Optimize swing-up trajectory for inverted double-pendulum on cart with SQP

Mingrui Li
Univ. of Michigan
Computer Science and Engineering
Ann Arbor, United States
cldflpr@umich.edu

Ruikai Yang
Univ. of Michigan
Computer Science and Engineering
Ann Arbor, United States
ruikai@umich.edu

***Abstract*—This paper investigates the application of Sequential Quadratic Programming (SQP) to optimize the swing-up trajectory of an inverted double-pendulum on a cart. The objective is to determine control inputs that drive the system from a downward equilibrium to an upright position while adhering to dynamic constraints, control effort limits, and state boundaries. A Lagrangian formulation is employed to derive the system's nonlinear dynamics, which are linearized and iteratively solved using the SQP algorithm. The approach is benchmarked against the Model Predictive Path Integral (MPPI) control method under varying experimental conditions, including different mass distributions (heavy vs. light cart) and initial states (downward vs. near-upright). Key results demonstrate that SQP achieves superior average pole heights and higher immediate success rates in reaching target positions. However, SQP exhibits challenges in long-term stabilization.**

***Index Terms*—Trajectory Optimization, SQP, Nonlinear Dynamics**

I. Introduction

Trajectory optimization is an important task in the fields of robotics. The common objective is to find a sequence of control inputs as the solution to an optimization problem. The control sequence is expected to lead the robot into a trajectory that avoids obstacles, minimizes costs, reaches the targets, and obeys other constraints.This problem becomes more serious when it comes to dynamic systems, as the robot's movement will be controlled by complicated non-linear dynamics. Under such situations, conventional techniques might not be efficient enough in delivering solutions that maintain the constraints of the system.

To tackle these challenges, it's possible to utilize linearization principles to apply a trajectory optimization algorithm that accommodates the dynamic system complexity. This report aims to investigate the utilization of a linearized system to create an optimization tactic to successfully perform the swing-up maneuver for the inverted double-pendulum on the cart. The objective is to determine a control input set that successfully moves the cart and swings the pendulum from the bottom position to the top position subject to constraints on control efforts, system dynamics, and state variables.

To achieve this, we use a Sequential Quadratic Programming (SQP) algorithm, a robust tool for the solution of general nonlinear optimization problems. SQP suits the task of trajectory optimization very well, as it gradually updates the control inputs by using linear approximations of the system's equations, in order to converge to the optimal solution of the trajectory. Below, the principles of SQP will be described, its connection to the problem at hand will be presented, and we will also demonstrate how the SQP algorithm will be applied to optimize the swing-up task of the inverted double-pendulum system.

This paper is funded by our arduous work into this project.

II. Implementation

A. SQP

Sequential Quadratic Programming (SQP) is an iterative optimization algorithm designed to solve non-linear constrained optimization problems as shown in Algorithm 1.

Algorithm 1 Sequential Quadratic Programming (SQP)

Require: Initial guess x_0, λ_0, μ_0; convergence tolerance ϵ
Ensure: Optimal solution x^*

1: Set $k = 0$
2: **while** $\|\nabla_x \mathcal{L}(x_k, \lambda_k, \mu_k)\| > \epsilon$ **do**
3: Build QP approximation at current iterate:

$$\begin{aligned} \min_d \ & \nabla f(x_k)^\top d + \frac{1}{2} d^\top H_k d \\ \text{s.t. } & h(x_k) + \nabla h(x_k)^\top d = 0 \\ & g(x_k) + \nabla g(x_k)^\top d \leq 0 \end{aligned}$$

4: Solve QP subproblem for direction d_k and multipliers $(\lambda_{k+1}, \mu_{k+1})$
5: Compute step size α_k via line search
6: Update iterate: $x_{k+1} = x_k + \alpha_k d_k$
7: Update Hessian approximation H_{k+1} (e.g., BFGS update)
8: $k \leftarrow k + 1$
9: **end while**=0

It is widely recognized for its efficiency and convergence properties, particularly for problems with smooth objective and constraint functions. The core idea of SQP is to approximate the original nonlinear problem as a sequence of quadratic

programming (QP) subproblems, each of which is solved to update the iterations toward the optimal solution. [1]–[3]

B. Lagrangian Method

In this study, the Lagrangian method is utilized to describe the motion of the system focusing on its trajectories. The Lagrangian approach is based on the principle of least action, which states that the path taken by a system between two states is the one that minimizes the action S. The action is defined as the integral of the Lagrangian $\mathcal{L}$ over time:

$$S = \int_{t_1}^{t_2} \mathcal{L}(q, \dot{q}, t)\, dt \tag{1}$$

where q represents the generalized coordinates of the system, $\dot{q}$ denotes the time derivatives of these coordinates, and t is time. The Lagrangian $\mathcal{L}$ is typically expressed as the difference between the kinetic energy T and the potential energy V of the system:

$$\mathcal{L} = T - V \tag{2}$$

By applying the Lagranian equations, we can derive the equations of motion for the system:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) - \frac{\partial \mathcal{L}}{\partial q_i} = 0 \tag{3}$$

Here, q_i represents the i-th generalized coordinate. This formulation allows us to account for both conservative and non-conservative forces acting on the system, making it a versatile tool for analyzing complex dynamics.

In our specific application, the Lagrangian method enables us to track the motion by making it easier to calculate the system by defining their positions, velocities, and accelerations in terms of the generalized coordinates. This approach provides a comprehensive understanding of the system's behavior over time. [4]

For the analytical dynamic performance of the environment, we adopt Lagrangian method, and set the generalized coordinate to be $< x, \theta, \varphi >$, and we have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{x}}\right) - \frac{\partial T}{\partial x} + \frac{\partial V}{\partial x} = F \tag{4}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{\theta}}\right) - \frac{\partial T}{\partial \theta} + \frac{\partial V}{\partial \theta} = 0 \tag{5}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{\phi}}\right) - \frac{\partial T}{\partial \phi} + \frac{\partial V}{\partial \phi} = 0 \tag{6}$$

By energy formula we can calculate the kinemic energy of all conponents to be:

$$T_c = \frac{1}{2} m_c \dot{x}^2 \tag{7}$$

$$T_1 = \frac{1}{2} m_p \left(\dot{x}^2 + l\dot{x}\dot{\theta}\cos\theta + \frac{l^2}{3}\dot{\theta}^2 \right) \tag{8}$$

$$T_2 = \frac{1}{2} m_p \Bigg[\dot{x}^2 + l^2\dot{\theta}^2 + \frac{l^2}{3}(\dot{\theta}+\dot{\phi})^2 + 2l\dot{x}\dot{\theta}\cos\theta + l\dot{x}(\dot{\theta}+\dot{\phi})\cos(\theta+\phi) + l^2\dot{\theta}(\dot{\theta}+\dot{\phi})\cos\phi \Bigg] \tag{9}$$

And set the COM of cart to zero height, the potential energy of the system is:

$$V = \frac{m_p g l}{2}\left(3\cos\theta + \cos(\theta+\phi)\right) \tag{10}$$

The lagrangian function

$$L = T_c + T_1 + T_2 - V \tag{11}$$

After some simple calculation, the lagrangian dynamic function becomes:

$$\begin{aligned} &(m_c + 2m_p)\ddot{x} \\ &+(\frac{3}{2}\cos\theta + \frac{1}{2}\cos(\theta+\phi))m_p l\ddot{\theta} \\ &+\frac{1}{2}\cos(\theta+\phi)m_p l\ddot{\phi} \\ &=F + \frac{3}{2}m_p l\dot{\theta}^2\sin\theta \\ &+\frac{1}{2}m_p l(\dot{\theta}+\dot{\phi})^2\sin(\theta+\phi) \end{aligned} \tag{12}$$

$$\begin{aligned} &(\frac{3}{2}\cos\theta + \frac{1}{2}\cos(\theta+\phi))m_p l\ddot{x} \\ &+(\frac{5}{3} + \frac{1}{2}\cos\phi)m_p l^2\ddot{\theta} \\ &+(\frac{1}{2}\cos\phi + \frac{1}{3})m_p l^2\ddot{\phi} \\ &=\frac{1}{2}m_p l^2(\dot{\theta}+\dot{\phi})\dot{\phi}\sin\phi \\ &+\frac{3}{2}m_p g l\sin\theta + \frac{1}{2}m_p g l\sin(\theta+\phi) \end{aligned} \tag{13}$$

$$\begin{aligned} &\frac{1}{2}\cos(\theta+\phi)m_p l\ddot{x} \\ &+(\frac{1}{2}\cos\phi + \frac{1}{3})m_p l^2\ddot{\theta} \\ &+\frac{1}{3}m_p l^2\ddot{\phi} \\ &=\frac{1}{2}m_p g l\sin(\theta+\phi) - \frac{1}{2}m_p l^2\dot{\theta}^2\sin\phi \end{aligned} \tag{14}$$

This can be regarded as a equation:

$$M(\theta,\phi)[\ddot{x},\ddot{\theta},\ddot{\phi}]^T = \vec{Q}(\theta,\phi,\dot{\theta},\dot{\phi}) \tag{15}$$

Then we can solve the 2nd order dynamics as follows:

$$[\ddot{x}, \ddot{\theta}, \ddot{\phi}]^T = M^{-1}\vec{Q} \tag{16}$$

The corresponding code is shown in get_ddot.py

C. Model structure

For this project, we defined a physical model in the Unified Robot Description Format (URDF), which is commonly used to describe the structure and motion characteristics of robotic or mechanical systems. Below are the main parts for the model. For parameters ,refer to Table I.

1) SlideBar: It is a rectangular prism which serves as the base of the entire model, to which other components are connected via joints. It is fixed.

2) Cart: It is a rectangular prism connected to the slide bar through a prismatic joint, which allows it to slide along the x-axis.

3) Pole 1: It is a rectangular prism connected to the cart through a continuous joint, which allows it to rotate freely around the y-axis.

4) Pole 2: It is a rectangular prism connected to Pole 1 at the end through a continuous joint, which allows it to rotate freely around the y-axis.

TABLE I
MODEL CONFIGURATION

Prism	Parameter	Value	Unit
SlideBar	Dimension	30×0.05×0.05	m*m*m
	Color	Green	
Cart	Dimension	0.5×0.5×0.2	m*m*m
	Position limit	[-15,15]	m
	Maximum Velocity	5	m/s
	Color	Blue	
Pole 1	Dimension	0.05×0.05×1	m*m*m
	Position limit	$[-\pi, \pi]$	rad
	Maximum Velocity	None	
	Color	Grey	
Pole 2	Dimension	0.05×0.05×1	m*m*m
	Position limit	$[-\pi, \pi]$	rad
	Maximum Velocity	None	
	Color	Grey	

D. Model Parameters

To more comprehensively compare the impact of differences in control algorithms on experimental performance, we designed the following sets of control experiments:

1) Initial states: Comparison between initial states that are nearly upright and initial states that are numerically down.

TABLE II
SYSTEM DYNAMICS PARAMETERS CONFIGURATION

Param.	Cart Mass	Pole Mass	Ini. Pos.
Unit	kg	kg	
1	1	0.1	Downright
2	1	0.1	Upright
3	0.1	0.1	Downright
4	0.1	0.1	Downright height

2) Mass: Comparison of the relationship between the mass of the cart and the mass of the pole. When the cart has a larger mass, the system is less sensitive to external forces. In contrast, when the cart and pole have similar masses, the system is highly sensitive to external forces and is prone to instability. Attached is the configuration of the experiment parameters for this system, as shown in II.

E. Environment Configuration

The simulation environment employs the PyBullet. PyBullet is a powerful physics engine and robotics simulator that provides efficient and realistic physics simulations, making it ideal for testing and developing robotic algorithms. It also supports various robot models and environments, allowing for flexible and customizable simulations.

The control is employed under a refresh rate of 240Hz. It can be observed that under the SQP algorithm, the system performs well, as shown in Figure 6, where the movement of system is shot every 0.125s, with the link approaching the target highest point multiple times. To quantitatively analyze the performance of the SQP algorithm, we compare it with the traditional MPPI control algorithm.

F. Controller Parameters

The controller parameters are listed in the table III.

TABLE III
CONTROLLER PARAMETERS CONFIGURATION

Controller	Param.	Value
MPPI	lambda	0.001
	noise sigma	Diag(10, 10, 10, 10, 10, 10)
	Q	Diag(1, 5, 5, 0.1, 0.1, 0.1)
SQP (Exp. 1-3)	horizon	50
	max iter	5
	Q	Diag(1, 6, 8, 5, 10, 10)
	R	0.1
SQP (Exp. 4)	horizon	50
	max iter	5
	Q	Diag(1, 6, 8, 0.5, 5, 10)
	R	21.5

III. RESULTS AND DISCUSSION

In this part we will compare the difference of SQP and a more traditional control algorithm, MPPI. We will start from the most traditional case but continue to dive in more specific cases for this behevior.

A. Heavy Cart, Downward Initial Position

For the traditional mass distribution where the cart significantly outweighs the poles starting from a downward initial position, the end heights of Pole 1 and Pole 2 over time are shown in Figure 1. Key differences between the SQP and MPPI control strategies are analyzed as follows, and for simplicity the successful atttempts refers to trails arrive at 1.5m, compared with the target 2m.

1) Mean Height: The SQP generates a significantly better performance for the most crucial criteria, i.e. average height of the two systems, as shown in the table IV

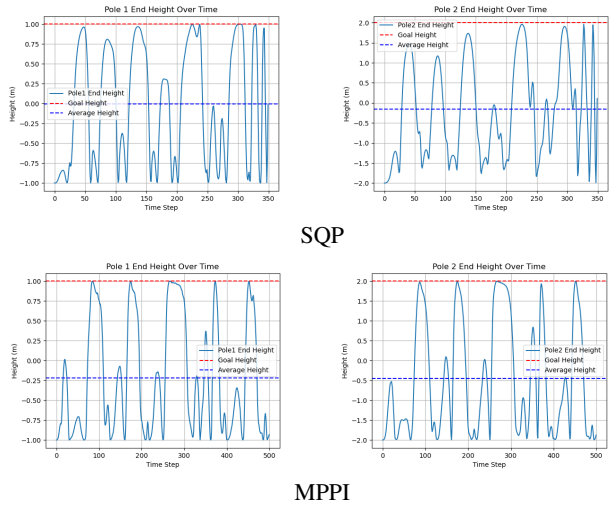

Fig. 1. Average endcap height for experiment 1

2) Attempts: The SQP algorithm demonstrates a higher immediate success rate in reaching the target height (Goal Height in Figure 1). However, the MPPI controller requires more attempts to achieve comparable results. This discrepancy arises because MPPI frequently terminates trials prematurely when suboptimal trajectories are detected. In contrast, SQP persists longer in individual trials, leading to more successful completions despite overshooting.

3) Coordination: MPPI exhibits superior coordination between the two poles in successful trials. As shown in the height trajectories (Figure 1), MPPI consistently aligns Pole 2's end height with the target of 2 m (Goal Height) in stable trials, whereas SQP struggles to synchronize both poles. For instance, SQP-controlled Pole 2 sometimes stabilizes at approximately -1.0 m (see 150–250 time steps), significantly below the target.

TABLE IV
MEAN HEIGHT COMPARISON FOR HEAVY CART, DOWNWARD INI. POS.

Algorithm	Pole 1 Avg. H/m	Pole 2 Avg. H/m
(Target)	1	2
SQP	0.0	-0.1
MPPI	-0.25	-0.5
(Natural)	-1	-2

4) Decay Pattern and Stability: Distinct decay patterns emerge at stabilization points. MPPI-controlled poles show *critically arrived* behavior, where velocity decays exponentially to zero at the target height (evident in Pole 1's near-flat trajectory after 200 time steps). This indicates effective energy dissipation. Conversely, SQP exhibits *underdamped overshoot*, manifesting as sinusoidal passing the target. These oscillations suggest residual kinetic energy and challenges in precise stabilization—a limitation that becomes more sensitive in upright initialization scenarios, as discussed later.

5) Long-Duration Oscillation: As depicted in Figure 1, the SQP-controlled system exhibits a notable anomaly during prolonged operation. Beyond approximately 300 time steps, the system displays a sudden increase in oscillation frequency, despite prior stabilization trends. The underlying mechanism driving this behavior remains unclear and warrants further investigation.

It is suspected that it results from the following mechanism:

After the pole passes its highest point, the SQP controller exhibits a counterintuitive behavior: instead of applying corrective actions (e.g., reversing cart motion) to stabilize the pole and return it to the peak position, it actively increases the pole's angular velocity to expedite a full rotation and reach the highest point again. This strategy inadvertently creates a positive feedback loop in control.

At each cycle, the controller prioritizes maximizing angular velocity to shorten the time between consecutive peaks. By not resisting or reducing the growing angular momentum, the system accumulates kinetic energy, further amplifying the angular velocity in subsequent cycles.

This behavior deviates from conventional stabilization strategies.

B. Heavy cart Upright

To illustrate the stability at the top point better, we provide the upright set to see whether the cart can stay stable starting from almost target point. Similarly, the end heights of Pole 1 and Pole 2 over time are shown in Figure 2.

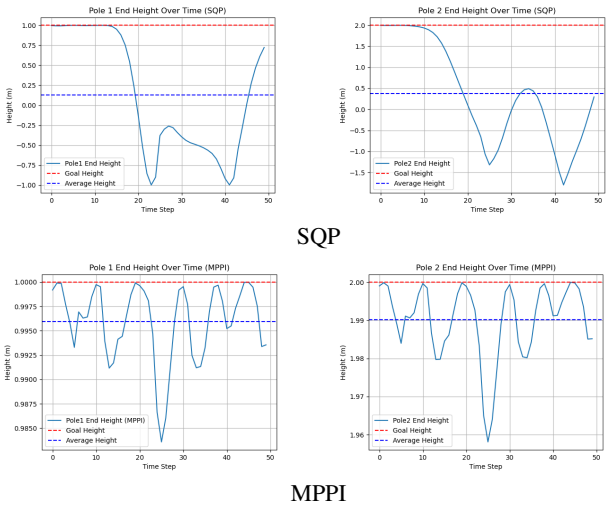

Fig. 2. Average endcap height for experiment 2

In this figure, it is found that the MPPI remains stable in the experiment. In contrast, the SQP algorithm fail to stay stable around the target point in the long term. In contrast, the MPPI controller stays around the target point very well.

To analyze this phenomenon, we perform the asymptotic analysis on the analytical dynamic equation:

If $\theta, \phi \to 0$, equation 1213 14becomes:

$$(m_c + 2m_p)\ddot{x} + 2m_p l\ddot{\theta} + \frac{1}{2}m_p l\ddot{\phi} \to F \tag{17}$$

$$2m_p l\ddot{x} + \frac{13}{6}m_p l^2\ddot{\theta} + \frac{5}{6}m_p l^2\ddot{\phi} \to 0 \tag{18}$$

$$\frac{1}{2}m_p l\ddot{x} + \frac{5}{6}m_p l^2\ddot{\theta} + \frac{1}{3}m_p l^2\ddot{\phi} \to 0 \tag{19}$$

In the asymptotic analysis of the analytical results, we observed that for systems starting with small perturbations, if the force provided by the SQP controller is small, the resulting acceleration in generalized coordinates remains correspondingly small. In other words, the excessive force exerted by the SQP controller is the root cause of its failure to stabilize the system at the apex. This may results from cost of x. If it starts from an deviated position, the force will drag it from x and break its stability.

C. Light Cart, Downward Initial Position

To figure out whether this results from too large R value, we change the mass of the cart to 0.1, and it turned out that the excess angular velocity still exists, which also results from too small regularization term R. The result is shown as below in figure 3.

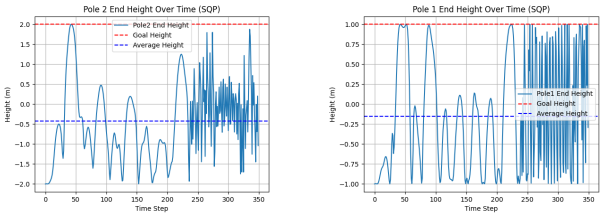

Fig. 3. Average endcap height for experiment 3

D. Trade-off between stability and max-reaching capacity

The experiments reveal a critical trade-off between stabilizing control and achieving maximum height. In the SQP framework, the weighting parameter R (penalizing control effort) directly influences this balance:

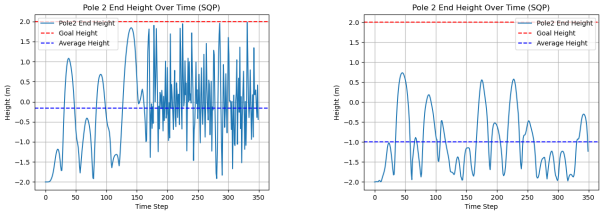

Fig. 4. Typical underdamped(left) and overdamped(case) for end height

1) Overdamped Case: When R is excessively large, the controller prioritizes minimizing force inputs, resulting in overly conservative actuation. As shown in Figure 4, the cart applies minimal forces, causing the pole to approach the target height sluggishly. The trajectory remains monotonically damped but fails to reach the Goal Height (2.0 m) in all trails.

2) Underdamped Case: Conversely, when R is too small, the controller aggressively exploits larger forces, leading to persistent oscillations. For instance, Pole 2's height (Figure 4, 0–350 time steps) exhibits rapid ±1.5 m swings after 200 time steps, with angular velocity escalating unchecked. This "runaway" instability reflects insufficient damping to dissipate kinetic energy.

E. SQP parameters adjusted

Based on the troubleshoot above, we adjusted the parameters R.We expect the R to be sensitive, so we perform a golden-section search on a logarithmic scale to rapidly converge to the desired R.

Initially it is determined that when R=10, it is underdamped and when R=100, it is overdamped. So by the search, we carried out following attempts, as shown in table V.

TABLE V
R ADJUSTMENT

Experiment	R	Result
1	10	Underdamped
2	100	Overdamped
3	50	Overdamped
4	30	Overdamped
5	19.5	Underdamped
6	25	Overdamped
7	22.7	Overdamped
8	21.5	Critically damped

Finally, we arrive at a perfect parameter configuration of the controller, where parameters is listed in table III, listed as experiment 4.

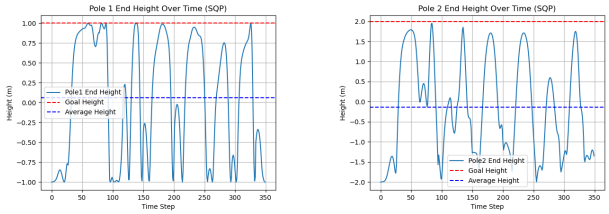

Fig. 5. Average endcap height for experiment 4

REFERENCES

[1] B. Goodman, "Sequential quadratic programming." https://optimization.cbe.cornell.edu/index.php?title=Sequential_quadratic_programming, 2016. Accessed: 2025-04-28.

[2] P. E. Gill and E. Wong, "Sequential quadratic programming methods," in *Mixed integer nonlinear programming*, pp. 147–224, Springer, 2011.

[3] I. M. Bomze, V. F. Demyanov, R. Fletcher, T. Terlaky, and R. Fletcher, "The sequential quadratic programming method," *Nonlinear Optimization: Lectures given at the CIME Summer School held in Cetraro, Italy, July 1-7, 2007*, pp. 165–214, 2010.

[4] P. LibreTexts, "8.5: The lagrangian formulation of classical physics." https://phys.libretexts.org/, 2024. Accessed: 2025-04-27.

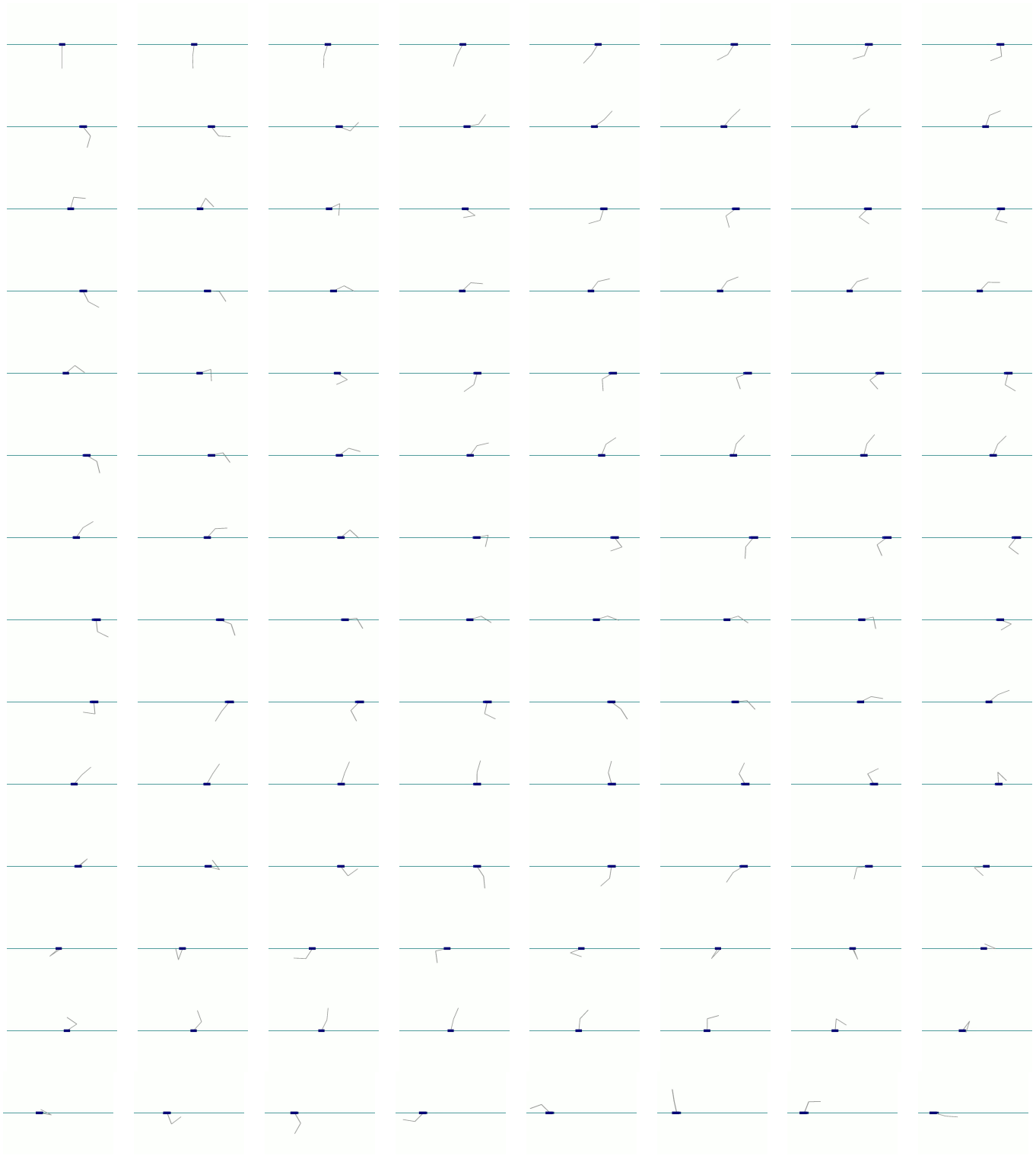

Fig. 6. 112 frames showing the evolution of the system