

Investigating Through Simulation the Mobility of Light Tracked Vehicles Operating on Discrete Granular Terrain

Dan Negrut, Daniel Melanz and Hammad Mazhar
University of Wisconsin - Madison

David Lamb, Paramsothy Jayakumar and Michael Letherwood
US Army TARDEC

ABSTRACT

This paper presents a computational framework for the physics-based simulation of light vehicles operating on discrete terrain. The focus is on characterizing through simulation the mobility of vehicles that weigh 1000 pounds or less, such as a reconnaissance robot. The terrain is considered to be deformable and is represented as a collection of bodies of spherical shape. The modeling stage relies on a novel formulation of the frictional contact problem that requires at each time step of the numerical simulation the solution of an optimization problem.

The proposed computational framework, when run on ubiquitous Graphics Processing Unit (GPU) cards, allows the simulation of systems in which the terrain is represented by more than 0.5 million bodies leading to problems with more than one million degrees of freedom. The numerical solution for the equations of motion is tailored to map on the underlying GPU architecture and is parallelized to leverage more than 1500 Scalar Processors available on modern hardware architectures.

As a demonstration of this technology, we present the simulation of a light tracked vehicle that negotiates several obstacles whose dimensions are comparable to those of the vehicle. The number of bodies used to represent the vehicle is larger than 100 and the terrain the vehicle operates on is considered to be made up of gravel.

CITATION: Negrut, D., Melanz, D., Mazhar, H., Lamb, D. et al., "Investigating Through Simulation the Mobility of Light Tracked Vehicles Operating on Discrete Granular Terrain," *SAE Int. J. Passeng. Cars - Mech. Syst.* 6(1):2013, doi: 10.4271/2013-01-1191.

INTRODUCTION

Engineers are increasingly relying on simulation to augment and, in some cases, replace costly and time consuming experimental work. However, current simulation capabilities are sometimes inadequate to capture phenomena of interest. In tracked vehicle analysis, for example, the interaction of the track with granular terrain has been difficult to characterize through simulation due to the prohibitively long simulation times associated with *many-body dynamics* problems. This is the generic name used here to characterize dynamic systems with a large number of bodies encountered, for instance, when one adopts a discrete representation of the terrain in vehicle dynamics problems. However, these many-body dynamics problems can now capitalize on recent advances in the microprocessor industry that are a consequence of Moore's law, of doubling the number of transistors per unit area roughly every 18 months. Specifically, until recently, access to massive computational

power on parallel supercomputers has been the privilege of a relatively small number of research groups in a select number of research facilities, thus limiting the scope and impact of high performance computing (HPC). This scenario is rapidly changing due to a trend set by general-purpose computing on graphics processing unit (GPU) cards. NVIDIA's CUDA library [21] allows one to use the streaming multiprocessors available in high-end graphics cards. In this setup, a latest generation NVIDIA GPU Kepler card will reach 1.5 Teraflops by the end of 2012 owing to a set of 1536 scalar processors working in parallel, each following a Single Instruction Multiple Data (SIMD) execution paradigm. Despite having only 1536 scalar processors, such a card is capable of managing tens of thousands of parallel threads at any given time. This overcommitting of the GPU hardware resources is at the cornerstone of a computing paradigm that aggressively attempts to hide costly memory transactions with useful computation, a strategy that has lead, in frictional

contact dynamics simulation, to a one order of magnitude reduction in simulation time for many-body systems [34, 20].

The challenge of using parallel computing to reduce simulation time and/or increase system size stems, for the most part, from the task of designing and implementing many-body dynamics specific parallel numerical methods. Designing parallel algorithms suitable for frictional contact many-body dynamics simulation remains an area of active research. Results reported in [16] indicate that the most widely used commercial software package for multibody dynamics simulation, which draws on a so called penalty or regularization approach, runs into significant difficulties when handling simple problems involving hundreds of contact events, and thus cases with thousands of contacts become intractable. Unlike these penalty or regularization approaches where the frictional interaction is represented by a collection of stiff springs combined with damping elements that act at the interface of the two bodies [11, 28, 29, 22], the approach embraced herein draws on a different mathematical framework. Specifically, the parallel algorithms rely on time-stepping procedures producing weak solutions of the differential variational inequality (DVI) problem that describes the time evolution of rigid bodies with impact, contact, friction, and bilateral constraints. When compared to penalty-methods, the DVI approach has a greater algorithmic complexity, but avoids the small time steps that plague the former approach.

The task of presenting this class of algorithms and their parallel implementation is organized as follows. Section provides a brief description of the general equations that capture the dynamics of many-body systems. This section also contains an outline of the parallel method embraced to numerically solve the equations of motion. One of the challenging components of the solution method is the collision detection step required to determine the set of contacts active in the many-body system. These contacts, crucial in producing the frictional contact forces at work in the system, are determined in parallel using an approach outlined in Section. A scalable rendering pipeline that can leverage thousands of CPU cores for visualization purposes is discussed in Section. The engineering application used to demonstrate this parallel simulation capability is that of a light tracked vehicle that operates on granular terrain and negotiates an obstacle course. To further illustrate the versatility of the simulation capability, the vehicle is assumed to be equipped with a drilling device used to penetrate the terrain. Both the vehicle dynamics and the drilling process are seamlessly analyzed within the same HPC-enabled simulation capability. A schematic of the vehicle is provided in Fig. 1. A cut-away image of the drilling tool is shown in isolation in Fig. 2.

THE MANY-BODY DYNAMICS PROBLEM

General Considerations

The modeling approach adopted in order to abstract and represent the dynamics of the vehicle/terrain interaction is based on a differential variational inequality (DVI) methodology. Compared to penalty or regularization approaches [11, 28, 29, 22], it allows for larger integration step sizes. The formulation of the equations of motion, that is, the equations that govern the time evolution of a multibody system, is based on the so-called absolute, or Cartesian, representation of the position and attitude of each rigid body in the system. The state of the system is denoted by the generalized positions $\mathbf{q} = [\mathbf{r}_1^T, \epsilon_1^T, \dots, \mathbf{r}_{n_b}^T, \epsilon_{n_b}^T]^T \in \mathbb{R}^{7n_b}$ and their time derivatives $\dot{\mathbf{q}} = [\dot{\mathbf{r}}_1^T, \dot{\epsilon}_1^T, \dots, \dot{\mathbf{r}}_{n_b}^T, \dot{\epsilon}_{n_b}^T]^T \in \mathbb{R}^{7n_b}$, where n_b is the number of bodies, $\mathbf{r}_j$ is the absolute position of the center of mass of the j th body, and the quaternions (Euler parameters) ϵ_j are used to represent rotation and to avoid singularities.

Instead of using quaternion derivatives in $\dot{\mathbf{q}}$, it is more advantageous to work with angular velocities expressed in the local (body-attached) reference frames; in other words, the method described will use the vector of generalized velocities $\mathbf{v} = [\dot{\mathbf{r}}_1^T, \bar{\omega}_1^T, \dots, \dot{\mathbf{r}}_{n_b}^T, \bar{\omega}_{n_b}^T]^T \in \mathbb{R}^{6n_b}$. Note that the generalized velocity can be easily obtained as $\dot{\mathbf{q}} = \mathbf{L}(\mathbf{q})\mathbf{v}$, where $\mathbf{L}$ is a linear mapping that transforms each $\bar{\omega}_i$ into the corresponding quaternion derivative $\dot{\epsilon}_i$ by means of the linear algebra formula $\dot{\epsilon}_i = \frac{1}{2}\mathbf{G}^T(\mathbf{q})\bar{\omega}_i$, with 3×4 matrix $\mathbf{G}(\mathbf{q})$ as defined in [12].

Bilateral constraints

Bilateral constraints represent kinematic relationships between two rigid bodies in the system. For example, spherical joints, prismatic joints, or revolute joints can be expressed as holonomic algebraic equations constraining the relative positions of two bodies. A set B of constraints leads to a collection of scalar equations

$$\Psi_i(\mathbf{q}, t) = 0, \quad i \in B, \quad (1)$$

the number of which depends on the type of constraints in set B . Bilateral constraints must also be satisfied at the velocity level,

$$\begin{aligned} \frac{d\Psi_i(\mathbf{q}, t)}{dt} = 0 & \Rightarrow \\ \frac{\partial \Psi_i}{\partial \mathbf{q}} \dot{\mathbf{q}} + \frac{\partial \Psi_i}{\partial t} &= \nabla_{\mathbf{q}} \Psi_i^T \dot{\mathbf{q}} + \frac{\partial \Psi_i}{\partial t} \\ &= \nabla_{\mathbf{q}} \Psi_i^T \mathbf{L}(\mathbf{q})\mathbf{v} + \frac{\partial \Psi_i}{\partial t} = 0, \end{aligned}$$

which is obtained by taking one time derivative of Equation 1.



Figure 1. Light autonomous vehicle negotiating a pile of rubble.

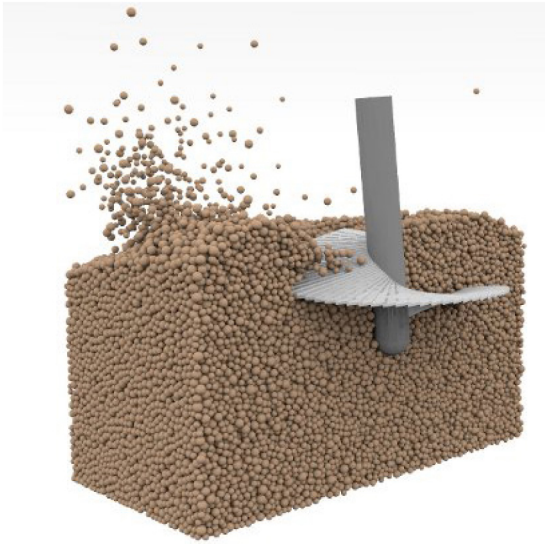


Figure 2. Cutaway view: Anchor penetrating granular material [18].

Unilateral constraints and friction

Unilateral constraints enforce contact constraints between rigid bodies in the system. It is assumed that a gap function, $\Phi(q)$, can be defined for each pair of near-enough bodies. This gap function describes the distance between the closest points on the two bodies of interest.

Unilateral contact constraints also introduce friction forces into the system. When a contact is active, or $\Phi_i(q) = 0$, a normal force acts on each of the two bodies at the contact point. When a contact is inactive, or $\Phi_i(q) > 0$, no normal force exists. This represents a complementarity condition. Consider two bodies A and B in contact as shown in Fig. 3. Let $\mathbf{n}_i$ be the normal at the contact pointing toward the exterior of the body of lower index, which by convention is considered to be body A . Let $\mathbf{u}_i$ and $\mathbf{w}_i$ be two vectors in the contact plane such that $\mathbf{n}_i, \mathbf{u}_i, \mathbf{w}_i \in \mathbb{R}^3$ are mutually

orthonormal vectors. The frictional contact forces are defined by the multipliers $\hat{\gamma}_{i,n} \geq 0$, $\hat{\gamma}_{i,u}$, and $\hat{\gamma}_{i,w}$, which lead to the normal component of the friction force, $\mathbf{F}_{i,N} = \hat{\gamma}_{i,n} \mathbf{n}_i$ and the tangential component of the force $\mathbf{F}_{i,T} = \hat{\gamma}_{i,u} \mathbf{u}_i + \hat{\gamma}_{i,w} \mathbf{w}_i$.

The Coulomb friction model, which draws for contact i on the friction coefficient μ_i , is used to write the following constraints:

$$\hat{\gamma}_{i,n} \geq 0, \quad \Phi_i(\mathbf{q}) \geq 0, \quad \Phi_i(\mathbf{q}) \hat{\gamma}_{i,n} = 0, \quad (2)$$

$$\mu_i \hat{\gamma}_{i,n} \geq \sqrt{\hat{\gamma}_{i,u}^2 + \hat{\gamma}_{i,w}^2}, \quad \|\mathbf{v}_{i,T}\| \left(\mu_i \hat{\gamma}_{i,n} - \sqrt{\hat{\gamma}_{i,u}^2 + \hat{\gamma}_{i,w}^2} \right) = 0, \\ \langle \mathbf{F}_{i,T}, \mathbf{v}_{i,T} \rangle = -\|\mathbf{F}_{i,T}\| \|\mathbf{v}_{i,T}\| \quad (3)$$

Equation 2 captures the complementarity condition previously described. The subsequent two equations relate the magnitude and direction of the friction force to the multipliers and tangential velocity of the contact. These remaining equations can be expressed in an equivalent manner using the maximum dissipation principle. This frames the Coulomb friction model as a minimization problem, which can be seen in Equation 4.

$$(\hat{\gamma}_{i,u}, \hat{\gamma}_{i,w}) = \underset{\sqrt{\hat{\gamma}_{i,u}^2 + \hat{\gamma}_{i,w}^2} \leq \mu_i \hat{\gamma}_{i,n}}{\operatorname{argmin}} \quad \mathbf{v}_{i,T}^T (\hat{\gamma}_{i,u} \mathbf{u}_i + \hat{\gamma}_{i,w} \mathbf{w}_i). \quad (4)$$

The nature of the friction cone can be seen if yet another form of the friction force equations is considered. The friction force of the i -th contact can be expressed as follows, where Υ is a cone in three dimensions whose slope is $\tan^{-1} \mu_i$.

$$\mathbf{F}_i = \mathbf{F}_{i,N} + \mathbf{F}_{i,T} = \hat{\gamma}_{i,n} \mathbf{n}_i + \hat{\gamma}_{i,u} \mathbf{u}_i + \hat{\gamma}_{i,w} \mathbf{w}_i \in \Upsilon, \quad (5)$$

Equations of motion of systems with frictional contact

The time evolution of the dynamical system is governed by the following differential variational inequality [7]:

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{L}(\mathbf{q}) \mathbf{v} \\ \mathbf{M} \dot{\mathbf{v}} &= \mathbf{f}(t, \mathbf{q}, \mathbf{v}) + \sum_{i \in \mathcal{B}} \hat{\gamma}_{i,b} \nabla \Psi_i + \\ &\quad + \sum_{i \in \mathcal{A}} (\hat{\gamma}_{i,n} \mathbf{D}_{i,n} + \hat{\gamma}_{i,u} \mathbf{D}_{i,u} + \hat{\gamma}_{i,w} \mathbf{D}_{i,w}) \\ i \in \mathcal{B} &: \Psi_i(\mathbf{q}, t) = 0 \\ i \in \mathcal{A} &: \hat{\gamma}_{i,n} \geq 0 \perp \Phi_i(\mathbf{q}) \geq 0, \quad \text{and} \\ (\hat{\gamma}_{i,u}, \hat{\gamma}_{i,w}) &= \underset{\mu_i \hat{\gamma}_{i,n} \geq \sqrt{\hat{\gamma}_{i,u}^2 + \hat{\gamma}_{i,w}^2}}{\operatorname{argmin}} \quad \mathbf{v}^T (\hat{\gamma}_{i,u} \mathbf{D}_{i,u} + \hat{\gamma}_{i,w} \mathbf{D}_{i,w}). \end{aligned} \quad (6)$$

The tangent space generators $\mathbf{D}_i = [\mathbf{D}_{i,n}, \mathbf{D}_{i,u}, \mathbf{D}_{i,w}] \in \mathbb{R}^{6n_b \times 3}$ are sparse and are defined given a pair of contacting bodies A and B as

$$\mathbf{D}_i^T = \begin{bmatrix} 0 & \dots & -\mathbf{A}_{i,p}^T & \mathbf{A}_{i,p}^T \mathbf{A}_A \tilde{\mathbf{s}}_{i,A} & 0 & \dots \\ 0 & \dots & \mathbf{A}_{i,p}^T & -\mathbf{A}_{i,p}^T \mathbf{A}_B \tilde{\mathbf{s}}_{i,B} & 0 & \dots \end{bmatrix}, \quad (7)$$

where $\mathbf{A}_A$ is the orientation matrix associated with body A , $\mathbf{A}_{i,p} = [\mathbf{n}_i, \mathbf{u}_i, \mathbf{w}_i]$ is the $\mathbb{R}^{3 \times 3}$ matrix of the local coordinates of the i th contact, the vectors $\tilde{\mathbf{s}}_{i,A}$ and $\tilde{\mathbf{s}}_{i,B}$ are the contact point positions in body coordinates (see Fig. 3). A tilde $\tilde{\mathbf{x}}$ over a vector $\mathbf{x} \in \mathbb{R}^3$ represents the skew symmetric matrix associated with the outer product of two vectors [12].

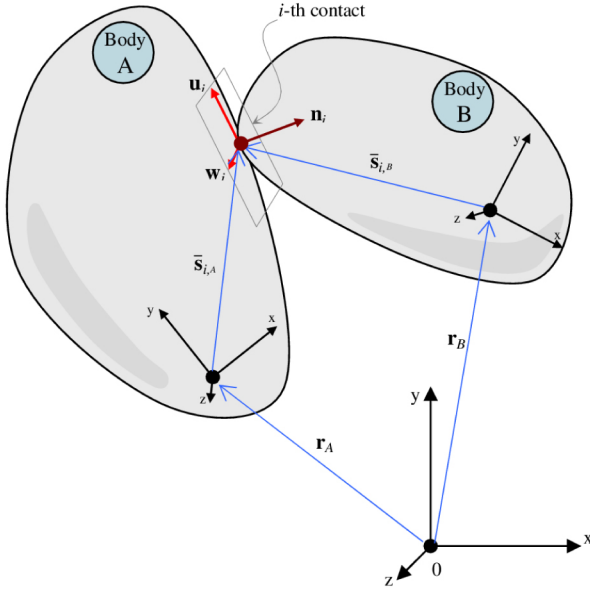


Figure 3. Contact i between two bodies $A, B \in \{1, 2, \dots, n_b\}$.

Discretization Scheme for Numerical Solution

Equation 2 expresses the complementarity condition between the normal force and the gap function in a contact event. The presence of these complementarity conditions is the trademark of a DVI formulation, whose numerical solution in the context of rigid body dynamics can be traced back to [19, 15, 17]. The DVI formulations have been classified by differential index in [23] and recent time-stepping schemes have included both acceleration-force linear complementarity problem (LCP) approaches [8, 24, 35] and velocity-impulse LCP-based time-stepping methods [31, 5, 6, 30]. The LCPs, obtained as a result of the introduction of inequalities in time-stepping schemes for DVI, coupled with a polyhedral approximation of the friction cone must be solved at each time step in order to determine the system state configuration as well as the Lagrange multipliers representing the reaction forces [15, 31]. If the simulation entails a large number of contacts and rigid bodies, as in the case of part feeders, packaging machines, and granular flows, the computational burden of classical LCP solvers can become

significant. Indeed, a well-known class of numerical methods for LCPs based on *simplex methods*, also known as *direct* or *pivoting methods* [10], may exhibit exponential worst-case complexity [9]. They may be impractical even for problems involving as few as several hundred bodies when friction is present [4, 33]. Moreover, the three-dimensional Coulomb friction case leads to a nonlinear complementarity problem (NCP): the use of a polyhedral approximation to transform the NCP into an LCP introduces artificial anisotropy in friction cones [31, 35, 5]. This discrete and finite approximation of friction cones is one of the reasons for the large dimension of the problem that needs to be solved in multibody dynamics with frictional contact.

In order to circumvent the limitations imposed by the use of classical LCP solvers and the limited accuracy associated with polyhedral approximations of the friction cone, a parallel fixed-point iteration method with projection on a convex set has been proposed, developed, and tested in [7]. The method is based on a time-stepping formulation that solves at every step a cone constrained optimization problem [2]. The time-stepping scheme, proved to converge in a measure differential inclusion sense to the solution of the original continuous-time DVI, sets off at time t_l by assuming that a set of contacts, $\mathcal{A}$, exists between bodies in the system, and a set of bilateral constraints, $\mathcal{B}$, is also active. The governing differential equations then assume the form of a DVI problem. The equation of motion is discretized so that an approximation to the solution can be found at discrete instants in time. Given a position $\mathbf{q}^{(l)}$ and velocity $\mathbf{v}^{(l)}$ at the time step $t^{(l)}$, the numerical solution is found at the new time step $t^{(l+1)} = t^{(l)} + h$ by solving the following optimization problem with equilibrium constraints [32]:

$$\mathbf{M}(\mathbf{v}^{(l+1)} - \mathbf{v}^{(l)}) = h\mathbf{f}(t^{(l)}, \mathbf{q}^{(l)}, \mathbf{v}^{(l)}) + \sum_{i \in \mathcal{B}} \gamma_{i,b} \nabla \Psi_i + \sum_{i \in \mathcal{A}} (\gamma_{i,n} \mathbf{D}_{i,n} + \gamma_{i,u} \mathbf{D}_{i,u} + \gamma_{i,w} \mathbf{D}_{i,w}), \quad (8)$$

$$i \in \mathcal{B} : \quad \frac{1}{h} \Psi_i(\mathbf{q}^{(l)}, t) + \nabla \Psi_i^T \mathbf{v}^{(l+1)} + \frac{\partial \Psi_i}{\partial t} = 0 \quad (9)$$

$$i \in \mathcal{A} : \quad 0 \leq \frac{1}{h} \Phi_i(\mathbf{q}^{(l)}) + \mathbf{D}_{i,n}^T \mathbf{v}^{(l+1)} \perp \gamma_n^i \geq 0, \quad (10)$$

$$(\gamma_{i,u}, \gamma_{i,w}) = \underset{\mu_i \gamma_{i,n} \geq \sqrt{\gamma_{i,u}^2 + \gamma_{i,w}^2}}{\operatorname{argmin}} \quad \mathbf{v}^{(l+1),T} (\gamma_{i,u} \mathbf{D}_{i,u} + \gamma_{i,w} \mathbf{D}_{i,w}) \quad (11)$$

$$\mathbf{q}^{(l+1)} = \mathbf{q}^{(l)} + h\mathbf{L}(\mathbf{q}^{(l)})\mathbf{v}^{(l+1)}. \quad (12)$$

Here, γ_s represents the constraint impulse of a contact constraint; that is, $\gamma_s = h\hat{\gamma}_s$, for $s = n, u, w$. The $\frac{1}{h} \Phi_i(\mathbf{q}^{(l)})$ term achieves constraint stabilization; its effect is discussed in [3]. Similarly, the term $\frac{1}{h} \Psi_i(\mathbf{q}^{(l)})$ achieves stabilization for