
Modeling Contact Friction and Joint Friction in Dynamic Robotic Simulation using the Principle of Maximum Dissipation

Evan Drumwright and Dylan A. Shell

¹ Evan Drumwright at George Washington University (edrumwri@gmail.com)

² Dylan A. Shell at Texas A&M University (dshell@cs.tamu.edu)

Summary. We present a unified treatment for modeling Coulomb and viscous friction within multi-rigid body simulation using the *principle of maximum dissipation*. This principle is used to build two different methods—an event-driven impulse-based method and a time stepping method—for modeling contact. The same principle is used to effect joint friction in articulated mechanisms. Experiments show that the contact models are able to be solved faster and more robustly than alternative models. Experiments on the joint friction model show that it is as accurate as a standard model while permitting much larger simulation step sizes to be employed.

1 Introduction

Rigid body dynamics is used extensively within robotics in order to simulate robots, learn optimal controls, and develop inverse dynamics controllers. The forward dynamics of rigid bodies in free space has been well understood for some time, and the recent advent of differential algebraic equation (DAE) based methods has made the dynamics of bodies in contact straightforward to compute as well. However, fast and stable robotic simulation remains somewhat elusive. The numerical algorithms used to compute contact forces run (on average) in time $O(n^3)$ in the number of contact points and are numerically brittle. In this paper, we present a class of methods for modeling contact with friction in multi-rigid body simulation that is not only faster empirically than existing methods, but is solvable with numerical robustness. We also extend our approach to modeling joint friction and show how it is at least as accurate as standard joint friction models, while permitting much larger simulation step sizes (and thus much greater simulation speed).

Our approach centers around the *principle of maximal dissipation*. Paraphrasing Stewart (2000), the principle of maximal dissipation states that, for bodies in contact, the friction force is the one force (of all possible friction forces) that maximizes the rate of energy dissipation. In this paper, we show that we can use the principle of maximal dissipation to implicitly solve for frictional forces, in contrast to prior approaches that set the frictional force by explicitly using the direction of relative motion. By employing this strategy, we are able to dispense with the complementarity constraints that are nearly universally employed and instead formulate the contact problem using a convex optimization model.

2 Methods for modeling contact with friction in multi-rigid body simulation

2.1 Background

As stated in the previous section, modeling contact with friction in multi-rigid body simulation has been extensively conducted using complementarity constraints. Such constraints, which take the form $\mathbf{a} \geq 0$, $\mathbf{b} \geq 0$, $\mathbf{a}^\top \mathbf{b} = 0$, have been utilized in multi-rigid body simulation to ensure that forces are not applied at contact points at which bodies are separating (Anitescu and Potra, 1997), that either sticking or sliding friction is applied (Anitescu and Potra, 1997), and that forces are applied only for joints at their limits (Miller and Christensen, 2003). A non-exhaustive survey of the literature dedicated to modeling contact with complementarity constraints includes Löstedt (1982); Moreau (1983); Löstedt (1984); Moreau (1985, 1988); Monteiro-Marques (1993); Baraff (1994); Anitescu et al (1997); Pfeiffer and Glocker (1996); Brogliato (1996); Trinkle et al (1997); Anitescu and Potra (1997); Anitescu et al (1999); Stewart and Trinkle (2000); Anitescu and Potra (2002); Anitescu and Hart (2004); Trinkle et al (2005); Anitescu (2006); Potra et al (2006); Erleben (2007); Anitescu and Tasora (2008); Petra et al (2009); Todorov (2010); Tassa and Todorov (2010). Initial efforts on modeling contact with friction (Löstedt, 1982, 1984; Baraff, 1994) attempted to solve a linear complementarity problem (LCP) for the unknown forces and accelerations. Later work—under which researchers realized that solving for forces and accelerations could be subject to *inconsistent configurations*³ (Baraff, 1991)—solved instead for unknown impulsive forces and velocities; this approach was able to avoid the problem of inconsistent configurations. Both iterative sequential impulse schemes (*e.g.*, Mirtich (1996b); Guendelman et al (2003)) and simultaneous impulse-based methods (*e.g.*, Anitescu and Potra (1997); Stewart and Trinkle (2000)) were employed. The former methods have been viewed as *splitting methods* (Cottle et al, 1992) for solving linear complementarity problems by Lacoursière, who reported slow convergence rates for coupled problems using such approaches (Lacoursière, 2003). Correspondingly, much of the multi rigid-body simulation community currently uses explicit linear or nonlinear complementarity problem formulations (Anitescu and Potra (1997), most prominently) with impulses and velocities in order to model contact with friction.

Figure 1 illustrates why complementarity conditions are necessary at the acceleration level; further discussion of their requirement at the acceleration level is present in Baraff (1994) and Chatterjee (1999). Given that using impulsive forces is necessary to avoid the problem of inconsistent configurations (exemplified by the problem of Painlevé (1895)), it must be discerned whether the complementarity conditions are accurate and appropriate in that new context (*i.e.*, at the velocity level). We stress that the context is indeed new, because only “resting” (zero relative normal velocity) contacts are treated with forces at the acceleration level⁴ while both resting and impacting contact are treated with impulses at the velocity level.

With respect to accuracy, Chatterjee (1999) argues that the complementarity conditions do not necessarily reflect reality. He conducts an experiment using carbon paper, a coin, and a hammer, that serves to prove his argument: the physical accuracy of complementarity-based models, at least in some scenarios, is lacking. Further physical experimentation is warranted,

³ An inconsistent configuration is a contact configuration that has no solution using non-impulsive forces.

⁴ Unless a penalty method is used; such methods are irrelevant to this discussion because we are assuming non-interpenetrating contact.

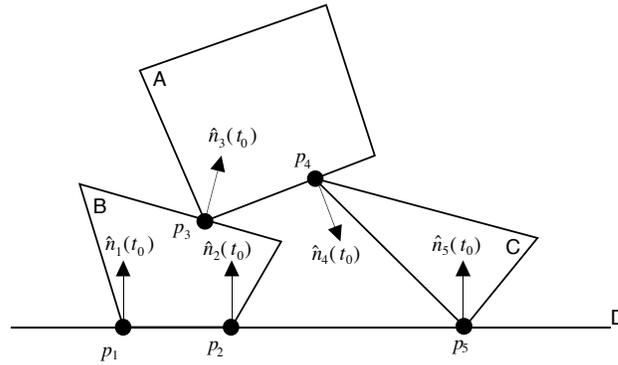


Fig. 1. A figure (taken from Baraff (1997)) illustrating the necessity of complementarity constraints for contact at the acceleration level. Complementarity constraints would ensure that body A does not move upward unnaturally fast if a strong wind were to move between B and A, accelerating A upward. In the absence of this wind, the complementarity condition keeps blocks A and B from interpenetrating.

especially given the attention to complementarity constraints at the velocity level in the literature, but if we accept that nature does not require complementarity constraints, do such conditions lead to models that are more advantageous in some other way (*e.g.*, computationally?)

Linear and nonlinear complementarity problem-based models are, in fact, inferior with respect to computation, at least relative to the models introduced in this paper. Solutions to bisymmetric LCPs—the form that many models take—are NP-hard in the worst case in the number of contact points (though the expected time solution is $O(n^3)$ (Cottle et al, 1992)). The contact models are non-convex, so only a few algorithms are capable of solving LCP-based contact models; codes for solving the NCP-based models are even more rare.⁵ The LCP-based models linearize the friction cone, and the fidelity of the friction cone approximation increases the size and, correspondingly, the computational cost of the model to be solved.

Given that complementarity conditions may not be physically warranted and that such models can be hard to solve computationally, what is the reason behind their strong popularity? We believe this quote from Chatterjee (1999) is instructive:

It is emphasized that many authors in this area are aware of the possible physical inaccuracies behind the complementarity assumption. For example, Moreau (1983) states that a primary benefit of such a formulation is internal mathematical consistency, and empirical corrections towards better accuracy can be accommodated later. Baraff (personal communication) mentions that there is no reason to think that real systems obey the complementarity conditions. However, in Pfeiffer and Glocker’s (1996) discussion of the “corner law of contact dynamics” there is unfortunately no explicit mention of the possible lack of physical realism behind the complementarity

⁵ The reader may question the desire to have multiple algorithms capable of solving a contact model. Optimization algorithms can be numerically brittle, so using multiple algorithms can reduce the likelihood of a simulation crashing.

conditions in the presence of impacts. Anitescu and Potra (1997) focus on mathematical aspects of their formulation, and also omit discussion of physical realism. The authors of the latter two authoritative works may therefore unintentionally convey an inaccurate impression to a reader who is new to the field.

2.2 Two representations for the contact models

Given that complementarity conditions are not a necessary feature of contact models, we now proceed to present our complementarity-free contact models. We will use two representations for these contact models. Both representations have been employed previously in the literature. We utilize the two representations here in order to make the community aware of their existence, to unify them, and, hopefully, to expose them to further study (for determination of computational or numerical advantages, for example).

A / \mathbf{b} representation

The first representation formulates the contact model in terms of matrices \mathbf{A} and \mathbf{b} (alternatively named \mathbf{K} and \mathbf{u} by Mirtich (1996b)). This representation has been employed by Baraff (1994), Mirtich (1996b), Kokkevis (2004), and Drumwright and Shell (2009). In this representation, \mathbf{b} is a vector of relative velocities in the $3n$ -dimensional contact space and \mathbf{A} is the $3n \times 3n$ sized *contact space inertia matrix* that transforms impulsive forces to relative velocities in contact space. \mathbf{A} is dense, symmetric, and positive-semi definite; the latter two properties were proven by Mirtich (1996b). The matrices \mathbf{A} and \mathbf{b} are related by the equation:

$$\mathbf{b}^+ = \mathbf{A}\mathbf{x} + \mathbf{b}^- \quad (1)$$

where \mathbf{b}^- is the vector of relative velocities pre-contact (external forces, such as gravity, are integrated into this vector), \mathbf{x} is the vector of impulses applied in the contact space, and \mathbf{b}^+ is the vector of relative velocities after impulses are applied.

The matrices \mathbf{A} and \mathbf{b} can be determined formulaically if the contacting bodies are not articulated, and via application of test impulses otherwise (*cf.*, Mirtich (1996b); Kokkevis (2004)). By using test impulses, the contact model can remain ignorant of whether the bodies are articulated and, if articulated, of whether the bodies are formulated using maximal or reduced coordinates. This is an advantage of the \mathbf{A} / \mathbf{b} representation; this representation also tends to produce simpler—though not necessarily computationally advantageous—objective and constraint function gradients and Hessians for solving the contact model.

Generalized coordinate representation

The generalized coordinate representation has been used in work by Anitescu and Potra (1997) and Stewart and Trinkle (2000), among others. This representation uses matrices \mathbf{M} (the generalized inertia matrix), \mathbf{N} (the contact normal Jacobian), \mathbf{J} (the joint constraint Jacobian); vectors \mathbf{c}_n (the contact normal impulse magnitudes), \mathbf{c}_j (the joint constraint impulse magnitudes), \mathbf{q} (the generalized coordinates), \mathbf{v} (the generalized velocities) and \mathbf{k} (the generalized external forces), and scalar h (the step size).

For models with complementarity constraints, additional matrices are used, both to enforce the complementarity constraints and to provide a linearized friction cone; we do not list such matrices here. Because our model does not employ complementarity constraints, it is

able to provide a true friction cone and still remain computationally tractable. The true friction cone is obtained using Jacobian matrices corresponding to the two tangential directions at each contact normal; we denote these matrices \mathbf{S} and \mathbf{T} and the corresponding contact tangent impulses as \mathbf{c}_s and \mathbf{c}_t .

All of the matrices described above are related using the following formulae:

$$\mathbf{v}^{t+1} = \mathbf{M}^{-1}(\mathbf{N}^T \mathbf{c}_n + \mathbf{S}^T \mathbf{c}_s + \mathbf{T}^T \mathbf{c}_t + \mathbf{J}^T \mathbf{c}_j + h\mathbf{k}) + \mathbf{v}^t \quad (2)$$

$$\mathbf{q}^{t+1} = \mathbf{q}^t + h\mathbf{v}^{t+1} \quad (3)$$

The second equation reflects that this representation is typically utilized in a semi-implicit integration scheme.

Note that \mathbf{M} is symmetric, positive-definite. \mathbf{M} , \mathbf{N} , \mathbf{S} , \mathbf{T} and \mathbf{J} are sparse and correspondingly make efficient determination of the objective and constraint gradients and Hessians conceptually more involved than with the \mathbf{A} / \mathbf{b} representation. Computing \mathbf{M} , \mathbf{N} , \mathbf{S} , \mathbf{T} and \mathbf{J} is quite simple, however, and computationally far more efficient than the test impulse method. We show below, however, that \mathbf{A} and \mathbf{b} can also be determined efficiently using the matrices from the generalized coordinate representation.

Unification of the \mathbf{A} / \mathbf{b} and generalized coordinate representations

The equations below show how \mathbf{b}^- and \mathbf{A} may be obtained using the generalized coordinate representation.

$$\mathbf{b}^- = \begin{bmatrix} \mathbf{N} \\ \mathbf{S} \\ \mathbf{T} \end{bmatrix} (\mathbf{v}^t + h\mathbf{M}^{-1}\mathbf{k}) \quad (4)$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{N} \\ \mathbf{S} \\ \mathbf{T} \end{bmatrix} \mathbf{C} [\mathbf{N}^T \ \mathbf{S}^T \ \mathbf{T}^T] \quad (5)$$

where $\mathbf{C} \triangleq \mathbf{M}^{-1} - \mathbf{M}^{-1}\mathbf{J}^T(\mathbf{J}\mathbf{M}^{-1}\mathbf{J}^T)^{-1}\mathbf{J}\mathbf{M}^{-1}$.

These equations show that recovering \mathbf{N} , \mathbf{S} , \mathbf{T} , and \mathbf{J} from \mathbf{A} and \mathbf{b}^- is not possible.

2.3 The contact models for event-driven and time-stepping simulation

Computer-based rigid body simulation methods have been categorized into three schemes by Brogliato et al (2002): penalty, event-driven, and time-stepping. Given that penalty methods necessarily cannot enforce non-interpenetration, we focus instead on event-driven and time-stepping methods. The former are able to utilize arbitrary integration schemes (time-stepping methods are frequently restricted to semi-implicit Euler integration) while the latter aim to be able to avoid (possibly nonexistent) accuracy problems due to continually restarting the integration process; Brogliato et al (2002) provide greater detail of the rationale behind development of the time-stepping approaches. We note that Drumwright (2010) has shown that one purported advantage of time-stepping approaches over event-driven simulations—avoidance of Zeno points—is nonexistent. Sections 2.4 and 2.5 present both event-driven and time-stepping methods in order to show that our method is applicable to both. Section 2.4 uses the \mathbf{A}/\mathbf{b} representation, while Section 2.5 uses the generalized coordinate representation.

2.4 Event-driven impulse-based method

Mirtich (1996b) defines the work done by collision impulses using the \mathbf{A}/\mathbf{b} representation as $\frac{1}{2}\mathbf{x}^\top(\mathbf{A}\mathbf{x} + 2\mathbf{b}^-)$. We can use the principle of maximal dissipation to determine the set of impulses that maximally dissipate kinetic energy. In particular, we can formulate and solve the following optimization problem.

Quadratic Program 1

Minimize	$\frac{1}{2}\mathbf{x}^\top(\mathbf{A}\mathbf{x} + 2\mathbf{b}^-)$	
Subject to:		
	$[\mathbf{A}_N \ \mathbf{B}] \mathbf{x} + \mathbf{b}_N^- \geq 0$	(Noninterpenetration constraint)
	$\mathbf{x}_N \geq 0$	(Compressive force constraint)
	$\mathbf{1}^\top \mathbf{x}_N \leq \kappa$	(Friction cone constraint)
	$\mu_c^2 \mathbf{x}_{N_i}^2 + \mu_v^2 (\mathbf{b}_{T_{1i}}^2 + \mathbf{b}_{T_{2i}}^2) \geq \mathbf{x}_{T_{1i}}^2 + \mathbf{x}_{T_{2i}}^2 \quad \forall i \in 1 \dots n.$	(Coulomb/viscous friction constraint)

where \mathbf{A} can be partitioned into $n \times n$ block \mathbf{A}_N , $n \times 2n$ block \mathbf{B} , and $2n \times 2n$ block \mathbf{A}_T (i.e., $\mathbf{A} = \begin{bmatrix} \mathbf{A}_N & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{A}_T \end{bmatrix}$) and, similarly, \mathbf{x} and \mathbf{b} can be partitioned into $\mathbf{x} = [\mathbf{x}_N \ \mathbf{x}_{T_1} \ \mathbf{x}_{T_2}]^\top$ and $\mathbf{b} = [\mathbf{b}_N \ \mathbf{b}_{T_1} \ \mathbf{b}_{T_2}]^\top$, respectively. Scalar κ is the sum of the normal impulses that describe the minimum kinetic energy solution when the tangential impulses are zero; this scalar is described further in Section 2.6.

2.5 Time stepping method

Equivalent in spirit to the event-driven method is the time-stepping method, which, like that of Anitescu and Potra (1997), is a semi-implicit scheme and, like that of Stewart and Trinkle (2000), adds contact constraint stabilization to the dynamics equations.

Quadratic Program 2

Minimize	$\frac{1}{2}\mathbf{v}^{t+1\top} \mathbf{M} \mathbf{v}^{t+1}$	
Subject to:		
	$\mathbf{N} \mathbf{v}^{t+1} \geq 0$	(Noninterpenetration constraint)
	$\mathbf{J} \mathbf{v}^{t+1} = 0$	(Bilateral joint constraint)
	$\mathbf{c}_n \geq 0$	(Compressive force constraint)
	$\mathbf{1}^\top \mathbf{c}_n \leq \kappa$	(Friction cone constraint)
	$\mu_c^2 \mathbf{c}_{n_i}^2 + \mu_v^2 [\mathbf{S}_i (\mathbf{v}^t + h\mathbf{M}^{-1}\mathbf{k})]^2 + \mu_v^2 [\mathbf{T}_i (\mathbf{v}^t + h\mathbf{M}^{-1}\mathbf{k})]^2 \geq \mathbf{c}_{s_i}^2 + \mathbf{c}_{t_i}^2 \quad \forall i \in 1 \dots n.$	(Coulomb/viscous friction constraint)

where \mathbf{S}_i and \mathbf{T}_i refer to the i^{th} row of \mathbf{S} and \mathbf{T} , respectively. Unlike the event-driven method, the time-stepping method includes a constraint for bilateral joints in case the method

is used with maximal-coordinate formulated articulated bodies; the \mathbf{A}/\mathbf{b} representation implicitly encodes such constraints into the matrix / vector formulations.

2.6 Solving the models

Both of the contact models introduced in the previous sections are convex and can be solved in polynomial time in the number of contacts using interior-point methods. Determining the value κ requires solving the models in an initial, frictionless phase (phase I). Next, a frictional phase (phase II) is solved. If normal restitution is necessary, a third phase is required as well.⁶ Fortunately, phase I is a quadratic programming model with box constraints, and can thus be solved extremely quickly using a gradient projection method (Nocedal and Wright, 2006). The model for phase I (\mathbf{A}/\mathbf{b} formulation) is:

Quadratic Program 3	
Minimize	$\frac{1}{2} \mathbf{x}_N^T \mathbf{A}_N \mathbf{x}_N + \mathbf{x}_N^T \mathbf{b}_N^-$
Subject to:	$\mathbf{x}_N \geq 0$ (Compressive force constraint)

Given that \mathbf{A}_N is symmetric and positive semi-definite (follows from symmetry and positive semi-definiteness of \mathbf{A}), this model describes a convex linear complementarity problem (see Cottle et al (1992), p. 5). As a result, the constraint $\mathbf{A}_N \mathbf{x}_N + \mathbf{b}_N^- \geq 0$ (which is equivalent to our non-interpenetration constraint) is automatically satisfied at the optimum. This convex LCP always has a solution and we prove that this solution does not increase the energy in the system (see Appendix). As phase II cannot increase the energy in the system our contact models are energetically consistent.

Phase I has two objectives: determine an energetically consistent, *feasible point*⁷ for the nonlinear Quadratic Program I and determine κ . The value κ is determined by calculating the sum $\mathbf{1}^T \mathbf{x}_N$, using the result from phase I. Observe that the friction cone inequality constraint in Quadratic Program I restricts the sum of the normal impulses to that determined in phase I; thus, phase II can reorder, transfer, or remove some normal force, but the friction cone is prevented from becoming enlarged arbitrarily (*i.e.*, the normal forces cannot be increased without bound to increase the amount of frictional force applicable) in order to decrease the kinetic energy more rapidly.

We note that, although we use a linear complementarity problem formulation to show that our contact models are energetically consistent, our contact models do not use complementarity constraints: neither phase I nor phase II of the contact models explicitly include any such constraint.

⁶ We omit details of this third phase but refer the reader to Anitescu and Potra (1997), which describes a Poisson-type restitution model that is applicable to our models also.

⁷ A feasible point is one which respects all inequality constraints. In the case of our contact model, a feasible point will respect non-interpenetration, compressive normal force, summed normal force, and Coulomb and viscous friction constraints.

3 Evaluating the contact models

There does not currently exist a standardized set of benchmark models for evaluating the performance and accuracy of contact models. In order to evaluate contact models, previous approaches have used either physical experimentation on a set of benchmark scenarios (*e.g.*, Stoianovici and Hurmuzlu (1996)) or pathological computer-based scenarios (*e.g.*, Mirtich (1996b); Sauer et al (1998); Ceanga and Hurmuzlu (2001)) that are known to exhibit certain behaviors in the real world.

Like Anitescu and Potra (1997), we do not provide exhaustive experimentation to indicate the predictive abilities of our contact model. Instead, we note that our model possesses the following properties, which are also possessed by leading alternative contact models that treat multiple contact points simultaneously (*e.g.*, (Anitescu and Potra, 1997; Stewart and Trinkle, 2000)):

1. Energy is not added to the system.
2. Interpenetration constraints are not violated.
3. Only compressive normal forces are applied.
4. The friction cone is not enlarged artificially.
5. The principle of maximal dissipation is obeyed.

Given that these are the main characterizations of both systems, we can expect the emergent behavior of both the complementarity-based methods and our non-complementarity-based methods to be similar. Indeed, as Figure 2 shows, the method of Anitescu and Potra (solved using Lemke’s algorithm (Lemke, 1965; Cottle et al, 1992)) produces identical results (to numerical precision) to our non-complementarity-based model on at least one scenario.

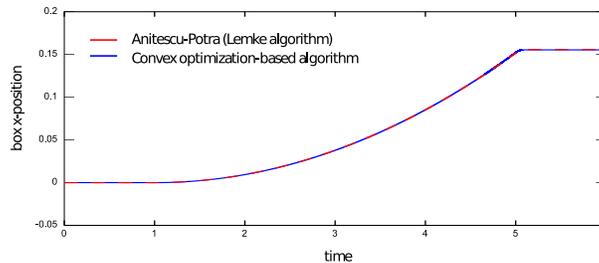


Fig. 2. Plot illustrating accuracy of a box sliding down a ramp under both the method of Anitescu and Potra (1997) and the convex optimization-based method introduced in this paper. Although the two models are formulated quite differently, the simulated results are identical.

We believe that our method is no more accurate than that of Anitescu and Potra or Stewart and Trinkle (notwithstanding the linearized friction cones generally employed by those two methods). The advantage of our method lies in the computational domain. Lacking complementarity constraints, our model runs faster than competing methods; additionally, the gradient projection method used to solve the first phase of our algorithm—recall that the first phase of our algorithm finds a point that respects properties (1)–(4) above—never fails to produce

a solution; failure in phase II of our method will only affect the accuracy of the solution and will not “break” the simulation.

The experiments below showcase these advantages of our method.

4 Experiments

4.1 Event-driven example: internal combustion engine

We constructed an inline, four cylinder internal combustion engine in order to illustrate the ability of our method to treat moderate numbers of contacts (between one and two hundred) far faster than complementarity-based models. All joints within the simulated engine were realized by contact constraints only. The engine was simulated by applying a torque to the crankshaft, which caused the cylinders to move upward and downward within the crankcase. The crankcase and crankshaft guides are stationary within the simulation—they possess infinite inertia, so they do not move dynamically—and the remaining parts are all of mass $1.0kg$. Although the masses do not reflect reality, the moment-of-inertia matrix for each part is calculated using its triangle mesh geometry via the method of Mirtich (1996a). Gravity acts along the vertical direction of the simulation at $9.8m/s^2$ and the engine surfaces are given zero coefficients of Coulomb and viscous friction: we initially wanted to judge how rapidly the polygonal-based geometric representation causes energy loss.

For purposes of comparison, we used the method of Anitescu and Potra (1997), modified to function as an event-driven method, with the simplest (*i.e.*, pyramidal) friction model. Using the pyramidal friction model resulted in six LCP variables per contact; thus, LCPs of order between 600 and 1200 were generated.

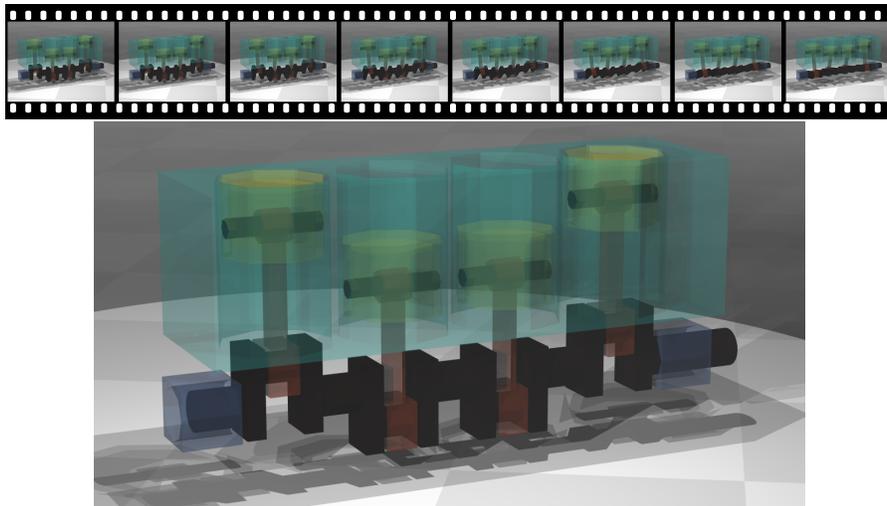


Fig. 3. Frames in sequence from a simulation of the rigid-body dynamics and interactions within an internal combustion engine

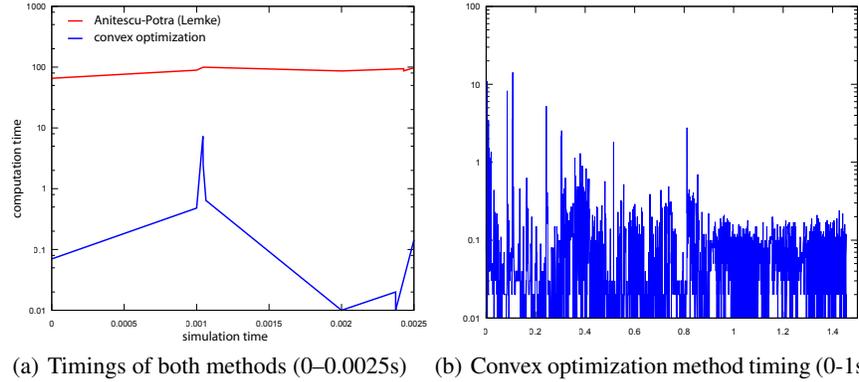


Fig. 4. Computation timings required to solve the contact model for a single cylinder of the internal combustion engine using both the method of Anitescu and Potra (1997) (implemented as an event-driven method) and the event-driven method introduced in this paper. Timings for the Anitescu-Potra method are only provided to 0.0025 seconds of simulation time; the Lemke-based solver (Fackler and Miranda, N/A) could not solve the LCP problem to sufficient tolerances to continue the simulation past that point. The PATH solver (Ferris and Munson, 2000) was able to complete only a single iteration of the simulation, and is thus not included in the comparison.

4.2 Time-stepping example: granular matter

We simulated 1001 spheres of radius 0.04m dropping into and settling within a box to illustrate the feasibility of our method on large scale simulations. We note that similar simulations have been conducted at even larger scales by Tasora and Anitescu ((in submission); however, that work is less general and exhibits several features (*e.g.*, permits interpenetration, non-Coulomb friction model) that limit its applicability outside of granular matter simulation. Figures 6 and 7 show several snapshots taken from the simulation and depict the rapid evolution of the system.

As Figure 5 indicates, we tested our time-stepping method against a time stepping implementation of the contact model of Anitescu and Potra (1997); we used two different LCP solvers (LEMKE (Fackler and Miranda, N/A) and PATH (Ferris and Munson, 2000)) though both exhibited issues with robustness. As in the previous experiment, the pyramidal friction model was used to effect minimum computation time.

5 Modeling joint friction for articulated bodies formulated in reduced coordinates

Many robotics applications simulate articulated bodies in reduced coordinates (Shabana, 2001) for several reasons. The reduced coordinate representation is more amenable to producing matrices useful to roboticists (*e.g.*, the joint space inertia matrix, the end-effector Jacobian matrix, *etc.*) and does not require tweaking parameters to minimize joint constraint violations. The reduced coordinate formulation admits generally simple formulae for modeling Coulomb

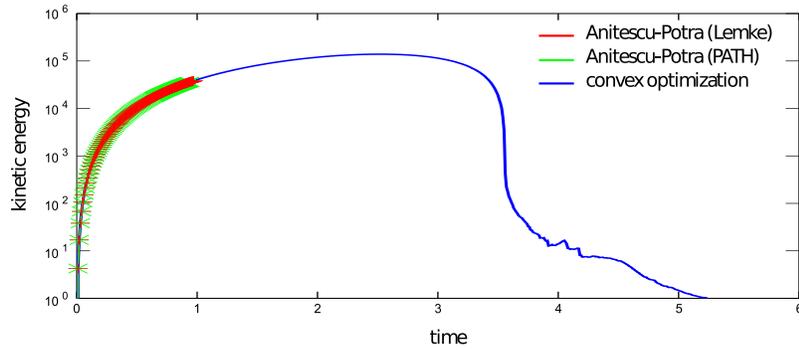


Fig. 5. Kinetic energy of the granular simulation over approximately five seconds of simulation time. The lack of robustness in the linear complementarity problem solvers is evident here, as neither solver for the Anitescu-Potra model was able to simulate to one second of simulation time.

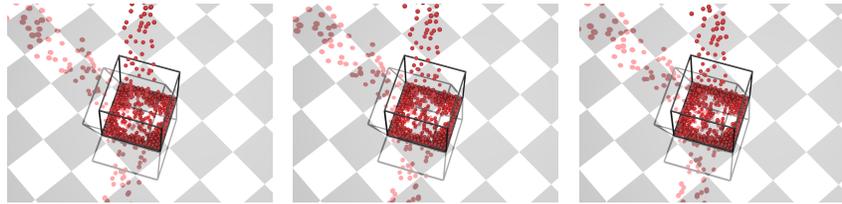


Fig. 6. Frames from time $t = 1.95s$, $t = 1.96s$, and $t = 1.97s$ show the granules dropping into the box.

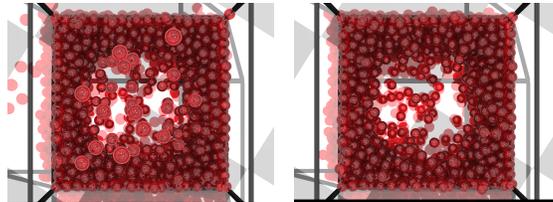


Fig. 7. Frames from time $t = 3.50s$, and $t = 3.55s$ depict the granules settling within the box.

and viscous friction at robot joints. From Sciavicco and Siciliano (2000), p. 141, the torques at joints due to joint friction can be modeled as:

$$\tau_{\mu} = \mu_v \dot{\mathbf{q}} + \mu_c \text{sgn}(\dot{\mathbf{q}}) \quad (6)$$

where μ_v and μ_c are the coefficients for viscous and Coulomb friction, respectively.

The issue with this model is that it tends to make the differential equations *stiff* (*i.e.*, difficult to solve numerically) using even moderately large values of μ_c and μ_v ; this statement is particularly true for μ_c , which uses the discontinuous signum function. The practical effect of this issue is that either extremely small integration steps must be taken or the joint friction must be poorly modeled. We can, however, use the principle of maximum dissipation and

velocity-level dynamics equations to model friction properly; Section 5.3 will show that our approach models the viscous component of Equation 6 exactly (for sufficiently small coefficients of friction or step sizes of the latter), and that the Coulomb component of Equation 6 asymptotically approaches our model as the integration step tends to zero.

5.1 Maximum dissipation-based joint friction model

Under the principle of maximum dissipation, we wish to find the impulses that minimize the new kinetic energy. The change in joint velocity is given by the formula $\Delta\dot{\mathbf{q}} = \mathbf{H}^{-1}\mathbf{x}$, where $\dot{\mathbf{q}}$ is the joint-space velocity, \mathbf{H} is the joint-space inertia matrix, and \mathbf{x} is the vector of applied impulses. Thus we wish to minimize the quantity $\frac{1}{2}(\dot{\mathbf{q}} + \Delta\dot{\mathbf{q}})^T \mathbf{H}(\dot{\mathbf{q}} + \Delta\dot{\mathbf{q}})$ subject to Coulomb and viscous constraints on the applied impulses \mathbf{x} .

Quadratic Program 4

Minimize	$\frac{1}{2}\mathbf{x}^T \mathbf{H}^{-1}\mathbf{x} + \mathbf{x}^T \dot{\mathbf{q}}$	
Subject to:	$-\mu_c \mathbf{1} - \mu_v \dot{\mathbf{q}} \leq \mathbf{x} \leq \mu_c \mathbf{1} + \mu_v \dot{\mathbf{q}}$	(Joint friction model)

5.2 Solving the joint friction model

The joint friction model provided above is a quadratic program with box constraints. Due to the symmetry and positive semi-definiteness of \mathbf{H} (see Featherstone (1987)), the quadratic program is convex, and thus a global minimum can be found in polynomial time. In fact, programs with hundreds of variables—well within the range of the joint spaces of all robots produced to date—can be solved within microseconds using a gradient projection method Nocedal and Wright (2006).

5.3 Empirical results

Figures 8 and 9 depict the efficacy of using the quadratic program defined above for effecting joint friction. We simulated an anthropomorphic, eight degree-of-freedom (DOF) manipulator arm acting only under the influence of gravity from an initial position. In Figure 8, coefficients of joint friction of $\mu_c = 5.0$ and $\mu_v = 0.0$ were used, while the coefficients of joint friction were $\mu_c = 0.25$ and $\mu_v = 0.1$ in Figure 9. Two methods were used to model joint friction: the acceleration level method described in Sciacicco and Siciliano (2000) (*i.e.*, Equation 6) and the quadratic programming method based on the principle of maximal dissipation described above.

Although Figures 8 and 9 depict only a single DOF of the robot arm (the bicep), the results in the plots are indicative of all the DOFs for the robot: the acceleration level method converges to the maximal dissipation-based method as the step size for the former becomes sufficiently small. Indeed, the top plots in the two figures (*i.e.*, the joint velocities) indicate the result for the acceleration level method when the step size is not sufficiently small: the simulation becomes unstable. Our results indicate that the quadratic programming model based on the principle of maximal dissipation is as accurate as an accepted model of joint friction but that the former admits simulation step sizes several orders of magnitude higher (and, correspondingly, simulation several orders of magnitude faster).

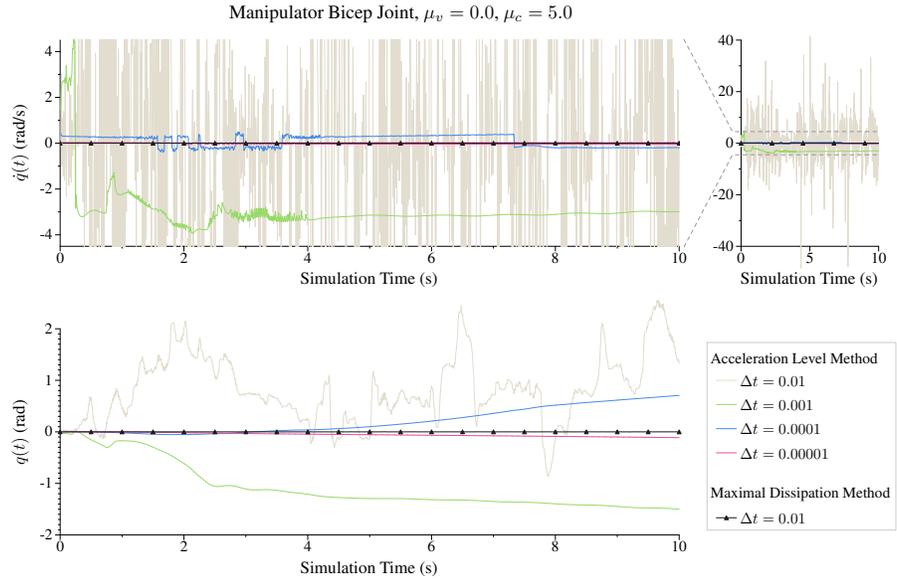


Fig. 8. Plots show that joint friction is comparable between the maximal-dissipation-based and acceleration level methods; the former method produces this behavior at a step size three orders of magnitude larger than the latter. Note that the correct behavior is for the joint position and velocity to remain at zero, given that the coefficient of Coulomb friction is so large (5.0).

6 Conclusions

Roboticians are very familiar with complementarity-based contact models like that of Anitescu and Potra (1997); such models have been incorporated into popular simulators like ODE (Smith, N/A). Consequently, the perception of the accuracy (and inaccuracy) of such models has been informed by considerable practice. The complementarity-free models that were introduced in this paper do not possess such a track record, and direct, empirical comparison between such models is the subject of future work. Nevertheless, the principle of maximal dissipation is accepted by the applied mechanics community, and we have shown evidence that—at minimum—this principle can be used to simulate plausibly mechanisms, granular matter, and joint friction. If the accuracy of the complementarity-free models proves acceptable, the computational advantages intrinsic to our models will yield considerable speedups in robotic simulation. Finally, we argue that, even if the models presented in this paper are found to be poorly predictive (compared to complementarity-based models), Chatterjee’s work makes it clear that non-complementarity-based contact models should be investigated further.

References

Anitescu M (2006) Optimization-based simulation of nonsmooth dynamics. *Mathematical Programming, Series A* 105:113–143

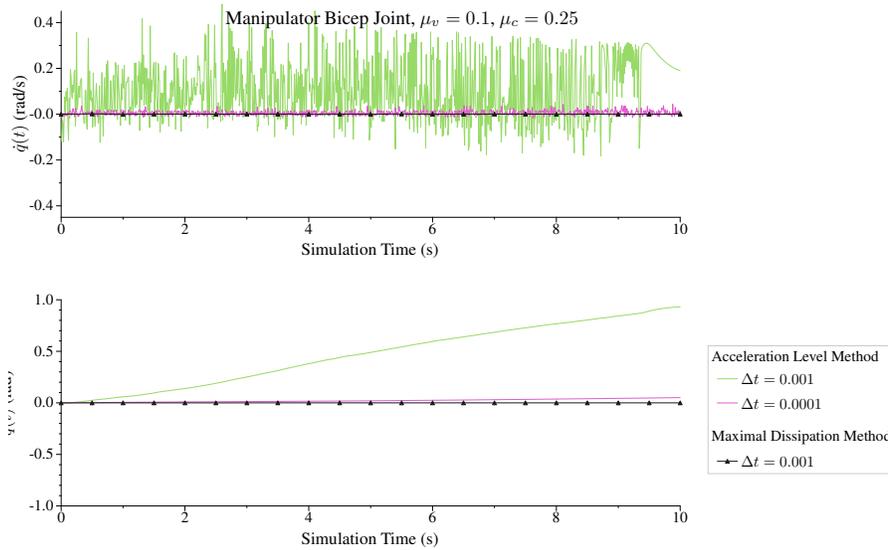


Fig. 9. Plots show that joint friction is comparable between the maximal-dissipation-based and acceleration level methods; the former method produces this behavior at a larger step size than the latter.

- Anitescu M, Hart G (2004) A constraint-stabilized time-stepping approach for rigid multi-body dynamics with joints, contacts, and friction. *Intl Journal for Numerical Methods in Engineering* 60(14):2335–2371
- Anitescu M, Potra FA (1997) Formulating dynamic multi-rigid-body contact problems with friction as solvable linear complementarity problems. *Nonlinear Dynamics* 14:231–247
- Anitescu M, Potra FA (2002) A time-stepping method for stiff multi-rigid-body dynamics with contact and friction. *Intl Journal for Numerical Methods in Engineering* 55:753–784
- Anitescu M, Tasora A (2008) An iterative approach for cone complementarity problems for nonsmooth dynamics. *Computational Optimization and Applications*
- Anitescu M, Cremer JF, Potra FA (1997) Properties of complementarity formulations for contact problems with friction. In: Ferris MC, Pang JS (eds) *Complementarity and Variational Problems: State of the Art*, SIAM, Philadelphia, PA, pp 12–21
- Anitescu M, Potra F, Stewart D (1999) Time-stepping for three dimensional rigid body dynamics. *Computer Methods in Applied Mechanics and Engineering* 177:183–197
- Baraff D (1991) Coping with friction for non-penetrating rigid body simulation. *Computer Graphics* 25(4):31–40
- Baraff D (1994) Fast contact force computation for nonpenetrating rigid bodies. In: *Proc. of SIGGRAPH*, Orlando, FL
- Baraff D (1997) An introduction to physically based modeling: Rigid body simulation II – constrained rigid body dynamics. Tech. rep., Robotics Institute, Carnegie Mellon University
- Brogliato B (1996) *Nonsmooth Impact Mechanics: Models, Dynamics, and Control*. Springer-Verlag, London

- Brogliato B, ten Dam A, Paoli L, Génot F, Abadie M (2002) Numerical simulation of finite dimensional multibody nonsmooth mechanical systems. *ASME Appl Mech Reviews* 55(2):107–150
- Ceanga V, Hurmuzlu Y (2001) A new look at an old problem: Newton’s cradle. *ASME J Appl Mech* 68:575–583
- Chatterjee A (1999) On the realism of complementarity conditions in rigid-body collisions. *Nonlinear Dynamics* 20:159–168
- Cottle RW, Pang JS, Stone R (1992) *The Linear Complementarity Problem*. Academic Press, Boston
- Drumwright E (2010) Avoiding Zeno’s paradox in impulse-based rigid body simulation. In: Proc. of IEEE Intl. Conf. on Robotics and Automation (ICRA), (to appear)
- Drumwright E, Shell DA (2009) A robust and tractable contact model for dynamic robotic simulation. In: Proc. of ACM Symp. on Applied Computing (SAC), pp 1176–1180
- Erleben K (2007) Velocity-based shock propagation for multibody dynamics animation. *ACM Trans on Graphics* 26(12)
- Fackler PL, Miranda MJ (N/A) LEMKE. http://people.sc.fsu.edu/burkardt/m_src/lemke/lemke.m
- Featherstone R (1987) *Robot Dynamics Algorithms*. Kluwer
- Ferris MC, Munson TS (2000) Complementarity problems in GAMS and the PATH solver. *J of Economic Dynamics and Control* 24(2):165–188
- Guendelman E, Bridson R, Fedkiw R (2003) Nonconvex rigid bodies with stacking. *ACM Trans on Graphics* 22(3):871–878
- Kokkevis E (2004) Practical physics for articulated characters. In: Proc. of Game Developers Conf.
- Lacoursière C (2003) Splitting methods for dry frictional contact problems in rigid multibody systems: Preliminary performance results. In: Ollila M (ed) Proc. of SIGRAD, pp 11–16
- Lemke CE (1965) Bimatrix equilibrium points and mathematical programming. *Management Science* 11:681–689
- Löstedt P (1982) Mechanical systems of rigid bodies subject to unilateral constraints. *SIAM Journal on Applied Mathematics* 42(2):281–296
- Löstedt P (1984) Numerical simulation of time-dependent contact friction problems in rigid body mechanics. *SIAM J of Scientific Statistical Computing* 5(2):370–393
- Miller AT, Christensen HI (2003) Implementation of multi-rigid-body dynamics within a robotic grasping simulator. In: Proc. of the IEEE Intl. Conf. on Robotics and Automation (ICRA), pp 2262–2268
- Mirtich B (1996a) Fast and accurate computation of polyhedral mass properties. *Journal of Graphics Tools* 1(2)
- Mirtich B (1996b) Impulse-based dynamic simulation of rigid body systems. PhD thesis, University of California, Berkeley
- Monteiro-Marques MDP (1993) Differential inclusions in nonsmooth mechanical problems: Shocks and dry friction. In: *Progress in Nonlinear Differential Equations and Their Applications*, vol 9, Birkhäuser Verlag, Basel
- Moreau JJ (1983) *Standard inelastic shocks and the dynamics of unilateral constraints*, Springer-Verlag, New York, pp 173–221
- Moreau JJ (1985) Standard inelastic shocks and the dynamics of unilateral constraints. In: del Piero G, Maceri F (eds) *C.I.S.M. Courses and Lectures*, vol 288, Springer-Verlag, Vienna, pp 173–221
- Moreau JJ (1988) Unilateral contact and dry friction in finite freedom dynamics. In: Moreau JJ, Panagiotopoulos PD (eds) *Nonsmooth Mechanics and Applications*, Springer-Verlag, Vienna, pp 1–82

- Nocedal J, Wright SJ (2006) Numerical Optimization, 2nd ed. Springer-Verlag
- Painlevé P (1895) Sur le lois du frottement de glissement. C R Académie des Sciences Paris 121:112–115
- Petra C, Gavrea B, Anitescu M, Potra F (2009) A computational study of the use of an optimization-based method for simulating large multibody systems. Optimization Methods and Software
- Pfeiffer F, Glocker C (1996) Dynamics of rigid body systems with unilateral constraints. John Wiley and Sons
- Potra FA, Anitescu M, Gavrea B, Trinkle J (2006) A linearly implicit trapezoidal method for stiff multibody dynamics with contact, joints, and friction. Intl Journal for Numerical Methods in Engineering 66(7):1079–1124
- Sauer J, Schoemer E, Lennerz C (1998) Real-time rigid body simulations of some “classical mechanics toys”. In: Proc. of European Simulation Symposium and Exhibition, Nottingham, UK
- Sciavicco L, Siciliano B (2000) Modeling and Control of Robot Manipulators, 2nd Ed. Springer-Verlag, London
- Shabana AA (2001) Computational Dynamics, 2nd edn. John Wiley & Sons
- Smith R (N/A) ODE: Open Dynamics Engine
- Stewart D, Trinkle J (2000) An implicit time-stepping scheme for rigid body dynamics with coulomb friction. In: Proc. of the IEEE Intl. Conf. on Robotics and Automation (ICRA), San Francisco, CA
- Stewart DE (2000) Rigid-body dynamics with friction and impact. SIAM Review 42(1):3–39
- Stoianovici D, Hurmuzlu Y (1996) A critical study of the applicability of rigid-body collision theory. ASME J Appl Mech 63:307–316
- Tasora A, Anitescu M ((in submission)) A convex complementarity approach for simulating large granular flow. J of Computational Nonlinear Dynamics
- Tassa Y, Todorov E (2010) Stochastic complementarity for local control of continuous dynamics. In: Proc. of Robotics: Science and Systems
- Todorov E (2010) Implicit nonlinear complementarity: a new approach to contact dynamics. In: Proc. of Intl. Conf. on Robotics and Automation ICRA
- Trinkle J, Pang JS, Sudarsky S, Lo G (1997) On dynamic multi-rigid-body contact problems with coulomb friction. Zeithschrift fur Angewandte Mathematik und Mechanik 77(4):267–279
- Trinkle JC, Berard S, Pang J (2005) A time-stepping scheme for quasistatic multibody systems. In: Proc. of IEEE Intl. Symp. on Assembly and Task Planning, pp 174–181

A Proof that feasible point determined via LCP is energetically consistent

We now prove that the solution determined in phase I in Section 2.6 is energetically consistent. Working from Mirtich (1996b), the equation for work done by collision impulses is:

$$\frac{1}{2} \mathbf{z}^T (\mathbf{A} \mathbf{z} + 2\mathbf{b}) \quad (7)$$

Given that $\mathbf{z} \triangleq \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix}$, we will prove that:

$$\frac{1}{2} \mathbf{y}^T (\mathbf{A}_N \mathbf{y} + 2\mathbf{b}_N) \leq 0 \quad (8)$$

The proof relies upon the linear complementarity solver finding the solution \mathbf{y} , which yields $\mathbf{y}^T (\mathbf{A}_N \mathbf{y} + \mathbf{b}_N) = 0$.

Equation 8 simplifies to:

$$\mathbf{y}^T (\mathbf{A}_N \mathbf{y} + 2\mathbf{b}_N) \leq 0 \quad (9)$$

or, equivalently:

$$\mathbf{y}^T (\mathbf{A}_N \mathbf{y} + \mathbf{b}_N) + \mathbf{y}^T \mathbf{b}_N \leq 0 \quad (10)$$

Given the linear complementarity condition $\mathbf{y}^T (\mathbf{A}_N \mathbf{y} + \mathbf{b}_N) = 0$, we are left with:

$$\mathbf{y}^T \mathbf{b}_N \leq 0 \quad (11)$$

The above equation must hold, because (again, due to the linear complementarity condition):

$$\mathbf{y}^T (\mathbf{A}_N \mathbf{y} + \mathbf{b}_N) = 0 \quad (12)$$

and because \mathbf{A}_N is symmetric, positive semi-definite, which implies that $\mathbf{y}^T \mathbf{A}_N \mathbf{y} \geq 0$ (and therefore, that $\mathbf{y}^T \mathbf{b}_N \leq 0$).