TRITA-NA -8214
Inst för Numerisk analys och datalogi
KTH
100 44 STOCKHOLM 70

Dept of Numerical Analysis and
Computing Science
The Royal Institute of Technology
S-100 44 STOCKHOLM 70, Sweden

NUMERICAL SIMULATION OF TIME-DEPENDENT CONTACT AND FRICTION PROBLEMS IN RIGID BODY MECHANICS.

by
Per Lötstedt

TRITA-NA-8214

This work has obtained financial support
from The National Swedish Board for
Technical Development under contract
dnr 80-3341.

Abstract.

A numerical method for solution of the system of ordinary differential equations and algebraic, unilateral constraints governing the motion of a mechanical system of rigid bodies, where contacts between the bodies are created and disappear in the time interval of interest. The ordinary differential equations are discretized by linear multistep methods. In order to satisfy the constraints, a quadratic programming problem is solved at each time-step. The fact that the variation of the objective function is small from step to step is utilized to save computing time. A discrete friction model, based on Coulomb's law of friction and suitable for efficient computation, is proposed for planar problems where dry friction cannot be neglected. The normal forces and the friction forces are the optimal solution to a quadratic programming problem. The methods are tested on four model problems. A data-structure and possible generalizations are discussed.

1. Introduction.

In many engineering applications there is an interest in the simulation on computers of the large scale motion of mechanical systems. These systems are often well approximated by a number of rigid bodies interacting with each other. The bodies may be interconnected by different kinds of joints, e.g. ball-and-socket joints or hinge joints, or in contact with each other at a point, along a curve or on a surface. The reason to choose a rigid body model instead of a more elaborate and perhaps accurate elastic or elastic-plastic model is that the governing equations are a system of ordinary differential equations and the number of degrees of freedom is much lower after discretization. Examples of areas of application where rigid body systems have been simulated are

(i) mechanisms and machinery, [33], [34], [38],

(ii) satellites, [12], [37],

(iii) biomechanics, the human body is represented by 10-15 rigid bodies in sports activities [36], vehicle crash simulations [20],

(iv) vehicles on the road [11] and on rails [24],

(v) rock mechanics, the motion of assemblages of rock blocks [7],

(vi) building construction, the collapse of a concrete building [28].

In the first four examples the joints between the bodies are assumed to be permanent during the time-interval of interest. Friction is usually ignored in the first three categories. These systems are simulated in both two and three space dimensions. In the examples (v) and (vi) it is of importance to use a realistic model for the creation and disappearance of contacts between the bodies and the friction between bodies sliding on each other. Due to the complicated geometry of the bodies in three space dimensions these systems have so far only been treated as planar systems.

The standard numerical methods for the solution of ordinary differential equations can often be applied directly to the systems in (i)-(iv), possibly with a modification to allow for algebraic constraints, Gear [15,p.226]. There are methods in analytical mechanics to reduce the dimension of the system of ordinary differential equations to the degree of freedom of the mechanical system, see Paul [34]. This technique is used in Sheth and Uicker [38] while in Orlandea and Calahan [33] extra variables, Lagrange multipliers, and associated algebraic constraints are introduced. For problems where the number of contacts is small, say 1-3, it is possible to reduce the size of the differential equation system as described above and to reformulate the system at each time a new contact is established or an old contact disappears. For mechanical systems with several contacts and dry friction the Lagrange multiplier approach is much more flexible and efficient. Furthermore, the multipliers are directly proportional to the contact and friction forces.

Coulomb's law is a simple and common model for dry friction between sliding bodies. In this paper we develop a numerical method for the simulation of time-dependent contact and Coulomb friction problems for rigid body systems based on a Lagrange multiplier approach. The method was outlined in Lötstedt [27]. An improved version is here analyzed and tested on model problems inspired by rock mechanical applications. Cundall [7], in his simulation of rock blocks, introduces springs and dampers at the points of contact to prevent the blocks from penetrating each other. In our method the Lagrange multipliers serve the same purpose. Cundall's idea is related to the penalty function method in nonlinear optimization and is analyzed by Lötstedt [26].

The analysis of two deformable bodies in contact is reviewed in Kalker [21] and methods of solution are indicated in Cottle [6]. Numerical methods for the elastostatical contact problem have been devised e.g. by Haug, Chand and Pan [19] for many bodies without friction and by Fredriksson [14] for two bodies with friction.

In the next section the system of equations and inequalities satisfied by the coordinates of the bodies and the Lagrange multipliers is formulated. Friction free contact problems are treated in §3. The system of ordinary differential equations is solved by a linear multistep method and the Lagrange multipliers are the optimal solution to a quadratic programming (QP) problem. A discrete interpretation of Coulomb´s law of friction is made in §4. The friction forces are also the solution to a QP problem. The performance of the method described in §3 and §4 is illustrated in four examples. In §5 a datastructure for the representation of the mechanical system in the computer is proposed. Possible generalizations are discussed in the final section.

2. The governing equations.

The governing equations for planar contact and Coulomb friction problems are derived in Lötstedt [29], [30]. The equations are stated and explained here for completeness and the notation used throughout the paper is introduced. The necessary modifications in three space dimensions are discussed in §6.

The system of ordinary differential equations satisfied by $q(t) \in R^m$, the vector of the coordinates of the bodies, is for $t \geq 0$

$$M\ddot{q} = f(q,\dot{q},t) + G(q)\lambda(t). \tag{2.1}$$

The position of a body is represented in q by the Cartesian coordinates (x,y) of its center of gravity and an angle of direction φ. $M \in \mathbb{R}^{m\times m}$ is the mass matrix, a constant, diagonal matrix with positive diagonal elements. M has the factorization

$$M = D^T D. \tag{2.2}$$

$f \in \mathbb{R}^m$ contains the explicity given driving forces and $G\lambda$ represents the contribution of the contact and friction forces to the equations of motion. λ is the Lagrange multiplier vector which is determined such that certain algebraic constraints and inequalities are fulfilled.

Let the components of $\phi(q) \in \mathbb{R}^p$ be the holonomic, unilateral constraint functions. ϕ_j may e.g. be the perpendicular distance between the corner of one body and the edge of another body. The unilateral algebraic constraints in the mechanical system are

$$\phi(q)\geq 0. \tag{2.3}$$

To simplify the presentation only holonomic constraints are discussed here. In the computational procedures in §3 holonomic and nonholonomic constraints are given equal treatment. Let $g_{Ni}= \partial\phi_i/\partial q$ and let the transposition of the vector or matrix C be denoted by C^T. If ϕ_j is the distance between a corner and an edge then $d\phi_j/dt = g_{Nj}^T \dot{q}$ is the relative velocity between the same corner and edge in a direction normal to the edge. λ_{Ni} is the non-negative Lagrange multiplier associated with ϕ_i. Define δ_N as follows

$$\delta_{Ni}= d^2\phi_i/dt^2= g_{Ni}^T\ddot{q}+\dot{g}_{Ni}^T\dot{q}. \tag{2.4}$$

Introduce the time-dependent, active constraint set J_N

$$J_N=\{i\,|\,\phi_i= 0\}. \tag{2.5}$$

The complementarity relations satisfied by ϕ_i, $g_{Ni}^T\dot{q}$, δ_{Ni} and λ_{Ni} are

$$\phi_i \geq 0,\ \lambda_{Ni} \geq 0,\ \lambda_{Ni}\phi_i = 0, \tag{2.6a}$$

$$\text{if } i \in J_N \text{ then } g_{Ni}^T\dot{q} \geq 0,\ \lambda_{Ni}g_{Ni}^T\dot{q} = 0, \tag{2.6b}$$

$$\text{if } g_{Ni}^T\dot{q} = 0 \text{ then } \delta_{Ni} \geq 0,\ \lambda_{Ni}\delta_{Ni} = 0. \tag{2.6c}$$

Since ϕ_i and λ_{Ni} are non-negative the conditions $\lambda_{Ni}\phi_i = 0$, $i=1,2,\ldots p$, are equivalent to $\lambda_N^T\phi = 0$. The physical interpretation of (2.6a) is that if there is no contact between the corner and the edge then $\phi_j > 0$ and the multiplier proportional to the normal force at the point of contact λ_{Nj} has vanished. If $\lambda_{Nj} > 0$ then we have a contact, $\phi_j = 0$.

The vectors $g_{Fi} \in R^m$, $i=1,2,\ldots k \leq p$, are determined such that $g_{Fi}^T\dot{q}$ is the relative velocity in the transverse direction between two bodies in contact at a point or along two edges. Let λ_{Fi} be the associated Lagrange multiplier proportional to the friction force acting on the bodies. δ_F has the definition

$$\delta_{Fi} = d(g_{Fi}^T\dot{q})/dt = g_{Fi}^T\ddot{q} + \dot{g}_{Fi}^T\dot{q}. \tag{2.7}$$

To each friction multiplier λ_{Fi} there is a corresponding normal multiplier λ_{Ni}. Introduce three index sets

$$\begin{aligned} J_{F0} &= \{i \mid i \in J_N,\ g_{Fi}^T\dot{q} = 0\}, \\ J_{FW} &= \{i \mid i \in J_N,\ g_{Fi}^T\dot{q} \neq 0\}, \\ J_F &= (J_{FW} \cup J_{F0}) \subset J_N. \end{aligned} \tag{2.8}$$

All the sets in (2.8) are time-dependent and are altered at discrete time points. If $i \in J_{F0}$ then no sliding takes place at this particular contact. According to Coulomb's law of friction the following relations are satisfied by $g_{Fi}^T\dot{q}$, δ_{Fi}, λ_{Ni}, and λ_{Fi}:

$$\text{if } i\in J_F \text{ then } |\lambda_{Fi}|\leq \mu\lambda_{Ni}\ , \tag{2.9a}$$

$$\text{if } i\in J_{FW} \text{ then } |\lambda_{Fi}|= \mu\lambda_{Ni},\ \lambda_{Fi}g_{Fi}^T\dot{q}\leq 0\ , \tag{2.9b}$$

$$\text{if } i\in J_{FO} \text{ then } (\mu\lambda_{Ni}-|\lambda_{Fi}|)\delta_{Fi}= 0,\ \lambda_{Fi}\delta_{Fi}\leq 0. \tag{2.9c}$$

The friction coefficient is μ, $\mu\geq 0$. The condition (2.9b) expresses the fact that if two bodies slide upon each other, then the friction force has the direction opposite to the relative velocity. It follows from (2.9c) that when there is no sliding then $\delta_{Fi}= 0$ and $|\lambda_{Fi}|\leq\mu\lambda_{Ni}$, but when the sliding has just begun then $\delta_{Fi}\neq 0$ and $|\lambda_{Fi}|= \mu\lambda_{Ni}$.

The columns of G in (2.1) are at least those vectors g_{Ni} and g_{Fi} for which $i \in J_N$ at a given time point. The elements in λ in (2.1) are the corresponding normal and friction multipliers λ_{Ni} and λ_{Fi}.

If $i \in J_{FW}$ then λ_{Fi} can be written in terms of λ_{Ni}, μ and $s_i= \pm 1$,

$$\lambda_{Fi}= s_i\mu\lambda_{Ni}\ . \tag{2.10}$$

For friction problems it is convenient to split G into two parts, G_1 and H. Let λ_1 contain the normal multipliers λ_{Ni} and the friction multipliers λ_{Fj}, $j \in J_{FO}$. Moreover, let the columns of G_1 consist of those g_{Ni} and g_{Fj} corresponding to the components in λ_1 and collect the remaining vectors g_{Fj}, $j \in J_{FW}$, in H. By (2.10) there exists a matrix U such that

$$\sum_{i\in J_{FW}} g_{Fi}\lambda_{Fi}= \sum_{i\in J_{FW}} g_{Fi}s_i\mu\lambda_{Ni}= HU\lambda_1. \tag{2.11}$$

The contribution of the working friction forces to the equations of motion (2.1) is (2.11). Thus, with the new notation, (2.1) is equivalent to

$$M\ddot{q}= f+(G_1+HU)\lambda_1,\ \ t\geq 0. \tag{2112}$$

Suppose that a new constraint j becomes active at t_*. Then $J_N(t\geq t_*)=J_N(t<t_*)\cup\{j\}$ and $\phi_j(t)>0$, $t<t_*$, but $\phi_j(t)=0$, $t\geq t_*$. In general, there is an impulse Λ in the system at t_* and $\dot{q}$ is discontinuous.

Let $\dot{q}_+$ and $\dot{q}_-$ denote $\dot{q}(t+0)$ and $\dot{q}(t-0)$, respectively, F is the vector of external impulses. Then one relation between $\dot{q}_+$, $\dot{q}_-$ and Λ is

$$M(\dot{q}_+-\dot{q}_-)=F+G(q)\Lambda. \tag{2.13}$$

G consists of the columns g_{Ni} and g_{Fi} such that $i\in J_N(t_*+0)$. A component Λ_{Ni} of Λ corresponds to a contact constraint ϕ_i and Λ_{Fj} to a friction constraint $g_{Fj}^T\dot{q}$. The complementarity relations for $\dot{q}_+$ and Λ depend on the material in the colliding bodies and on the impulse friction model. Here the collision is assumed to be inelastic and Λ is the optimal solution to the QP problem

$$\min \frac{1}{2}\Lambda^T G^T M^{-1} G\Lambda+\Lambda^T G^T(\dot{q}_-+M^{-1}F),$$
$$\Lambda_{Ni}\geq 0,\ -\mu_I\Lambda_{Ni}\leq\Lambda_{Fi}\leq\mu_I\Lambda_{Ni}. \tag{2.14}$$

The impulse friction coefficient is μ_I, $\mu_I\geq 0$. For a friction free system $\mu_I=0$ and $\Lambda_{Fi}=0$. By (2.13) and the Kuhn-Tucker conditions [13,p.51] at the optimum of (2.14) we have

$$g_{Ni}^T\dot{q}_+\geq 0,\ \Lambda_{Ni}\geq 0,\ \Lambda_{Ni}g_{Ni}^T\dot{q}_+=0. \tag{2.15}$$

The implications of (2.14) on Λ_{Fi} and $g_{Fi}^T\dot{q}_+$ when $\mu_I>0$ are discussed in §4.3.

Another source of discontinuities is when a friction constraint j is transferred from J_{FW} to J_{F0}, i.e. when $g_{Fj}^T\dot{q}\to 0$. Then $\ddot{q}$ and λ are in general discontinuous. The time derivatives $\dddot{q}$ and $\dot{\lambda}$ are usually discontinuous when λ_{Nj} reaches its lower bound 0 and j is removed from J_N and when $|\lambda_{Fj}|$ attains its upper bound $\mu\lambda_{Nj}$ and j is transferred from J_{F0} to J_{FW}.

In order to solve (2.1) or (2.12) in combination with the relations (2.6) and (2.9), the problem is discretized by two linear multistep methods, Gear [15]. An approximation $(q_i,\dot{q}_i)$ to $(q(t_i), \dot{q}(t_i))$ is determined at $t = t_i$, $i=0,1,\ldots$, where $t_0= 0$ and $t_{i+1}= t_i+ h_{i+1}$. Assume that $(q_i,\dot{q}_i)$, $i=0,1,\ldots n-1$, are known. Then the formulas for computation of $(q_n,\dot{q}_n)$ with a constant step size $h_i= h$ and two r-step methods are

$$\sum_{i=0}^{r} \alpha_i^1 q_{n-i}=h \sum_{i=0}^{r} \beta_i^1 \dot{q}_{n-i}, \tag{2.16a}$$

$$\sum_{i=0}^{r} \alpha_i^2 \dot{q}_{n-i}=h \sum_{i=0}^{r} \beta_i^2 M^{-1}(f_{n-i}+G_{n-i}\lambda_{n-i}), \tag{2.16b}$$

where $f_{n-i}= f(q_{n-i},\dot{q}_{n-i},t_{n-i})$, $G_{n-i}= G(q_{n-i})$ and the coefficients α_i^j, β_i^j, $j=1,2$, are constant. Sum all the quantities in (2.16a) and (2.16b) that are known from previous steps in b_n^1 and b_n^2, respectively, so that q_n and $\dot{q}_n$ are the solutions to the nonlinear equations

$$\begin{aligned} q_n&=(h\beta_0^1\dot{q}_n+b_n^1)/\alpha_0^1, \\ \dot{q}_n&=(h\beta_0^2M^{-1}(f_n+G_n\lambda_n)+b_n^2)/\alpha_0^2. \end{aligned} \tag{2.17}$$

In a friction free problem, $\mu=0$, q_n and λ_n satisfy in addition to (2.17) the discrete version of (2.6a)

$$\phi(q_n)\geq 0, \ \lambda_n\geq 0, \ \lambda_n^T\phi(q_n)=0. \tag{2.18}$$

The system (2.17), (2.18) is referred to as a nonlinear complementarity problem (NLCP) in Cottle [6]. The prospects of having to solve an NLCP at each time-step are not so encouraging. Therefore, in the next section by a suitable choice of linear multistep methods in (2.16) and by replacing the condition (2.6a) by (2.6b), we arrive at a linear complementarity (LCP) problem whose solution is calculated with a QP algorithm.

The norm $||\cdot||$ in the sequel always denotes the Euclidean vector norm or the subordinate, spectral matrix norm. The weighted norm $||x||_M$ of the vector x has the definition

$$||x||_M^2=x^TMx=||Dx||^2. \tag{2.19}$$

3. Contact problems.

The difference equations for friction free contact problems fulfilled by $(q_n,\dot{q}_n)$ are chosen in this section. A numerical procedure to obtain λ_n is suggested for time points t_n where the active constraint set is constant, $J_N(t_{n-1}) = J_N(t_n)$, and when contacts disappear, $J_N(t_{n-1})\supset J_N(t_n)$, $\phi_j(t_{n-1}) = 0$ but $\phi_j(t_n)>0$ for some j. When a new contact is established at t_*, the impulse in the system and the jump in $\dot{q}$ satisfy (2.13) and (2.14). The section is concluded by two examples.

3.1 Advancing the solution in time.

The coefficients α_i^1, β_i^1 in (2.16a) are determined by a member of the Adams-Bashforth family of explicit formulas denoted by AB-r [15,p.104]. The method in (2.16b) is a backward difference formula [15,p.217] with $\beta_0^2 = 1$, $\beta_j^2 = 0$, j=1,2,...r, abbreviated BDF-r. Replace (2.18) by the condition (2.6b) in intervals where J_N is constant or indices are dropped from J_N. Let g_{Ni}, $i \in J_N$, form G. Since $\beta_0^1 = 0$ we have by (2.17) and (2.6b) that q_n and $\dot{q}_n$ are the solution to the system

$$q_n = b_n^1/\alpha_0^1, \tag{3.1a}$$

$$\dot{q}_n = (h_n M^{-1}(f_n + G_n\lambda_n) + b_n^2)/\alpha_0^2, \tag{3.1b}$$

$$G_n^T\dot{q}_n \geq 0,\ \lambda_n \geq 0,\ \lambda_n^T G_n^T \dot{q} = 0. \tag{3.1c}$$

First compute q_n in (3.1a). If we let c_n denote $b_n^2 + h_n M^{-1} f_n$ and temporarily assume that $f = f(q,t)$, then ((3.1b), (3.1c)) is a linear complementarity problem. The system ((3.1b), (3.1c)) is equivalent to the QP problem, Cottle [6],

$$\min \tfrac{1}{2}\ \lambda_n^T G_n^T M^{-1} G_n \lambda_n + \lambda_n^T G_n^T c_n / h_n, \tag{3.2}$$

$$\lambda_n \geq 0.$$

Since

$$c_n = h_n M^{-1} f_n - \sum_{i=1}^{r} \alpha_i^2 \dot{q}_{n-i},$$

we have $||G_n^T\ c_n|| = O(h_n)$ if h_i varies smoothly and $||\lambda_n|| = O(1)$ in (3.2). The dual problem to (3.2) is

$$\min \tfrac{1}{2}\ v_n^T G_n^T M^{-1} G_n v_n,$$

$$h_n G_n^T M^{-1} G_n v_n + G_n^T c_n \geq 0, \tag{3.3}$$

where $G_n\lambda_n = G_n v_n$, see Dorn [10]. Substitute

$$G_n\lambda_n = M(\alpha_0\dot{q}_n - c_n)/h_n$$

in (3.3) and after dropping unnecessary constants in the objective function and the constraints we have

$$\min ||\sum_{i=0}^{r} \alpha_i^2 \dot{q}_{n-i} - h_n M^{-1} f_n||_M, \tag{3.4}$$

$$G_n^T \dot{q}_n \geq 0.$$

There are two advantages using the complementarity condition on the derivative of the constraint (2.6b) or (3.1c) instead of (2.6a) or (2.18). The nilpotency of the system (3.1) is lower, see Petzold [35] for a treatment of these matters, and the constraint function $G^T\dot{q}$ is linear in $\dot{q}$. The disadvantage is that $\phi_i(q) = 0$, $i \in J_N$, is not satisfied exactly. A possible remedy is the stabilization procedure developed by Baumgarte [1].

How to compute λ_n is treated in §3.2. Here we shall continue with a discussion of the properties of the time discretization. In most steps from t_{n-1} to t_n the active constraint set J_N remains unaltered. Then the constraint on $\dot{q}_n$ in (3.1c) is an equality constraint, $G_n^T\dot{q}_n = 0$. It is shown in [27] that the global accuracy in q_n, $\dot{q}_n$, λ_n and $\phi(q_n)$ for a constant step size h is $O(h^r)$ when $G_n^T\dot{q}_n = 0$. The recursion in (3.1) is restarted using the first order method (AB-1, BDF-1) for two steps after a discontinuity in $\dot{q}$ or $\ddot{q}$. In order to simplify the solution process, the discontinuities in $\ddot{q}$ stemming from the removal of a constraint from J_N are ignored. Errors proportional to h^3 and h^2 are introduced in q_n and $\dot{q}_n$, respectively. It does not pay to use methods of higher order than two, r=2, if the order of the global error is to be the same at points when J_N is constant and when constraints are dropped from J_N. We assume that the alterations of J_N are independent of h. Since these take place at discret points the global accuracy of $O(h^2)$ in $\dot{q}_n$ is maintained.

The stepsize h_n is chosen such that a weighted norm of the local error ℓ_n in q_n per unit step is less than a prescribed tolerance ε,

$$||\ell_n||_M \leq h_n\varepsilon. \tag{3.5}$$

The local error at t_n in AB-1 (the forward Euler method) is proportional to $h_n^2\ddot{q}_n$ and in AB-2 to $h_n^3\dddot{q}_{n-1} = h_n^2(\ddot{q}_{n-1}-\ddot{q}_{n-2}) + O(h_n^4)$ if $h_n=h_{n-1}$. The estimate of $\dddot{q}_{n-1}$ is less accurate when a constraint is dropped from J_N but is for simplicity used there also.

The stability of the compound methods (AB-1, BDF-1) and (AB-2, BDF-2) is studied by means of the scalar test equation

$$\ddot{u}+d\dot{u}+ku = 0 \tag{3.6}$$

in [27]. This spring and damper equation has stable solutions if and only if $k>0$, $d\geq 0$ or $d>0$, $k\geq 0$. The stability regions for the first and second order methods are depicted in fig.1.

Suppose that $\phi_j(t_{n-1})>0$ but $\phi_j(t_n)\leq 0$ for some j. Then a collision has occurred at t_*, $\phi_j(t_*)= 0$, $t_{n-1}<t_* \leq t_n$, where $\dot{q}$ is discontinuous. The point t_* is determined by inverse linear interpolation between $\phi_j(t_{n-1})$ and $\phi_j(t_n)$. The error introduced by this approximate computation of t_* is of $O(h_n^2)$. An example where $\ddot{q}$ is discontinuous is when one body in contact with another body slides off the support of the other body. Another reason for a jump in $\ddot{q}$ is when f is discontinuous. The crucial point t_* is also here determined by inverse linear interpolation.

The assumption $f = f(q,t)$ is not always satisfied. A linear damper gives rise to an external force proportional to $\dot{q}$, $f = f(q,\dot{q},t)$. Sufficient conditions for the existence of a unique solution to ((3.1b), (3.1c)) when f is velocity dependent are stated below.

Proposition 1.

Let f be continuously differentiable in $\dot{q}$ and let A be defined by

$$A(\dot{q})= M- \frac{h_n}{\alpha_0^2}\frac{\partial f}{\partial \dot{q}}(q_n,\dot{q},t_n).$$

Assume that A is invertible in a sufficiently large neighborhood Q of $\dot{q}_{n-1}$ and let

$$B(\dot{q}) = G_n^T A^{-1} G_n.$$

If

$$x^T B x \geq \kappa x^T x, \ \kappa > 0, \ \dot{q} \in Q, \qquad (3.7)$$

then there is a unique solution $(\dot{q}_n, \lambda_n)$ to ((3.1b), (3.1c)).

<u>Proof.</u>

Since A^{-1} exists when $\dot{q} \in Q$, $\dot{q}_n$ in (3.1b) can be expressed uniquely as $\dot{q}_n = u_n(\lambda_n)$, [9, thm. 10.2.2].

Furthermore,

$$\frac{\partial \dot{q}_n}{\partial \lambda_n} = \frac{h_n}{\alpha_0^2} A^{-1}(u_n(\lambda_n)) G_n = \frac{h_n}{\alpha_0^2} A_n^{-1} G_n.$$

Consider the NLCP

$$\lambda_n \geq 0, \ G_n^T \dot{q}_n = G_n^T u_n(\lambda_n) \geq 0, \ \lambda_n^T G_n^T u_n = 0. \qquad (3.8)$$

We find that

$$\frac{\partial}{\partial \lambda_n} G_n^T \dot{q}_n = \frac{h_n}{\alpha_0^2} G_n^T A_n^{-1} G_n = \frac{h_n}{\alpha_0^2} B.$$

If B has the property (3.7) then there exists a unique λ_n solving (3.8) according to [22, thm. 3.2]. Hence, $\dot{q}_n$ is also unique. ■

<u>Remark.</u>

Assume that G_n has full column rank. If h_n is sufficiently small then $A^{-1} = M^{-1} + h_n C$, $||C|| = O(1)$, $x^T G_n^T M^{-1} G_n x \geq \kappa_0 x^T x$, $\kappa_0 > 0$, and (3.7) is satisfied. It can be shown that if $\partial f/\partial \dot{q}$ is symmetric and the eigenvalues λ_i of $D^{-T} \partial f/\partial \dot{q} \ D^{-1}$ satisfy $1 > \kappa_1 \geq h_n \lambda_i / \alpha_0$ then the sufficient conditions in the proposition are fulfilled.

The proposed solution process for ((3.1b), (3.1c)) when $f = f(q,\dot{q},t)$ is based on functional iteration. The value $\dot{q}_n^{(0)}$ is predicted by $\dot{q}_{n-1}$, the AB-1 or the AB-2 formula depending on the number of $\dot{q}_i$ available with sufficient continuity. Then $q_n^{(j)}$, $j=1,2,\ldots$ is determined by the corrector equation

$$\dot{q}_n^{(j)}=(h_n M^{-1}(f(q_n,\dot{q}_n^{(j-1)},t_n)+G_n\lambda_n^{(j)})+b_n^2)/\alpha_0^2 ,$$

$$G_n^T\dot{q}_n^{(j)}\geq 0,\ \lambda_n^{(j)}\geq 0,\ \lambda_n^{(j)}G_n^T\dot{q}^{(j)}= 0. \tag{3.9}$$

The criterion for interruption of the iteration is discussed after the following proposition.

Proposition 2.

Let $f_n= f(q_n,\dot{q},t_n)$ be continuous and differentiable in $\dot{q}$. Let c_1 and c_2 be constants such that $||M^{-1}\partial f_n/\partial\dot{q}||_M\leq c_1$ in $||\dot{q} - \dot{q}_n^{(0)}||_M<c_2$. If h_n is so small that $h_n c_1/\alpha_0^2 <1$ and $||\dot{q}_n^{(1)}- \dot{q}_n^{(0)}||_M<c_2(1-h_n c_1/\alpha_0^2)$ then the proposed iterative process (3.9) converges to the unique solution $\dot{q}_n$ of ((3.1b), (3.1c)).

Proof.

Define F_0 and F_1 by

$$F_0(y)= \frac{1}{2}\alpha_0^2\ y^TMy-h_n y^T f_n^0-y^TMb_n^2,$$

$$F_1(y)= \frac{1}{2}\alpha_0^2\ y^TMy-h_n y^T f_n^1-y^TMb_n^2,$$

where $f_n^i= f(q_n,x_i,t_n)$. Define the minimization problems

$$\min F_0(y),\ G_n^Ty\geq 0, \tag{3.10a}$$

$$\min F_1(y),\ G_n^Ty\geq 0. \tag{3.10b}$$

Let y_0 be the solution to (3.10a) and $y_1= y_0+ \delta y$ the solution to (3.10b). It is shown in [31,(3.5)-(3.7)] that y_0 and y_1 satisfy

$$\delta y^T(\nabla F_1(y_1) - \nabla F_1(y_0)) \leq \delta y^T(\nabla F_0(y_0) - \nabla F_1(y_0)), \tag{3.11}$$

where ∇F_i, $i=1,2$, denotes the gradient of F_i. Then

$$||\delta y||_M^2 \leq \delta y^T(h_n/\alpha_0^2(f_n^1-f_n^0))=(D\delta y)^T(h_n/\alpha_0^2 DM^{-1}(f_n^1-f_n^0)).$$

Therefore, by (2.19)

$$||\delta y||_M \leq h_n/\alpha_0^2||M^{-1}(f_n^1-f_n^0)||_M.$$

It follows from [9, 8.5.4] that if $||x_0-\dot{q}_n^{(0)}||_M<c_2$ and $||x_1-\dot{q}_n^{(0)}||_M<c_2$ then

$$||y_1-y_0||_M \leq h_n c_1/\alpha_0^2||x_1-x_0||_M<||x_1-x_0||_M. \tag{3.12}$$

If $x_i=\dot{q}_n^{(j-1+i)}$ then by (3.4) $\dot{q}_n^{(j)}=y_i$ in the iteration process (3.9). The fixed point theorems 10.1.1 and 10.1.2 in [9] and (3.12) prove the convergence of the procedure (3.9) to the unique solution of ((3.1b), (3.1c)). ∎

Using the suggested predictor methods it can be shown that for the difference between the predictor $\dot{q}_n^{(0)}$ and the first corrector solution $\dot{q}_n^{(1)}$ we have

$$||\dot{q}_n^{(1)}-\dot{q}_n^{(0)}||_M=O(h_n^k),\ k\geq 1.$$

The sufficient conditions in the proposition can always be met by choosing h_n small enough. However, for problems where c_1 is large the step size h_n may be intolerably short. The conditions are the same in a problem without the constraints $G^T\dot{q}\geq 0$, cf. Gear [15,p.114].

The error $||\dot{q}_n-\dot{q}_n^{(s)}||_M$ in the accepted solution $\dot{q}_n^{(s)}$ of (3.9) is allowed to be at most 0.1ε. This error is multiplied by $\beta_1^1 h$ in the computation of q_{n+1}. The factor 0.1 is chosen in order to make $0.1\beta_1^1 h_{n+1}\varepsilon$ small in comparison with the local error (3.5) in q_{n+1}. Let the error $G_n\delta\lambda_n^{(j)}$ in $G_n\lambda_n^{(j)}$ satisfy

$$||D^{-T}G_n\delta\lambda_n^{(j)}||=||M^{-1}G_n\delta\lambda_n^{(j)}||_M\leq\varepsilon_0. \tag{3.13}$$

The iterative error in $\dot{q}_n^{(j)}$ is estimated by

$$||\dot{q}_n-\dot{q}_n^{(j)}||_M \leq (\rho||\dot{q}_n^{(j)}-\dot{q}_n^{(j-1)}||_M+\varepsilon_0)/(1-\rho), \tag{3.14}$$

Dahlquist and Björck [8,p.239]. An approximate rate of convergence $\rho \approx h_n c_1/\alpha_0^2$ is determined during the iteration. When the error bound in (3.14) is less than 0.1ε for $j=s$ then the iteration is terminated. If f is dependent on the velocity $\dot{q}$ then $\varepsilon_0 = 0.01\varepsilon$. Otherwise, $\dot{q}_n$ is computed directly in (3.1b) and the required accuracy ε_0 in $G_n\lambda_n$ is 0.1ε.

In conclusion, we give the algorithm for advancing the solution one step from t_{n-1} to t_n.

Algorithm 1.

1. Compute q_n by AB-1 or AB-2 with a step length h_n such that the local error satisfies (3.5).
2. If necessary, predict the value $\dot{q}_n^{(0)}$ and then compute $f_n^{(0)} = f_n(q_n,\dot{q}_n^{(0)},t_n)$. Determine $G_n\lambda_n$ to the prescribed accuracy ε_0 and calculate $\ddot{q}_n = M^{-1}(f_n + G_n\lambda_n)$ and $\dot{q}_n$ by BDF-1 or BDF-2. If f is a function of $\dot{q}$ this step is performed iteratively (3.9).
3. Test whether
   (i) $\dot{q}$ is discontinuous between t_{n-1} and t_n because there is an external impulse $F \neq 0$ in (2.13) or a new constraint becomes active i.e. $\phi_j(t_{n-1})>0$ but $\phi_j(t_n)\leq 0$ for some j,
   (ii) $\ddot{q}$ is discontinuous between t_{n-1} and t_n because f or $G\lambda$ is discontinuous, e.g. due to the loss of support of a body.

   If (i) or (ii) has occurred then the time t_* where $\dot{q}$ or $\ddot{q}$ is discontinuous is computed by inverse linear interpolation. If there are several detected jumps in $\dot{q}$ or $\ddot{q}$ in the interval at $t_*^1, t_*^2, \ldots t_*^j$, then $t_* = \min(t_*^1, \ldots t_*^j)$. After a discontinuity in $\dot{q}$ the new velocity $\dot{q}_+$ in (2.13) is first determined. Then restart the integration at step 1 by AB-1 with $q(t_*)$, $\dot{q}(t_*+0)$ and the new active constraint set $J_N(t_*+0)$.

4. Test whether there are any elements λ_{nj} in λ_n such that $\lambda_{nj} = 0$ and the corresponding Δ_{nj} in

$$\Delta_n = G_n^T \dot{q}_n$$

satisfies $\Delta_{nj} > \varepsilon$. If that is the case then remove j from J_N.

5. End of algorithm.

Initially, J_N at t=0 is defined by $\phi(q(0))$ and $G^T\dot{q}(0)$ as follows

$$J_N = \{ | \phi_i = 0 \text{ and } g_i^T \dot{q} = 0 \text{ at } t=0 \}.$$

We exclude the possibility of a j such that $\phi_j(q(0))<$, i.e. the initial conditions are compatible with the constraints. It may be necessary to introduce an impulse in the system at t=0 so that $g_j^T\dot{q} \geq 0$ for every j with $\phi_j = 0$. In the restart step $\ddot{q}_{n-1}$ is not availble for the local error estimate. The first step size h_n in step 1 is instead taken to be h_{n-1}, the old step size, and when step 2 has been completed, the local error in q_n is estimated using $\ddot{q}_n$. If (3.5) is not satisfied then the steps 1 and 2 are repeated with a new step size. The threshold value ε in step 4 is chosen to avoid unnecessary zigzagging, i.e. a constraint that left J_N at t_{n-1} is reintroduced in J_N between t_{n-1} and t_n, cf. nonlinear optimization in Fletcher [13,p.113]. If $\Delta_{nj} \sim \varepsilon$, then $\lambda_{nj} = 0$ and λ_{nj} does not affect the equations of motion, even if j is not dropped from J_N.

3.2. The quadratic programming problem.

The velocity $\dot{q}_n$ can be calculated directly as the solution to (3.4) without introducing λ_n. It is, however, advantageous to solve (3.2) for λ_n first and then insert the result into (3.1b) to obtain $\dot{q}_n$. The merits of (3.2) in comparison with (3.4) are

1. An initial, feasible λ_n^0 in an active set algorithm for QP problems is easily found.
2. The QP problem for friction free systems is a special case of the QP problem for systems with Coulomb friction.

3. The multiplier vector λ_n is sometimes of interest itself, e.g. in friction problems.

It is well-known from classical mechanics that in time-independent problems, the contact forces cannot be uniquely determined due to the fact that the columms of G are linearly dependent, ("statical indeterminacy"). This can of course also be the case in dynamical problems. In Lötstedt [30] it is shown that $G\lambda$ in (2.1) is always unique even if G does not have full column rank.

We shall consider a slightly generalized QP problem (3.2) and study the uniqueness of the solution and the sensitivity to perturbations. The generalization is motivated by our Coulomb friction model in §4. The new QP problem is

$$\begin{aligned} &\min \frac{1}{2} x^T A^T A x + x^T A^T d, \\ &a_i \le x_i \le b_i, \quad i=1,2,\ldots \ell, \\ &a_i \le x_i \quad , \quad i=\ell+1,\ell+2,\ldots k, \end{aligned} \tag{3.15}$$

where $A \in \mathbb{R}^{m\times k}$, $d \in \mathbb{R}^m$, a, $x \in \mathbb{R}^k$ and $b \in \mathbb{R}^\ell$. If the objective function in (3.15) is replaced by $||Ax + d||$, we have transformed (3.15) to a linear least squares problem with the same solution x_*. If x_* is not unique then on the manifold of solutions choose x_* satisfying

$$\begin{aligned} &\min \frac{1}{2} x^T x, \\ &\ x \text{ solves } (3.15). \end{aligned} \tag{3.16}$$

In the perturbed problem ((3.15), (3.16)), denoted by $((3.15)_\delta$, $(3.16)_\delta)$, the quantities A, d, a, and b are perturbed by δA, δd, δa, and δb. The next proposition characterizes the solutions to ((3.15), (3.16)) and $((3.15)_\delta, (3.16)_\delta)$. In order to state the proposition we need some notation. $R(B)$ denotes the range and $N(B)$ the null space of the matrix B. P_X is the ortogonal projector on the subspace X.

Furthermore, let S and S_δ denote the sets of constraints on $x_N \in N(A)$ that must be satisfied as equalities at the solutions x_* and $x_* + \delta x_*$ to (3.15) and $(3.15)_\delta$, respectively. Let us consider an example. Suppose that $\ell=0$, $n=2$, $a^T=(0,0)$, $x_*^T=(0,0)$,

$$R(A^T)=\begin{pmatrix}1\\1\end{pmatrix}z_1 \text{ and } N(A)=\begin{pmatrix}1\\-1\end{pmatrix}z_2, \tag{3.17}$$

where z_1 and z_2 are arbitrary. Then $S=\{1,2\}$. Conversely, if the definitions of $R(A^T)$ and $N(A)$ in (3.17) are interchanged then $S=\emptyset$. If $N(A)=\emptyset$ then of course $S=\emptyset$.

Proposition 3.

A solution x_* to (3.15) always exists and $P_{R(A^T)}x_*$ and Ax_* are unique. If A has full column rank or if x_* also satisfies (3.16), then x_* is unique. Define $\eta = \max\{||\delta A||, ||\delta d||, ||\delta a||, ||\delta b||\}$. Assume that

$$\text{rank}(A)= \text{rank}(A+\delta A),$$

$a_i<b_i$, $i=1,2,\ldots\ell$, and $S=S_\delta$. Then there exist positive c and η_0 such that if $\eta\le\eta_0$ then

$$||\delta x_*||\le c\eta. \tag{3.18}$$

Proof.

The existence and uniqueness results follow from [31, thm.1]. The conditions (i) and (iii) in [31, thm.3] are satisfied and the perturbation bound (3.18) is a consequence of the theorem ∎

The proposition is directly applicable to (3.2) with $A = D^{-T}G_n$, $x = \lambda_n$, $d = Dc_n/h_n$, $\ell\neq 0$ and $a=0$. In contact problems where we are not explicitly interested in a unique λ_n, it is sufficient to compute the unique $G_n\lambda_n$.

The QP problem (3.2) is solved by an algorithm exploiting the fact that the objective function does not vary much from step to step. The active set method by Gill and Murray [18] is modified to handle a positive semi-definite A^TA and in the step where a new search direction is calculated.

An algorithm similar to that in [18] is described in Lawson and Hanson [25]. The procedure for solving (3.15) with $a=0$ and $\ell=0$ is presented below. The algorithm is justified and a comparison is made in a separate paper Lötstedt [32]. The index of a variable belongs to one of the sets F, X or P. The restriction A_F consists of the columns of A corresponding to an index in F. F is the set of free variables, $x_i \geq 0$, such that a QR decomposition of A_F with a non-singular R is available. X is the index set of the fixed variables $x_i = 0$ and P, the passive set, contains the rest of the indices. The parameter ε_1 is an error tolerance.

Algorithm 2.

Initialization.

1. If necessary compute a QR decomposition of the maximal number of linearly independent columns in A. Let F contain the indices of the columns represented by the QR decomposition and P the remaining indices.
2. Compute the residual $r = Ax + d$.

Main iteration loop.

3. If $F \neq \emptyset$ then compute a new descent direction p by solving

$$A_F^T A_F p + A_F^T r = 0. \tag{3.19}$$

   Otherwise go to 7.
4. Let θp, $0 \leq \theta \leq 1$, be the maximal step possible in the direction p without violating any constraints. θ satisfies

$$\tau_i = -x_i/p_i,\ i \in F_* = \{j \mid j \in F,\ p_j < 0\},$$

$$\Theta = \min(\min_{i \in F_*} \tau_i,\ 1).$$

5. Update x and r

$$x_i := x_i + \Theta p_i,\ i \in F,$$

$$r := r + \Theta A_F p.$$

6. If $\tau_i > 1$, $i \in F_*$, in step 4 then goto 7. For each x_j reaching its lower bound in this step, transfer j from F to X and remove the corresponding column in the QR decomposition of A_F.
   Go to 3.
7. Compute the gradient of the objective function

$$\Delta = A^T r = A^T Ax + A^T d.$$

8. Find

$$\gamma = \max(-\min_{i \in X} \Delta_i, \max_{i \in P} |\Delta_i|).$$

   and an index j such that $\gamma = -\Delta_j$, $j \in X$, or $\gamma = |\Delta_j|$, $j \in P$. If $\gamma > \varepsilon_1$ then the objective function can be decreased further if x_j becomes a free variable. Transfer j from X or P to F and add the corresponding column to the QR decomposition. Go to 3. If $\gamma \leq \varepsilon_1$ then an approximate optimum has been found.
9. End of algorithm.

On termination of the algorithm for the problem (3.2) we have

$$r = D^{-T} G_n \lambda_n + Dc_n / h_n$$

Thus,

$$\dot{q}_n = h_n / \alpha_0^2 D^{-1} r \text{ and } G_n^T \dot{q}_n = h_n / \alpha_0^2 \Delta.$$

The algorithm is initialized with $x = \lambda_{n-1}$, the feasible solution from the previous step. Another possibility is to use polynomial extrapolation to obtain a start value. At most time points t_n, the optimum is found after the computation of one descent direction in (3.19). Only in exceptional steps the lower bound on a variable x_i is attained and more than one system (3.19) has to be solved.

The matrix $A = D^{-T} G_n$ varies continuously in time and not very much from step to step. Therefore, A_F is not factorized at each step, but the old R-matrix is used to accelerate a conjugate gradient algorithm for solution of (3.19) as described by Björck and Elfving [3, p.156].

The convergence rate for $A_F R^{-1}$ in (3.19) is superior to that of A_F if R is part of a recent factorization. The old QR decomposition is updated in the steps 6 and 8. The column added in step 8 is a fresh column of A computed at t_n. If $j \in X$ at step 9, then the corresponding constraint ϕ_j is most likely dropped from J_N at step 4 in algorithm 1. The QR decomposition is already prepared in algorithm 2 for this change in J_N. The factorization of A is obtained by Householder transformations, Businger and Golub [4], and the determination of its rank is based on Karasalo [23]. Methods for updating the decomposition are described in Gill et al [16].

The criterion for computing a new factorization is derived as follows. Denote the total number of calls of algorithm 2 by T, i.e. the number of steps taken, the cost in floating point operations of a factorization by C_{QR}, and the total number of decompositions performed by n. Assume that the cost due to the iterative solution of (3.19) is $C_0 + C_{IT} \cdot s^r$, C_0 C_{IT}, $r>0$, where s is the number of steps since the last decomposition and s^r is the number of cg-iterations per step. This is a reasonable assumption if (3.19) is solved only once per step. In a steady-state situation, we have that the cost C_c in one cycle from one factorization up to the next is

$$C_c \approx C_{QR} + \int_0^{T/n} (C_0 + C_{IT} \cdot s^r) ds.$$

Thus, the total cost is

$$C(n) \approx C_{QR} n + C_0 T + C_{IT} T^{r+1} n^{-r}/(r+1).$$

The optimal number of steps per cycle is

$$\frac{T}{n} \approx \left(\frac{r+1}{r} \frac{C_{QR}}{C_{IT}}\right)^{1/(r+1)}$$

rendering C(n) a minimum.

Here we also assume that r=1. A new QR decomposition is computed when the number of iterations in the cg-algorithm exceeds $(2\ C_{QR}/C_{IT})^{1/2}$.

The conjugate gradient iterations are terminated when the residual $v = A_F^T A_F p + A_F^T r$ in (3.19) is sufficiently small. The error δp in p satisfies

$$R\delta p \approx R^{-T} A_F^T A_F \delta p = -R^{-T} v.$$

If the loop in algorithm 2 is repeated only once then

$$||M^{-1} G_n \delta\lambda_n||_M = ||D^{-T} G_n \delta\lambda_n|| = ||A_F \delta p|| \approx ||R\delta p|| \leq ||v||\ ||R^{-1}||. \tag{3.20}$$

Hence, if $||v||$ in (3.20) is less than $\varepsilon_0/||R^{-1}||$ then the accuracy requirements in (3.13) are satisfied. $||R^{-1}||$ is estimated according to Cline et al [5]. The parameter ε_1 in step 8 is taken to be $\varepsilon_0/||R^{-1}||$.

If the number of active constraints always is small, say 1-3, there is of course always the straightforward method to compute λ_n by guessing which of the components will attain 0 and then solve a system of linear equations for the remaining multipliers. If the complementarity conditions (3.1c) are not satisfied then try another probable combination. This approach can be very slow for larger problems without more information on which multipliers we can expect to reach 0.

3.3 Computation of impulses.

The treatment of the friction free impulse problem is almost equal to the treatment of the time- dependent contact problem. The QP-problem (2.14) with $\mu_I = 0$ corresponds to (3.2) and the equation (2.13) to (3.1b). An impulse Λ is computed in step 3 of algorithm 1 at each time the external impulse vector $F \neq 0$ or a constraint j is included in J_N. A modification of algorithm 2 that can handle the constraints in (2.14) when $\mu_I > 0$ is used to detetmine Λ, Gill and Murray [18]. This modified version is for $\mu_I = 0$ equivalent to algorithm 2. The algorithm always works with the exact QR decomposition.

A constraint j is dropped from J_N if $g_{Nj}^T \dot{q}_+ > \varepsilon$, the same ε as in (3.5).

3.4. Examples

The proposed algorithms 1 and 2 have been tested on two friction free model problems. More details about the implementation are given in §5. In the figures the x-axis is horizontal and x increases to the right. The y-axis is vertical and y is increasing upwards. In the first example in fig. 2, the boxes are of dimension 1.5×1 and their density is 1. Since the only external force is the gravity force, f in (2.1) is constant. The bodies have no initial velocities except for body 2 where $\dot{x}(0) = 2$. The numbers below each picture indicate the time passed from the initial state. When the simulation was interrupted, body 1 moved slowly to the right and the bodies 4 and 5 slowly to the left. In fig.3, a homogeneous hexagon is suspended in a linear spring and damper. Its shortest side is 1.5 and its mass is 7.5. The spring and damper constants (cf.(3.6)) are both 6. The external forces are the gravity force and the spring and damper force, $f = f(q,\dot{q})$. The body has no initial velocity. On the ground the hexagon oscillates back and forth until all kinetic energy has been dissipated by the damper. The velocity in the x-direction at t = 4.60 is 0.69. The error tolerance ε was 0.001 in both examples.

4. Coulomb friction problems.

Coulomb's friction model for the transverse force at a contact between two rigid bodies has several undesirable properties, Lötstedt [29]. The existence of a solution to (2.1), (2.6) and (2.9) cannot be guaranteed, and if a solution exists it is not always unique even if the columns of G are linearly independent. A solution may also be very sensitive to small perturbations in the data. We shall here develop a numerical friction model based on Coulomb's law. The equations in our model always possess a unique solution which is reasonably insensitive to small perturbations. The Lagrange multipliers are determined by a QP problem such that the friction free problem is a special case of the friction problem. A model for friction impulses is also presented. Two examples illustrate the method.

4.1 Advancing the solution in time.

The discretization methods for the numerical solution of the system of ordinary differential equations are the same as those selected in §3. Apply AB-r and BDF-r, r=1,2, to (2.12) to obtain

$$q_n = b_n^1/\alpha_0^1, \tag{4.1a}$$

$$\dot{q}_n = (h_n M^{-1}(f_n+(G_{1n}+ H_n U_n)\lambda_{1n})+b_n^2)/\alpha_0^2, \tag{4.1b}$$

where b_n^i, i=1,2, are defined as in (2.17). The discrete equations (4.1) and the relations (2.6) and (2.9) suffer from the same inconsistencies as the analytical equations (2.12) and the relations (2.6), (2.9), i.e. non-existence and non-uniqueness of solutions and sensitivity to small perturbations. The sufficient conditions in [29] for existence and uniqueness are not always satisfied by physical systems.

In order to circumvent the difficulties the computation of $(G_{1n}+ H_n U_n)\lambda_{1n}$ is split into two parts.

The first part, $G_{1n}\lambda_{1n}$, is the contribution to the equations of motion of the contact forces and the friction forces at contacts where no sliding takes place. We describe how $G_{1n}\lambda_{1n}$ is determined by a QP algorithm in §4.2. The second part, $H_n U_n \lambda_{1n}$, is the term of the working friction forces in the equations of motion. In the step from t_{n-1} to t_n the columns of H_n are those $g_{Fj}(t_n)$ for which $j \in J_{FW}$ at t_{n-1}. The signs s_i in (2.10) are such that (2.9b) is satisfied at t_{n-1}. It follows from (2.10) and (2.11) that the only components of λ_{1n} involved in the computation of $H_n U_n \lambda_{1n}$ are those normal multipliers associated with a sliding contact. They are computed by polynomial extrapolation. If $\ddot{q}$ is continuous then λ can be chosen continuous [30]. Let λ_{1j}^{i} denote the approximate value of $\lambda_{Ni}(t_j)$, $i \in J_F$, solving the QP problem at t_j. If $\ddot{q}$ is not discontinuous between t_{n-1} and t_n and $\lambda_{1,n-1}^{i}$ and $\lambda_{1,n-2}^{i}$ are available then λ_{1n}^{i} is approximated by λ_{*n}^{i} in

$$\lambda_{*n}^{i} = \lambda_{1,n-1}^{i} + h_n(\lambda_{1,n-1}^{i} - \lambda_{1,n-2}^{i})/h_{n-1}. \tag{4.2}$$

If only $\lambda_{1,n-1}^{i}$ is available due to a restart of the integration algorithm at t_{n-2}, see algorithm 2, then

$$\lambda_{*n}^{i} = \lambda_{1,n-1}^{i}. \tag{4.3}$$

In the very first step or in the first step after a restart because of a discontinuity in $\ddot{q}$ the procedure to obtain λ_{*n}^{i} is as follows:

1. Compute approximate positions at t_n of the bodies involved in sliding contacts ignoring the contact constraints by the formula:

$$q_{*} = q_{n-1} + h_n\dot{q}_{n-1} + 0.5\ h_n^2 f_{n-1}.$$

2. Compute the corresponding constraint functions $\phi_i(q_*)$ at t_n.

Then let λ_{*n}^{i} be the force from a spring and damper at the point of contact

$$\lambda^i_{*n} = \max\{0, -(dg^T_{Ni}(q_*)\dot{q}_{n-1} + k\phi_i(q_*))\}, \tag{4.4}$$

with suitably chosen spring and damper constants k and d. After a jump in $\ddot{q}$ (4.3) is used. Now that we have λ^i_{*n}, substitute

$$\sum_{i\in J_{FW}} g_{Fi}(q_n)s_i\mu\lambda^i_{*n}, \quad s_i = -\,\mathrm{sign}(g^T_{Fi}\dot{q}(t_{n-1})),$$

for $H_nU_n\lambda_{1n}$ in (4.1b). If this term from the working friction forces is included in the sum of the terms explicitly given by the previous time-steps b^2_n, then (4.1b) has the same form as (3.1b)

$$\dot{q}_n = (h_nM^{-1}(f_n + G_{1n}\lambda_{1n}) + b^2_n)/\alpha^2_0. \tag{4.5}$$

Assuming that $\lambda_1(t)$ and λ_{1n} are uniquely determined by (2.12), (4.1), (2.6) and (2.9) initially or after a discontinuity, we shall examine the errors introduced by replacing λ_{1n} by λ^i_{*n}, i=1,2,.... . Observe that if both $\lambda_1(t)$ and $\lambda_2(t)$ fulfill (2.12), (2.6) and (2.9) and $\lambda_1 - \lambda_2 \notin N(G_1 + HU)$, then there are two different solutions q_1 and q_2 to (2.12). If $\ddot{q}$ and $\dot{\lambda}$ are continuous, then the additional local errors in $\dot{q}_n$ caused by λ^i_{*n} in (4.2) are of $\mathcal{O}(h^3_n)$ when the quotient h_{n-1}/h_n is of $\mathcal{O}(1)$. The additional local error in $\dot{q}_n$ is of $\mathcal{O}(h^2_n)$ when (4.3) is used or when (4.2) is used and $\dot{\lambda}$ is discontinuous. The order of the local error due to the discretization methods employed in combination with (4.2) or (4.3) is the same as the order of the local error due to the approximations (4.2) or (4.3). Hence, the order of the local and global errors in $\dot{q}_n$ and q_n is not altered. The local errors introduced in $\dot{q}_n$ immediately after a discontinuity in $\dot{q}$ or $\ddot{q}$ are of $\mathcal{O}(h_n)$. The inconsistencies in the analytical problem exemplified in [29] occurred after a discontinuity in $\dot{q}$. Since the solution may fail to exist after such an event, it is not always meaningful to discuss rates of convergence for the numerical solution there.

The time t_*, $t_* \le t_n$, such that $g_{Fj}^T\dot{q}(t_{n-1}) \neq 0$, $j \in J_{FW}(t_{n-1})$, and $g_{Fj}^T\dot{q}(t_*)= 0$, $j \in J_{F0}(t_*)$, is calculated by inverse linear interpolation. Since $\ddot{q}$ in general has a jump at t_*, the integration algorithm is restarted from that point. The additions to algorithm 1 in step 3 and 4 motivated by the friction algorithm are:

3. Test whether

   (ii)´ $\ddot{q}$ is discontinuous between t_{n-1} and t_n because $g_{Fj}^T\dot{q}$ has reached 0 in the interval.

   If (ii)´ has occurred then update J_{FW} and J_{F0}.

4. Test whether there is a multiplier λ_{Fj} in λ_{1n} such that $j \in J_{F0}(t_{n-1})$, but $|\lambda_{Fj}|= \mu\, \lambda_{Nj}$ and $|g_{Fj}^T\dot{q}_n| > \varepsilon$. If that is the case then transfer j from J_{F0} to J_{FW}.

4.2 The quadratic programming problem.

The remaining task in the time marching algorithm is show how to compute $G_{1n}\lambda_{1n}$ in (4.5). Since we wish the friction free algorithm to be a special case of the friction algorithm we accept λ_{1n} as the solution to

$$
\begin{aligned}
&\min \tfrac{1}{2}\, \lambda^T G_{1n}^T M^{-1} G_{1n}\lambda + \lambda^T G_{1n}^T c_n/h_n \\
&\qquad \lambda_{Ni} \ge 0, \quad i \in J_N(t_{n-1}), \\
&-a_i \le \lambda_{Fi} \le a_i, \quad i \in J_{F0}(t_{n-1}).
\end{aligned}
\tag{4.6}
$$

The bound on $|\lambda_{Fi}|$ is $a_i = \mu\lambda_{*n}^i$. The Kuhn-Tucker conditions [13] satisfied at the optimum of (4.6) are

$$
\begin{aligned}
&\lambda_{Ni} \ge 0,\ g_{Ni}^T\dot{q}_n \ge 0,\ \lambda_{Ni} g_{Ni}^T\dot{q}_n = 0, \\
&|\lambda_{Fi}| \le a_i,\ \lambda_{Fi} g_{Fi}^T\dot{q}_n \le 0,\ (a_i - |\lambda_{Fi}|) g_{Fi}^T\dot{q}_n = 0.
\end{aligned}
\tag{4.7}
$$

These conditions on the friction multipliers are the same as those imposed by Coulomb´s law in (2.9b) and (2.9c) except for the approximate upper bound on $|\lambda_{Fi}|$.

It follows from proposition 3 that if λ_{1n} solves (4.6) then $G_{1n}\lambda_{1n}$ is unique. Unfortunately, if G_{1n} does not have full column rank then the components λ_{Nj} of λ_{1n} are not necessarily unique. Any solution λ_{2n} satisfying the constraints and such that $\lambda_{1n}-\lambda_{2n} \in N(G_{1n})$ is a possible minimizer. Thus, by (4.2) and (4.3) a_i in (4.6) and the substitution for $H_nU_n\lambda_{1n}$ are not necessarily unique.
As a remedy we introduce the principle of choosing the λ_{1n} among the solutions of (4.6) rendering $\lambda_{1n}^T\lambda_{1n}$ a minimum. The actual computation is performed in two stages. Firstly, a λ_* solving (4.6) determined by algorithm 2 modified in the steps 4,6 and 8 to allow for upper and lower bounds on the variables, see Gill and Murray [17]. Secondly, with λ_* as the initial feasible starting point, $\lambda^T\lambda$ is minimized by reducing the component of λ_* in $N(G_{1n})$. This is achieved by the following method.

Compute the QR decomposition of G_{1n}^T by Householder transformations [4]. Let G_{1n} have the dimension $m\times\ell$ and $r=\text{rank}\ (G_{1n})$. If $r < \ell$ then Q can be partitioned $Q=(Q_1,Q_2)$, $Q_1 \in \mathbb{R}^{\ell\times r}$, $Q_2 \in \mathbb{R}^{\ell\times(\ell-r)}$, where Q_2 spans $N(G_{1n})$ [18]. Let $Q_2z = P_{N(G_{1n})}\lambda$ and $\lambda_R = P_{R(G_{1n}^T)}\lambda_*$. The projection λ_R is unique according to proposition 3. Then seek the optimum of

$$\min \frac{1}{2}\, z^TQ_2^TQ_2z = \frac{1}{2}\, z^Tz, \tag{4.8}$$

$$\lambda = Q_2z + \lambda_R \text{ satisfies the constraints in (4.6).}$$

The problem (4.8) is a QP problem in z termed the least distance programming problem [25,p.159]. The solution z_* is calculated by Gill and Murray's method [18] adapted to the special properties of the problem. Note that the initial $z = Q_2^T\lambda_*$ is consistent with the constraints. The final solution λ_{1n} to (4.6) and (4.8) is obtained directly from (4.8) as $Q_2z_* + \lambda_R$. The quantities $G_{1n}\lambda_{1n}$ and $\dot{q}_n$ in (4.5) are results available on termination of the algorithm for solving (4.6), cf.

the friction free case. The formulation (4.8) of the problem of minimizing $\lambda^T\lambda$ requires one QR decomposition per time-step. A method, where old decompositions are utilized to accelerate a conjugate gradient procedure for the problem, is described and compared with the approach in (4.8) in Lötstedt [32].

If $f = f(q,\dot{q},t)$ in (4.5), then an iterative procedure for solution of (4.5) and (4.6) corresponding to (3.9) in the friction free case converges to the unique solution $\dot{q}_n$ of (4.5) under the same assumptions as in proposition 2. We shall only sketch the proof of this assertion here. Let y_0, $y_1 = y_0 + \delta y$, f_n^0, and $f_n^1 = f_n^0 + \delta f$ be defined by

$$y_i = (h_n M^{-1}(f_n^i + G_{1n}\lambda_{1n}^i) + b_n^2)/\alpha_0^2,$$

$$f_n^i = f(q_n, x_i, t_n),\ i=0,1,$$

where λ_{1n}^0 and $\lambda_{1n}^1 = \lambda_{1n}^0 + \delta\lambda$ solve the associated problems (4.6). Then apply (3.11) to (4.6) to obtain

$$(D^{-T}G_{1n}\delta\lambda)^T(D^{-T}G_{1n}\delta\lambda) \leq -(D^{-T}G_{1n}\delta\lambda)^T D^{-T}\delta f.$$

It follows from the definition of δy and the above inequality that

$$||\delta y||_M \leq h_n/\alpha_0^2 ||\ M^{-1}\delta f||_M.$$

Hence, (3.12) holds true and the rest of the proof is identical to the last part of the proof of proposition 2.

The conclusion from proposition 3 applied to (4.6) and (4.8) is that under reasonable assumptions, the solution λ_{1n} is well-behaved when the data is subject to small perturbations. It is clear from the definition of λ_{*n}^i in (4.4) after a discontinuity in $\dot{q}$ at t_*, (4.6), and (4.8) that the solution $\dot{q}_n$ at $t_n = t_* + h_n$ does not blow up as $h_n \to 0$ even if the analytical solution fails to exist for $t>t_*$. Finally, it should be pointed out that the algorithm presented above for computation of time-dependent systems obeying Coulomb's law is not the only possible modification of the law.

4.3 Computation of impulses.

The model for friction impulses is defined by the QP problem (2.14). By [31,thm.1] $G\Lambda$ and $\dot{q}_+$ in((2.13) are unique. The problem (2.14) is solved by the modified version of algorithm 2 mentioned in §3.3.

We shall study the optimal solution of (2.14) by means of the Kuhn-Tucker conditions when $\mu_I \geq 0$. Suppose that $J_N=\{1,2,\ldots n\}$ and $J_F=\{1,2,\ldots k\}$, $k \leq n$. Let the variables Λ_{Ni}, Λ_{Fi} be ordered in Λ such that

$$\Lambda^T=(\Lambda_{N1},\Lambda_{N2},\ldots\Lambda_{Nn},\ \Lambda_{F1},\ \Lambda_{F2},\ldots\Lambda_{Fk}).$$

Then the constraints in (2.14) can be written in matrix form

$$W\Lambda \geq 0,\quad W=\begin{pmatrix} I_1 & 0 \\ \mu_I E & -I_2 \\ \mu_I E & I_2 \end{pmatrix},$$

where I_1 and I_2 are the n×n and k×k identity matrices and $E \in \mathbb{R}^{k\times n}$, $E_{ij}=\delta_{ij}$, the Kronecker delta. The dual problem to (2.14) is, Dorn [10],

$$\min \tfrac{1}{2}\, u^T G^T M^{-1} G u,$$

$$W^T v = G^T M^{-1} G u + G^T \dot{q}_- + G^T M^{-1} F, \tag{4.9}$$

$$v \geq 0.$$

Since $Gu = G\Lambda$, [10], (4.9) is equivalent to

$$\min \tfrac{1}{2}\, (\dot{q}_+ - (\dot{q}_- + M^{-1}F))^T M (\dot{q}_+ - (\dot{q}_- + M^{-1}F)), \tag{4.10a}$$

$$W^T v = G^T \dot{q}_+, \tag{4.10b}$$

$$v \geq 0. \tag{4.10c}$$

By virtue of (4.10b) and the Kuhn-Tucker conditions [13], Λ and $\dot{q}_+$ fulfill

$$v^T W \Lambda = \Lambda^T G^T \dot{q}_+ = 0. \tag{4.11}$$

Let us interpret (4.10b) and (4.11) for the individual contacts. For i=k+1, k+2,...n, the relations (2.15) are obtained.

For $i \leq k$, $\Lambda_{Ni} > 0$ and $|\Lambda_{Fi}| < \mu_I \Lambda_{Ni}$ we have $v_i = v_{n+i} = v_{n+k+i} = 0$ and $g^T_{Ni}\dot{q}_+ = g^T_{Fi}\dot{q}_+ = 0$. If $\Lambda_{Ni} = 0$ then $v_j \geq 0$ and $g^T_{Ni}\dot{q}_+ \geq 0$. The case that differs from the friction force case in §4.2 is when $|\Lambda_{Fi}| = \mu_I \Lambda_{Ni} > 0$. If $\Lambda_{Fi} = \mu_I \Lambda_{Ni}$ then $v_i = v_{n+k+i} = 0$, $v_{n+i} \geq 0$, and $g^T_{Fj}\dot{q}_+ = -v_{n+i} \leq 0$, $g^T_{Ni}\dot{q}_+ = \mu_I v_{n+i} \geq 0$. If $\Lambda_{Fi} = -\mu_I \Lambda_{Ni}$ then $v_i = v_{n+i} = 0$, $v_{n+k+i} \geq 0$, and $g^T_{Fi}\dot{q}_+ = v_{n+k+i} \geq 0$, $g^T_{Ni}\dot{q}_+ = \mu_I v_{n+k+i} \geq 0$. Hence, if $\mu_I > 0$ and the relative velocity in the tangential direction between the bodies involved in a contact is non-zero after the collision, then the contact disappears immediately. The bodies merely "bounce" on each other. The effect on the kinetic energy $T = \frac{1}{2}\dot{q}^T M\dot{q}$ before and after the collision is, taking (4.11) into account,

$$
\begin{aligned}
2\Delta T &= 2(T_+ - T_-) = \dot{q}_+^T M\dot{q}_+ - \dot{q}_-^T M\dot{q}_- = (\dot{q}_+ + \dot{q}_-)^T M(\dot{q}_+ - \dot{q}_-) = \qquad (4.12)\\
&= (\dot{q}_+ + \dot{q}_-)^T (F + G\Lambda) = \dot{q}_-^T G\Lambda + (\dot{q}_+ + \dot{q}_-)^T F = \\
&= \sum_{i\in J_N} \Lambda_{Ni} g^T_{Ni}\dot{q}_- + \sum_{i\in J_F} \Lambda_{Fi} g^T_{Fi}\dot{q}_- + (\dot{q}_+ + \dot{q}_-)^T F.
\end{aligned}
$$

The first term in the resulting expression for $2\Delta T$ in (4.12) is always non-positive since $\Lambda_{Ni} \geq 0$, $g^T_{Ni}\dot{q}_- \leq 0$, $i \in J_N$. The second term can be positive if there is a j such that $g^T_{Fj}\dot{q}_- \neq 0$, and sign $(g^T_{Fj}\dot{q}_-) =$ sign (Λ_{Fj}), e.g. if sign $(g^T_{Fj}\dot{q}_+) = -$ sign $(g^T_{Fj}\dot{q}_-)$. However, if $g^T_{Fi}\dot{q}_- = 0$ or sign $(g^T_{Fi}\dot{q}_-) = -$ sign (Λ_{Fi}), $i \in J_F$, and $F=0$ then $\Delta T \leq 0$.

This friction impulse model was chosen for reasons of simplicity and consistency with the other contact and friction models in §§3.2, 3.3 and 4.2. It depends on the particular application if it is acceptable or not. The choice of model here is important since the difference in $\dot{q}_+$ for different models is of $\mathcal{O}(1)$.

4.4 Examples.

The method for systems obeying Coulomb´s friction law has been tested on two model problems. The dimension of the boxes in fig.4 is 3×1 and their density is 1. The friction coefficient $\mu = 0.5$ and there is no friction impulse $\mu_I = 0$. The gravity force is the only external force in both fig. 4 and 5. Initially, an external impulse $F_1^T = (F_{1x}, F_{1y}, F_{1\varphi}) = (10, 0, -10)$ is applied to body 1 simulating a push in the positive x-direction in the upper left corner. The bodies were at rest in the final picture in fig. 4. The system a [illegible] 2.0073 is statically indeterminate, since the G-matrix has 9 rows (three coordinates per body) and 11 columns (two multipliers per point contact and three per edge-to-edge contact, see §5). The configurations in the first and the last picture of fig. 5 are also indeterminate: G does not have full column rank. The dimension of the boxes is 2×1 and their density is 1. The values of the friction coefficients are $\mu = 0.5$ and $\mu_I = 0.8$. The initial velocity of body 2 is $\dot{x}(0) = 4$. All other initial velocities are zero. The system is at rest in the last picture in fig. 5. The error tolerance ε was 0.001 in both examples. The first example was also run with $\mu_I = 0.8$. With this choice of μ_I more zigzagging was observed due to the effects of the friction impulse model explained in §4.3.

5. Practical considerations.

In this section we shall discuss two details of general interest in the program that produced the figs. 2, 3, 4 and 5. The first issue is how the constraint functions ϕ_i and $g_{Fi}^T \dot{q}$ are defined. Then the data structure is described for the system of rigid bodies subject to contact and friction constraints.

Suppose that the number of corners of each body is k. Since there are k edges on each body, there are k^2 potential point contacts corner-edge for each body with respect to every other body.

In a system with n bodies there are $k^2n(n-1)$ possible constraint functions ϕ_i between the bodies and then there are kn possible ϕ_i between the bodies and the ground. Therefore, the introduction of a constraint in the program must be done with selection. Initially, one constraint function ϕ_i is defined for each point of contact between a corner and an edge. A contact between an edge of one body and an edge of another body is consequently described by ϕ_i at the corner in one end of the line of contact and by ϕ_{i+1} at the other corner. The user is monitoring the progress of the simulation interactively on a grafical terminal and introduces a new constraint $\phi_j > 0$ by a simple command when he discovers a collision risk. A constraint j is removed only when ϕ_j passes a positive threshold value. The bookkeeping of the constraints is done automatically in Cundall [7]. In order to register a collision a corner must hit the line defining an edge within the two ends of the edge. Observe that the condition $\phi_j(t_*)= 0$ is only a necessary condition for collision. If $\mu > 0$ and two bodies have collided at t_*, $\phi_j(t)> 0$ for $t < t_*$ but $\phi_j(t_*)= 0$, then the corresponding friction constraint function $g_{Fj}^T\dot{q}$ is introduced by the program at t_*. In an edge-to-edge contact the friction constraint j is associated with two contact constraints $\phi_j= \phi_{j+1}= 0$. The upper bound a_j for $|\lambda_{Fj}|$ in (4.6) is $\mu(\lambda_{Nj}+\lambda_{N,j+1})$

The data structure for rigid body systems such as those in the examples must contain easily accessible information on the properties of the bodies, the properties of the active constraints and relations between the bodies and the constraints. It should also be flexible enough to allow for simple creation and deletion of active constraints. From this data the discretized equations of motion (3.1) and (4.1) are assembled at each step. The data structure for a simple system is displayed in fig. 6. Each body and each active constraint are represented by a vertex in a directed graph.

The ground is a separate vertex. The arcs in the graph correspond to connections between the bodies and the constraints. The vertex A is the first vertex in the linked list of body vertices. Similarly, the vertex B is the head of the list of active constraints. For every constraint vertex there is a Lagrange multiplier to compute in (3.1) or (4.1). Constraints belonging to the same contact are connected by arcs. There are two arcs from a constraint to the two bodies involved in the constraint. It is a simple task to update the datastructure when constraints are to be inserted or deleted. The program was implemented in SIMULA [2]. SIMULA is a general purpose language where graph structures are manipilated conveniently. With the class-concept it is natural to represent the bodies and the constraints as instances of different classes. Each vertex in fig. 6 contains information on the body (geometry, mass, moment of inertia etc.) or the constraint (geometry of the contact, sign of μ etc.). The arcs are implemented as reference variables.

6. Generalizations.

The numerical procedure developed in §§3 and 4 was tested on two dimensional rigid body systems where the bodies had polyhedral boundary. It is a straight-forward generalization to introduce wheels in the systems. Then the constraint function ϕ_i denotes the shortest distance between the periphery of the wheel and a corner or an edge of a polyhedral body or the periphery of another wheel. The friction constraint function $g_F^T\dot{q}$ is the relative velocity in the direction tangential to the periphery of the wheel between the points of contact on the wheel and the other body. Extensions to arbitrary bodies are of course possible, but from a computational point of view, bodies with a simple geometry of their boundary are preferred.

Other kinds of constraints are possible in the framework of §§3 and 4. Since the holonomic constraint function ϕ_i never appears explicitly in the numerical formulation (3.1), we can obtain solutions to systems with non-holonomic constraint functions $g_i^T(q)\dot{q}$ by the same method. There are no upper or lower bounds on the Lagrange multipliers λ_i corresponding to bilateral constraints $\phi_i = g_i^T\dot{q} = 0$, $t \geq 0$. In the simple plasticity model for reinforcement steel bars in [28] there is only an upper bound on the Lagrange multiplier proportional to the force in the bar. After a few minor modifications of the algorithms in §§3.2 and 4.2, the last two examples of constraints can also be treated. The third space dimension is essential in many applications. In three dimensions M in (2.1) is still symmetric and positive definite almost everywhere but not constant, $M = M(q)$. The algorithms in §3 do not rely on M being constant and diagonal but on the existence of a factorization (2.2). Wittenburg [39] has developed a computational method for assembling M and f in (2.1) such that multipliers corresponding to bilateral constraints are eliminated. The method can be extended to include also systems with unilateral constraints. The constraint functions ϕ_i are defined in a way similar to the planar case, but the geometry of the bodies in three dimensions is more complex. An example of such a mechanical system is the human body, see §1. The angles between the over-arm and the fore-arm and the thigh and the lower part of the leg belong to the interval $[0,\pi)$ and the chin cannot penetrate the chest. Another potential area is the simulation of industrial robots. Coulomb friction in three space dimensions requires a non-trivial extension of algorithm 2. The natural way to generalize (2.9) is to introduce $\lambda_{Fi} = (\lambda_{Fi}^1, \lambda_{Fi}^2) \in R^2$. λ_{Fi}^1 and λ_{Fi}^2 are proportional to the two orthogonal components of the friction force on the slip surface. The condition on the modulus of λ_{Fi} that is reduced to (2.9a) in two dimensions is

$$(\lambda_{Fi}^{1})^2+(\lambda_{Fi}^{2})^2 \leq \mu^2\lambda_{Ni}^2. \qquad (6.1)$$

The constraint (6.1) on λ_{Fi} is quadratic whereas the constraints in (4.6) are linear. In special cases, the structure of the problem is such that the introduction of quadratic constraints on the multipliers is not necessary.

Finally, we wish to mention another area of engineering, where the governing equations resemble the equations and relations (2.1) and (2.6): the simulation of electrical networks. As an example, a simple model for a diode is considered. Suppose that in a branch of a network we have a diode, a resistor R and an inductance L. The branch voltage is v and i is the branch current. Then i satisfies

$$L\,\frac{di}{dt} + Ri = v+\lambda, \qquad i\geq 0,\ \lambda\geq 0,\ \lambda i = 0. \qquad (6.2)$$

The Lagrange multiplier λ can be regarded as the voltage drop over the diode. Electrical networks, whose components have this and similar behavior, can be simulated by the methods in §§3 and 4. In (6.2) the complementarity condition is particularly simple.

Figure captions.

Fig.1. Stability regions for the methods (AB-1, BDF-1) and (AB-2, BDF-2) applied to the test equation (3.6) are plotted in picture I and II. The methods are stable in the unshaded areas except for the corners where the difference equation has a double root $\zeta_1 = \zeta_2$, $|\zeta_1| = 1$.

Fig.2. The motion of five boxes piled on each other is simulated when the initial velocity of box 2 is 2.0 in the horizontal x-direction, $\mu = \mu_I = 0$, $\varepsilon = 0.001$.

Fig.3. The motion of a hexagon is studied. The body is suspended in a linear spring and damper with one end fixed in the inertial frame, $\mu = \mu_I = 0$, $\varepsilon = 0.001$.

Fig.4. The motion of three boxes colliding with each other is simulated when body 1 is given an initial push in the upper, left corner, $\mu = 0.5$, $\mu_I = 0$, $\varepsilon = 0.001$.

Fig.5. Body 2 has the initial velocity 4.0 in the horizontal x-direction. The motion of the system is simulated with $\mu = 0.5$, $\mu_I = 0.8$ and $\varepsilon = 0.001$.

Fig.6. The data structure in the lower half of the figure represents the mechanical system in the upper half. The constraints are numbered by roman numerals. The contact constraints are denoted by 'N' and the friction constraints by 'F'.

References.

1. J. Baumgarte, Stabilization of constraints and integrals of motion in dynamical systems, Comp. Meth. Appl. Mech. Eng., 1(1972), 1-16.

2. G.M. Birtwhistle, O.-J. Dahl, B. Myrhaug and K. Nygaard, SIMULA BEGIN, Auerbach, Philadelphia, PA, 1973.

3. Å. Björck and T. Elfving, Accelerated projection methods for computing pseudoinverse solutions of systems of linear equations, BIT, 19(1979), 145-163.

4. P. Businger and G.H. Golub, Linear least squares solutions by Householder transformations, Num. Math., 7(1[illegible]), 269-276.

5. A.K. Cline, C.B. Moler, G.W. Stewart and J.H. Wilkinson, An estimate for the condition number of a matrix SIAM J. Numer. Anal., 16(1979), 368-375.

6. R.W. Cottle, Numerical methods for complementarity problems in engineering and applied science, Computing methods in applied sciences and engineering, 1977, I, R. Glowinski and J.L. Lions, eds., Springer, Berlin, 1979, 37-52.

7. P.A. Cundall, Rational design of tunnel supports, AD/A-00162, Dept. of Civil and Mining Eng., Univ. of Minnesota, Minneapolis, MN, 1974.

8. G. Dahlquist and Å. Björck, Numerical methods, Prentice-Hall, Englewood Cliffs, NJ, 1974.

9. J. Dieudonné, Foundations of modern analysis, Academic Press, New York, 1969.

10. W.S. Dorn, Duality in quadratic programming, Quart. Appl. Math., 18(1960), 155-162.

11. H. Dugoff and R.W. Murphy, The dynamic performance of articulated highway vehicles-A review of the state of the art, SAE paper 710223, 1971.

12. G.E. Fleischer, Multibody system applications and simulations at the Jet Propulsion Laboratory, Dynamics of multibody systems, K. Magnus, ed., Springer, Berlin, 1978, 36-47.

13. R. Fletcher, Practical methods of optimization, vol. 2, Constrained optimization, John Wiley & Sons, Chichester-New York, 1981.

14. B. Fredriksson, On elastostatic contact problems with friction, Ph.D. thesis, Dept. of Mech. Eng., Linköping Inst. of Techn., Linköping, 1976.

15. C.W. Gear, Numerical initial value problems in ordinary differential equations, Prentice-Hall, Englewood Cliffs, NJ, 1971.

16. P.E. Gill, G.H. Golub, W. Murray and M.A. Saunders, Methods for modifying matrix factorizations, Math. Comp., 28(1974), 505-535.

17. P.E. Gill and W. Murray, Minimization of a non-linear function subject to bounds on the variables, NAC-72, National Physical Laboratories, Teddington, Middlesex, 1976.

18. P.E. Gill and W. Murray, Numerically stable methods for quadratic programming, Math. Prog., 14(1978), 349-372.

19. E. Haug, R. Chand and K. Pan, Multibody elastic contact analysis by quadratic programming, J. Opt. Theory Appl., 21(1977), 189-198.

20. R.L. Huston, R.E. Hessel and J.M. Winget, Dynamics of a crash victim-A finite segment model, AIAA J., 14(1976), 173-178.

21. J.J. Kalker, A survey of the mechanics of contact between solid bodies, ZAMM , 57(1977), T3-T17..

22. S. Karamardian, The nonlinear complementarity problem with applications, part I, J. Opt. Theory Appl., 4(1969), 87-98.

23. I. Karasalo, A criterion for truncation of the QR-decomposition algorithm for the singular linear least squares problem, BIT, 14(1974), 156-166.

24. A. Kerr and J.L. Blair, Simulation of the longitudinal dynamics of a train, Dynamics of multibody systems, K. Magnus, ed., Springer, Berlin, 1978, 108-119.

25. C.L Lawson and R.J. Hanson, Solving least squares problems, Prentice-Hall, Englewood Cliffs, NJ, 1974.

26. P. Lötstedt, On a penalty function method for the simulation of mechanical systems subject to constraints, TRITA-NA-7919, Dept. of Num. Anal. and Comp. Science, Royal Inst. of Techn., Stockholm, 1979.

27. P. Lötstedt, A numerical method for the simulation of mechanical systems with unilateral constraints, TRITA-NA-7920, Dept. of Num. Anal. and Comp. Science, Royal Inst. of Techn., Stockholm, 1979.

28. P. Lötstedt, Interactive simulation of the progressive collapse of a building, revisited, TRITA-NA-7921, Dept. of Num. Anal. and Comp. Science, Royal Inst. of Techn., Stockholm, 1979.

29. P. Lötstedt, Coulomb friction in two-dimensional rigid body systems, ZAMM, 61(1981), 605-615.

30. P. Lötstedt, Mechanical systems of rigid bodies subject to unilateral constraints, SIAM J. Appl. Math., 42(1982), 281-296.

31. P. Lötstedt, Perturbation bounds for the linear least squares problem subject to linear inequality constraints, TRITA-NA-8205, Dept. of Num. Anal. and Comp. Science, Royal Inst. of Techn., Stockholm, 1982. (accepted for publication in BIT)

32. P. Lötstedt, Solving the minimal least squares problem subject to bounds on the variables, TRITA-NA-8215, Dept. of Num. Anal. and Comp. Science, Royal Inst. of Techn., Stockholm 1982. (to appear)

33. N. Orlandea and D.A. Calahan, A sparsity-oriented approach to the design of mechanical systems, Problem analysis in science and engineering, F.H. Branin, Jr. and K. Huseyin, eds., Academic Press, New York, 1977, 361-389.

34. B. Paul, Analytical dynamics of mechanisms-A computer oriented overview, Mech. Mach. Theory, 10(1975), 481-507.

35. L. Petzold, Differential/algebraic equations are not ODEs, SAND 81-8668, Sandia National Laboratories, Livermore, CA, 1981.

36. M.R. Ramey and A.T Yang, A simulation procedure for human motion studies, J. Biomechanics, 14(1981), 203-213.

37. R. Schwertassek, Der Roberson/Wittenburg Formalismus and das Programmsystem MULTIBODY zur Rechnersimulation von Mehrkörper - systemen, DFVLR-FB 78-08, Deutsche Forschungs und Versuchsanstalt für Luft und Raumfahrt, Oberpfaffenhofen, 1978.

38. P.N. Sheth and J.J. Uicker, Jr. , IMP (Integrated Mechanisms Program), A computer-aided design analysis system for mechanisms and linkage, Trans. ASME, J. Eng. Indust., Ser. B, 94(1972), 454-464.

39. J. Wittenburg, Dynamics of systems of rigid bodies, Teubner, Stuttgart, 1977.

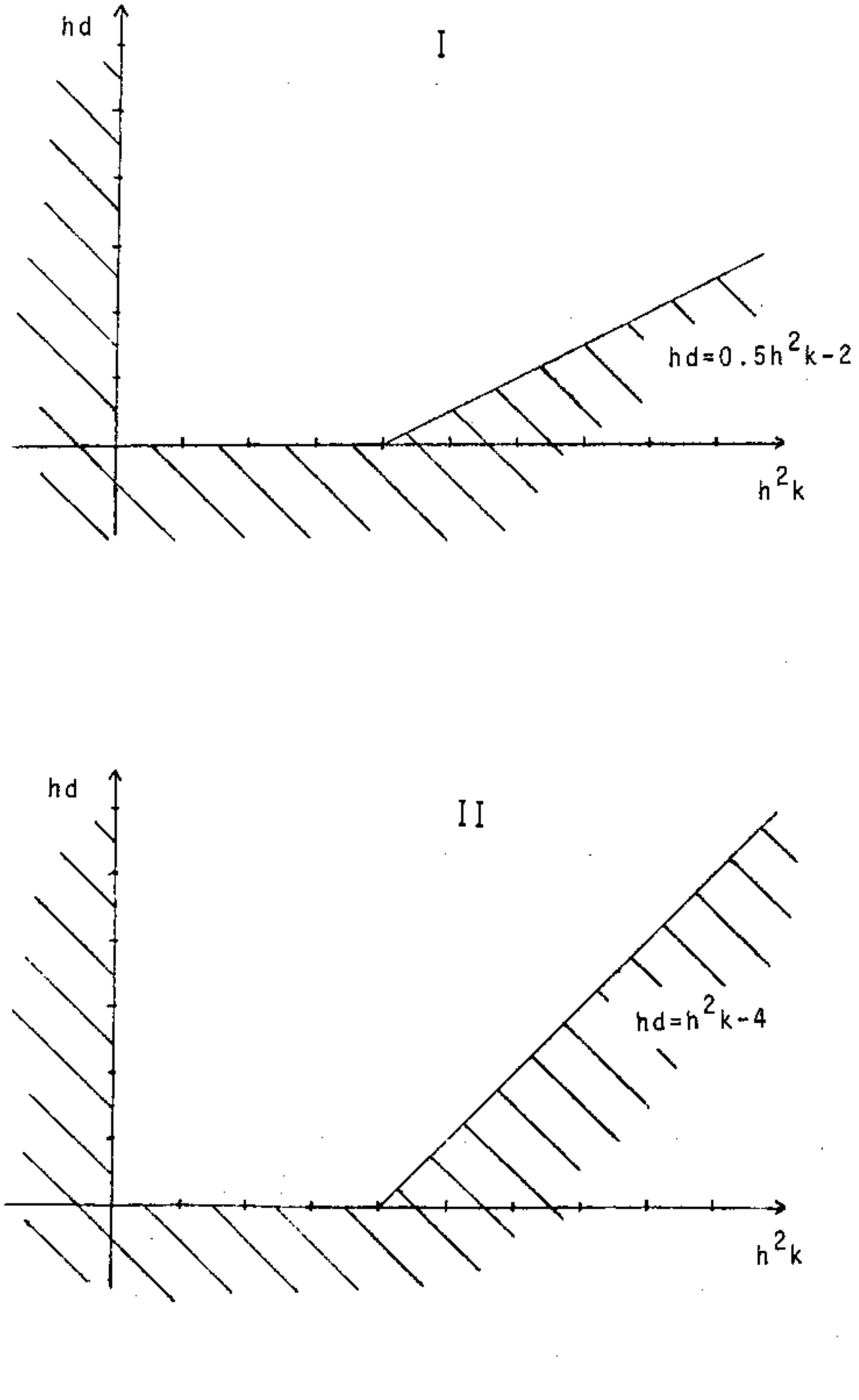

Fig. 1.

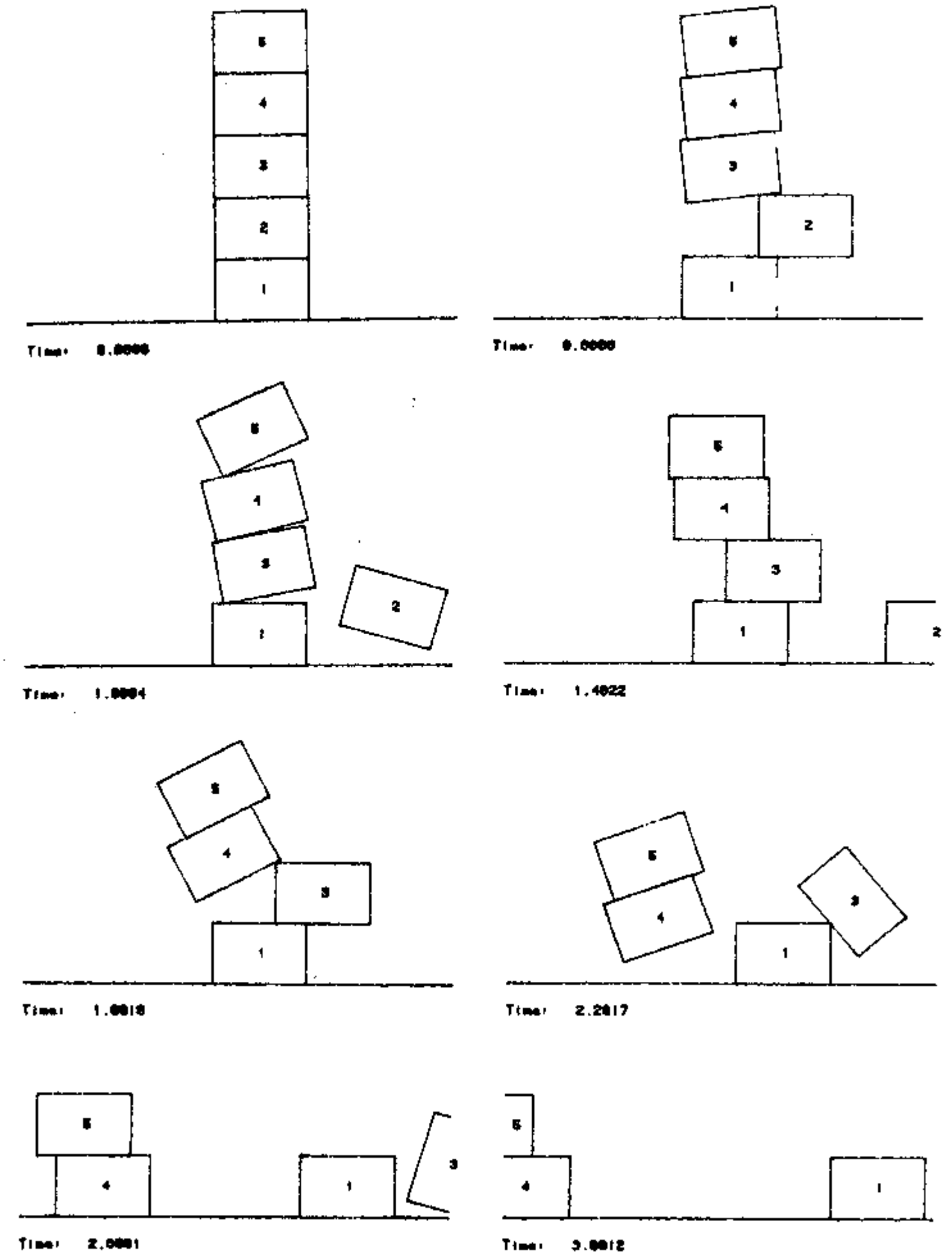

Fig. 2.

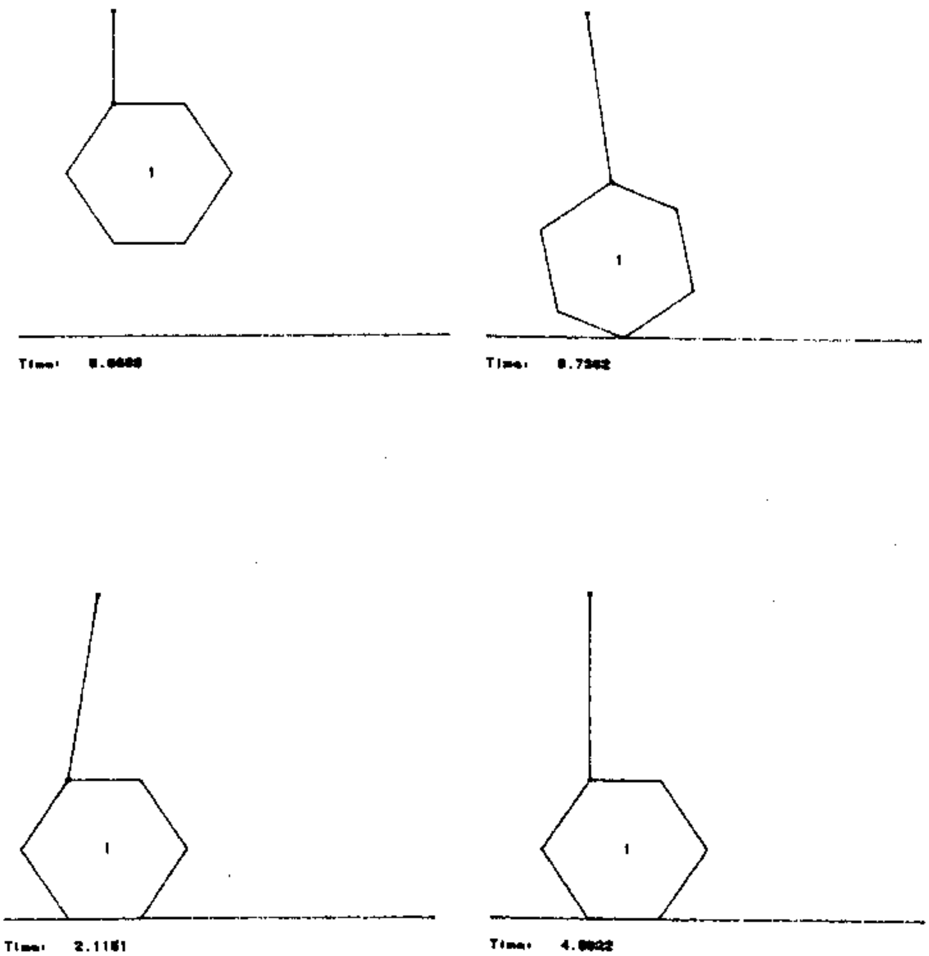

Fig. 3.

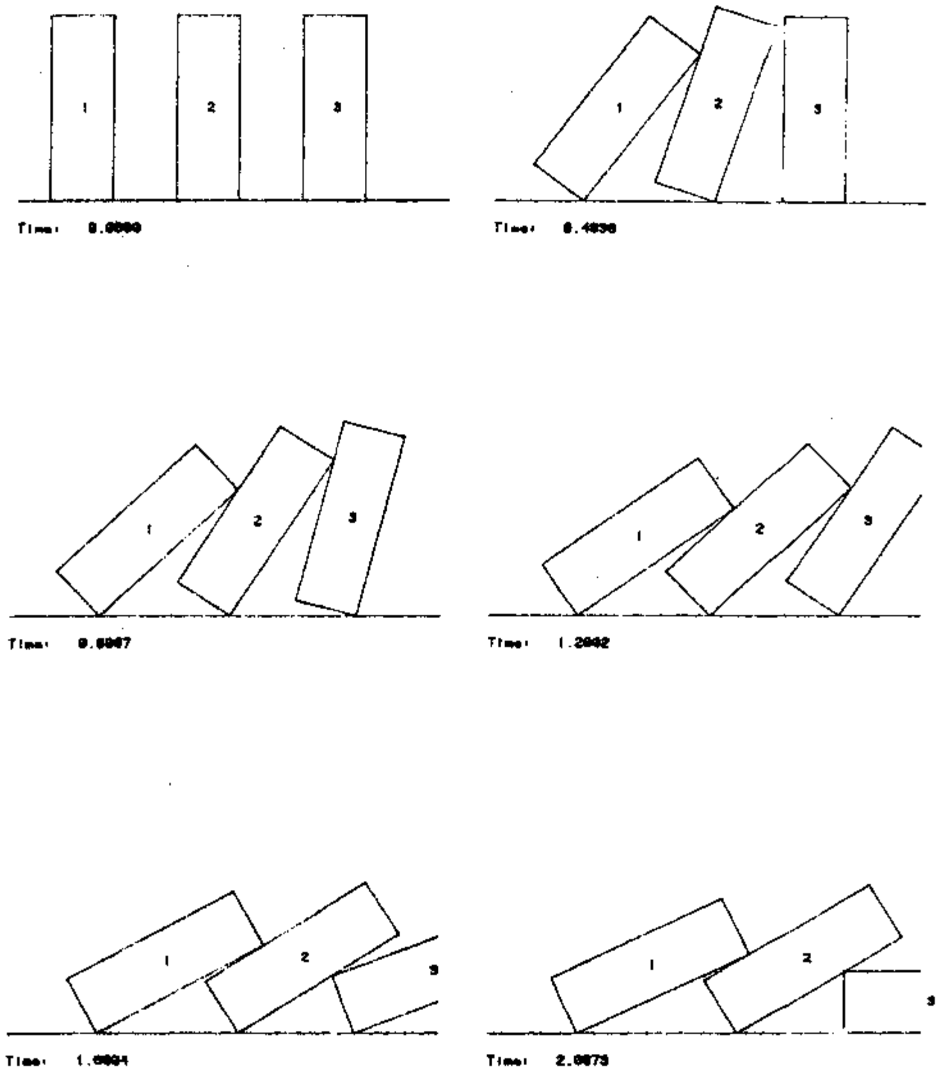
1
2
3
Time: 0.0000
Time: 0.4836
Time: 0.8807
Time: 1.2842
Time: 1.6804
Time: 2.0873

Fig. 4.

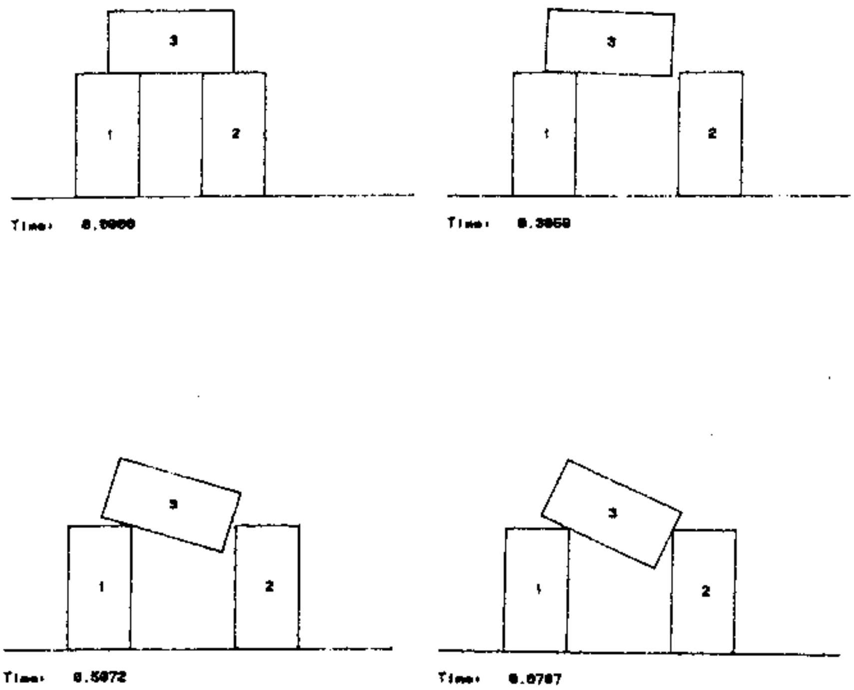

Fig. 5.

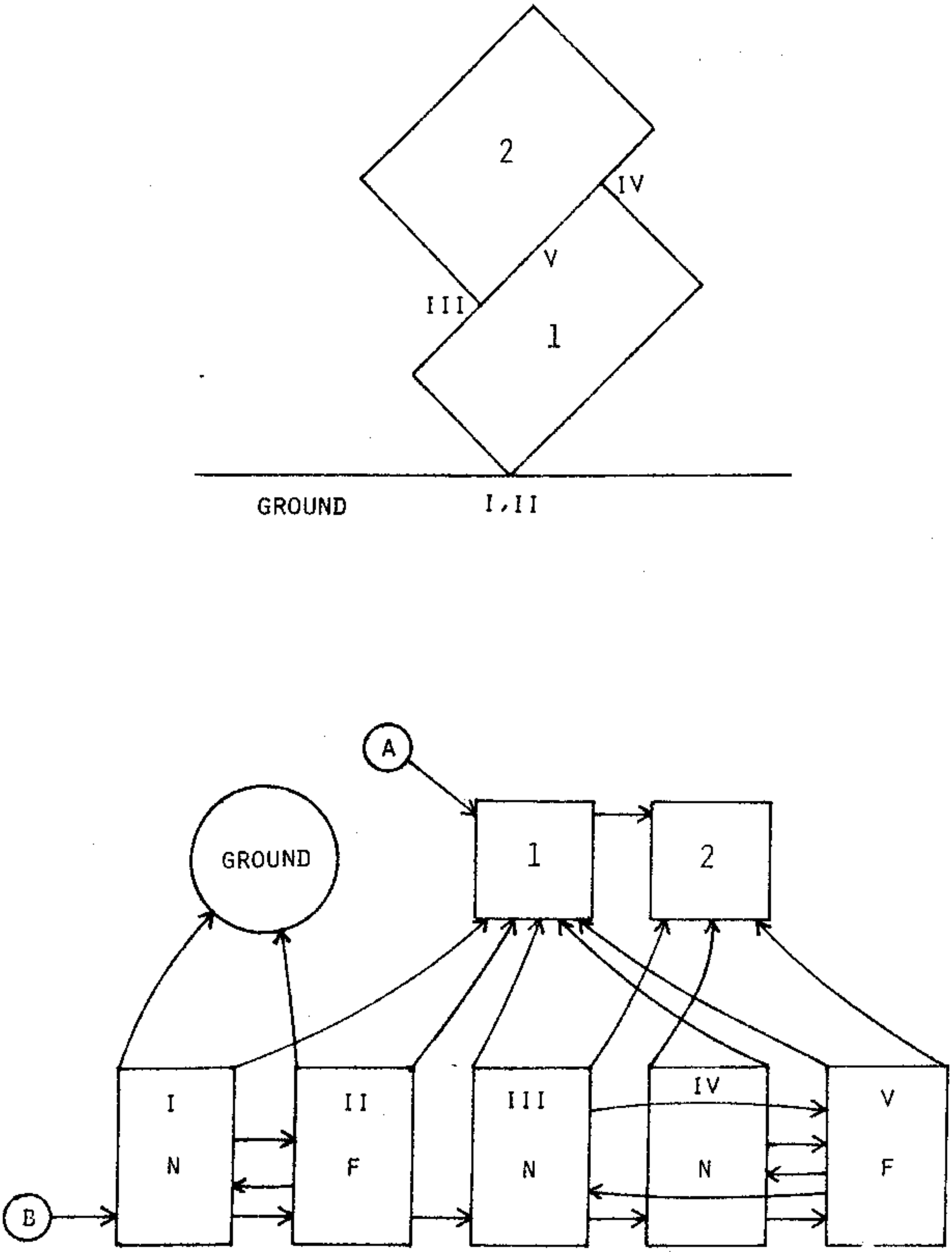

Fig. 6.