1 Derivations

1.1

We are investigating systems with control-affine deterministic dynamics with additive white Gaussian noise. The continuous dynamics are written as:

$$\dot{x} = f(x) + g(x)u + (\mu(x) + \sigma(x)w_t) \tag{1}$$

where w_t at time t is sampled from a standard Gaussian distribution ($\mu = 0$, $\sigma = 1$). Each random variable w_t is independent and identically distributed throughout all time t. w_t can be interpreted as a white Gaussian noise process over time. Note that our system dynamics are described by a stochastic differential equation, rather than the typical deterministic differential equation.

For the sake of derivations, we will be analyzing a discrete approximation of this continuous system and then take the limit as Δt goes to zero. For a more rigorous, strictly continuous treatment of stochastic differential equations with additive white noise seeThe mathematical treatment covered here is for intuition and understanding only.

The discrete approximation to our system is:

$$\begin{aligned}
x_{k+1} - x_k &= \int_{t=k\Delta t}^{(k+1)\Delta t} (f(x) + g(x)u + (\mu(x) + \sigma(x)w_t))dt && (2) \\
&= \int_{t=k\Delta t}^{(k+1)\Delta t} (f(x) + g(x)u + \mu(x))dt + \int_{t=k\Delta t}^{(k+1)\Delta t} \sigma(x)w_t dt && (3) \\
&\simeq (f(x_k) + g(x_k)u_k + \mu(x_k))\Delta t + \sigma(x_k) \int_{t=k\Delta t}^{(k+1)\Delta t} w_t dt && (4)
\end{aligned}$$

Where we've separated out the drift due to the additive noise and x_k denotes the state at time $t = k\Delta t$. We could have approximated the change due to the white noise in the same manner as we did the change due to the deterministic dynamics; namely, assume the noise is constant over the time interval Δt and approximate as $\sigma(x_k)w_{k\Delta t}\Delta t$.

However, a better approximation can be made that leverages random process theory. The integral $\int_{t=k\Delta t}^{(k+1)\Delta t} w_t dt$ is interpreted in the Ito sense. Taking the integral of i.i.d white noise over time is one method for generating the Wiener Process (which is responsible for Brownian motion, amongst other things). The value of the Wiener process W_t at time t is distributed according to another Gaussian with mean zero and variance equal to the length of time the integral is take over Δt. It is important to note that although the Wiener process is also Guassian distributed at each time point, it is not the same as the white noise it was generated from. The Wiener process is not i.i.d. That is, each value of $W_t = \int_{t_{start}}^{t} w_t dt$ is dependent on previous values of W_s (with $s < t$). This dependence makes the trajectory of W_t continuous, in contrast to the discontinuous, noisy plot of w_t.

We take the Ito integral from start time $k\Delta t$ to $(k+1)\Delta t$, and get a new random variable for each k interval that represents the drift in the state due to Brownian motion. Replacing the integral with this new random variable W_k:

$$x_{k+1} - x_k \simeq (f(x_k) + g(x_k)u_k + \mu(x_k))\Delta t + \sigma(x_k)W_k \tag{5}$$

Noting that this variable is Gaussian distributed with variance Δt, we can replace this variable with a new standard normal distribution scaled by the standard deviation:

$$x_{k+1} - x_k \simeq (f(x_k) + g(x_k)u_k + \mu(x_k))\Delta t + \sigma(x_k)\sqrt{\Delta t}b_k \tag{6}$$

$$\to x_{k+1} \simeq x_k + (f(x_k) + g(x_k)u_k + \mu(x_k))\Delta t + \sigma(x_k)\sqrt{\Delta t}b_k \tag{7}$$

$$= x_k + h(x_k, u_k)\Delta t + \sigma(x_k)\sqrt{\Delta t}b_k \tag{8}$$

where:

$$h(x_k, u_k) = f(x_k) + g(x_k)u_k + \mu(x_k)$$

1.2 Expected Terminal Reward

We want to calculate:

$$\mathbf{E}(r(x(T)))$$

where $r(x)$ is the terminal payoff function.

We will start our treatment by examining a previous framework for stochastic reachability presented in Abate et. al. Using Abate's work as a starting point is a convenient method for introducing established, foundational concepts. In the next section, we will expand outside the assumptions Abate made about payoff functions to leverage more powerful methods. For now, we will use his payoff function $r(x) = \mathbf{1}_\mathbf{A}(x)$ where A is the safe set, and $\mathbf{1}_\mathbf{A}(x)$ is the indicator function. An indicator function for A is 1 inside the set A and 0 outside.

Note that our system is a purely continuous state system (no discrete mode transitions) and is therefore not strictly a hybrid system. However, we can still use Abate's framework by considering our continuous system as a type of hybrid system with only one mode.

Abate's equation () describes the update equation for the value function for backwards reachability. Initialize $V_0(x) = \mathbf{1}_\mathbf{A}(x)$ and take

$$V_{k-1}(x) = \mathbf{1}_\mathbf{A}(x)\mathrm{E}_{z \sim P(x_k = z | x_{k-1} = x)}(V_k(z)) \tag{9}$$

The value function at discrete time $k-1$ is just the expected value at the next time step if $x \in A$ and is 0 if $x \in A^c$ since we've already failed to stay inside the safe set for all time. Here the probability of $x_k = z$ is governed by the discrete, stochastic dynamics described in Eq. 7. Plugging in these discrete dynamics and expecting over b_{k-1} instead yields:

$$
\begin{aligned}
V_{k-1}(x) &= \mathbf{1_A}(x)\mathrm{E}_{z\sim P(x_k=z|x_{k-1}=x)}(V_k(z)) && (10)\\
&= \mathbf{1_A}(x)\mathrm{E}_{b_k\sim N}(V_k(x+(f(x)+g(x)u_k+\mu(x))\Delta t+\sigma(x)b_k\sqrt{\Delta t}) && (11)\\
&= \mathbf{1_A}(x)\int_{-\infty}^{\infty} V_k(x+(f(x)+g(x)u_k+\mu(x))\Delta t+\sigma(x)b_k\sqrt{\Delta t})\frac{1}{\sigma\sqrt{2\pi\Delta t}}e^{-\frac{x^2}{2\sigma^2\Delta t}}db && (12)
\end{aligned}
$$

In his development, Abate proves that this returns the expectaction of never leaving the safe set. Actually using formula 12 to calculate the expected safety function, would be rather arduous. Formula 12 with a one-dimensional state space and noise vector corresponds to a convolution between the value function and the Gaussian PDF.

Abate also proves that the optimal control policy for maximizing the safety expectation is a time-pointwise maximization for each time step. Although his proof is more rigorous, the time-pointwise maximization can also be (more simply) derived using the Principle of Optimality.

1.3 Differential Form

To ameliorate the processing burden, we will use tricks learned from deriving deterministic reachability in class. Namely, we will cast the value function update into a partial differential equation, whereupon we can leverage previous advances in viscosity solution solvers to obtain the stochastically safe sets. Let H denote the Hessian of V_k.

$$
\begin{aligned}
V_{k-1}(x) &= \mathbf{1_A}(x)\mathrm{E}_{z\sim P(x_k=z|x_{k-1}=x)}(V_k(z))\\
&= \mathbf{1_A}(x)\mathrm{E}_{b_k\sim N}(V_k(x+h(x,u_k)\Delta t+\sigma(x)\sqrt{\Delta t}b_k)\\
&= \mathbf{1_A}(x)\mathrm{E}_{b\sim N}(V_k(x)+\nabla V_k^T(h(x,u_k)\Delta t+\sigma(x)\sqrt{\Delta t}b_k)\\
&+ (h(x,u_k)\Delta t+\sigma(x)\sqrt{\Delta t}b_k)^T H(h(x,u_k)\Delta t+\sigma(x)\sqrt{\Delta t}b_k))\\
&= \mathbf{1_A}(x)[V_k(x)+\nabla V_k^T h(x,u_k)\Delta t+\nabla V_k^T\sigma(x)\sqrt{\Delta t}\mathrm{E}_{b\sim N}(b_k)\\
&+ \mathrm{E}_{b\sim N}(h(x,u_k)\Delta t+\sigma(x)\sqrt{\Delta t}b_k)^T H(h(x,u_k)\Delta t+\sigma(x)\sqrt{\Delta t}b_k)]\\
&= \mathbf{1_A}(x)[V_k(x)+\nabla V_k^T h(x,u_k)\Delta t\\
&+ \Delta t^2 h^t Hh+\Delta t^{3/2}\mathrm{E}_{b\sim N}[h^T H\sigma b_k+\sigma^T Hhb_k]\\
&+ \sigma^T H\sigma\mathrm{E}_{b\sim N}[b_k^2]\Delta t]
\end{aligned}
$$

Multiplying by the indicator function $\mathbf{1_A}(x)$ on each of these terms, just says that we should only update points inside the safe set. Points outside the safe set, should never change from zero. This behavior can also be attained by ensuring that the updates are always non-positive by taking the max with zero. Making this substitution allows us to focus on the term previously multiplying the indicator function:

$$
\begin{aligned}
V_{k-1}(x) &= V_k(x)+\nabla V_k^T h(x,u_k)\Delta t && (13)\\
&+ \Delta t^2 h^t Hh+\Delta t^{3/2}\mathrm{E}_{b\sim N}[h^T H\sigma b_k+\sigma^T Hhb_k] && (14)\\
&+ \sigma^T H\sigma\cdot 1\Delta t && (15)
\end{aligned}
$$

Subtract the V_k from both sides and divide by Δt to obtain:

$$\frac{V_{k-1}(x) - V_k(x)}{\Delta t} = \nabla V_k^T h(x, u_k) \tag{16}$$

$$+ \Delta t h^t H h + \Delta t^{1/2} \mathrm{E}_{b \sim N}[h^T H \sigma b_k + \sigma^T H h b_k] \tag{17}$$

$$+ \sigma^T H \sigma \tag{18}$$

Taking the limit as $\Delta t \to 0$ produces the differential form and collapses the discrete approximation back into the true continuous dynamics with:

$$-\frac{\delta V(x)}{\delta t} = \nabla V_k^T h(x, u(t)) + \sigma^T H \sigma \tag{19}$$

$$= \nabla V^T (f(x) + g(x)u(t)) + \nabla V^T \mu(x) + \sigma^T H \sigma \tag{20}$$

This update has three components. The first term depends solely on the deterministic control-affine dynamics and is identical to the Hamiltonian for a deterministic system. The second term is the change in V due to where the noise will move on average. The third term is the most interesting. It is second-order with space like in the heat equation. It represents a diffusion of expected value over time due to random variations in the disturbance. Perhaps most intriguingly is the connection between this Heat-equation-esque term and our underlying disturbance. Our disturbance is a Brownian motion which is also used to predict how heated particles will move.

To maximize our expected value, we should choose $u(t)$ at each time step to maximize $\frac{\delta V(x)}{\delta t}$. Since the first (deterministic) term is the only term dependent on u and it is isolated from the stochastic terms, maximizing u is accomplished identically to how it would be done in the deterministic case.

Eq. 20 is of a form compatible with Ian Mitchell's level set toolbox for solving Hamilton Jacobi partial differential equations.