a parabola, the normal should be perpendicular to the edge. As this is the best approximation to the surface we can hope for with the minimal number of degrees of freedom, and it is a second-order accurate approximation of the normal, we eliminate one degree of freedom for the normal by requiring it to be perpendicular to the edge. The remaining degree of freedom can be represented in a natural way by measuring angular deflection from the average normal of two adjacent triangles. Thus we introduce an extra angular variable for each edge. The mean curvature estimate in this case is quite similar to (16):

$$H_{MN}(T) =$$
$$(1/A)\left(l_1(\frac{\theta_1}{2} \pm \varphi_1) + l_2(\frac{\theta_2}{2} \pm \varphi_2) + l_3(\frac{\theta_3}{2} \pm \varphi_3)\right)$$
$$(19)$$

where φ_i are angular variables associated with edges and the plus or minus sign for each term is chosen according to an arbitrary fixed choice of positive direction for φ_i for each edge. We see that in this case additional variables can be regarded as corrections to the curvature estimates for each triangle edge.

Once this expression is linearized, it is possible to make a simple change of variables which results in elimination of dependence of the linearized energy on the points outside the triangle: the only degrees of freedom involved are the midpoint rotations and corners of the triangle itself, so we obtain a simple finite element for plates. The element obtained in this way is called the Morley element [13, 24], and is known to satisfy the patch test. Thus quite remarkably, it proves to be possible to obtain an element satisfying both conditions, with only six degrees of freedom per element in the linear case, if we introduce midpoint rotations.

Of course, this increases the total number of degrees of freedom for the mesh: assuming simple mesh topology, the number of edges is three times larger than the number of vertices; as there are 3 positional degrees of freedom for each vertex the total number of degrees of freedom approximately doubles. However, this appears to be one of the simplest possible discrete energies satisfying conditions C1 and C2 at the same time.

5. Conclusions

We conclude with several remarks on possible ways of applying the conditions on discrete energy functionals and how these conditions can be extended to other types of functionals. Violations of conditions C1 and C2 need to be considered in the context of a specific application. The absence of convergence in the limit for general meshes does not render a functional useless. For fixed-resolution computations on coarse meshes, interactive simulation or fairing of control meshes for splines or subdivision surfaces, convergent functionals may actually yield higher errors. The extent to which conditions C1 and C2 are violated for a given type of triangles (e.g. with an upper bound on aspect

ratio) can be used to estimate how far the solution energy may deviate from the continuum, and to what extent the energy is mesh-dependent. For example, one can observe (Figure 1) that for a quadric the maximum error for the one-dimensional functional with incorrect weights for local curvature estimates does not exceed 1.5%. This suggests that the same is true for any polygonal line, so if this error is acceptable, the functional is likely to be adequate for an application not requiring higher precision.

On the other hand, adaptive resolution computations may require a convergent method, as otherwise increasing resolution does not result in decreased errors; the mesh-dependent nature of the discrete functional makes it difficult to relate results for different meshes. Convergence may be essential if the surface is defined as a solution of a continuous energy minimization problem; in this case, using convergent approximations is the only way to obtain the true surface.

It is possible to extend geometric and convergence conditions to other types of functionals, using the same principles. The geometric invariance and scaling of the discrete functional should match that of the continuum functional. For convergence, one can consider action on explicit solutions of the linearized problem (typically the ones that keep the integrand constant) to obtain conditions on linearized discrete energy.

Finally, recent work [4] and the midpoint normal energy briefly described here, suggest that for general meshes using additional degrees of freedom, corresponding to higher-order continuous quantities (normals, curvatures etc), may lead to efficient and accurate discrete energy functionals for arbitrary meshes.

References

[1] D. Baraff and A. P. Witkin. Large steps in cloth simulation. In *Proceedings of SIGGRAPH 98*, Computer Graphics Proceedings, Annual Conference Series, pages 43–54, July 1998.

[2] R. Bridson, S. Marino, and R. Fedkiw. Simulation of clothing with folds and wrinkles. In D. Breen and M. Lin, editors, *Proceedings of ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, 2003.

[3] P. B. Canham. The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell. *J. Theor Biol.*, 36(1):68–81, Jan 1970.

[4] U. Clarenz, U. Diewald, G. Dziuk, M. Rumpf, and R. Rusu. A finite element method for surface restoration with smooth boundary conditions. *Comput. Aided Geom. Design*, 21(5):427–445, 2004. 0167-8396.

[5] D. Cohen-Steiner and J.-M. Morvan. Restricted delaunay triangulations and normal cycle. In *Proc. 19th Annu. ACM Sympos. Comput. Geom.*, pages 237–246, 2003.

[6] C. A. Felippa, B. Haugen, and C. Militello. From the individual element test to finite element templates: evolution of the patch test. *Internat. J. Numer. Methods Engrg.*, 38(2):199–229, 1995. ø 0029-5981.

[7] S. Germain. *Recherches sur la theorie de surfaces elastiques*. Mme. V. Courcier, Paris, 1821.

[8] G.Kirchhoff. Über das Gleichgewicht und die Bewegung einer elastichen Scheibe. *Crelles J.*, 40:51–88, 1850.

IEEE COMPUTER SOCIETY

[9] J. Goldfeather and V. Interrante. A novel cubic-order algorithm for approximating principal direction vectors. *ACM Transactions on Graphics*, 23(1):45–63, Jan. 2004.

[10] G. Greiner. Variational design and fairing of spline surfaces. *Computer Graphics Forum*, 13(3):143–154, 1994.

[11] E. Grinspun, M. Desbrun, A. Hirani, and P. Schröder. Discrete shells. In D. Breen and M. Lin, editors, *Proceedings of ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, 2003.

[12] W. Helfrich. Elastic properties of bilayers: Theory and possible experiments. *Z. Naturforsch.*, 28c:693–703, 1973.

[13] L. R. Hermann. Finite element bending analysis of plates. *J. Eng. Mech. Div.*, 93:13–25, 1967.

[14] K. Hildebrandt, K. Polthier, and M. Wardetzky. On the convergence of metric and geometric properties of polyhedral surfaces. Submitted for publication, 2005.

[15] L. Hsu, R. Kusner, and J. Sullivan. Minimizing the squared mean curvature integral for surfaces in space forms. *Experiment. Math.*, 1(3):191–207, 1992. 1058-6458.

[16] B. Irons and M. Loikkanen. An engineers' defence of the patch test. *Internat. J. Numer. Methods Engrg.*, 19(9):1391–1401, 1983. 0029-5981.

[17] B. M. Irons and A. Razzaque. Experience with the patch test for convergence of finite elements. *The mathematical foundations of the finite element method with applications to partial differential equations (Proc. Sympos., Univ. Maryland, Baltimore, Md., 1972)*, pages 557–587, 1972. Academic Press New York.

[18] T. Kármán. Festigkeistprobleme im Maschinenbau. In *Encyklopädie der Mathematischen Wissenschaften*, volume 4, pages 311–385. 1910.

[19] L. Kobbelt. Discrete fairing. *The mathematics of surfaces, VII (Dundee, 1996)*, pages 101–129, 1997. Info. Geom., Winchester.

[20] W. T. Koiter. A consistent first approximation in the general theory of thin elastic shells. In *Proc. Sympos. Thin Elastic Shells (Delft, 1959)*, pages 12–33. North-Holland, Amsterdam, 1960.

[21] A. E. H. Love. *A treatise on the Mathematical Theory of Elasticity*. Dover Publications, New York, 1944.

[22] A. E. H. Love. The small free vibrations and deformation of a thin elastic shell. *Phil. Trans.*, 179, 1988.

[23] M. Meyer, M. Desbrun, P. Schröder, and A. H. Barr. Discrete differential-geometry operators for triangulated 2-manifolds. In H.-C. Hege and K. Polthier, editors, *Visualization and Mathematics III*, pages 35–57. Springer-Verlag, Heidelberg, 2003.

[24] L. S. D. Morley. On the constant moment plate bending element. *Journal of Strain Analysis*, 6:10, 1971.

[25] R. A. Nay and S. Utku. An alternative to the finite element method. *Variational Methods Engineering*, 1, 1972.

[26] E. Oñate and M. Cervera. Derivation of thin plate bending elements with one degree of freedom per node. *Engineering Computations*, 10:543, 1993.

[27] E. Oñate and F. Zrate. Rotation-free triangular plate and shell elements. *Internat. J. Numer. Methods Engrg.*, 47(1-3):557–603, 2000. ñ á Richard H. Gallagher Memorial Issue 0029-5981.

[28] U. Pinkall and K. Polthier. Computing discrete minimal surfaces and their conjugates. *Experiment. Math.*, 2(1):15–36, 1993. 1058-6458.

[29] R. Schneider and L. Kobbelt. Geometric fairing of irregular meshes for free-form surface design. *Comput. Aided Geom. Design*, 18(4):359–379, 2001. 0167-8396.

[30] R. Schneider, L. Kobbelt, and H.-P. Seidel. Mesh fairing based on harmonic mean curvature surfaces. *Hierarchical and geometrical methods in scientific visualization*, pages 243–267, 2003. Math. Vis. Springer Berlin.

[31] G. Strang and G. J. Fix. *An analysis of the finite element method*. Prentice-Hall Inc., Englewood Cliffs, N. J., 1973. Prentice-Hall Series in Automatic Computation.

[32] F. Stummel. The generalized patch test. *SIAM J. Numer. Anal.*, 16(3):449–471, 1979. 0036-1429.

[33] D. Terzopoulos, J. Platt, A. Barr, and K. Fleischer. Elastically deformable models. In *Computer Graphics (Proceedings of SIGGRAPH 87)*, volume 21, pages 205–214, July 1987.

[34] M. Wang. On the necessity and sufficiency of the patch test for convergence of nonconforming finite elements. *SIAM J. Numer. Anal.*, 39(2):363–384 (electronic), 2001. 1095-7170.

[35] W. Welch and A. Witkin. Variational surface modeling. In *Computer Graphics (Proceedings of SIGGRAPH 92)*, volume 26, pages 157–166, July 1992.

[36] W. Welch and A. Witkin. Free-form shape design using triangulated surfaces. In *Proceedings of SIGGRAPH 94*, Computer Graphics Proceedings, Annual Conference Series, pages 247–256, July 1994.

[37] T. J. Willmore. Note on embedded surfaces. *An. Şti. Univ. "Al. I. Cuza" Iaşi Secţ. I a Mat. (N.S.)*, 11B:493–496, 1965.

[38] T. J. Willmore. *Riemannian geometry*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1993.

[39] G. Xu. Convergence of discrete Laplace-Beltrami operators over surfaces. *Comput. Math. Appl.*, 48(3-4):347–360, 2004.

[40] S. Yoshizawa and A. Belyaev. Fair triangle mesh generations via discrete elastica. In *Geometric Modeling and Processing*, Saitama, Japan, 2002.

[41] O. C. Zienkiewicz and R. C. Taylor. *The finite element method*. McGraw Hill, 1989. 1.

[42] D. Zorin and E. Grinspun. Computing discrete curvature-based energy. in preparation.

IEEE
COMPUTER
SOCIETY