## Metastable Walking Machines

| | |
|---|---|
| **Citation** | Byl, Katie, and Russ Tedrake. "Metastable Walking Machines." The International Journal Of Robotics Research 28.8 (2009) : 1040 -1064. Copyright © 2009, SAGE Publications |
| **As Published** | http://dx.doi.org/10.1177/0278364909340446 |
| **Publisher** | Sage |
| **Version** | Author's final manuscript |
| **Citable link** | http://hdl.handle.net/1721.1/61973 |
| **Terms of Use** | Creative Commons Attribution-Noncommercial-Share Alike 3.0 |
| **Detailed Terms** | http://creativecommons.org/licenses/by-nc-sa/3.0/ |

# Metastable Walking Machines

Katie Byl, *Member, IEEE* and Russ Tedrake, *Member, IEEE*

*Abstract*— **Legged robots that operate in the real world are inherently subject to stochasticity in their dynamics and uncertainty about the terrain. Due to limited energy budgets and limited control authority, these "disturbances" cannot always be canceled out with high-gain feedback. Minimally-actuated walking machines subject to stochastic disturbances no longer satisfy strict conditions for limit-cycle stability; however, they can still demonstrate impressively long-living periods of continuous walking. Here, we employ tools from stochastic processes to examine the "stochastic stability" of idealized rimless-wheel and compass-gait walking on randomly generated uneven terrain. Furthermore, we employ tools from numerical stochastic optimal control to design a controller for an actuated compass gait model which maximizes a measure of stochastic stability - the mean first-passage-time - and compare its performance to a deterministic counterpart. Our results demonstrate that walking is well-characterized as a metastable process, and that the stochastic dynamics of walking should be accounted for during control design in order to improve the stability of our machines.**

## I. Introduction

The dynamics of legged locomotion are plagued with complexity: intermittent ground interactions, variations in terrain and external perturbations all have a significant (and stochastic) impact on the evolution of dynamics and ultimate stability of both animals and machines which walk. Detailed analytical and computational investigations of simplified models have captured much that is fundamental in the dynamics of walking (Coleman, Chatterjee, & Ruina, 1997; Garcia, Chatterjee, Ruina, & Coleman, 1998; Goswami, Thuilot, & Espiau, 1996; Koditschek & Buehler, 1991; McGeer, 1990). These analyses reveal the limit cycle nature of ideal walking systems and often employ Poincaré map analysis to assess the stability of these limit cycles. However, the very simplifications which have made these models tractable for analysis can limit their utility.

Experimental analyses of real machines based on these simple models (Collins, Ruina, Tedrake, & Wisse, 2005) reveal that the real devices differ from their idealized dynamics in a number of important ways. Certainly the dynamics of impact and contact with the ground are more subtle than what is captured by the idealized models. But perhaps more fundamental is the inevitable stochasticity in the real systems. More than just measurement noise, robots that walk are inherently prone to the stochastic influences of their environment by interacting with terrain which varies at each footstep. Even in a carefully designed laboratory setting, and especially for passive and minimally-actuated walking machines, this stochasticity can have a major effect on the long-term system dynamics. In practice, it is very difficult (and technically incorrect) to apply deterministic limit cycle stability analyses to our experimental walking machines - the real machines do not have true limit cycle dynamics.

In this paper, we extend the analysis of simplified walking models toward real machines by adding stochasticity into our models and applying mathematical tools from stochastic analysis to quantify and optimize stability. We examine two classic models of walking: the rimless wheel (RW) and the compass gait (CG) biped. Although we have considered a number of sources of uncertainty, we will focus here on a compact and demonstrative model - where the geometry of the ground is drawn from a random distribution. Even with mild deviations in terrain from a nominal slope angle, the resulting trajectories of the machine are different on every step, and for many noise distributions (e.g., Gaussian) the robot is guaranteed to eventually fall down (with probability one as $t \to \infty$). However, one can still meaningfully quantify stochastic stability in terms of expected time to failure, and maximization of this metric in turn provides a principled approach to controller design for walking on moderately rough, unmodeled terrain.

Stochastic optimization of a controlled compass gait model on rough terrain reveals some important results. Modeling walking as a metastable limit cycle changes the optimal control policy; upcoming terrain knowledge, or knowledge of the terrain distribution, can be exploited by the controller to enhance long-term "stochastic" stability. Using our newly defined stability metrics, we demonstrate that these risk-adjusted modifications to the control policy can dramatically improve the overall stability of our machines.

The paper proceeds as follows, Section II provides a quick background on metastable stochastic processes. Section III applies metastability analysis to limit cycle dynamics on the Poincaré map. Section IV numerically investigates the stochastic stability of simple *passive* walking models. The concepts and methodologies presented in Sections II through IV were originally introduced in (Byl & Tedrake, 2008c). Section V extends this stochastic analysis of walking devices by employing the same tools demonstrated for the evaluation of the MFPT of purely passive walkers toward (approximately) optimal stochastic control of an *actuated* compass gait model; stochastic stability of the resulting systems is then investigated. Finally, Section VI discusses some important implications of viewing our robots as "metastable walking machines".

## II. Background

Many stochastic dynamic systems exhibit behaviors which are impressively long-living, but which are also guaranteed to exit these locally-attractive behaviors (i.e., to fail) with probability one, given enough time. Such systems cannot be classified as "stable", but it is also misleading and incomplete to classify them as "unstable". Physicists have long used the term *metastable* to capture this interesting phenomenon
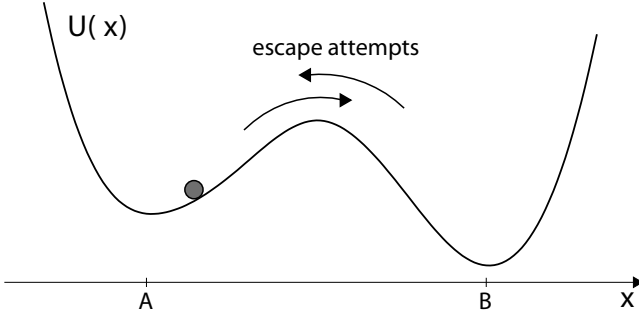
Fig. 1. Cartoon of a particle subject to Brownian motion in a potential $U(x)$ with two metastable states, $A$ and $B$.

and have developed a number of tools for quantifying this behavior (Hanggi, Talkner, & Borkovec, 1990; Kampen, 2007; Muller, Talkner, & Reimann, 1997; Talkner, Hanggi, Freidkin, & Trautmann, 1987). Many other branches of science and engineering have also borrowed this terminology to describe dynamic systems in a wide variety of fields. Familiar metastable systems include crystalline structures (e.g. diamonds), flip-flop circuits, radioactive elements, oscillatory wave patterns in the brain, and ferromagnetic materials, such as spin glass or magnetic tape film (which explains why a taped recording sitting in storage still inevitably fades over time).

The canonical example of metastability is a particle in a potential well subject to Brownian motion, as cartooned in Figure 1. These systems have local attractors which tend to keep the dynamics within a particular neighborhood in state space. In the limit as such systems become deterministic (no noise), these local attractors are fixed points, and the system is truly stable whenever the dynamics begin with an initial condition somewhere inside the basin of attraction of the fixed point. In contrast, stochasticity constantly pushes the dynamics about within this neighborhood, and for some systems and noise types, this turns a stable system into a metastable one. Occasionally but repeatedly, such systems will deviate particularly far from a metastable attractor in state space (making "escape attempts"), and eventually, they will successfully exit (by which we mean entering a region where a different attractor is now a far more dominating influence).

The study of metastable phenomena is reasonably well-developed. For readers searching for additional references on metastability, we suggest the following small, representative set of the literature on metastable system dynamics (Hanggi et al., 1990; Huisinga, Meyn, & Schutte, 2004; Muller et al., 1997; Gaveau & Schulman, 1998), tools for estimating stochastic stability in discrete cases (Markov chains) (Talkner et al., 1987; Bovier, Eckhoff, Gayrard, & Klein, 2000; Bovier, 2004; Boyd, Diaconis, & Xiao, 2004; Jain & Jain, 1994; Larralde & Leyvraz, 2005; Weber, Kube, Walter, & Deuflhard, 2006), and issues of model order reduction (Horenko, Dittmer, Fischer, & Schutte, 2006; Vijayakumar, D'Souza, & Schaal, 2005; Au, 2004). Additionally, two recommended texts on stochastic processes are (Gardiner, 2004; Kampen, 2007).

## III. METASTABLE LIMIT CYCLE ANALYSIS

The dynamics of walking systems are continuous, but they are punctuated by discrete impact events when a foot comes into contact with the ground. These impacts provide a natural time-discretization of a gait onto a Poincaré map. Therefore, we will consider walking systems governed by the discrete, closed-loop return-map dynamics:

$$\mathbf{x}[n+1] = \mathbf{f}(\mathbf{x}[n], \gamma[n]), \qquad (1)$$

where $\mathbf{x}[n]$ denotes the state of the robot at step $n$ and $\gamma[n]$ represents the slope of the ground, which is a random variable drawn independently from a distribution $P_\gamma$ at each $n$. This model for stochastically rough terrain dramatically simplifies our presentation in this paper, but it also restricts our analysis to strictly forward walking[1]. These state evolution equations naturally represent a discrete-time, continuous-state Markov process (or infinite Markov chain). For computational purposes, we will also discretize the continuous-state dynamics, approximating the continuous state space with a finite set of states, $x_i$. Defining the state distribution vector, $\mathbf{p}[n]$, as

$$p_i[n] = \Pr(\mathbf{X}[n] = x_i), \qquad (2)$$

we can describe the state distribution (master) equation in the matrix form:

$$\mathbf{p}[n+1] = \mathbf{p}[n]\mathbf{T}, \quad T_{ij} = \Pr(\mathbf{X}[n+1] = x_j \mid \mathbf{X}[n] = x_i). \qquad (3)$$

$\mathbf{T}$ is the (stochastic) state-transition matrix; each row must sum to one. The $n$-step dynamics are revealed by the Chapman-Kolmogorov equation,

$$\mathbf{p}[n] = \mathbf{p}[0]\mathbf{T}^n.$$

We obtain the transition matrix numerically by integrating the governing differential equations forward from each mesh point, using barycentric interpolation (Munos & Moore, 1998) to represent the transition probabilities; details of that discretization are in the Appendix.

For walking, we will designate one special state, $x_1$, as an absorbing state representing all configurations in which the robot has fallen down. Transitions to this state can come *from* many regions of the state space, but there are no transitions *away* from this absorbing, failure state. Assuming that it is possible to get to this failure state (possibly in multiple steps) from any state, then this absorbing Markov chain will have a unique stationary distribution, with its entire probability mass in the absorbing state.

The dynamics of convergence to the absorbing state can be investigated using an eigenmode analysis (Byl & Tedrake, 2006, 2008b). Without loss of generality, let us order the eigenvalues, $\lambda_i$, in order of decreasing magnitude, and label the corresponding (left) eigenvectors, $\mathbf{v}_i$, and characteristic times, $\tau_i = \frac{-1}{\log(\lambda_i)}$. The transition matrix for an absorbing Markov chain will have $\lambda_1 = 1$, while $\mathbf{v}_1$ represents the stationary distribution, with all probability mass in the absorbing state. The magnitudes of the remaining eigenvalues

---

[1]Including backward steps is straightforward, but it requires the model to include spatio-temporal correlations in the slope angle.

$(0 \leq |\lambda_i| < 1, \forall i > 1)$ describe the transient dynamics and convergence rate (or mixing time) to this stationary distribution. Transient analysis on the walking models we investigate here will reveal a general phenomenon: $\lambda_2$ is very close to 1, and $\tau_2 \gg \tau_3$. This is characteristic of metastability: initial conditions (in eigenmodes 3 and higher) are forgotten quickly, and $\mathbf{v}_2$ describes the long-living (metastable) neighborhood of the dynamics in state space. In metastable systems, it is useful to define the *metastable distribution*, $\phi$, as the stationary distribution conditioned on having not entered the absorbing state:

$$\phi_i = \lim_{n \to \infty} \Pr(\mathbf{X}[n] = x_i \mid \mathbf{X}[n] \neq x_1).$$

This is easily computed by zeroing the first element of $\mathbf{v}_2$ and normalizing the resulting vector to sum to one.

Individual trajectories in the metastable basin are characterized by random fluctuations around the attractor, with occasional "exits", in which the system enters a region dominated by a different attractor. For walking systems this is equivalent to noisy, random fluctuations around the nominal limit cycle, with occasional transitions to the absorbing (fallen) state. The existence of successful escape attempts suggests a natural quantification of the relative stability of metastable attractors in terms of first-passage times. The *mean* first-passage time (MFPT) to the fallen absorbing state describes the time we should expect our robot to walk before falling down, measured in units of discrete footsteps taken.

Let us define the mean first-passage time vector, $\mathbf{m}$, where $m_i$ is the expected time to transition from the state $x_i$ into the absorbing state. Fortunately, the mean first-passage time is particularly easy to compute, as it obeys the relation:

$$m_i = \begin{cases} 0 & i = 1 \\ 1 + \sum_{j>1} T_{ij} m_j & \text{otherwise} \end{cases}$$

(the expected first-passage time must be one more than the expected first-passage time after a single transition into a non-absorbing state). In matrix form, this yields the one-shot calculation:

$$\mathbf{m} = \begin{bmatrix} 0 \\ (\mathbf{I} - \hat{\mathbf{T}})^{-1} \mathbf{1} \end{bmatrix}, \tag{4}$$

where $\hat{\mathbf{T}}$ is $\mathbf{T}$ with the first row and first column removed. The vector of state-dependent mean first-passage times, $\mathbf{m}$, quantifies the relative stability of each point in state space.

One interesting characteristic of metastable systems is that the mean first-passage time around an attractor tends be very flat; most system trajectories rapidly converge to the same metastable distribution (forgetting initial conditions) before escaping to the absorbing state. Therefore, it is also meaningful to define a *system mean first-passage time*, $M$, by computing the expected first-passage time over the entire metastable distribution,

$$M = \sum_i m_i \phi_i. \tag{5}$$

When $\tau_2 \gg \tau_3$, we have $M \approx \tau_2$, and when $\lambda_2 \approx 1$, we have

$$M \approx \tau_2 = \frac{-1}{log(\lambda_2)} \approx \frac{1}{1 - \lambda_2}.$$

## IV. NUMERICAL MODELING RESULTS

This section uses two simple, classic walking models to demonstrate use of the methodology presented in Section III and to illustrate some of the important characteristics typical for metastable walking systems more generally. The two systems presented here are the rimless wheel and the passive compass gait walker, each of which is illustrated in Figure 2.
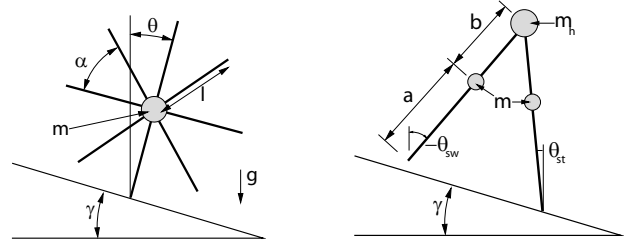


Fig. 2. The Rimless Wheel (left) and Compass Gait Walker (right) models.

### A. Rimless Wheel Model

The rimless wheel (RW) model consists of a set of $N$ massless, equally-spaced spokes about a point mass. Kinetic energy is added as it rolls downhill and is lost at each impulsive impact with the ground. For the right combination of constant slope and initial conditions, a particular RW will converge to a steady limit cycle behavior, rolling forever and approaching a particular velocity at any (Poincaré) "snapshot" in its motion (e.g., when the mass is at a local apex, with the stance leg oriented at $\theta = 0$ in Fig. 2). The motions of the rimless wheel on a constant slope have been studied in depth (Coleman et al., 1997; Tedrake, 2004).

In this section, we will examine the dynamics of the RW when the slope varies stochastically at each new impact. To do this, we discretize the continuous set of velocities, using a set of 250 values of $\omega$, from 0.01 to 2.5 (rad/s). We also include an additional absorbing failure state, which is defined here to include all cases where the wheel did not have sufficient velocity to complete an additional, downhill step. Our wheel model has $N = 8$ spokes ($\alpha = \frac{\pi}{4}$). At each ground collision, we assume that the slope between ground contact points of the previous and new stance leg is drawn from an approximately[2] Gaussian distribution with a mean of $\bar{\gamma} = 8°$.

For clarity, we will study only wheels which begin at $\theta = 0$ with some initial, downhill velocity, $\omega_o$, and we consider a wheel to have failed on a particular step if the mass does not reach an apex in travel ($\theta = 0$) with $\omega > 0$. (Clockwise rotations go downhill, as depicted in Fig. 2, and have positive values of $\omega$.) Note that the dynamic evolution of angular velocity over time does not depend on the choice of a particular magnitude of the point mass, and we will use spokes of unit length, $l = 1$ meter, throughout.

On a constant slope of $\gamma = 8°$, any wheel which starts with $\omega_o > 0$ has a deterministic evolution over time and is

---

[2]To avoid simulating pathological cases, the distribution is always truncated to remain within $\pm 10°$, or roughly $6\sigma$, of the mean.

guaranteed to converge to a fixed point of $\omega = 1.2097$ (rad/s). The return map defining the step-to-step transitions from $\omega_n$ to $\omega_{n+1}$ for the Poincaré section taken when the stance leg is precisely vertical is given as:

$$\omega_{n+1} = \sqrt{\cos^2\alpha\left(\omega_n^2 + \frac{2g}{L}\left(1 - cos\beta_1\right)\right) - \frac{2g}{L}\left(1 - cos\beta_2\right)}$$

where $\beta_1 = \frac{\alpha}{2} + \gamma$ and $\beta_2 = \frac{\alpha}{2} - \gamma$, with $\gamma > 0$ as the downhill slope. A plot of this return function is shown in Figure 3.
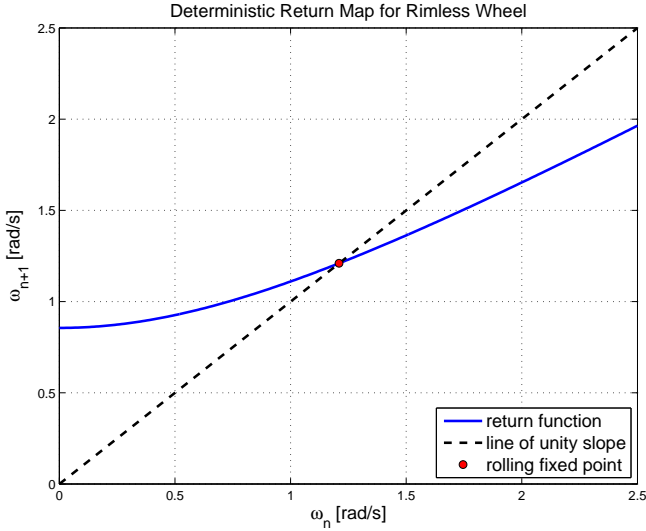


Fig. 3. Return map and fixed point for an 8-spoke rimless wheel on constant, downhill slope of $8°$. Here, $\omega_n$, is defined as the angular velocity when the support spoke is exactly vertical.
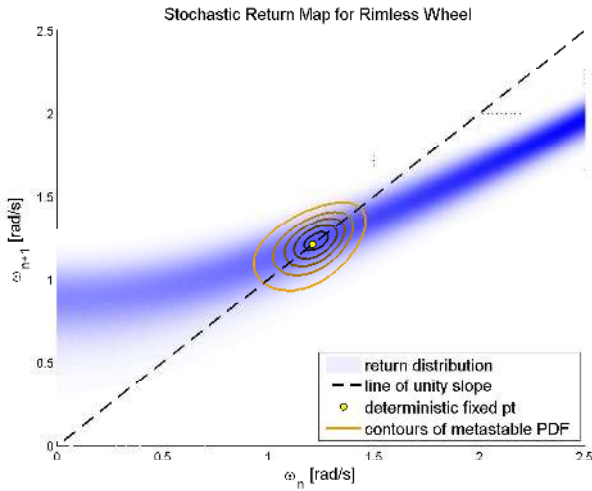


Fig. 4. Return distribution and metastable "neighborhood" for an 8-spoke rimless wheel on downhill terrain with a mean step-to-step slope of 8 degrees and $\sigma = 1.5°$. There is now a probability density function describing the transition from $\omega_n$ to $\omega_{n+1}$.

When the slope between successive ground contacts is drawn from a stochastic distribution, the function given in Figure 3 is now replaced by a probabilistic description of the transitions, as illustrated in Figure 4. Given the current state is some particular $\omega_n$, there is a corresponding probability

mass function (PMF) to describe what the next state, $\omega_{n+1}$, will be. Because the actual state space of the true system dynamics are continuous, we have chosen here to illustrate these probabilities using the more intuitive probability *density* function (PDF), instead of using the discrete PMF.

Figure 5 shows the set of PDF's describing the stochastic dynamics more clearly; it is a 3D plot of the same probabilistic return map shown from overhead in Figure 4. Each height value in Figure 5 is proportional to element $T_{ij}$ of the transition matrix, where $i$ is the state we are coming from ($\omega_n$, on the x-axis) and $j$ is the state we are going to ($\omega_{n+1}$, on the y-axis); Figures 4 and 5 provide a graphical representation of the transition matrix describing this metastable dynamic system.
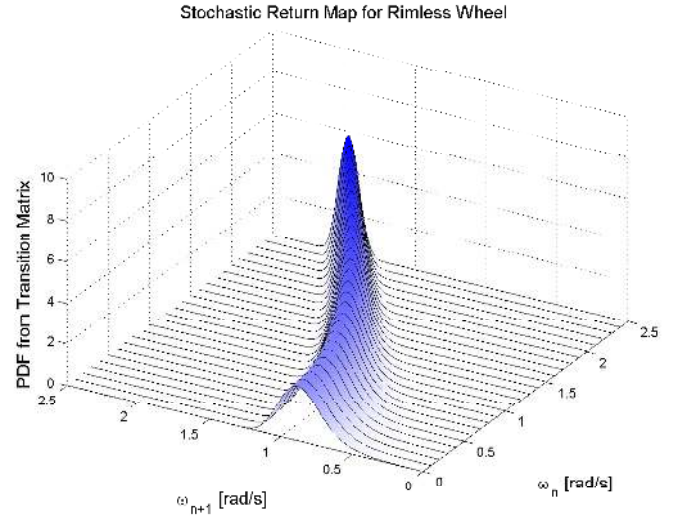


Fig. 5. 3D view of the return distribution for the stochastic rimless wheel system. This is a smoothed rendering of the step-to-step transition matrix, $T$, with the probability density functions for some particular values of $\omega_n$ overlaid as lines for greater clarity.

To generate the discrete transition matrix, we calculate $\omega_{n+1} = f(\omega_n, \gamma)$ for each of a discrete set of 601 possible $\gamma$ values, in the range of $\pm 10$ degrees from the mean. Each new state is then represented in the mesh using barycentric weighting interpolation (Munos & Moore, 1998), which (we note) inherently adds a small level of additional (unintended) noise to the modeled dynamics.

In Figures 4 and 5, the terrain slope variation has a standard deviation of $\sigma = 1.5°$. Using MATLAB to take the 3 largest eigenvalues of the transpose of the transition matrix for this case, we find that the largest eigenvalue, $\lambda_1$, is within $10^{-14}$ of being exactly unity, which is within the mathematical accuracy expected. This eigenvalue corresponds to the absorbing failure state, and the corresponding eigenvector sums to 1, with all values except the failure state having essentially zero weight[3] in this vector (since all rimless wheels will eventually be at this state, as $t \rightarrow \infty$). All other eigenvectors sum to zero (within numerical limits), since they must die away as $t \rightarrow \infty$. The second-largest eigenvalue is $\lambda_2 = 0.999998446$. Using the

[3] All states except the failure state had a magnitude less than $10^{-10}$, numerically.

methods presented in Section III, this corresponds to a *system-wide* MFPT of about $1/0.000001554 = 643,600$ steps.

Each initial condition also has a particular *state-dependent* MFPT, $m(\omega)$, which is obtained from Eq. 4 and plotted in Figure 6. Note that the value of the mean first-passage time is nearly flat throughout a large portion of state space. This is characteristic for metastable systems, which justifies the notion of a "system-wide" MFPT, $M \approx 1/(1 - \lambda_2)$, quantifying the overall stochastic stability of a particular dynamic system. For this particular case, there are no regions in state space (except the failure state) with MFPT significantly lower than the system-wide value, which is *not* typical more generally; the passive compass gait walker in Section IV-B is highly sensitive to initial conditions, for example, although it too has regions of state space which share a nearly uniform MFPT value.
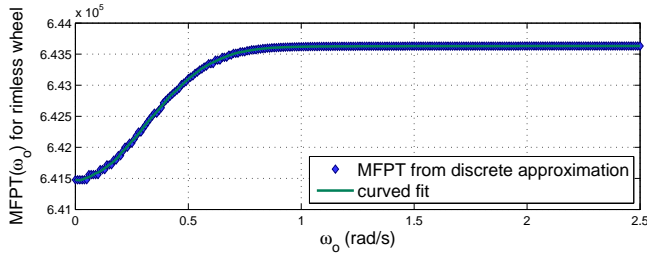


Fig. 6. Mean first-passage time as a function of the initial condition, $\omega_o$. Data are for a rimless wheel on stochastic terrain with mean slope of 8 deg and $\sigma = 1.5°$. Points show the approximation obtained through eigen-analysis of the discretized system, and a smoothed line is overlaid. Note that MFPT is largely constant over a large portion of state space.

The eigenvector associated with $\lambda_2$ yields the PDF of the metastable dynamic process – the relative probability of being in any particular location in continuous state space, given initial conditions have been forgotten and the walker has not yet failed. Figure 7 shows the resulting probability distribution functions for the metastable rimless wheel for each of several levels of terrain noise. Pictorially, each system-wide PDF for a metastable system is analogous to the fixed point for a stable, deterministic system. In the deterministic case, the probability of being exactly at the fixed point approaches unity as $t \to \infty$.

The third-largest eigenvalue of the transition matrix, $\lambda_3$, quantifies the characteristic time scale in which initial conditions are forgotten, as the dynamics evolve toward the metastable distribution (or toward failure). For the case illustrated in Figures 4, 5 and 6 (i.e., $\sigma = 1.5°$), $\lambda_3 \approx 0.50009$, which means almost half of the contribution to the probability function at the initial condition which is due to this eigenvector is lost ("forgotten") with each, successive step; an even larger fraction-per-step is lost for all remaining eigenvectors (as they will have even smaller values of $\lambda$). Within a few steps, initial conditions for any wheel beginning in our range of analysis, $0 < \omega_o \leq 2.5$, have therefore predominantly evolved into the metastable PMF (or have failed). If we multiply the (discretized) metastable PMF, $\phi(\omega)$, by the transition matrix, we obtain the joint probability, $Pr(\omega_n, \omega_{n+1})$, of having just transitioned from $\omega_n$ to $\omega_{n+1}$, given the wheel has not failed by step $n+1$. This joint probability is illustrated as a smoothed
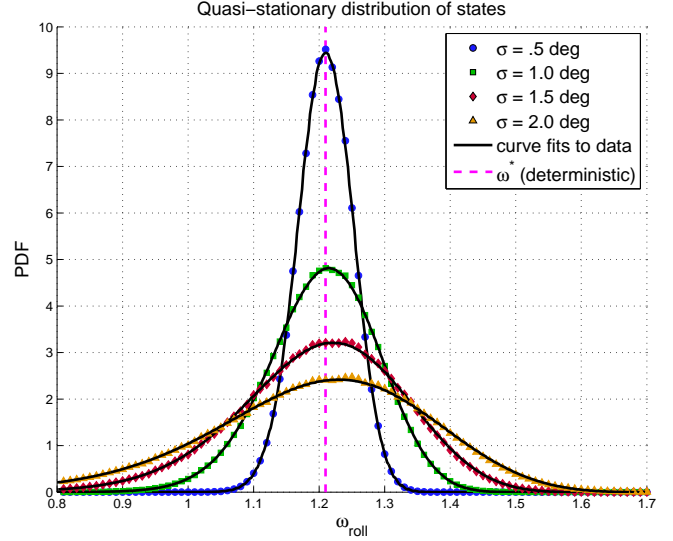


Fig. 7. Quasi-stationary probability density functions for the stochastic rimless wheel for each of several values of terrain noise, $\sigma$. Each distribution is estimated by renormalizing the eigenvector associated with the second-largest eigenvalue of the transpose of the transition matrix. Note that meshing inherently adds noise to the dynamic system; smoothed lines are drawn on top of the raw data (shown as points) from the scaled eigenvectors.

PDF in a 3D plot over continuous state space in Figure 8 and as a set of overlaid contour lines in Figure 4.
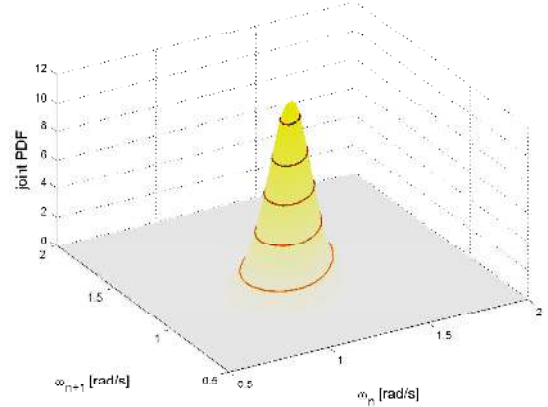


Fig. 8. 3D view of the metastable "neighborhood" of state-to-state transitions, $(\omega_n, \omega_{n+1})$. If a rimless wheel starts from some arbitrary initial condition and has not fallen after several steps, this contour map represents the joint probability density function of being in state $\omega_n$ now and transitioning to $\omega_{n+1}$. The contour lines drawn are identical to those overlaid in Figure 4. They correspond to the neighborhood of likely $(\omega_n, \omega_{n+1})$ pairings, analogous to the unique fixed point of the deterministic case.

This particular system has a beautiful simplicity which allows us to extract some additional insight from the conditional probability in Figure 8. Because of the definition of $\omega_n$ as being the velocity when the mass is at its apex in a given step, the value of $\omega_{n+1} = 0$ represents the boundary to the absorbing failure state in this example. If we visualize the contours of the conditional probability as they extend toward $\omega_{n+1} = 0$ in Figure 4, we see that most failures do not occur because we transition from a very slow state ($\omega_n$ close to zero) to failure but are more typically due to sudden transitions from

more dominant states in the metastable distribution to failure.

Finally, when this methodology is used to analyze the rimless wheel for each of a variety of noise levels ($\sigma$), the dependence of system-wide MFPT on $\sigma$ goes as shown in Figure 9. For very low levels of noise, MATLAB does not find a meaningful solution (due to numerical limits). As the level of noise increases, the MFPT decreases smoothly but precipitously. (Note that the y-axis is plotted on a logarithmic scale.) The stochastic stability of each particular system can be quantified and compared by calculating this estimate of MFPT which comes from $\lambda_2$ of the transition matrix.
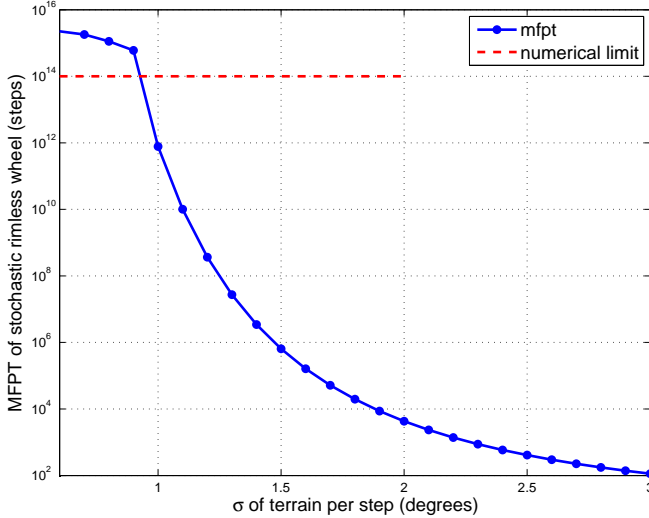


Fig. 9. Mean first-passage time (MFPT) for the rimless wheel, as a function of terrain variation, $\sigma$. Estimates above $10^{14}$ correspond to eigenvalues on the order of $1 - 10^{-14}$ and are beyond the calculation capabilities of MATLAB.

### B. Passive Compass Gait Walker

The second metastable dynamic system we analyze in this paper is a passive compass gait (CG) walker. This system consists of two, rigid legs with distributed mass. In our model, there are three point masses: at the intersection of the legs ("the hip") and partway along each leg. The dynamics of the compass gait have been studied in detail by several authors, e.g., (Garcia et al., 1998; Goswami, Thuilot, & Espiau, 1996; Spong & Bhatia, 2003). Referring to Figure 2, the parameters used for our metastable passive walker[4] are $m = 5$, $m_h = 1.5$, $a = .7$, and $b = .3$. Given an appropriate combination of initial conditions, physical parameters and *constant* terrain slope, this ideal model will walk downhill forever.

When each step-to-step terrain slope is instead selected from a stochastic distribution (near-Gaussian, as in Section IV-A), evolution of the dynamics becomes stochastic, too, and we can analyze the stochastic stability by creating a step-to-step transition matrix, as described in detail for the rimless wheel. The resulting system-wide MFPT as a function of terrain noise, $M(\sigma)$, is shown in Figure 10. Note that it is similar in shape to the dependence shown in Figure 9.

[4]This particular mass distribution was chosen based on empirical results so that it provides good passive stability on rough terrain.
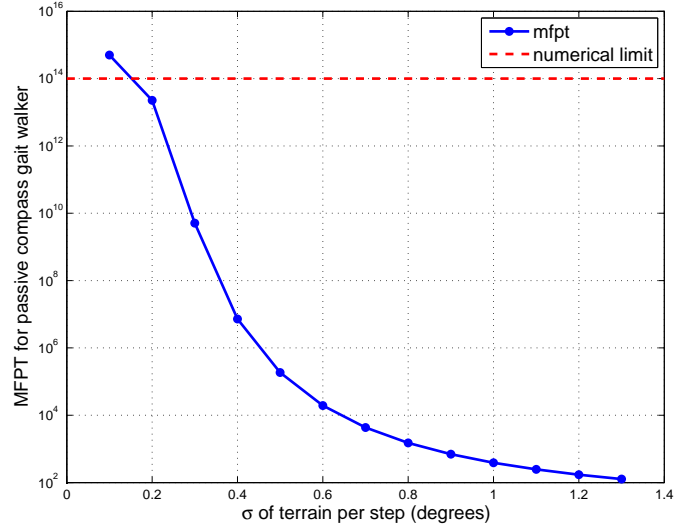


Fig. 10. Mean first-passage time as a function of terrain variation. Results for analysis of a compass gait walker using a discretized (meshed) approximation of the transitions. Average slope is 4 degrees, with the standard deviation in slope shown on the x-axis. Noise is a truncated Gaussian distribution, limited to between 0 and 8 degrees for all cases.

To analyze this system, our discretized mesh is defined using the state immediately after each leg-ground collision. The state of the walker is defined completely by the two leg angles and their velocities. On a constant slope, these four states are reduced to three states, since a particular combination of slope and inter-leg angle will exactly define the orientation of both the stance and swing leg during impact. Although the slope is varying (rather than constant) on stochastic terrain, we still use only three states to define our mesh. To do so, we simulate the deterministic dynamics (including impacts) a short distance forward or backward in time to find the robot state at the Poincaré section where the slope of the line connecting the "feet" of the legs is equivalent to our desired, nominal slope. Because the dynamics between collisions are entirely deterministic, these two states are mathematically equivalent for the stochastic analysis. If such a state does not exist for a particular collision (which occurs only very rarely), we treat this as a member of the absorbing failure state. This approximation allows us to reduce the dimensionality from 4 states to 3, which improves numerical accuracy significantly. Specifically, it has allowed us to mesh finely enough to capture near-infinite MFPT for low-noise systems, while using four states did not. The three states we use in meshing are: (1) absolute angular velocity of the stance leg, $X_3$, (2) relative velocity of the swing leg, $X_4$, and (3) the inter-leg angle, $\alpha$.

Figure 11 shows a slice of the basin of attraction (BoA) for this compass gait on a constant slope (top), along with regions in state space with nearly-constant MFPT (bottom two) for two different magnitudes of noise ($\sigma$) in terrain. Each slice is taken at the same inter-leg angle, $\alpha \approx 25.2°$. In the deterministic case, the basin of attraction defines the set of all states with infinite first-passage time: all walkers beginning with an initial condition in this set will converge toward the fixed point with probability 1. For stochastic systems which result in metastable

dynamics, there is an analogous region which defines initial conditions having MFPT very close to the system-wide value, $M$. Interestingly, the deterministic and stochastic basin shapes are quite similar here; we expect this may often be the case for systems such as this with discrete jumps in state space.
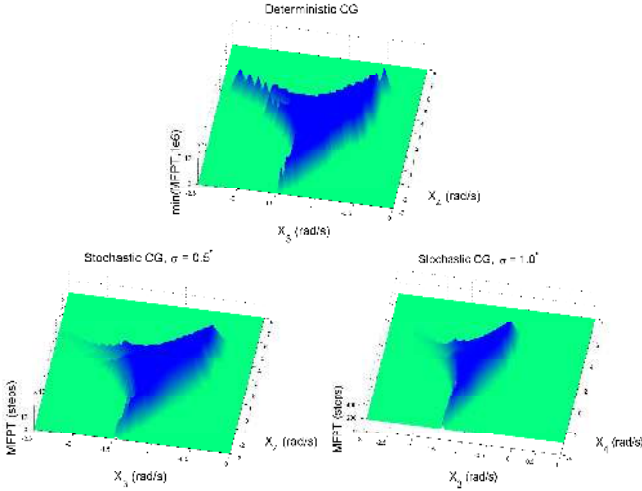


Fig. 11. Basin of attraction (top) for deterministic CG walker and map of MFPT for lower-noise ($\sigma = 0.5°$, lower left) and higher-noise ($\sigma = 1.0°$, lower right) examples. To aid in visual comparison, all 3 plots use the same mesh. The "near-constant MFPT basin" for each stochastic system is essentially a low-pass filtered version of the deterministic basin of attraction, and its shape does not change significantly, even when the magnitude of the MFPT itself varies greatly (e.g., 180,000 steps [left] vs 390 [right]). This region represents a boundary in the volume in state space within which a walker is likely to be pulled into the metastable distribution.

The image at the top of Figure 12 shows the deterministic BoA for this CG walker more clearly. This plot was generated by sampling carefully over the state space and simulating the dynamics. By contrast, the plot at the top of Figure 11 intentionally uses the same mesh discretization used for the stochastic system, to provide a better head-to-head comparison of the change in shape due to the addition of terrain noise (as opposed to the noise of the discretization itself). The second image in Figure 12 shows the deterministic basin of attraction for a different set of physical parameters ($m = m_h$; $a = b = .5$) on the same, constant slope of $4°$. This basin looks qualitatively more delicate and the resulting performance of this walker on stochastic terrain is in fact much worse (e.g., MFPT of about 20 steps when $\sigma = 0.5°$, where we find $M = 180,000$ for the other walker).

Just as in the case of the rimless wheel, the fixed point (for our deterministic compass gait system) is now replaced (in the stochastic case) by a probability density function, defining the likelihood of being in any particular state (conditioned on not having fallen) as $t \to \infty$. Figure 13 shows 2D contour plot sections of the approximate PDF, obtained from an eigen-analysis of the stochastic compass gait. The outermost contour defines a boundary containing 0.999 of the probability distribution in state space. The distribution spreads over more of state space as the level of noise increases, in a manner analogous to the widening of the probability distribution with noise seen in Figure 7.

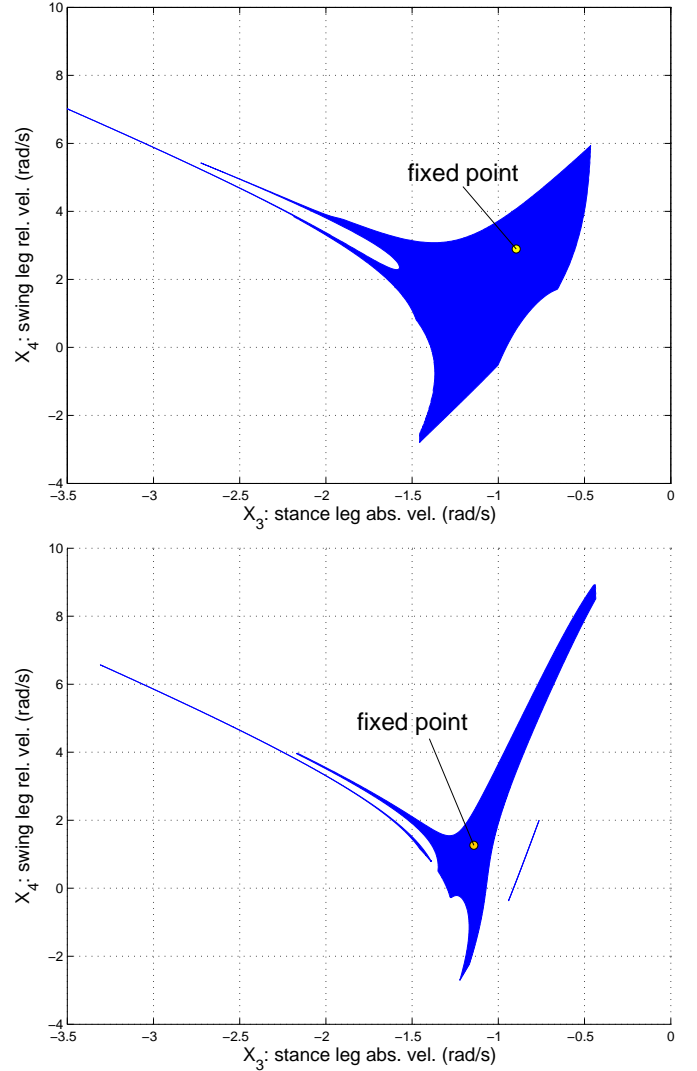Finally, we note that the relationship in state space between



Fig. 12. Basins of attraction (blue regions) and fixed point for two compass gait walkers, each on a constant slope of $4°$. Walker with basin at top is more stable and uses the parameters defined for the stochastic system described throughout Section IV-B; for the other walker, $m = m_h$ and $a = b = .5$. MFPT is infinite inside the shaded region and is small (1-4 steps) outside of it. This image shows only a slice of the 3D basin, taken at the inter-leg angle of the fixed point for each respective walker. The fixed point is at $X_3 = -.89$ (rad/s), $X_4 = 2.89$ (rad/s), $\alpha = 25.2°$ for the first walker, and it is at $X_3 = -1.14$ (rad/s), $X_4 = 1.26$ (rad/s), $\alpha = 33.4°$ for the lower one. The deterministic basin of attraction for the second walker is narrower in shape, and this walker is significantly less stable on stochastic terrain.

the PDF of the metastable dynamics, shown in Figure 13, and the region of nearly-uniform mean first-passage time, $M$, shown at the bottom of Figure 11, hints at where successful "escape attempts" (i.e., failures) are most likely to occur over time. Figure 14 overlays these two regions across a different dimensional slice of the 3D space for both $\sigma = .5°$ and for $\sigma = 1.0°$. As the tails of the metastable PDF (shown in yellow) approach the boundary of the near-uniform MFPT basin (shown in blue), there is a higher probability of failing on any given step during the metastable process, resulting in turn in a less stochastically stable system (i.e., one with a lower system-wide value of $M$).
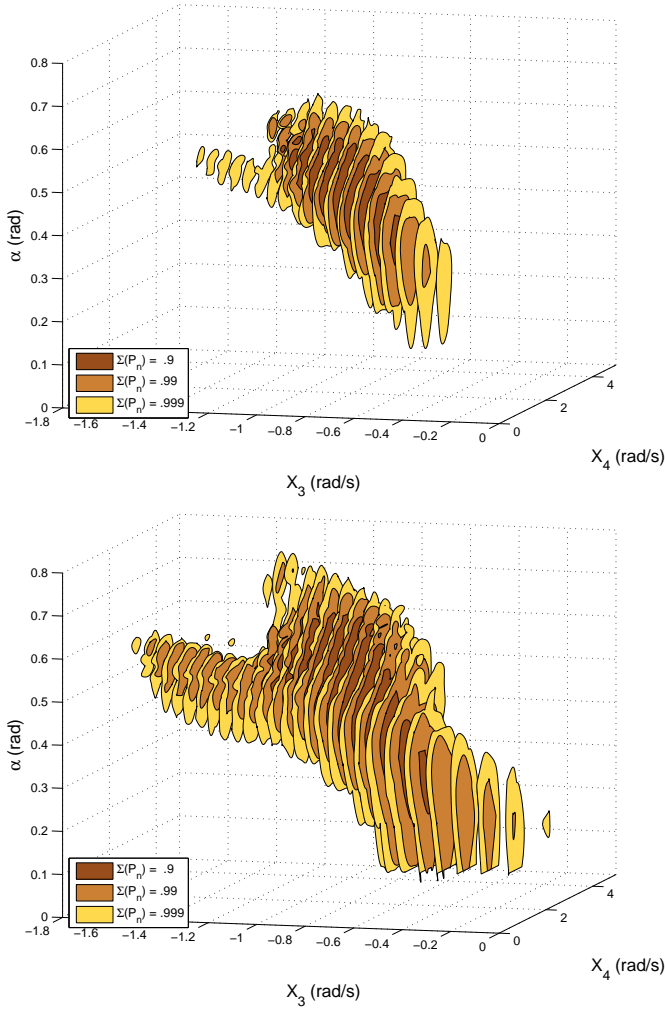
Fig. 13. On stochastic terrain, there is no fixed point for the compass gait walker. Instead, there are metastable "neighborhoods" of state space which are visited most often. As time goes to infinity, if a walker has not fallen, it will most likely be in this region. The contours shown here are analogous to the PDF magnitude contours in Figure 7; they are drawn to enclose regions capturing 90%, 99%, and 99.9% of walkers at any snapshot during metastable walking. Top picture corresponds to $\sigma = 0.5°$. Larger noise ($\sigma = 1.0°$, bottom) results in larger excursions in state space, as expected.



Fig. 14. Metastable system: Contours of the stochastic "basin of attraction" are shown where MFPT is $0.5M$, $0.9M$ and $0.99M$ (blue) versus contours where the integral of the PDF accounts for .9, .99, and .999 of the total metastable distribution (yellow). The metastable dynamics tend to keep the system well inside the "yellow" neighborhood. As the tails of this region extend out of the blue region, the system dynamics become less stochastically stable (lower $M$). The axis into the page represents the swing leg relative velocity, $X_4$, and a slice is taken at $X_4 = 2.33$ (rad/s). Terrain variability for the top plot is $\sigma = 0.5$ degrees (with $M \approx 180,000$ steps). For the noisier system at bottom ($\sigma = 1.0$ degrees), $M$ is only 20 steps or so.

## C. Mean-first Passage Time as Policy Evaluation

The numerical analysis performed on these simple (low-dimensional) models, based on a fine discretization of the state space, will not scale to more complicated systems. The analysis, however, can be generalized to higher dimensional systems by observing that the mean-first passage time calculation can be recast into a "policy evaluation"(Sutton & Barto, 1998) - estimating the expected cost-to-go of executing a fixed policy - using the one-step cost function:

$$g(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \in \text{fallen} \\ -1 & \text{otherwise.} \end{cases} \quad (6)$$

If this cost function is evaluated on every visit to the Poincare map, then evaluating the infinite horizon cost-to-go,

$$V(\mathbf{x}_0) = \sum_{n=1}^{\infty} g(\mathbf{x}[n]), \quad \mathbf{x}[0] = \mathbf{x}_0 \quad (7)$$
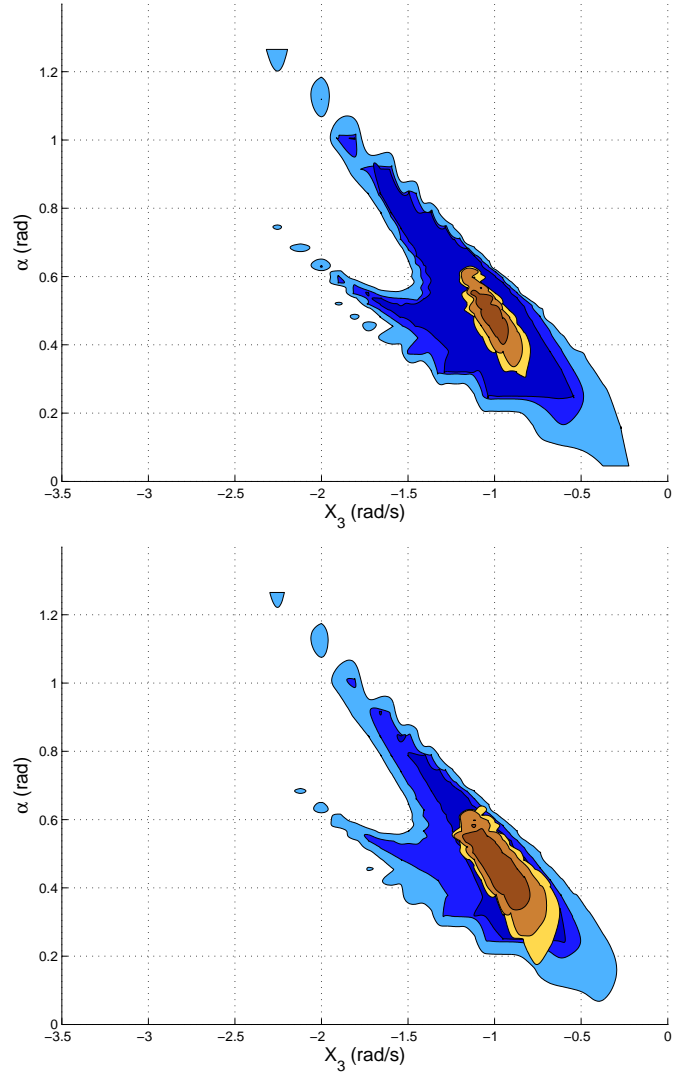
is equivalent to evaluating the (negative of the) mean-first passage time of the system (in units of number of steps). If we wished to evaluate the MFPT in seconds, we could replace the $-1$ on each step with the time taken for the step. The same idea works to evaluate distance traveled. Although the cost function is deceptively simple, the resulting policy evaluation computation is quite rich due to the stochasticity in the plant model.

Recasting the MFPT calculation in terms of optimal control has a number of important implications. First, in the case of a Markov chain with known transition probabilities, computing $V(\mathbf{x}_0)$ exactly reduces to equation 4. But in the case where the state space is too large to be effectively
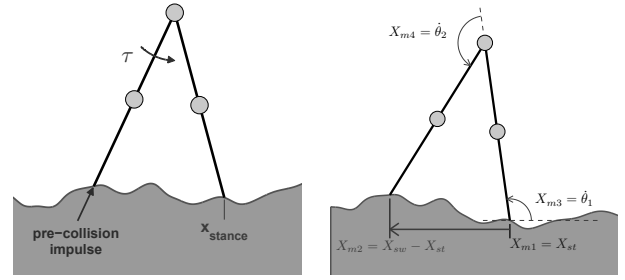
Fig. 15. Compass gait on wrapping terrain. At left, the actuation for this model includes a torque at the hip and a constant magnitude impulse which occurs at each step. At right, the model is labeled with the four states which define each post-collision element in the discrete mesh.

discretized, this policy evaluation can also be accomplished very efficiently with function approximation, using methods from reinforcement learning such as least-squares temporal-difference learning (LSTD)(Boyan, 2002). Although reinforcement learning, in general, struggles with high-dimensional systems, it is important to remember that policy evaluation is considerably easier and more efficient than policy optimization and can be implemented effectively in medium-scale problems. Furthermore, these tools can potentially allow the MFPT to be evaluated on a real robot, even without a model of the plant.

The connection between mean first passage times and optimal control cost functions also makes it natural to consider the problem of designing a controller which maximizes the MFPT. This optimization, in general, will require the tools for policy optimization from optimal control and reinforcement learning. In this paper, with a focus on simple models, we have the luxury of examining the implications of the optimal control formulation using brute-force dynamic programming methods operating on the discretized dynamics - which in the control case are a Markov Decision Process (MDP).

## V. Approximate optimal control on rough terrain

In this section, we study a minimally-actuated version of the compass gait walker introduced in Section IV-B on rough terrain. Two types of rough terrain are explored: wrapping terrain (allowing the walker to travel an arbitrary distance while using a finite representation of a particular terrain) and stochastic terrain (where the change in terrain height at each footstep is drawn from a random distribution). While our goal is to maximize the MFPT of the walker, we limit our control policy here to a hierarchical approach, so that we can employ the tools introduced in Section III directly to further illustrate the characteristics of metastable walking, now for the case of an actively-controlled system. Correspondingly, this section presents results for the *approximate* optimal control of an actuated compass gait walker on rough terrain.

In Section V-A, we begin our presentation on optimization of the mean first-passage time metric for the actuated compass gait walker by examining the performance of a baseline case, where we have access to perfect terrain knowledge. In Sections V-B and V-C, we then quantify and discuss the performance that is achieved when only limited or statistical information about upcoming rough terrain is available in solving for the optimal control policy.

A full description of the actuated model, the general control strategy over which we perform optimization, and the optimization algorithm itself are presented in detail in the Appendix.

### A. Controlled CG on Wrapping Rough Terrain

Figure 15 illustrates both example terrain (at a somewhat exaggerated scale) and the four meshing states used in discretizing the system dynamics. The actuation strategy used throughout our presentation of the controlled compass gait is described in Section IX-A. Briefly, it employs 1) a torque source at the hip joint where the legs meet one another and 2) an impulsive toe-off, delivered axially at each step. The

torque is set by a low-level PD controller, which attempts to drive the interleg angle to a particular set-point angle, selected once per step. It is this choice of a desired interleg angle which constitutes the control policy for our model. Section IX-B presents our implementation of the value iteration algorithm for deriving the approximate optimal stochastic control solution for this actuated walking model; the same actuated model and value iteration algorithms are used throughout all work presented in Section V.
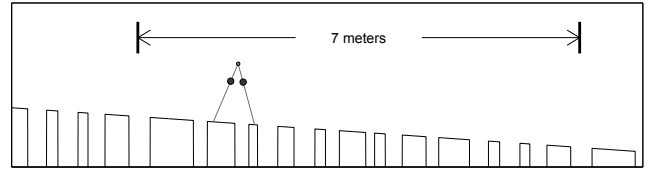


Fig. 16. Intermittent-foothold terrain. Terrain profile repeats every 7 meters, allowing the numerical mesh to "wrap".

*Results on wrapping terrain:* Below are simulation results for the control strategy of a constant magnitude toe-off and PD regulation of desired interleg angle, described in detail in Section IX-B. We attempted to solve the optimal control problem on a variety of both continuous and discontinuous[5] wrapping terrains. Examples are shown in Figures 16, 17, 18 and 20.

Summarized briefly, the PD control contribution can be effective in regulating foot placement on terrain, although it also contributes importantly to the coupled dynamic motions of both legs. Interestingly, use of PD control alone, (i.e., without the impulsive toe-off at each step) allows for only very limited variations in step width or height. Use of an impulse control action alone (without PD control) cannot regulate the upcoming step length effectively in our simulations, even if we allow for the selection of a different magnitude of impulse at each step. Combining a constant magnitude toe-off with the PD controller provides significantly better performance than either actuation component demonstrates alone, allowing for significantly greater variations in both step length and step height during continuous walking, through the combined

---

[5]Our tests of the hip-actuated compass gait walker on discontinuous terrain were inspired by the results from (Hodgins & Raibert, 1991) in using a simple, intuitive strategy to control the Raibert hopper on intermittent terrain.

actions of energy addition (from the toe-off) and step length regulation (via hip torque).
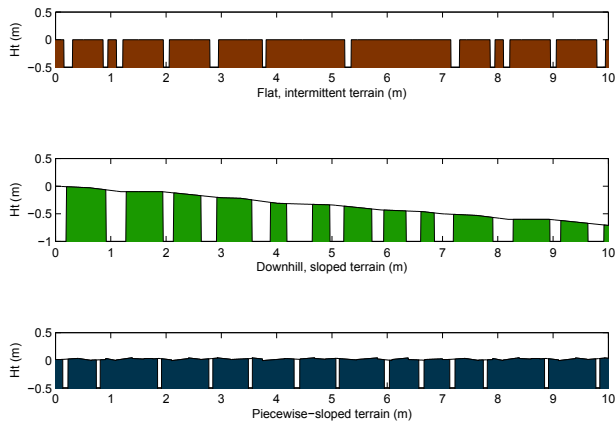


Fig. 17. Examples of terrain which were successfully negotiated using PD control alone. Without the added energy contribution of an impulsive toe-off, the actuated walker is limited to either mild variations in height or to terrain with an average gradient downhill. Terrain at top is flat except for a set of bottomless no-go gaps. The middle and bottom examples consist of piecewise slopes. Each was successfully solved both as a continuous terrain, shown as a solid black line, and as an intermittent terrain, represented by the solid regions. Figure 18 shows detail from the bottom plot.
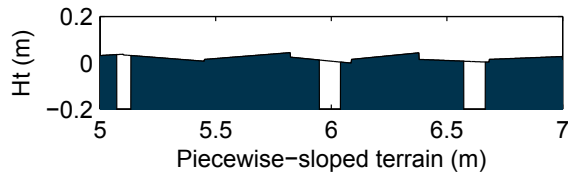


Fig. 18. A close-up of the bottommost terrain in Figure 17.

For our combined actuation strategy of toe-off and PD regulation of the interleg angle at each step, the terrain shown in Figure 20 was iteratively scaled (vertically) to find the most extreme variations that could still be negotiated. The most extreme scaling for which the optimal control policy yields continuous walking had a standard deviation in terrain height of 4.7 cm and a max-to-min difference in height of 19.1 cm. The deviations in the width and height of the actual footsteps taken using the optimal policy vary in height with a standard deviation of 3.3 cm, and the SD of the inter-leg angle at each step is about 5.9°.

### B. Controlled CG on Stochastically Rough Terrain

In addition to the wrapping terrain models described in Section V-A, we also solved for the approximate optimal policy for the same, actuated walker on stochastically rough terrain. In the stochastic case, the height difference between successive footholds is drawn from a Gaussian distribution. Figure 21 illustrates the nature of the resulting terrain. This model for the terrain is somewhat artificial, since height would typically be a function of step length on real terrain. By contrast, the next upcoming step in our simulation occurs at a particular pre-selected height, regardless of step length. However, this compact terrain model can still produce arbitrarily
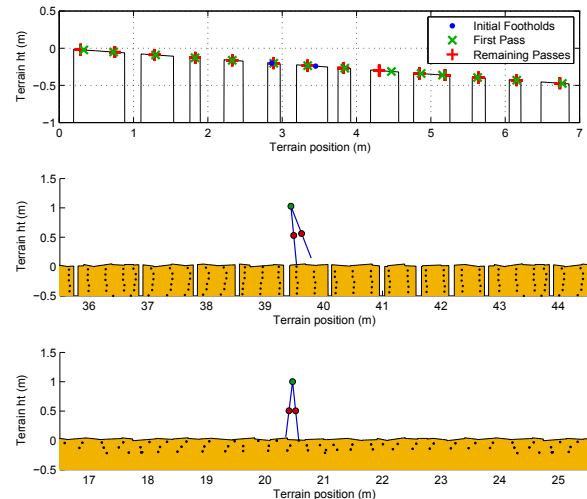


Fig. 19. Foothold patterns from optimal control. The optimal footholds on extreme terrain quickly converge to fixed pattern (top). On easier terrain, no fixed pattern emerges (bottom). In intermediate terrain, we see an intermediate level of organization in the pattern of footholds (middle). Footholds taken are plotted as dots in the lower two figures, with the earliest steps plotted lowest and with dots repeating as the terrain repeats, every 7 meters.



Fig. 20. Dots represent the horizontal locations of footsteps taken during a 60-second trial using the optimal control policy from value iteration.



Fig. 21. Compass gait on stochastic terrain. At left, the actuation for this model again includes a torque at the hip and a constant magnitude impulse which occurs at each step. At right, the four states which define each post-collision elements in the discrete mesh are given.

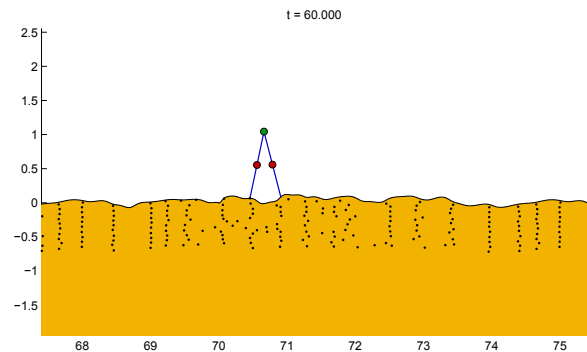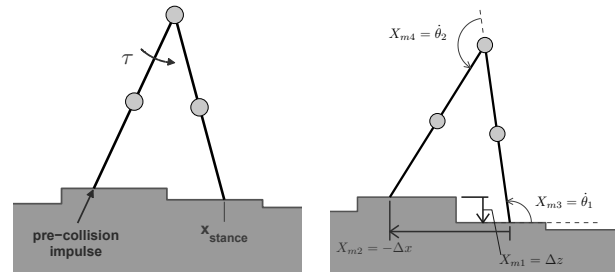extreme terrain (by increasing the standard deviation of the step heights, as desired). Based on the results obtained on wrapping terrain, we included both the low-level PD controller and the constant-magnitude impulsive toe-off in all simulations on stochastic terrain. As in the wrapping terrain case, details on the discretization of meshing are given in the Appendix.

Optimizing the MFPT cost function over the post-collision state (which again used four state variables, as for the wrapping terrain) on stochastic terrain results in a policy that requires no information about *specifics* of the terrain but does (through the transition model used in the optimization) know the *statistics* of what may happen. Not surprisingly, the optimal control solutions for this "blind" walker were quite poor when compared with the performance obtained on wrapping terrain with statistically similar variations in terrain height.

Intuitively, including some information about the immediately-upcoming terrain should improve the performance of our underactuated biped significantly. For example, it would allow the biped walker to execute a shorter step (which loses less energy) if going uphill or a longer step going downhill. We tested this hypothesis by implementing a one-step lookahead. This required enhancing the four-variable representation of the post-collision state of the walker with an additional, fifth state variable: the $\Delta z$ value giving the height of the next, immediate step.

This one-step lookahead improved the mean first-passage on stochastic terrain dramatically. Figure 22 compares the MFPT for the one-step and no lookahead cases. For example, on terrain where the value of $\Delta z$ is known to be drawn from a Gaussian with zero mean and a standard deviation of 1 cm, the MFPT is about 76 with no lookahead, versus 12,000 with the additional knowledge of the particular height of the next, immediate step.

Comparisons between the passive walker and the same, two control policies are even more dramatic, as one would expect. We note first that the passive walker can only locomote continually on *downhill* terrain, so that our comparisons only have any meaning at all if we allow the passive walker to use gravity to compensate for energy lost at each ground collision. Given this limitation, a passive walker with the same physical parameters (i.e., masses and lengths) walking down an average $4°$ slope with a standard deviation of 0.5 cm in terrain per step has a mean first-passage time (MFPT) of only about 9.3 steps. When the active walker is tested on terrain with the same variability (but on terrain that is flat on average - not downhill), no lookahead and one-step policies yield MFPT's on the order of $10^4$ and $> 10^9$ steps, respectively. Although the controlled system is still significantly underactuated, its performance is dramatically improved over that of the purely passive dynamics when we optimize for rough-terrain walking.

Given these results, it is natural to enquire how a two-step or (more generally) an n-step lookahead would compare on a similar plot. Unfortunately, obtaining such results would require multiplying the size of the total mesh by a factor of 19 for each additional lookahead, because our representation of the upcoming terrain is discretized to be drawn from one of 19 particular values. To compare the one-step lookahead with a longer lookahead, we will instead return (in Section V-

C, ahead) to the case of wrapping terrain, where arbitrary lookahead can be obtained without adjusting the size of our state space (nor our mesh).
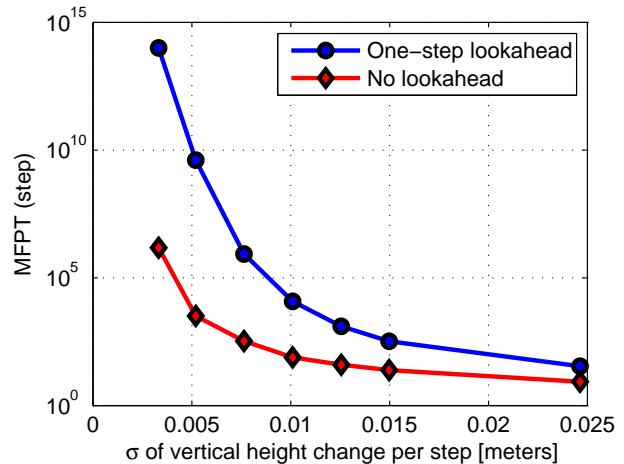


Fig. 22.   MFPT for one-step vs. no lookahead on stochastic terrain

### C. Efficacy of a Limited Lookahead on Wrapping Terrain

For our passive walking examples, we noted that initial conditions are forgotten very quickly when walking on rough terrain. Correspondingly, we posit that an actively-controlled walker may often require only a short lookahead to negotiate rough terrain with near-optimal results. Intuitively, if we assume that a particular walker has not yet fallen down, a quick mixing time (to forget initial conditions rapidly) implies that the consequences of decisions made now should not have a significant impact on long-term stability, given that at least a few steps will in fact be successfully taken.

To explore this hypothesis, we began with an estimate of the value function for a model of *stochastic* terrain that had approximately the same standard deviation in step-to-step terrain height as our known, *wrapping* terrain. Using this initialization of the value function, we then ran the value iteration algorithm for one additional pass, now using the known, wrapping terrain in our dynamic model (instead of the stochastic approximation). This strategy provides the optimal policy with one-step lookahead knowledge of the particular terrain. Similarly, running a second pass results in the optimal policy for a two-step lookahead, and so on.

This approach results in surprisingly effective walking, but it has important limitations. We found that even the one-step lookahead solution was able to walk continuously on terrain which was intentionally scaled until the optimal solution (with an essentially infinite lookahead) was only barely able to achieve continuous walking. As one would expect, any $n$-step policy (where $n > 1$) resulted in the same performance. Essentially, because the terrain was wrapping, it was impossible to distinguish between two different policies which both had very large MFPTs. This became clear when we observed that simulations of the one-step lookahead policy on wrapping terrain resulted in a repeating pattern of footholds,

automatically selected over time. If the walker successfully completed its first pass around the wrapping terrain, it would often converge to a particular, repeating pattern of twenty of so steps that resulted in continuous walking, forever. Figure 23 shows such a convergence of states over time. As observed in Section V-A, we notice that selected footsteps are more apt to repeat themselves on wrapping terrain which is more extreme; the foothold selection appears more haphazard on "easier" terrain. We suspect this occurs because particular regions of the terrain have a locally-optimal "goal" foothold, which acts to regulate an entire, emergent pattern on wrapping terrain.

Finally, we should emphasize, briefly, that resetting the value function with the optimal values found for control on stochastic terrain is a critical step: if the value function is initiated as zero everywhere (for example), then a single pass of value iteration only results in a value of either 0 or -1 at every mesh state, for each possible action. This means the only information we have is whether or not a single step will be successful. Not surprisingly, the resulting policy does not work well in our simulations. The walker has no notion of the stability of its internal dynamic state when the value function is initiated with such a "blank slate"; using the solution from stochastic terrain to preset the value function intrinsically provides some information about the dynamic capability to take future steps.
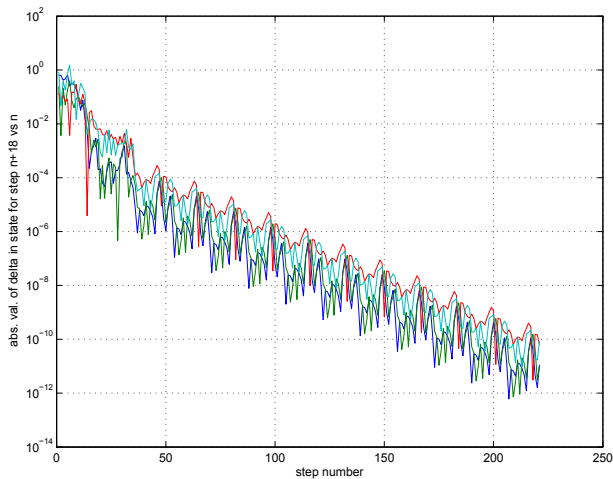


Fig. 23. One-step policy shows convergence on wrapping terrain. Even when only a one-step lookahead is used to fix a policy on wrapping terrain, the states converge over time to a particular, repeating pattern. This logarithmic plot shows the deviation in the value in each state, comparing step $n$ to step $n+18$, i.e., after one complete 18-step "lap" about the wrapping terrain. The plot illustrates an exponential convergence of the controlled system to a fixed trajectory through state space.

The fact that the one-step control policy is nearly as effective as the optimal policy on our wrapping terrain simulations provides some evidence that use of a limited lookahead (e.g., 1 to 3 steps) results in near-optimal performance. We find this particularly compelling, as it correspondingly implies that near-optimal results should also be possible on real-world legged robots having only a short-sighted knowledge of upcoming terrain. We anticipated this result after analyzing our results on passive CG walkers, where we noted a sur-

prisingly fast mixing dynamics of the metastable limit cycle generated by a passive compass gait walker on statistically rough terrain (Byl & Tedrake, 2006, 2008c), and further work is currently planned to test this hypothesis on a real compass-gait robot, as discussed in Section VI-C.

## VI. DISCUSSION

This section briefly discusses our use of the stochastic methods presented in designing controllers for walking systems. Additionally, we provide further observations on the properties of metastable systems which result in multiple attractors (e.g., period-$n$ gaits). Finally, we note work is currently underway to apply our methods and results to real walking machines.

### A. Impacts on Control Design

One of the primary goals of a controller is to enhance the dynamic stability of a system. For walking systems, we propose throughout this paper that this should be defined as increasing the *stochastic stability* of a walking machine. We would like time-to-failure to be long, and we would like a system to converge toward its long-term, metastable distribution quickly from a large set of initial conditions. The methods described here can be used in optimizing controllers with either or both of these two aims in mind.

One key point to observe is that the perspective that walking is (by its nature) a *metastable process* results in optimal control policies which automatically adapt to exploit knowledge about the upcoming terrain characteristics; as the statistics and particular local features of the environment change, the optimal strategy for walking changes as well. Ideally, our performance metrics for a robot should not be dictated by nominal distances from particular reference trajectories (e.g., in state space) but should instead fully reflect the ultimate goal of the device: maximizing the longevity of successful locomotion. Correct formulation of our optimization problem can achieve this.

As a motivating example, we have presented results for an active compass gait model, with a torque source at the hip but with the ankles still unactuated at the ground contact during mid-step. By designing a low-level PD controller to regulate inter-leg angle, we simplify the control problem to a create a tractable model which can be solved using standard numerical methods from machine learning. Our results for such a control methodology allow this underactuated compass gait model to walk continuously over impressively rough terrain. Some of our initial results for control on wrapping terrain have previously appeared in (Byl & Tedrake, 2008a).

### B. Periodicity of Stable Limit Cycles

Metastable dynamic systems sometimes have an inherent periodicity. We expect this may be the case on a slightly steeper slope, for instance, where compass gait models experience period-doubling bifurcations (Goswami, Thuilot, & Espiau, 1996). Another case where periodicity arises is for wrapping terrain, such as the terrain for the controlled walker in Figure 24. Wrapping is a realistic model for many in-laboratory walking robots, as they are often confined to walk
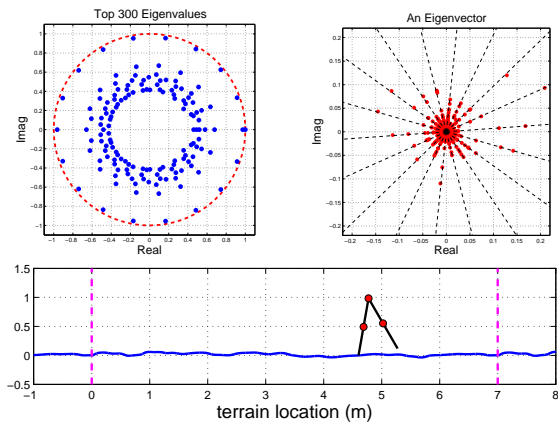
Fig. 24. Controlled compass gait walker, with torque at the hip. To solve for an optimal policy using value iteration, the terrain wraps every 7 meters. The optimization maximizes the MFPT from any given state. An eigen-analysis reveals a complex set of eigenvalues (top), spaced evenly about (but strictly inside of) the unit circle. Corresponding eigenvectors are also complex.



Fig. 25. Compass gait robot posed on rough terrain.

on a boom – repeatedly covering the same terrain again and again. In our simulation of a hip-actuated CG walker on wrapping terrain, we observe that a repeating, n-step cycle results in multiple eigenvalues, $\lambda_2$ through $\lambda_{n+1}$, all with magnitude just under unity. They are complex eigenvalues, as are the corresponding eigenvectors. The top left image in Figure 24 shows such a set of eigenvalues, all lying just within the unit circle. The next-smallest set of eigenvalues are all significantly smaller in this example. The complex eigenvalues and eigenvectors mathematically capture an inherent periodicity, in which the probability density function changes over time in a cyclical manner.

### C. Verification of Compass Gait Results on a Real Robot

Given the success of our simulations, we plan to implement a similar control strategy on a real compass-gait robot. This robot is shown in Figure 25 and was designed and built by labmates at the Robot Locomotion Group at MIT. The design is intentionally similar to the idealized model studied in our investigations here of active control. Specifically, it has a direct-drive motor at the hip and a constant-magnitude (post-collision) toe-off at each sensed ground collision. It is mounted on a boom, providing lateral stability but also introducing some additional, unmodeled dynamics. We operate this robot in a motion capture environment on well-characterized terrain, allowing us to know both the state of the robot and the upcoming, rough terrain profile with good accuracy, and have completed a number of initial control experiments(Iida & Tedrake, 2009; Manchester, Mettin, Iida, & Tedrake, 2009).

### D. Implications for Development of Highly Dynamic Robots

We conclude by highlighting the implications our results have toward the development of dynamic, autonomous robots in the coming years. First, we reiterate that global stability will not typically exist for our walking machines, and that our goal should be to optimize *stochastic stability*.

The success of short-sighted strategies, discussed in Section V-C, has important implications in kinodynamic planning

for legged robots. It means that near-optimal strategies may simply require good low-level control, and that selecting a "greedy" short-term action may often be a good policy for long-term stability.

Finally, we note that much of this work applies more generally to a broader class of highly dynamic robots (flying, swimming, etc.) in realworld environments, and that we have presented powerful tools which can be adapted quite naturally for machine learning.

## VII. Conclusions

The goal of this paper has been to motivate the use of stochastic analysis in studying and (ultimately) enhancing the stability of walking systems. Robots that walk are inherently more prone to the stochastic influences of their environment than traditional (e.g., factory) robots. Locomotory systems capable of interacting with the real world must deal with significant uncertainty and must perform well with both limited energy budgets and limited control authority.

The stochastic dynamics of walking on rough terrain fit nicely into the well-developed study of metastability. The simplified models studied here elucidate the essential picture of metastable limit cycle dynamics which make occasional escape attempts to the fallen-down state. We anticipate that metrics for stochastic stability, such as the mean first-passage time, will provide potent metrics for quantifying both the relative stability across state-space and the overall system stability for real walking systems.

## VIII. Acknowledgements

## IX. Appendix:Compass Gait Implementation Details

This appendix provides details on the compass gait model simulations on both stochastic and (known) wrapping terrain.

Included are a description of the control strategy implementation for the actuated model and of the value iteration algorithm used to obtain an approximate optimal control solution for a given terrain type. This is followed by the equations of motion that describe the dynamic system and details on meshing and on implementation of interpolation for post-collision states in state space.
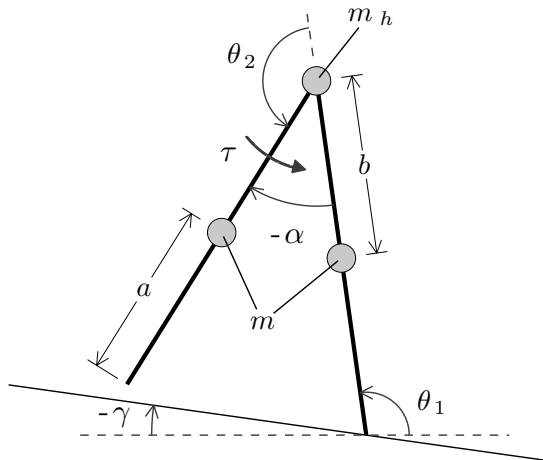
## A. Actuated Compass Gait Model



Fig. 26. Actuated compass gait model with torque source at hip. Simulations for the actuated compass gait use the following parameters: $m = m_h$ = 2kg, $a = b$ = 0.5m

The basic compass gait model, introduced by (Goswami, Espiau, & Keramane, 1996), is used in all simulations and is depicted in Figure 26. In addition to the modeling details previously given for the passive compass gait (in Section IV-B), the actuated model now also includes a pure torque source at the "hip" linkage where the legs meet. The stance leg is oriented at an absolute angle $\theta_1$ and contacts the ground at a zero-torque, frictionless joint, called the "toe". The swing leg position is measured relative to the stance leg at angle $\theta_2$. Parameters of the model used for the controlled walker are given in the caption of Figure 26, and the direction of walking for all simulations is (once again) strictly from the left to the right.

In addition to the hip torque actuation, we also include an ideal impulse actuation at each toe which we assume can be delivered at a prescribed magnitude, axially at the stance leg contact at the moment that the swing leg collides with the ground, when support is instantaneously transferred from one leg to the other. Inclusion of an impulse source is inspired by the work of (Kuo, Donelan, & Ruina, 2005), who illustrate that this can be an efficient method for imparting energy to the system to make up for energy lost in ground collisions[6].

## B. Approximate Stochastic Optimal Control

To solve for the approximate optimal control solution for the compass gait walker given a particular definition of rough

[6]Note that (Kuo et al., 2005) use the simplified version of the compass gait, with all mass concentrated at the hip and infinitesimal mass at each toe. However, the effect is similar in the model with distributed mass.

terrain, we use the same methods as before both to discretize and to interpolate the dynamics. Now, however, instead of the discretization resulting in a Markov Chain, the result is a Markov Decision Process, where the transition matrix, **T**, is parameterized by one or both of the following possible control actions: (1) an instantaneous impulse at the toe and (2) the hip torque. To avoid using an enormous action space of possible torque profiles for a given step for this illustrative control example, we apply hip torques based on a proportional-derivative (PD) controller which regulates the inter-leg angle, $\alpha$. The choice of control action is the desired inter-leg angle, $\alpha_{des}$, which can be selected from a discretized set of values. We employ the value iteration algorithm (Sutton & Barto, 1998) to find an optimal step-to-step feedback policy for the discretized system. We present results for the system with hip-torque alone, with the impulsive toe-off alone, and with both sources of actuation together.

To address the practical concerns of defining and detecting ground collisions as the swing leg moves across variable-height terrain, we assume the swing leg shortens in length immediately after a ground collision occurs and remains so until an interleg angle within $10°$ of the desired value is achieved. At this point, the swing leg extends to full length, instantaneously. Note that this retraction does not affect the continuous dynamics of the walker, since the lower portion of the leg is massless. If, after simulating the extension of the swing leg to its full length, we discover its tip ("foot") would be below the ground height, we consider the walker to have failed to have taken a successful step: it then enters the absorbing failure state. For successful steps, the identities of the stance and swing legs are swapped when the (extended) swing leg toe finally collides with the ground.

*1) Value iteration algorithm:* Our goal is to find a policy which will maximize the number of steps taken before falling. To determine the full capabilities of the walker, one would wish to discretize with greater resolution over time and to select a torque magnitude at each $dt$. We chose instead to discretize on a step-to-step basis, to reduce the size of the state space (by one dimension) in order to keep the problem computationally tractable. We were pleased to find that such a strategy still produces continuous walking on significantly rough terrain in our simulations.

The post-collision state of the walker is represented in most plots using four meshing state variables: $X_{m1}$: $x$ location of the stance leg; $X_{m2}$: the distance $\Delta x$ from the stance leg ground contact to the swing leg contact; $X_{m3}$: the angular velocity of the stance leg, $\dot{\theta}_{st}$; and $X_{m4}$: the angular rate of change of the inter-leg angle, $\dot{\alpha}$. Details on the discretization used are given in the Appendix (IX). Note that the inter-leg angle is increasing when the swing leg rotates "forward". Counter-clockwise is positive for all angle measurements, as illustrated by Figure 26.

In solving for the optimal policy, we begin by assigning a value $V_o(s) = 0$ for each state ($s$) in the mesh. Through iteration, we update these values to reflect which states can achieve the greatest estimated number of steps before falling. Equation 8 shows the update for the estimated cost of starting

in state $i$ and performing control action option $a$.

$$Q_n(i,a) = \sum_j T_{ij}^a \left[ g(j) + \gamma V_{n-1}(j) \right], \qquad (8)$$

where $g(j)$ is the cost function from equation 6, now defined as a function of discrete state $j$.

After this update is done for every possible *control action*, we can then update $V_n(s)$ for each state, so that it is the lowest of all $Q_n(s,a)$ values (corresponding to the best action found so far):

$$V_n(s) = \min_a Q_n(s,a) \qquad (9)$$

and our optimal n-step policy is the set of actions at each step (after $n$ iteration steps) which gives this lowest cost:

$$\pi_n(s) = \arg\min_a Q_n(s,a) \qquad (10)$$

We include a discount factor, $\gamma = 0.9$, in Equation 8 to ensure that the value function converges to a finite value everywhere as the number of iterations goes to infinity. As previously discussed, our choice of one-step cost function in this work results in a policy that maximizes MFPT, as measured in expected number of steps taken.

*2) Hierarchical controller design:* Our simulations use a two-part actuation strategy: PD control of a desired inter-leg angle and a constant magnitude impulsive stance-foot toe-off, applied just before the ground collision occurs for each new step. At each step, the control policy dictates a high-level control action, which is the set-point for the inter-leg angle, $\alpha_{des}$, to be used by the PD controller for this step. Below, we describe the qualitative contribution of each of the two actuations toward negotiating rough terrain.

The primary purpose of the PD controller is to regulate the step length of the walker, which in turn selects upcoming foot placement on the terrain. However, the dynamics are coupled, so that the controller also affects the entire dynamic state of the walker. There is currently no closed-form description for the step-to-step transitions resulting from this coupling, which is precisely why simulation of the dynamics provides a practical methodology for investigating the system.

The main goal in employing the impulsive toe-off action is to compensate for the energy that is lost at each ground collision. This in turn allows the walker to take larger steps than would otherwise be possible, since more energy is of course lost for larger step angles (Coleman, 1998).

*3) PD control of inter-leg angle:* A low-level PD controller regulates the inter-leg angle, $\alpha$, which is defined as:

$$\alpha = \theta_{sw} - \theta_{st} = \theta_2 - \pi \qquad (11)$$

Our PD controller was designed by hand, to obtain swing leg motions which do not overpower the stance leg dynamics entirely but which are still effective in regulating the desired step length approximately. The PD controller is active throughout the course of any particular step. Equation 12 gives the commanded controller torque at the hip:

$$\tau = K_p(\alpha_{des} - \alpha) + K_d(0 - \dot{\alpha}) \qquad (12)$$

where $K_p = 100$ and $K_d = 10$.

*4) Impulsive toe-off at impact:* In this control strategy, a particular magnitude impulse is applied axially from the stance leg downward, primarily to add energy at each step. Toward eventual implementation on a real robot, we note here that it is well-known that applying the impulse immediately *before* collision is more efficient in imparting energy to the compass gait walker (Kuo et al., 2005). We assume the impulse is imparted instantaneously in our simulations. As a result, a given toe-off simply results in a particular, prescribed, instantaneous change in the angular velocity of each leg, which is a function of the interleg angle. Correspondingly, no additional dynamic simulations are necessary to implement value iteration with this when this control actuation is included.

We note briefly that the velocity of the upcoming stance leg must be tested after the impulse is applied to ensure that this velocity is still directed toward the ground. A large enough impulse could theoretically send the entire walker airborne! However, this has not been a practical concern for the impulse magnitudes we have tested in our simulations, which are never large enough to overcome the acceleration of gravity.

Our initial tests using value iteration allowed for the selection of one of 21 values of impulse, ranging from 0 to 2 (kg-m/s). In practice, however, the largest allowable magnitude of impulse was almost always selected. The resulting optimal control policy performs almost identically to one in which the impulse value is set to a constant magnitude of 2 at every step.

*5) Scaling of wrapping terrain:* To test the performance limits of each control strategy analyzed, each terrain was scaled to make its features more dramatic until value iteration failed to converge on a stable walking solution. Each terrain consisted of a particular profile which repeats every 7 meters, as shown in Figure 16. This allows value iteration to converge on a fixed policy using a finite representation for the terrain. A repeating terrain may also be thought to represent the terrain of a compass gait walker circling endlessly on a fixed boom, such as the boom-mounted robot described in Section VI-C.
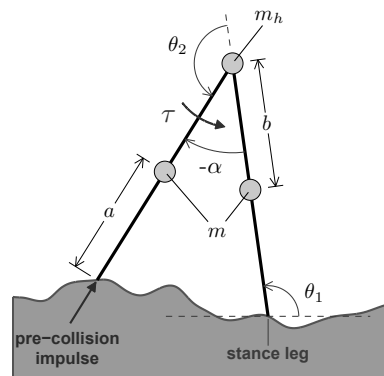
### C. Details of Dynamics and Meshing



Fig. 27. Four states of the CG walker during continuous dynamics. The biped model is depicted here immediately after an impact. A pre-collision impulse occurs at each step, and only the torque source at the hip actuates the walker until the next step occurs.
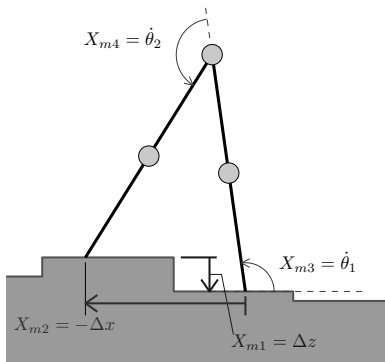
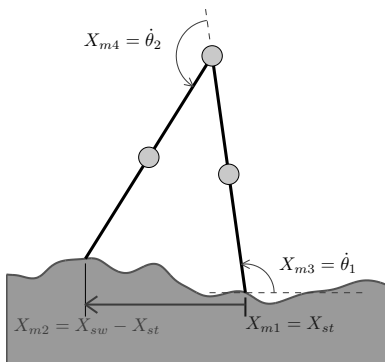Fig. 28.   Compass gait meshing parameters for stochastic terrain.



Fig. 29.   Compass gait meshing parameters for wrapping terrain.

If one considers only the continuous dynamics *between* two steps, the dynamics of the compass gait walker are *identical* to those of the underactuated acrobot. Figure 27 depicts the state variables and model parameters which define the continuous phase of the dynamics. With the stance leg of the compass gait modeled as a passive pivot joint, the entire state of the biped is defined by the two angles and two angular velocities of the legs. We use the absolute angle of the stance leg, $\theta_1$, and the relative angle of the swing leg (with respect to the stance leg), $\theta_2$, along with their time derivatives, $\theta_3 = \dot{\theta}_1$, $\theta_4 = \dot{\theta}_2$. The model includes a point mass $m_h$ at the "hip" and a point mass of magnitude $m$ along each of the (otherwise massless) legs, all at the geometric locations shown.
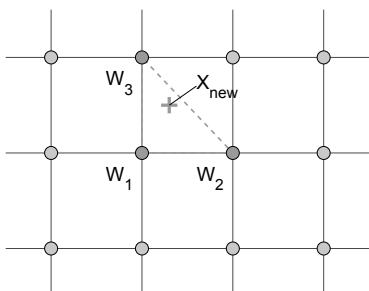
*Noise effects induced by meshing*



Fig. 30.   Illustration of meshing approximation.

The value function and feedback policy between the discretized points of our mesh must be approximated, since values are only stored at the particular mesh nodes. Figure 30 illustrates the effect of meshing; each new post-collision state, $X_{new}$, which we observe after a simulated step, must be approximated by a weighted sum of some subset of nodes which exist in our pre-determined mesh. The minimal number of nodes to create a meshed volume in an N-dimensional mesh is $N+1$; for our 4-dimensional mesh of post-collision states, this means each exact new state is modeled as if transitions occurs to 5 neighboring mesh nodes (each with corresponding weight, $W_k$). The individual transition probabilities are set through barycentric interpolation (Munos & Moore, 2002) (and must of course sum to unity). Any post-collision states falling outside the defined mesh range were automatically binned into the absorbing failure (fallen) state.

We note briefly that the optimal *policy* for walking on rough terrain often involves discontinuities. In these same regions, however, the cost function describing the *value* of being at a particular location in state space generally varies rather smoothly. As a specific example, there are situations where either a small or large step is acceptable, while a middle-sized step would result in stumbling (e.g., into a pothole). In the corresponding regions of our mesh, the best choice of step length may change suddenly from one of taking a "big" step to the choice of a "small" step. However, although the transition between selecting a large or small step may be sudden, the difference in value between one (rather extreme) action and the other may be relatively small and will tend to vary smoothly.
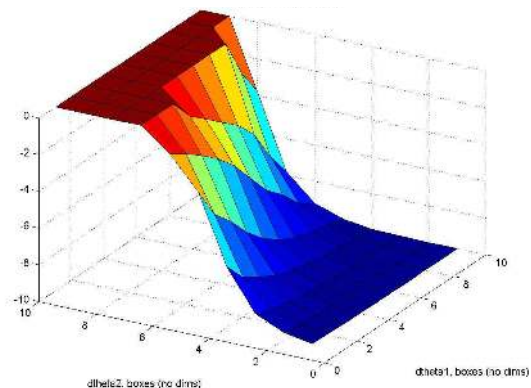


Fig. 31.   Typical smoothness of the cost function over state space.

*Implications of meshing on numerical accuracy*

The meshing noise discussed above acts to increase the overall level of noise above that of the intended stochasticity (if any) of the terrain. Toward ensuring reasonable numerical accuracy in our estimates of MFPT, our meshing implementation was first tested in a worst-case situation, where these discretization effects would be expected to have the most significant impact: a passive compass gait walker on a constant, downhill terrain (no noise). We consider this an extreme case because (1) the effects of meshing noise are more significant for the passive compass gait walker than for

TABLE I

MESHING FOR COMPASS GAIT MODEL ON STOCHASTIC TERRAIN

| Parameter | num. of elements | min | max | [units] |
|---|---|---|---|---|
| $X_{m1}$ | 19 | -.1 | .1 | [m] |
| $X_{m2}$ | 10 | -.7 | -.16 | [m] |
| $X_{m3}$ | 10 | -2.1 | -1.1 | [rad/s] |
| $X_{m4}$ | 10 | -1 | 1.5 | [rad/s] |
| $\alpha_{des}$ | 11 | 15 | 40 | [deg] |

TABLE II

MESHING FOR COMPASS GAIT MODEL ON WRAPPING TERRAIN

| Parameter | num. of elements | min | max | [units] |
|---|---|---|---|---|
| $X_{m1}$ | 140 | 0 | 7 | [m] |
| $X_{m2}$ | 15 | -.85 | -.15 | [m] |
| $X_{m3}$ | 14 | -3.0 | -.4 | [rad/s] |
| $X_{m4}$ | 14 | -.1 | 5.1 | [rad/s] |
| $\alpha_{des}$ | 13 | 10 | 40 | [deg] |

the controlled models, since active control increases the rate of convergence toward a metastable neighborhood and (2) the relative contribution of unwanted meshing noise toward the overall model noise goes down as the level of terrain noise increases.

For this test case of a passive walker on flat terrain, the numerically calculated value of $\lambda_2$ was within the numerical limits[7] of its expected value of exactly one (i.e., corresponding to an infinite MFPT). We believe the overall effects of meshing noise on numerical accuracy are correspondingly not quantitatively significant and that the effect of meshing noise is to provide a lower bound on the actual MFPT.

*Meshing details for stochastic terrain modeling*

Figure 28 illustrates the states used in meshing for optimal control on *stochastic* terrain. Only *post-collision* states are used in meshing, meaning both legs are in contact with the ground. For our stochastic terrain model, each upcoming step height is drawn from a Gaussian distribution. The figure illustrates the somewhat artificial nature of the terrain. The entire post-collision state on stochastic terrain is defined by the four *meshing* parameters shown. As compared with the state definitions in Figure 27, the first two states are now the horizontal and vertical distances between the two feet (rather than the two angles of the legs). The meshing states for stochastic terrain are also defined in the equations below:

$$X_{m1} = \Delta z = z_{st} - z_{sw}$$
$$X_{m2} = -\Delta x = x_{sw} - x_{st}$$
$$X_{m3} = \dot{\theta}_1$$
$$X_{m4} = \dot{\theta}_2$$

Table I gives the meshing discretization used for the stochastic terrain case. The spacing in the mesh for Table I is *not* linear. The exact values are listed below:

$$X_{m1} = [-.1, -.075, -.04, -.03 : .005 : .03, .04, .06, .1]$$
$$X_{m2} = [-.7 : .06 : -.16]$$
$$X_{m3} = [-2.1 : .1 : -1.4, -1.25, -1.1]$$
$$X_{m4} = [-1, -.7, -.5 : .25 : .75, 1.1, 1.5]$$

*Meshing details for wrapping terrain modeling*

Figure 29 illustrates the states used in meshing for optimal control on *wrapping* terrain. On wrapping terrain, the terrain

[7]in MATLAB, approximately $10^{-15}$

profile over a repeating, 7-meter long stretch of terrain is known exactly and is included in the simulation of step-to-step dynamics. Therefore, the entire post-collision state of the walker can be defined using the absolute $x$ position of the (upcoming) stance foot on terrain, the relative $x$ distance to the swing foot (which also touches the terrain instantaneously at collision), and the two angular velocities (given in both Figures 27 and 28). Meshing states for wrapping terrain are given below:

$$X_{m1} = x_{st}$$
$$X_{m2} = -\Delta x = x_{sw} - x_{st}$$
$$X_{m3} = \dot{\theta}_1$$
$$X_{m4} = \dot{\theta}_2$$

Table II gives the meshing discretization used for the stochastic terrain case. Unlike the stochastic case, the spacing in the mesh for the wrapping case (given in Table II) is linear. These simulations were done *before* the simulations on stochastic terrain, and the stochastic mesh was refined based on the results on wrapping terrain (to refine more-frequently visited regions in state space).

The rest of the modeling details given in this appendix apply to *both* the stochastic and wrapping terrain cases.

*Swing phase dynamics*

During the swing phase of the leg, the dynamics of the system are the so-called "acrobot" dynamics; this is a classic underactuated model which has been studied in depth (Spong, 1997; Boone, 1997; Brown & Passino, 1997; Spong, 1994). We intentionally choose to present the two equations of motion below using a familiar definition of the parameters and states:

$$d_{11}\ddot{\theta}_1 + d_{21}\ddot{\theta}_2 + h_1 + p_1 = 0 \qquad (13)$$

$$d_{12}\ddot{\theta}_1 + d_{22}\ddot{\theta}_2 + h_2 + p_2 = \tau \qquad (14)$$

where

$$d_{11} = m_1 l_{c1}^2 + m_2 \left( l_1^2 + l_{c2}^2 + 2l_1 l_{c2} \cos\theta_2 \right) + I_1 + I_2$$
$$d_{12} = d_{21} = m_2 \left( l_{c2}^2 + l_1 l_{c2} \cos\theta_2 \right) + I_2$$
$$d_{22} = m_2 l_{c2}^2 + I_2$$
$$h_1 = -m_2 l_1 l_{c2} \sin\theta_2 \dot{\theta}_2^2 - 2m_2 l_1 l_{c2} \sin\theta_2 \dot{\theta}_2 \dot{\theta}_1$$
$$h_2 = m_2 l_1 l_{c2} \sin\theta_2 \dot{\theta}_1^2$$
$$p_1 = \left( m_1 l_{c1} + m_2 l_1 \right) g \cos\theta_1 + m_2 l_{c2} g \cos\left(\theta_1 + \theta_2\right)$$
$$p_2 = m_2 l_{c2} g \cos\left(\theta_1 + \theta_2\right)$$

and

$$
\begin{aligned}
m_1 &= m_2 = m + \frac{1}{2} m_h \\
l_1 &= l_2 = a + b \\
l_{c1} &= L - \frac{bm}{m_1} \\
l_{c2} &= L - l_{c1} \\
I_1 &= I_2 = m (b - l_{c2})^2 + \frac{1}{2} m_h l_{c2}^2
\end{aligned}
$$

Solving explicitly for the time derivatives of the state variables, we obtain:

$$
\ddot{\theta}_1 = \frac{d_{22} \eta_1 - d_{12} \eta_2}{d_{11} * d_{22} - d_{12}^2} \tag{15}
$$

$$
\ddot{\theta}_2 = \frac{-d_{12} \eta_1 + d_{11} \eta_2}{d_{11} * d_{22} - d_{12}^2} \tag{16}
$$

where

$$
\begin{aligned}
\eta_1 &= m_2 l_1 l_{c2} \sin \theta_2 \dot{\theta}_2^2 + 2 m_2 l_1 l_{c2} \sin \theta_2 \dot{\theta}_2 \dot{\theta}_1 + \ldots \\
&\quad m_2 l_{c2} g \cos (\theta_1 + \theta_2) + (m_1 l_{c1} + m_2 l_1) g \cos \theta_1 \\
\eta_2 &= -m_2 l_1 l_{c2} \sin \theta_2 \dot{\theta}_1^2 - m_2 l_{c2} g \cos (\theta_1 + \theta_2) + \tau
\end{aligned}
$$

*Collision dynamics*

We assume collisions are instantaneous and inelastic. For convenience and for the reader to compare with typical notation, e.g. (Goswami, Thuilot, & Espiau, 1996), we will define absolute angular velocities of the legs:

$$
\dot{\theta}_{ns} = \dot{\theta}_3 + \dot{\theta}_4 \tag{17}
$$

$$
\dot{\theta}_s = \dot{\theta}_3 \tag{18}
$$

Geometry and conservation of angular momentum yield the following relationships between the pre-collision and post-collision states, which are given the superscripts $-$ and $+$, respectively.

$$
\theta_1^+ = \theta_1^- + \theta_2^- - \pi \tag{19}
$$

$$
\theta_2^+ = 2\pi - \theta_2^- \tag{20}
$$

$$
Q_{11}^- \dot{\theta}_{ns}^- + Q_{12}^- \dot{\theta}_s^- = Q_{11}^+ \dot{\theta}_{ns}^- + Q_{12}^+ \dot{\theta}_s^- \tag{21}
$$

$$
Q_{21}^- \dot{\theta}_{ns}^- + Q_{22}^- \dot{\theta}_s^- = Q_{21}^+ \dot{\theta}_{ns}^- + Q_{22}^+ \dot{\theta}_s^- \tag{22}
$$

where the last two equations can be expanded immediately, as given below. We also present an expanded form below, which has been useful for (vectorized) MATLAB implementation; many collision relationships can be calculated at once by using expanded forms of the matrix inverse and other algebraic expressions. Eqn. 21, which represents the angular momentum balance about the swing toe immediately before collision, will be modified later to account for pre-collision impulse later in the appendix.

$$
Q_{11}^- \left( \dot{\theta}_1^- + \dot{\theta}_2^- \right) + Q_{12}^- \dot{\theta}_1^- = Q_{11}^+ \left( \dot{\theta}_1^+ + \dot{\theta}_2^+ \right) + Q_{12}^+ \dot{\theta}_1^+ \tag{23}
$$

$$
Q_{21}^- \left( \dot{\theta}_1^- + \dot{\theta}_2^- \right) + Q_{22}^- \dot{\theta}_1^- = Q_{21}^+ \left( \dot{\theta}_1^+ + \dot{\theta}_2^+ \right) + Q_{22}^+ \dot{\theta}_1^+ \tag{24}
$$

where

$$
\begin{aligned}
Q_{11}^- &= -mab \\
Q_{12}^- &= -mab + \left( m_h L^2 + 2 m a L \right) \cos (2\alpha) \\
Q_{21}^- &= 0 \\
Q_{22}^- &= -mab \\
Q_{11}^+ &= mb (b - L \cos (2\alpha)) \\
Q_{12}^+ &= mL (L - b \cos (2\alpha)) + ma^2 + m_h L^2 \\
Q_{21}^+ &= mb^2 \\
Q_{22}^+ &= -mbL \cos (2\alpha)
\end{aligned}
$$

where $2\alpha = \pi - \theta_2^-$ is the interleg angle. We can solve for the two new angular velocities by using the matrix inverse of $\mathbf{Q}^+$, $\mathbf{R}^+ = (\mathbf{Q}^+)^{-1}$, which is simply:

$$
R_{11}^+ = \frac{Q_{22}^+}{Q_{11}^+ Q_{22}^+ - Q_{12}^+ Q_{21}^+}
$$

$$
R_{12}^+ = \frac{-Q_{12}^+}{Q_{11}^+ Q_{22}^+ - Q_{12}^+ Q_{21}^+}
$$

$$
R_{21}^+ = \frac{-Q_{21}^+}{Q_{11}^+ Q_{22}^+ - Q_{12}^+ Q_{21}^+}
$$

$$
R_{22}^+ = \frac{Q_{11}^+}{Q_{11}^+ Q_{22}^+ - Q_{12}^+ Q_{21}^+}
$$

Now let $\mathbf{A} = \mathbf{R}^+ \mathbf{Q}^-$:

$$
\begin{aligned}
A_{11} &= R_{11}^+ Q_{11}^- + R_{12}^+ Q_{21}^- \\
A_{12} &= R_{11}^+ Q_{12}^- + R_{12}^+ Q_{22}^- \\
A_{21} &= R_{21}^+ Q_{11}^- + R_{22}^+ Q_{21}^- \\
A_{22} &= R_{21}^+ Q_{12}^- + R_{22}^+ Q_{22}^-
\end{aligned}
$$

The matrix $\mathbf{A}$ now relates the *absolute* angular velocities of the two legs. That is, the matrix $\mathbf{A}$ defines the following two relationships:

$$
A_{11} \left( \dot{\theta}_s^- \dot{\theta}_{ns}^- \right) + A_{12} \left( \dot{\theta}_s^- \right) = \dot{\theta}_s^+ + \dot{\theta}_{ns}^+
$$

$$
A_{21} \left( \dot{\theta}_s^- \dot{\theta}_{ns}^- \right) + A_{22} \left( \dot{\theta}_s^- \right) = \dot{\theta}_s^+
$$

We can simply rearrange these relationships to find a direct transformation, given $\dot{\theta}_2 = \dot{\theta}_{ns} - \dot{\theta}_s$ is actually a *relative* velocity. Note that these equations have also accounted for the fact that the two legs "swap roles" during impact (with stance becoming swing leg, and vice versa). We will call this final transformation matrix $\mathbf{B}$:

$$
\begin{aligned}
B_{11} &= A_{21} + A_{22} \\
B_{12} &= A_{21} \\
B_{21} &= A_{11} + A_{12} - A_{21} - A_{22} \\
B_{22} &= A_{11} - A_{21}
\end{aligned}
$$

These final relationships relating pre- and post-collision velocities are then:

$$
\dot{\theta}_1^+ = B_{11} \dot{\theta}_1^- + B_{12} \dot{\theta}_2^- \tag{25}
$$

$$
\dot{\theta}_2^+ = B_{21} \dot{\theta}_1^- + B_{22} \dot{\theta}_2^- \tag{26}
$$

*PD controller*

During a step, the interleg angle is regulated with respect to a desired set-point, which is selected as a high-level control action, once per step.

$$\alpha = \theta_{sw} - \theta_{st} \tag{27}$$

$$\tau = K_p(\alpha_{des} - \alpha) + K_d(0 - \dot{\alpha}), \qquad \theta_{st} < 0 \tag{28}$$

where $K_p = 100$ and $K_d = 10$. These two equations appear as Equations 11 and 12.

*Pre-collision Impulse*

Our control was much more successful if we used a pre-collision impulse in addition to a PD-controlled torque at the hip, as described in Section V-A (p. 9). We assume the impulse occurs instantaneously and happens exactly before the step-to-step collision takes place.

If the imparted impulse is very large, the entire compass gait walker will become airborne. Calculation of a new state in this "ballistic" case simply involves solving two equations for conservation of angular momentum and another two to conserve linear momentum. For the magnitudes of impulse we use in simulation, there is typically not enough momentum added for the upcoming stance leg to lose contact with the ground. In this "non-ballistic" case, we simply assume that momentum directed toward the ground is perfectly absorbed in an instantaneous collision.

For the case where the upcoming stance foot does not become ballistic, we simply add the effect of the impulse as an additional term in the equations of conservation of angular momentum for the pre-to-post collision dynamics. Specifically, we rewrite Equation 23 to include a term which scales the impulse to give its contribution to the angular momentum about the pre-collision swing toe. Defining our impulse as having magnitude $p$, directed axially from the toe to the hip of the pre-collision stance leg, Equation 23 then becomes:

$$Q_{11}^-\dot{\theta}_{ns}^- + Q_{12}^-\dot{\theta}_s^- - 2p\,(\sin\alpha)\,(\cos\alpha) = Q_{11}^+\dot{\theta}_{ns}^- + Q_{12}^+\dot{\theta}_s^- \tag{29}$$

Propagating these terms and solving algebraically, Equations 25 and 26 are now replaced by the following:

$$\dot{\theta}_1^+ = B_{11}\dot{\theta}_1^- + B_{12}\dot{\theta}_2^- - R_{21}\,(2\cos\alpha\sin\alpha)\,p \tag{30}$$
$$\dot{\theta}_2^+ = B_{21}\dot{\theta}_1^- + B_{22}\dot{\theta}_2^- - (R_{11}-R_{21})\,(2\cos\alpha\sin\alpha)\,p \tag{31}$$

### REFERENCES

Au, S. K. (2004, Mar). Probabilistic failure analysis by importance sampling markov chain simulation. *Journal of Engineering Mechanics*, 303–311.

Boone, G. (1997). Minimum-time control of the acrobot. In *IEEE International Conference on Robotics and Automation (ICRA)* (Vol. 4, p. 3281-3287).

Bovier, A. (2004). Metastability and ageing in stochastic dynamics. In A. Maas, S. Martinez, & J. S. Martin (Eds.), *Dynamics and Randomness II* (pp. 17–81). Kluwer.

Bovier, A., Eckhoff, M., Gayrard, V., & Klein, M. (2000, Nov). Metastability and small eigenvalues in markov chains. *J. Phys. A: Math. Gen.*, *33*(46), L447–L451.

Boyan, J. A. (2002). Technical update: Least-squares temporal difference learning. *Machine Learning*, *49*, 233246.

Boyd, S., Diaconis, P., & Xiao, L. (2004, Dec). Fastest mixing markov chain on a graph. *SIAM Review*, *46*(4), 667-689.

Brown, S. C., & Passino, K. M. (1997). Intelligent control for an acrobot. *Journal of Intelligent and Robotic Systems*, *18*, 209–248.

Byl, K. (2008). *Metastable legged-robot locomotion*. Unpublished doctoral dissertation, MIT.

Byl, K., & Tedrake, R. (2006, May). Stability of passive dynamic walking on uneven terrain. In A. Kuo (Ed.), *Proceedings of Dynamic Walking II*.

Byl, K., & Tedrake, R. (2008a). Approximate optimal control of the compass gait on rough terrain. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*.

Byl, K., & Tedrake, R. (2008b). Control of the compass gait on rough terrain. In M. Wisse (Ed.), *Proceedings of Dynamic Walking IV*.

Byl, K., & Tedrake, R. (2008c). Metastable walking on stochastically rough terrain. In *Proceedings of Robotics: Science and Systems IV*.

Coleman, M. J. (1998). A stability-study of a three-dimensional passive-dynamic model of human gait (Doctoral dissertation, Cornell University). *Dissertations & Theses: A&I database. (Publication No. AAT 9831144)*.

Coleman, M. J., Chatterjee, A., & Ruina, A. (1997). Motions of a rimless spoked wheel: a simple 3D system with impacts. *Dynamics and Stability of Systems*, *12*(3), 139-160.

Collins, S. H., Ruina, A., Tedrake, R., & Wisse, M. (2005, February 18). Efficient bipedal robots based on passive-dynamic walkers. *Science*, *307*, 1082-1085.

Garcia, M., Chatterjee, A., Ruina, A., & Coleman, M. (1998, Apr). The simplest walking model: Stability, complexity, and scaling. *Journal of Biomechanical Engineering – Transactions of the ASME*, *120*(2), 281–288.

Gardiner, C. (2004). *Handbook of stochastic methods for physics, chemistry and the natural sciences* (Third ed.). Springer-Verlag.

Gaveau, B., & Schulman, L. S. (1998, Mar). Theory of nonequilibrium first-order phase transitions for stochastic dynamics. *Journal of Mathematical Physics*, *39*(3), 1517–1533.

Goswami, A., Espiau, B., & Keramane, A. (1996). Limit cycles and their stability in a passive bipedal gait. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)* (pp. 246–251).

Goswami, A., Thuilot, B., & Espiau, B. (1996, October). *Compass-like biped robot part I : Stability and bifurcation of passive gaits* (Tech. Rep. No. RR-2996). INRIA.

Hanggi, P., Talkner, P., & Borkovec, M. (1990, Apr). Reaction-rate theory: fifty years after kramers. *Reviews of Modern Physics*, *62*(2), 251–342.

Hodgins, J., & Raibert, M. (1991, June). Adjusting step

length for rough terrain locomotion. *IEEE Transactions on Robotics and Automation*, 7(3), 289-298.

Horenko, I., Dittmer, E., Fischer, A., & Schutte, C. (2006). Automated model reduction for complex systems exhibiting metastability. *Multiscale Model. Simul.*, 5(3), 802–827.

Huisinga, W., Meyn, S., & Schutte, C. (2004). Phase transitions and metastability in Markovian and molecular systems. *Annals of Applied Probability*, 14(1), 419-458.

Iida, F., & Tedrake, R. (2009). Minimalistic control of a compass gait robot in rough terrain. In *Proceedings of the ieee/ras international conference on robotics and automation (icra)*.

Jain, S., & Jain, R. K. (1994, Nov). Problems of estimating meantime system failure. *Microelectronics and reliability*, 34(11), 1755–1760.

Kampen, N. V. (2007). *Stochastic processes in physics and chemistry* (Third ed.). Elsevier.

Koditschek, D. E., & Buehler, M. (1991, Dec). Analysis of a simplified hopping robot. *International Journal of Robotics Research*, 10(6), 587-605.

Kuo, A. D., Donelan, J. M., & Ruina, A. (2005). Energetic consequences of walking like an inverted pendulum: Step-to-step transitions. *Exerc. Sport Sci. Rev.*, 33(2), 88–97.

Larralde, H., & Leyvraz, F. (2005, Apr). Metastability for markov processes with detailed balance. *Phys. Rev. Lett.*, 94(16).

Manchester, I. R., Mettin, U., Iida, F., & Tedrake, R. (2009). Stable dynamic walking over rough terrain: Theory and experiment. In *Under review.*

McGeer, T. (1990, April). Passive dynamic walking. *International Journal of Robotics Research*, 9(2), 62-82.

Muller, R., Talkner, P., & Reimann, P. (1997, Jun). Rates and mean first passage times. *Physica A*, 247(247), 338–356.

Munos, R., & Moore, A. (1998). Barycentric interpolators for continuous space and time reinforcement learning. In M. S. Kearns, S. A. Solla, & D. A. Cohn (Eds.), *Advances in neural information processing systems* (Vol. 11, pp. 1024–1030). MIT Press.

Munos, R., & Moore, A. (2002, November/December). Variable resolution discretization in optimal control. *Machine Learning*, 49(2/3), 291-323.

Spong, M. W. (1994). Swing up control of the acrobot. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (p. 2356-2361).

Spong, M. W. (1997). Underactuated mechanical systems. In B. Siciliano & K. P. Valavanis (Eds.), *Control problems in robotics and automation.* Springer-Verlag.

Spong, M. W., & Bhatia, G. (2003). Further results on control of the compass gait biped. In *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS)* (p. 1933-1938).

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction.* MIT Press.

Talkner, P., Hanggi, P., Freidkin, E., & Trautmann, D. (1987). Discrete dynamics and metastability: Mean first passage times and escape rates. *J. of Stat. Phys.*, 48(1/2), 231–254.

Tedrake, R. L. (2004). Applied optimal control for dynamically stable legged locomotion (Doctoral dissertation, Massachusetts Institute of Technology). *Dissertations & Theses: A&I database. (Publication No. AAT 0807532).*

Vijayakumar, S., D'Souza, A., & Schaal, S. (2005). Incremental online learning in high dimensions. *Neural Computation*, 17, 2602-2634.

Weber, M., Kube, S., Walter, L., & Deuflhard, P. (2006, Nov). *Stable computation of probability densities for metastable dynamical systems* (Tech. Rep. No. ZIB-Report 06-39). Konrad-Zuse-Zentrum fur Informationstechnik Berlin.