1 Hamiltonian of the System without Adjoint

The Hamiltonian of the system without adjoint variables is given by:

$$H = \frac{1}{2}m\boldsymbol{v}^2 - \frac{GM_1 m}{|\boldsymbol{r} - \boldsymbol{r}_1|} - \frac{GM_2 m}{|\boldsymbol{r} - \boldsymbol{r}_2|} \tag{1}$$

where $\boldsymbol{r}$ is the position of the planet, $\boldsymbol{v}$ is its velocity, G is the gravitational constant, M_1 and M_2 are the masses of the two suns, and $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are their positions.

The state equations of the system are:

$$\frac{d\boldsymbol{r}}{dt} = \boldsymbol{v} \tag{2}$$

$$\frac{d\boldsymbol{v}}{dt} = -\frac{GM_1(\boldsymbol{r} - \boldsymbol{r}_1)}{|\boldsymbol{r} - \boldsymbol{r}_1|^3} - \frac{GM_2(\boldsymbol{r} - \boldsymbol{r}_2)}{|\boldsymbol{r} - \boldsymbol{r}_2|^3} \tag{3}$$

1.1 Finite Differencing

To compute the gradients using finite differencing, we approximate the derivatives as:

$$\frac{\partial H}{\partial \boldsymbol{r}} \approx \frac{H(\boldsymbol{r} + \Delta\boldsymbol{r}, \boldsymbol{v}) - H(\boldsymbol{r}, \boldsymbol{v})}{\Delta\boldsymbol{r}} \tag{4}$$

$$\frac{\partial H}{\partial \boldsymbol{v}} \approx \frac{H(\boldsymbol{r}, \boldsymbol{v} + \Delta\boldsymbol{v}) - H(\boldsymbol{r}, \boldsymbol{v})}{\Delta\boldsymbol{v}} \tag{5}$$

1.2 Gradient Descent for Optimization

Using gradient descent, we update the position and velocity at every n-th timestep:

$$\boldsymbol{r}_{n+1} = \boldsymbol{r}_n - \eta\nabla a(\boldsymbol{r}_n) \tag{6}$$

where η is the learning rate and $\nabla a(\boldsymbol{r})$ is the gradient of the objective function with respect to $\boldsymbol{r}$.

2 Hamiltonian of the System with Adjoint

The Hamiltonian of the system with adjoint variables $\boldsymbol{\lambda_r}$ and $\boldsymbol{\lambda_v}$ is given by:

$$H = a(\boldsymbol{r}) + \boldsymbol{\lambda_r} \cdot \boldsymbol{v} + \boldsymbol{\lambda_v} \cdot \left(-\frac{GM_1(\boldsymbol{r} - \boldsymbol{r}_1)}{|\boldsymbol{r} - \boldsymbol{r}_1|^3} - \frac{GM_2(\boldsymbol{r} - \boldsymbol{r}_2)}{|\boldsymbol{r} - \boldsymbol{r}_2|^3} \right) \tag{7}$$

where $a(\boldsymbol{r}) = \|\boldsymbol{r} - \boldsymbol{r}_{\text{figure8}}(t)\|^2$.

The adjoint equations are:

$$\frac{d\boldsymbol{\lambda_r}}{dt} = -\left(\nabla a(\boldsymbol{r}) + \boldsymbol{\lambda_v} \cdot (\boldsymbol{H_1} + \boldsymbol{H_2})\right) \tag{8}$$

$$\frac{d\boldsymbol{\lambda_v}}{dt} = -\boldsymbol{\lambda_r} \tag{9}$$

where $\boldsymbol{H_1}$ and $\boldsymbol{H_2}$ are the Hessians of the gravitational potential from each sun.

2.1 Hessians of the Gravitational Potential

The Hessians $\boldsymbol{H_1}$ and $\boldsymbol{H_2}$ are given by:

$$\boldsymbol{H_1} = \begin{bmatrix} \frac{-GM_1(3(\boldsymbol{r}_x - \boldsymbol{r}_{1x})^2 - |\boldsymbol{r} - \boldsymbol{r}_1|^2)}{|\boldsymbol{r} - \boldsymbol{r}_1|^5} & 0 \\ 0 & \frac{-GM_1(3(\boldsymbol{r}_y - \boldsymbol{r}_{1y})^2 - |\boldsymbol{r} - \boldsymbol{r}_1|^2)}{|\boldsymbol{r} - \boldsymbol{r}_1|^5} \end{bmatrix} \tag{10}$$

$$\boldsymbol{H_2} = \begin{bmatrix} \frac{-GM_2(3(\boldsymbol{r}_x - \boldsymbol{r}_{2x})^2 - |\boldsymbol{r} - \boldsymbol{r}_2|^2)}{|\boldsymbol{r} - \boldsymbol{r}_2|^5} & 0 \\ 0 & \frac{-GM_2(3(\boldsymbol{r}_y - \boldsymbol{r}_{2y})^2 - |\boldsymbol{r} - \boldsymbol{r}_2|^2)}{|\boldsymbol{r} - \boldsymbol{r}_2|^5} \end{bmatrix} \tag{11}$$

2.2 Update Based on Adjoint Simulation

The update of the positional term based on the adjoint simulation is given by:

$$\frac{d\boldsymbol{\lambda_r}}{dt} = -\left(\nabla a(\boldsymbol{r}) + \boldsymbol{\lambda_v} \cdot (\boldsymbol{H_1} + \boldsymbol{H_2})\right) \tag{12}$$

To compute $\boldsymbol{\lambda_r}$ and $\boldsymbol{\lambda_v}$ at each timestep:

$$\boldsymbol{\lambda}_{\boldsymbol{r}}^{n+1} = \boldsymbol{\lambda}_{\boldsymbol{r}}^{n} - \left(\nabla a(\boldsymbol{r}^n) + \boldsymbol{\lambda}_{\boldsymbol{v}}^{n} \cdot (\boldsymbol{H_1} + \boldsymbol{H_2})\right) \Delta t \tag{13}$$

$$\boldsymbol{\lambda}_{\boldsymbol{v}}^{n+1} = \boldsymbol{\lambda}_{\boldsymbol{v}}^{n} - \boldsymbol{\lambda}_{\boldsymbol{r}}^{n} \Delta t \tag{14}$$

2.3 Adjoint Gradient Descent for Optimization

Using the adjoint variables, we compute the gradient and update the parameters:

$$\frac{\partial J}{\partial \boldsymbol{r}} = \boldsymbol{\lambda}_{\boldsymbol{r}} \tag{15}$$

The positions are updated at every n-th timestep as:

$$\boldsymbol{r}_{n+1} = \boldsymbol{r}_n - \eta \boldsymbol{\lambda}_{\boldsymbol{r}}^n \tag{16}$$