Dynamic Simulation of Non-penetrating Flexible Bodies

David Baraff

Program of Computer Graphics
Cornell University
Ithaca, NY 14853

Andrew Witkin

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

A model for the dynamic simulation of flexible bodies subject to non-penetration constraints is presented. Flexible bodies are described in terms of global deformations of a rest shape. The dynamical behavior of these bodies that most closely matches the behavior of ideal continuum bodies is derived, and subsumes the results of earlier Lagrangian dynamics-based models. The dynamics derived for the flexible-body model allows the unification of previous work on flexible body simulation and previous work on non-penetrating rigid body simulation. The non-penetration constraints for a system of bodies that contact at multiple points are maintained by analytically calculated contact forces. An implementation for first- and second-order polynomially deformable bodies is described. The simulation of second-order or higher deformations currently involves a polyhedral boundary approximation for collision detection purposes.

1. Introduction

In this paper we present a new formulation for the dynamics of flexible bodies that covers collisions and continuous contact as well as free motion. The model, which draws on the flexible-body model proposed by Witkin and Welch[9] and on the analytical contact force model for rigid bodies presented by Baraff[1,2], centers on the idea that flexible body simulation can be greatly simplified through the introduction of a suitable geometric approximation. By restricting the body's changes of shape to those that can be represented by a global parametric deformation, we solve two problems that plague conventional local formulations: first, the dimensionality of the simulation is reduced, and second, the severe numerical problems to which local interactions can give rise are eliminated.

Because it is restricted by the geometric approximation, the flexible body's behavior generally exhibits error when compared to the behavior of an ideal continuum body under like conditions. Whether this approximation error is acceptable depends on factors such as the materials and forces being modeled, and the purpose of the simulation. In any case, once an approximation has been adopted it is clearly desirable to minimize the resulting error.

Our formulation is derived from the criterion of minimal approximation error. Following a brief discussion of local models and the problems they introduce, we will review the global model of Witkin and Welch, and show that their formulation for free body motion is in fact the minimal error solution. Then we will consider the problem of collision-response for flexible bodies. Collisions between rigid bodies are generally modeled as instantaneous events through the use of impulses. In the case of flexible bodies, this would be inappropriate, since we are interested in seeing them compress and rebound. The minimal error criterion leads us to a two-phase model in which an initial impulse prevents inter-penetration, followed by a non-impulsive phase of continuous contact in which the body compresses due to its own momentum, then rebounds due to the buildup of internal strain energy. Finally, we extend Baraff's analytic contact force model to flexible bodies, discuss implementation issues for polynomial deformations, and present simulation results.

This is an electronic reprint. Permission is granted to copy part or all of this paper for noncommercial use provided that the title and this copyright notice appear. This electronic reprint is ©1994 by CMU. The original printed paper is ©1992 by the ACM.

Author address (May 1994): David Baraff, School of Computer Science, Carnegie Mellon University, Pittburgh, PA 15213, USA.
Email: baraff@cs.cmu.edu, aw@cs.cmu.edu

1.1 Local models and the problem of stiffness

Traditional models for flexible bodies, including finite difference, finite element, and mass-and-spring lattice models, approximate the deviation of a continuum body from its rest shape in terms of displacements at a finite number of points called *nodal points*. Given a sufficient density of nodal points, these formulations can represent essentially any deformation. At the other end of the spectrum, continuum bodies whose deformations are considered small enough to be neglected can be approximated as rigid bodies, which are free only to translate and rotate.

Obviously, the rigid body formulation cannot be applied to bodies whose deformation we do not *want* to neglect. This is unfortunate because nodal formulations tend to give rise to stiff differential equations which are difficult to solve numerically. The basic problem is that nodal formulations model global phenomena—for example, acceleration of the whole body due to a point force—only via local interactions among adjacent nodes. The compression waves that translate these local interactions into large-scale effects involve deformations that are generally far too small and fast to play a significant role in computer animation, the more so as the mechanical stiffness of the flexible body is increased. Even so, if ordinary numerical methods are used, it is these high-speed effects rather than the phenomena of primary interest that dictate the size of the time steps, with potentially disastrous effects on performance[7]. Where it applies, the rigid body approximation sweeps these problems away by eliminating local interactions altogether: the only degrees of freedom a rigid body has are the global ones that govern its position and orientation.

Modeling of flexible body collisions and contact raises additional problems for the nodal formulation. In particular, collisions are difficult to handle because they usually involve extremely large transient forces and accelerations which, like any others, must operate strictly through local interactions. Hence the usual stiffness problem is exacerbated. Things are made still worse if penalty methods[1,3] (another local model) are used to enforce non-penetration constraints, since these introduce stiffness problems of their own.

1.2 Global models

Witkin and Welch[9] present a flexible-body model that represents a compromise between the extremes of the nodal and rigid formulations. Changes in the body's shape are approximated neither by nodal displacements nor by rigid body transformations, but by global deformations that apply a parametric "space warp" to all points in the body. This global formulation constrains the range of deformations that the body can undergo to those that can be represented by the chosen deformation function. However, because the shape parameters are global in their effect, the stiffness problems due to local interactions are eliminated, as in the rigid case. In a different flexible-body formulation, Pentland and Williams[6] achieve much the same goal by linearizing the dynamic model, then using modal analysis to eliminate high-frequency vibrations. We prefer the formulation used in Witkin and Welch primarily because, as we will see in the next section, it minimizes the error due to the geometric approximation, whereas the consequences of the dynamics-based approximations of Pentland and Williams are more difficult to assess.

1.3 Collision and contact

Baraff[1,2] presents an analytical method for computing contact forces between configurations of rigid bodies. Contact forces are computed by solving simultaneous equations that reflect all of the forces acting on all of the bodies. Thus, an external force acting on one body has an immediate effect on any contacting bodies. The formulation of the contact forces as a global problem eliminates the stiffness problems encountered by the penalty method for contact forces. To complete our global formulation, we will adapt this method to deal with flexible bodies.

2. Global Deformations

In this section, we will define the basic global formulation for flexible bodies used throughout this paper. As stated in the introduction, rather than describe the shape of a deformed body in terms of some number of nodal points, a body's shape is described in terms of a global deformation function D_q, where q is a vector of parameters that controls the deformation. The function D_q maps $\mathbf{R}^3$ onto itself; D_q describes the deformed shape of a body by mapping each point p of the body's rest shape to the point $D_q(p)$.

2.1 The dynamics of global deformations

Our goal in deriving the dynamical behavior of our flexible-body model is to relate the forces acting on a body to the acceleration of the control parameters q, in a generalization of the familiar equation $f = ma$. That is, given the control parameters q, and the first derivative $\dot{q}$, we want to express $\ddot{q}$ in terms of q, $\dot{q}$, and any forces acting on the body. We want the motion prescribed by this $\ddot{q}$ to have the minimum possible deviation from the motion of an ideal continuum body. The expression for $\ddot{q}$ derived in this paper that minimizes the deviation matches the expression derived by the earlier work of Witkin and Welch[9] using Lagrangian dynamics; thus, we show that the Lagrangian dynamics formulation satisfies our minimal approximation error criterion. The model of collision-response derived in section 3.2 is also based on a minimum error criterion.

In the derivations in this paper, bodies are parameterized by a coordinate p in body space, which ranges over some fixed volume. The density of the body at any point p in body space is denoted as $\rho(p)$. If we let $q(t)$ describe the *state* of the body as a function of time, then at time t, $D_{q(t)}$ specifies how the body is mapped from body space into world space. Specifically, at time t the deformed body has density $\rho(p)$ at the world space point $D_{q(t)}(p)$.

In an ideal continuum body, the acceleration at any point is given by the standard equation $f = ma$. An ideal continuum body is therefore completely unconstrained in its motion, and can be deformed into any arbitrary shape, by appropriate forces. In contrast, global deformations represent a constraint, limiting the allowable deformations, and thus the motion, of a body. Because of this constraint, the acceleration a of a point with mass m, in response to a force f, will not in general satisfy $a = f/m$. The best we can do is to minimize the deviation of the actual acceleration a, from the ideal acceleration f/m. To measure the total acceleration deviation for the entire body, we will integrate the acceleration deviation at each point over the entire volume of the body.

Let the net force in world space acting on a body be described as a vector-valued function $f(p)$ over the body. The force function f includes external forces such as gravity, internal forces due to deformation, and forces due to contact. If we let $a(p)$ denote the acceleration at each point of the body, then $a(p)$ would be "ideal" if it always satisfied $a(p) = f(p)/\rho(p)$, or equivalently, $a(p) - f(p)/\rho(p) = 0$. We will measure the net acceleration error E from this ideal acceleration for the entire body by writing

$$E = \frac{1}{2}\int \rho(p)|a(p) - f(p)/\rho(p)|^2\,dV \tag{1}$$

where p ranges over the volume of the body in world space. The deviation is mass-weighted, since the error contributed by an acceleration deviation in some region of the body should grow linearly with the density in that region. Using E, we can relate the acceleration of the control parameters to the net force function f by requiring that $\ddot{q}$ be chosen so as to minimize E.

However, the error E defined by equation (1) is the same as a quantity named by Gauss as the "constraint" of a system. Gauss formulated a principle called the *principle of least constraint* that asserts that the motion of a system subject to constraints always minimizes the "constraint", namely equation (1). The principle of least constraint yields the same result for constrained motion as is given by Lagrange's equation of motion[5]; thus, our notion of motion satisfying a minimal error criterion is equivalent to treating bodies as mechanical systems and defining their motion in terms of the very well known Lagrangian dynamics.

2.2 Linear deformations

The previous section describes the dynamics when the deformation function $D_{q(t)}$ is arbitrary. For the remainder of this paper, we will limit ourselves to deformations $D_{q(t)}(p)$ that depend linearly on q, but may vary non-linearly with respect to p. Rather than describe the state in terms of a vector $q(t)$, we will switch notation and define the state as a matrix $R(t)$; correspondingly, $D_{R(t)}$ denotes the deformation function specified by $R(t)$. The fact that the deformation $D_{R(t)}$ is linear with respect to $R(t)$ means that $D_{R(t)}$ must have the form

$$D_{R(t)}(p) = R(t)Z(p) \tag{2}$$

where $Z(p)$ is a vector-valued function that does not depend on either t, or R. This restriction greatly simplifies some of the dynamics equations, while still allowing much latitude in the choice of the deformation function D. Note the $Z(p)$ need not be a linear function of p. A deformation that is quadratic in the undeformed coordinates could be specified by

$$Z(p) = [p_x^2, p_y^2, p_z^2, p_xp_y, p_xp_z, p_yp_z, p_x, p_y, p_z, 1]^T \tag{3}$$

with $R(t)$ a 3×10 matrix. Clearly, polynomial deformations of any order can be expressed in terms of deformations that are linear with respect to the state. Second-order polynomial deformations in particular are fairly liberal in terms of the allowable deformations of a body. Section 4 will discuss the details of implementing first- and second-order polynomial deformation functions.

From the previous section, we know that we can use Lagrangian dynamics to relate $\ddot{R}(t)$ to $R(t)$, $\dot{R}(t)$, and the force $f(p)$ with minimal error. Witkin and Welch[9] derive the relation

$$\ddot{R}(t) = Q(t)M^{-1} \tag{4}$$

where M is a constant square matrix and $Q(t)$ is a matrix of the generalized net force acting on the body. The matrix M is defined by

$$M = \int \rho(p)Z(p)Z(p)^T\, dV \tag{5}$$

which yields a square matrix of size n, where n is the dimension of the column vector $Z(p)$. Additionally, M is both symmetric and positive definite. The matrix $Q(t)$ has dimension $3 \times n$. Generalized force Q is related to force in $\mathbf{R}^3$ as follows: a force f in $\mathbf{R}^3$ acting on the point $D_{R(t)}(p)$ in world space yields the matrix of generalized force $fZ(p)^T$. Note that both f and $Z(p)$ are column vectors, so that the outer product $fZ(p)^T$ yields a $3 \times n$ matrix.

For bodies with a complicated rest shape, the integral of equation (5) is difficult to calculate analytically. Since M is constant however, it can be calculated off-line of the simulation process by numerical techniques; currently, we use Monte-Carlo integration to precompute M. Thus, all that is required to simulate the motion of these flexible bodies is the solution of equation (4). This is done by converting equation (4) to the coupled first-order ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{dt}} \begin{pmatrix} R(t) \\ \dot{R}(t) \end{pmatrix} = \begin{pmatrix} \dot{R}(t) \\ Q(t)M^{-1} \end{pmatrix}. \tag{6}$$

The remaining task is to evaluate the generalized force $Q(t)$, which subsumes external forces, internal forces, and any contact forces acting on the body. External forces are by definition forces that are known to us at time t, such as gravity, or viscous damping. In the next section, we show how internal forces due to deformation may be calculated. Section 3 discusses the computation of the contact forces included in $Q(t)$.

2.3 Potential energy functions of linear deformations

The internal forces that act to restore a flexible body to its rest shape are specified in terms of the derivative $\partial V/\partial R$ of a potential energy function V[9]. For most physical bodies, potential energy functions can be described in terms of the *metric tensor* of a body. (Internal damping forces are calculated similarly in terms of the metric tensor and its time derivative.) For affine deformations, the metric tensor (and its time derivative), at any instant of time, are constant over the volume of a body; thus, evaluating integrals of the metric tensor over the body volume is trivial. However, for non-affine deformations, the metric tensor can vary over the volume of a body; if the body's shape is complex, the integrals needed to compute $\partial V/\partial R$ can be difficult to evaluate.

Our choices are to either calculate the integrals analytically, or use a numerical approximation. For polynomial deformations, integrals involving the metric tensor can be expressed as a linear combination of integrals that are independent of $R(t)$. These integrals are weighted by functions of $R(t)$ and summed to compute $\partial V/\partial R$. Internal damping forces are computed similarly as a linear combination of precomputed integrals, weighted by functions of both $R(t)$ and $\dot{R}(t)$. However, even for the second-order polynomial deformations we have implemented, the necessary expansion of precalculated integrals is quite large. An alternate technique is to simply approximate the integrals by evaluating the integrand at some finite number of points scattered throughout the body. This has the virtue of being trivial to implement, and may also be more general in dealing with complex energy functions that allow for plasticity or fracture[8]. For second-order polynomial deformations, a small number of sample points (on the order of fifty) yields adequate results.

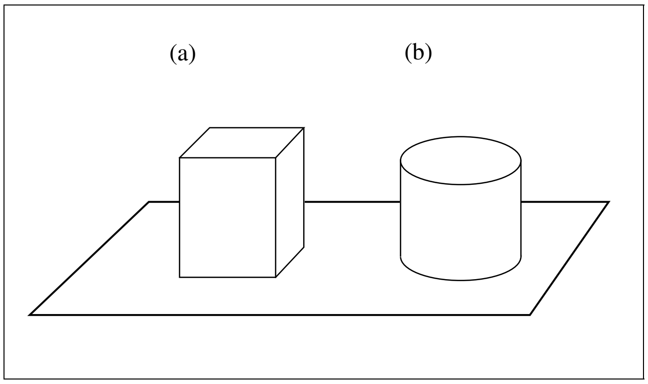

Figure 1: (a) Non-penetration is enforced by preventing only the four contacting vertices from moving below the plane. (b) The entire lower rim of the cylinder must be prevented from moving below the plane.

3. Non-penetration Constraints

The derivation for the contact forces between flexible bodies that prevent inter-penetration closely parallels the derivation of contact forces between rigid bodies in Baraff[1,2]. The most notable difference arises when considering collisions between bodies, and is detailed in section 3.2.

3.1 Contact geometry restrictions

In considering contact between bodies, we will make the assumption that contact between bodies can be described in terms of finitely many contact points. That is, we will restrict ourselves to configurations in which inter-penetration can be prevented globally by enforcing a finite number of local constraints. As an example, consider figure 1. If all the bodies involved are rigid, then in figure 1a, inter-penetration is prevented globally between the cube and the plane by enforcing the four local constraints that each vertex of the cube in contact with the plane remain on or above the plane. However, to prevent the cylinder in figure 1b from dipping below the plane and inter-penetrating, while still allowing it to tip over arbitrarily, we must constrain *each* of the infinitely many boundary points on the lower cylinder to remain on or above the plane.

Thus, we do not allow configurations such as figure 1b. For bodies whose deformations are affine functions of the material coordinates (as well as rigid bodies), polyhedral contact regions are allowed (figure 2a). However, for all other deformations, configurations such as figure 1a cannot be allowed. A deformation which caused the contact face of the cube to curve could result in inter-penetration even if the vertices of the face were prevented from inter-penetrating the plane (figure 2b). Section 4.2 describes a geometric polyhedral boundary discretization that permits the simulation of configurations with one- or two-dimensional polygonal contact regions when arbitrary deformation functions are allowed. This approximation is somewhat unsatisfactory; a non-discretized method for dealing with situations like figure 2b would be preferable.

3.2 Colliding contact

Having described the geometry of contact, we can now consider the dynamics of contact. When bodies initially come into contact at a point p_c, we say a collision has occurred. In order to maintain

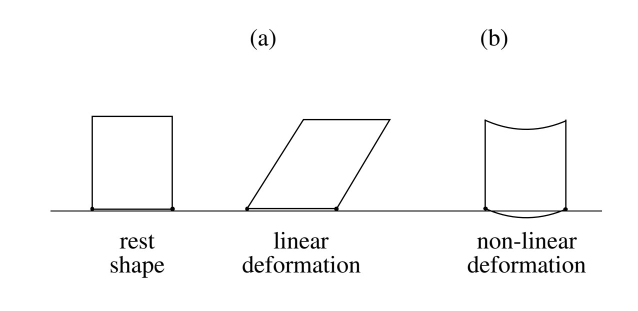

Figure 2: (a) Affine deformations allow polyhedral contact regions, because inter-penetration can be prevented with finitely many local constraints. (b) For non-linear deformations, preventing the vertices from moving below the plane does not necessarily prevent inter-penetration.

the non-penetration constraint between the two bodies, at least one of the two contact surfaces must undergo a velocity discontinuity at the contact point. Otherwise, some amount of inter-penetration would occur in the vicinity of the contact point. Since rigid bodies cannot deform at all, in the case of a collision between two rigid bodies, every point on the rigid body experiences a velocity discontinuity. This implies an abrupt change in momentum, which can only be accomplished by an impulse. Collisions can be quantified in terms of the energy lost: a collision can be completely elastic, meaning no energy is lost, or completely inelastic, meaning that the kinetic energy of both bodies is completely dissipated (measured with respect to a center of mass coordinate system). For rigid body collisions, this energy loss is instantaneous. Immediately after a collision, bodies bounce apart with a velocity that depends on the elasticity of the collision.

In contrast, the collision process for flexible bodies takes some non-zero amount of time. During that time, the colliding bodies remain in contact with each other. To derive a collision-response model for our flexible bodies, we will use a variational principle. To motivate such a derivation, let us consider the case of a one-dimensional rod colliding with a fixed obstacle. The rod is discretized into n mass points. Since this example is one-dimensional, each mass point may undergo displacement either left or right. The mass points are numbered left to right, from 1 to n (figure 3).

When the first mass point collides with the fixed obstacle, its velocity must discontinuously change to zero, to prevent inter-penetration. However, the velocity of the other mass points of the rod are unchanged. In particular, as the second mass point continues with its original velocity, the distance between the first and second mass point decreases. As this distance decreases, an internal force acts to repel the two mass points. This has the effect of deaccelerating the second mass point, which means that the distance between the second and third mass point decreases; clearly, this effect propagates throughout the entire rod. While this is happening, the fixed obstacle has been exerting a (non-impulsive) contact force on the first mass point, to prevent the repulsive internal force from accelerating the first mass point leftwards. After some finite period of contact, the internal forces cause the rod to bounce away from the obstacle. If there is no damping, the only energy lost will be due to the dissipation of the kinetic energy of the first mass point. However, in the limit as n goes to infinity, the mass of this point, and thus its initial kinetic energy, both go to zero. Even though a velocity discontinuity occurs (at the left end of the bar), the actual change in momentum is zero, and thus no energy is lost. If we wish to model collisions with some amount of inelasticity, we must impose damping forces on the body that dissipate energy during the collision.

The collision-response model for our constrained flexible bodies

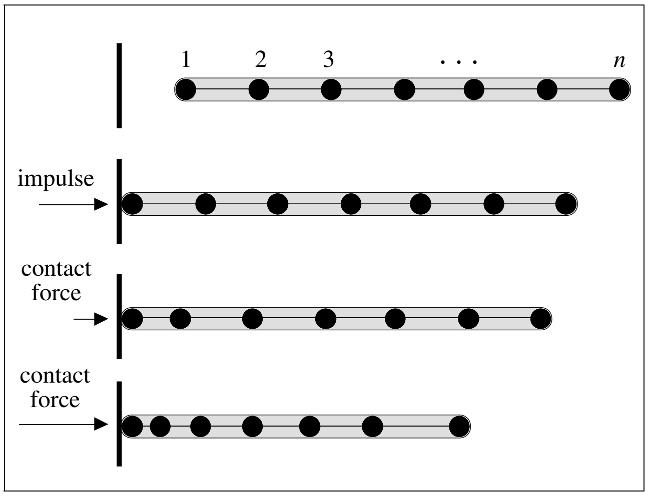

Figure 3: A discretized rod collides with an immovable obstacle. Mass 1's velocity changes instantaneously to zero upon impacting the obstacle. As mass 2 approaches mass 1, an internal force between the two acts, pushing mass 1 leftwards. A contact force acts rightwards on mass 1 to prevent inter-penetration.

is based on this analysis. As in the rigid body case, we will need to apply an impulse at the contact point to prevent immediate inter-penetration. This means that some kinetic energy must be lost, and we cannot attain a perfectly elastic collision. This is a consequence of limiting the allowable deformations of the body. After the impulse has been applied, the bodies will no longer be colliding, and a non-impulsive force (described in the next section) will act at the contact point to prevent inter-penetration.[1]

As in section 2.1, we use an error measure to compare the collision-response of our flexible-body model with the collision-response of an ideal continuum body. To measure the net collision-response, we generalize equation (1) to obtain a new error measure E' given by

$$E' = \tfrac{1}{2}\int \rho(p)|\Delta v(p)|^2\, dV \tag{7}$$

where $\Delta v(p)$ is the change in velocity in response to an impulse. For an ideal continuum body, $\Delta v(p)$ is zero everywhere except at the point of collision and E' is zero. For our flexible-body model, E' measures the deviation between the actual change in velocity, $\Delta v(p)$, and the ideal change in velocity, which is zero. The correct instantaneous change $\Delta\dot{R}(t)$ in the state velocity $\dot{R}(t)$ is the one which minimizes E', subject to the kinematic constraints of the collision.

From the description of the collision in figure 3, it seems intuitive that the correct course of action is to apply an impulse between two colliding bodies at the point of contact, such that they just come to rest (relative to each other) at the contact point. In the absence of friction, the impulse should act normal to the contact surfaces. A standard constrained-minimization principle applied to equation (7) shows that the change in velocity $\Delta\dot{R}(t)$ from such an impulse is in fact exactly the $\Delta\dot{R}(t)$ which minimizes E', subject to the kinematic constraints of the collision.

The actual computation of the impulse is trivial. For simplicity, let us consider a collision between a flexible body A and an immovable, undeformable obstacle B with a well defined surface normal

[1]In general, the energy loss can be decreased by altering the deformation function D so that it allows the body more degrees of freedom. The limiting case, when D can represent any deformation, is the same as the limiting case for the discretized bar of figure 3 when n goes to infinity. Obviously, neither case can be simulated without an infinite amount of computation.

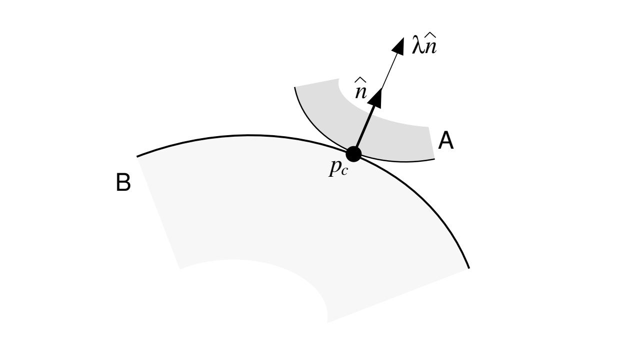

Figure 4: Impact between a flexible body A and an immovable undeformable body B.

at the contact point p_c. The vector $\hat{n}$ denotes the outwards pointing unit surface normal of B at the contact point p_c (figure 4). The generalization to the case when B is an ordinary rigid or deformable body is straightforward. As is the case for generalized forces, an impulse f in $\mathbf{R}^3$ acting on a body at the point $D_{R(t)}(p)$ produces a generalized impulse $Q = fZ(p)^T$. A flexible body's state velocity $\dot{R}(t)$ changes discontinuously to $\dot{R}(t) + QM^{-1}$ when subject to a generalized impulse Q.

To calculate the required impulse, let $p_c = D_{R(t)}(p)$. We know that the impulse occurs at p_c, and has direction $\hat{n}$ for a frictionless impact. Thus, we can express the impulse as $\lambda\hat{n}$, where λ is an unknown scalar. The velocity of the contact point on the moving body before the collision is

$$\frac{\mathrm{d}}{\mathrm{dt}} D_{R(t)}(p) = \dot{R}(t)Z(p). \tag{8}$$

After the collision, the velocity of the contact point is $(\dot{R}(t) + QM^{-1})Z(p)$. Since body B is fixed, the velocity of the contact point in the $\hat{n}$ direction after the collision must be zero; that is, we require

$$\hat{n} \cdot (\dot{R}(t) + QM^{-1})Z(p) = 0. \tag{9}$$

Substituting $Q = (\lambda\hat{n})Z(p)^T$, we have the constraint

$$\hat{n} \cdot (\dot{R}(t) + \lambda\hat{n}Z(p)^T M^{-1})Z(p) = 0 \tag{10}$$

which yields

$$\lambda = \frac{-\hat{n} \cdot \dot{R}(t)Z(p)}{\hat{n} \cdot \hat{n}Z(p)^T M^{-1} Z(p)} = \frac{-\hat{n} \cdot \dot{R}(t)Z(p)}{Z(p)^T M^{-1} Z(p)} \tag{11}$$

since $\hat{n}$ has unit length. Since M is positive definite, the denominator of equation (11) is non-zero and positive. Moreover, since $\hat{n} \cdot \dot{R}(t)Z(p)$ is the initial approach speed in the normal direction (which is negative), λ is positive as one would expect from figure 4. For collisions involving more than one contact point, the collisions at the contact point can be considered as a sequence of collisions, slightly staggered in time. If collisions involving multiple contact points are modeled as occurring simultaneously, the approach taken by Baraff[1] can be used.

3.3 Resting contact

The derivation for the resting contact forces between flexible bodies is almost the same as the derivation for resting contact forces between rigid bodies. The non-penetration constraints for rigid bodies described in Baraff[2] are restricted to handle situations with only finitely many contact points, as described in section 3.1. Since we are restricted to situations in which inter-penetration can be prevented by considering only finitely many contact points, we can index the contact points from 1 to n. At each contact point between two bodies A and B, we will write down a constraint of the form

$$\ddot{\chi}_i(t_0) \geq 0 \tag{12}$$

where $\chi_i(t)$ is a measure of the separation between A and B in the normal direction at time t. Since $\chi_i(t)$ is a spatial measure, $\ddot{\chi}_i(t)$ measures the relative normal acceleration between A and B. In particular, if $\ddot{\chi}_i(t_0) < 0$, then the bodies are accelerating so as to inter-penetrate at the ith contact point. Conversely, if $\ddot{\chi}_i(t_0) > 0$, then the bodies are accelerating apart, and contact will be broken at the ith contact point immediately after time t_0. If $\ddot{\chi}_i(t_0) = 0$, then contact is not broken at the ith contact point. Thus, to prevent inter-penetration, we must enforce equation (12).

The relation $\ddot{\chi}_i(t) \geq 0$ is maintained at each contact point by a time-varying contact force, acting normal to the contact surface. As in the case of colliding contact, we need to calculate the magnitudes of these contact forces. Because $\ddot{\chi}_i(t_0)$ measures acceleration, $\ddot{\chi}_i(t_0)$ depends linearly upon all the forces acting on bodies A and B. The contact force at each contact point is required to be repulsive and conservative; that is, it must not add energy to the system of bodies. Since the normal force at one contact may affect the acceleration of one or both of the bodies at another contact, satisfying $\ddot{\chi}_i(t) \geq 0$ at all of the contact points involves satisfying a system of simultaneous linear inequalities. The constraint that the contact forces act conservatively can be expressed in terms of a quadratic constraint on the contact forces. Contact forces satisfying these constraints can be computed by quadratic programming[1,2].

Exactly the same formulation is used to prevent inter-penetration between flexible bodies. The expression derived for $\ddot{\chi}_i(t_0)$ in Baraff[2] requires the spatial and temporal derivatives of functions describing the contact surfaces. Suppose we express the undeformed rest shape of our contact surface in body space in terms of a real-valued function $F_0(p)$; that is, a point p_b in body space is on the undeformed surface if and only if $F_0(p_b) = 0$. Then the deformed contact surface at time t consists of those points p in world space for which

$$F(p,t) = F_0(D^{-1}_{R(t)}(p)) = 0. \tag{13}$$

Deriving expressions for the various derivatives of $F(p,t)$ necessary to symbolically evaluate $\ddot{\chi}_i(t_0)$ becomes mostly an exercise in applying the chain rule of calculus. For any deformation function $D_{R(t)}(p) = R(t)Z(p)$, $\ddot{\chi}_i(t_0)$ is a linear function of $\ddot{R}(t_0)$, so $\ddot{\chi}_i(t_0)$ depends linearly upon the contact forces. The contact forces can be extended to include friction as described by Baraff[3].

4. Implementing Polynomial Deformations

We have implemented flexible bodies for the cases of first- and second-order polynomial deformations. There is no difficulty in performing simulations that involve bodies with a mix of differing deformation functions. Simulations can also mix rigid and deformable bodies. In this section, we will discuss a number of implementation details.

4.1 Collision detection

Currently, our bodies are limited to unions of convex primitives, where a primitive is either a polyhedron or a convex closed curved surface. When deformations are limited to first-order deformations, convexity is conserved under deformation. Thus, the collision detection method described in Baraff[2] can be used without alteration. Additionally, when flexible bodies are determined to contact at some point p_c in world space, it is necessary to compute p_c in the body space of both of the bodies; that is, $D^{-1}_{R(t)}(p_c)$ must be computed for each body. This is trivial for first-order deformations since $D_{R(t)}$ is simply an affine transformation whose inverse is easily computed.

Since polygons in body space are mapped into polygons in world space, one- or two-dimensional contact regions are allowed, as described in section 3.1.

Matters are more complicated for second-order or higher deformations. Computing $D^{-1}_{R(t)}(p_c)$ is not a severe problem. A coarse mesh of control points with known body space coordinates can be transformed into world space for each body. Given a point p_c in world space, the body space coordinates of the control points closest to p_c in world space can be interpolated to provide a rough estimate of $D^{-1}_{R(t)}(p_c)$. Starting with this estimate, $D^{-1}_{R(t)}(p_c)$ can be computed numerically using iterative techniques[4]. However, the initial determination of the contact points in world space is a difficult problem. Even given a suitable collision detection algorithm, our simulations would be severely restricted because of the contact geometry restrictions of section 3.1. It is our hope to eventually deal with these problems; for now however, we will describe an approximation method that removes the contact geometry restriction and lets us use previously developed collision detection algorithms.

4.2 Polyhedral approximation

Our approximation method consists of discretizing contact surfaces into a triangular mesh. The undeformed contact surface of the body is decomposed into some number of triangular patches that completely cover the surface. In body space, a given triangular patch can be described as the triple of vertices (p_0, p_1, p_2). At time t, this triangle is transformed into world space as the triangle $(D_{R(t)}(p_0), D_{R(t)}(p_1), D_{R(t)}(p_2))$. When performing collision detection and enforcing non-penetration constraints, the contact surface is considered to be the collection of these transformed triangles. The use of a coherence based culling step results in a collision detection algorithm that is nearly linear in the number of polygons[3].

The non-penetration constraints are written in terms of the deformed triangles. Each triangle is treated as a plane in solving for contact forces. Since the plane equation for each triangle can be written in terms of the vertices, the derivatives needed for section 3.3 can be computed in terms of the derivatives of $D_{R(t)}(p_0)$, $D_{R(t)}(p_1)$ and $D_{R(t)}(p_2)$. Since the contact surfaces always remain planar, the contact geometry restriction of finitely many contact points is not an issue. As in the first-order case, the non-penetration constraint can be formulated in terms of finitely many constraints, even if one- or two-dimensional contact regions result.

If the results of a simulation are displayed using the actual deformed shape of a body, instead of the polyhedral approximation used for the dynamics computations, visual anomalies can occur. If the discretization of the polyhedral mesh is very low compared to the curvature of the body, bodies may appear to inter-penetrate somewhat. We have found however that a fairly coarse mesh produces quite reasonable results for second-order deformations. Presumably, higher-order deformations would require finer meshes to avoid visual artifacts. Figure 5 shows a deformed rectangular block; for display purposes, the block was meshed sufficiently to appear as a smooth curved surface. For simulation purposes, the (deformed) block was subdivided into a $3 \times 3 \times 5$ cubic mesh, after which each exterior square face was split into two triangles (for a total of 156 triangles). This was sufficient to remove any suggestion of inter-penetration throughout the entire simulation.

5. Conclusions

We have demonstrated a new technique for simulating flexible bodies subject to non-penetration constraints. We have implemented both first- and second-order polynomial deformable bodies. The use of second-order deformations requires a polyhedral approximation for collision detection purposes. The model presented unifies previous work on deformable bodies, non-penetration constraints, and friction.

Acknowledgements

Research at Cornell University was supported in part by an AT&T Bell Laboratories PhD Fellowship, and grants from the NSF. Research at Carnegie Mellon University was supported in part by a grant from Apple Computer, Inc. Simulations were performed on equipment generously donated by the Hewlett Packard Corporation and the Digital Equipment Corporation. We would like to thank Gene Baraff for several helpful conversations concerning variational principles.

References

[1] D. Baraff. Analytical methods for dynamic simulation of non-penetrating rigid bodies. In *Computer Graphics* (*Proc. SIGGRAPH*), volume 23, pages 223–232. ACM, July 1989.

[2] D. Baraff. Curved surfaces and coherence for non-penetrating rigid body simulation. In *Computer Graphics* (*Proc. SIGGRAPH*), volume 24, pages 19–28. ACM, August 1990.

[3] D. Baraff. *Dynamic Simulation of Non-penetrating Rigid Bodies*. PhD thesis, Cornell University, May 1992.

[4] J.E. Dennis, Jr. and R.B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice Hall, Inc., 1983.

[5] C. Lanczos. *The Variational Principles of Mechanics*. Dover Publications, Inc., 1970.

[6] A. Pentland and J. Williams. Good vibrations: Modal dynamics for graphics and animation. In *Computer Graphics* (*Proc. SIGGRAPH*), volume 23, pages 215–222. ACM, July 1989.

[7] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes*. Cambridge University Press, 1986.

[8] D. Terzopoulos and K. Fleischer. Modeling inelastic deformation: Viscoelasticity, plasticity, fracture. In *Computer Graphics* (*Proc. SIGGRAPH*), volume 22, pages 269–278. ACM, August 1988.

[9] A. Witkin and W. Welch. Fast animation and control of nonrigid structures. In *Computer Graphics* (*Proc. SIGGRAPH*), volume 24, pages 243–252. ACM, August 1990.

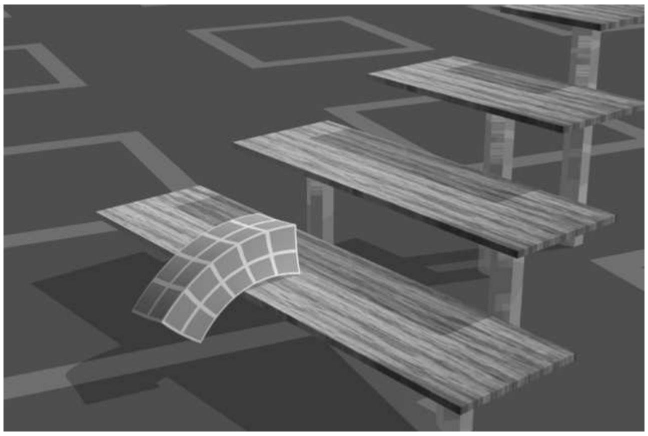

Figure 5: Quadratically deformable block falling down stairs.