

# I Derivations

## I.1

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

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

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  $\Delta t$  goes to zero. For a more rigorous, strictly continuous treatment of stochastic differential equations with additive white noise see The mathematical treatment covered here is for intuition and understanding only.

The discrete approximation to our system is:

$$x_{k+1} - x_k = \int_{t=k\Delta t}^{(k+1)\Delta t} (f(x) + g(x)u + (\mu(x) + \sigma(x)w_t))dt \quad (2)$$

$$= \int_{t=k\Delta t}^{(k+1)\Delta t} (f(x) + g(x)u + \mu(x))dt + \int_{t=k\Delta t}^{(k+1)\Delta t} \sigma(x)w_t dt \quad (3)$$

$$\simeq (f(x_k) + g(x_k)u_k + \mu(x_k))\Delta t + \sigma(x_k) \int_{t=k\Delta t}^{(k+1)\Delta t} w_t dt \quad (4)$$

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  $\Delta t$  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  $\Delta t$ . It is important to note that although the Wiener process is also Gaussian 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  $W_k$ :

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

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

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

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

## I.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  $r(x) = \mathbf{1}_A(x)$  where  $A$  is the safe set, and  $\mathbf{1}_A(x)$  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  $V_0(x) = \mathbf{1}_A(x)$  and take

$$V_{k-1}(x) = \mathbf{1}_A(x) \mathbf{E}_{z \sim P(x_k=z|x_{k-1}=x)}(V_k(z)) \quad (8)$$

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:

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

$$= \mathbf{1}_A(x) \mathbf{E}_{b_k \sim N}(V_k(x + (f(x) + g(x)u_k + \mu(x))\Delta t + \sigma(x)\sqrt{\Delta t}b_k)) \quad (10)$$

$$= \mathbf{1}_A(x) \int_{-\infty}^{\infty} V_k(x + (f(x) + g(x)u_k + \mu(x))\Delta t + \sigma(x)\sqrt{\Delta t}b_k) \frac{1}{\sigma\sqrt{\Delta t}\sqrt{2\pi}} e^{-\frac{b^2}{2\Delta t\sigma^2}} db \quad (11)$$

In his development, Abate proves that this returns the expectation of never leaving the safe set. Actually using formula ?? to calculate the expected safety function, would

be rather arduous. Formula ?? 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.

### I.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.

$$V_{k-1}(x) = \mathbf{1}_A(x) \mathbb{E}_{z \sim P(x_k=z|x_{k-1}=x)}(V_k(z)) \quad (12)$$

$$= \mathbf{1}_A(x) \mathbb{E}_{b_k \sim N}(V_k(x + (f(x) + g(x)u_k + \mu(x))\Delta t + \sigma(x)\sqrt{\Delta t}b_k)) \quad (13)$$

$$= \mathbf{1}_A(x) \mathbb{E}_{b \sim N}(V_k(x) + \nabla V_k^T((f(x) + g(x)u + \mu(x))\Delta t + \sigma(x)\sqrt{\Delta t}b_k)) \quad (14)$$

$$= \mathbf{1}_A(x) V_k(x) + \mathbb{E}_{b \sim N}(\nabla V_k^T(\nabla V_k^T((f(x) + g(x)u + b)\Delta t + \sigma(x)\sqrt{\Delta t}b_k))) \quad (15)$$

$$= \mathbf{1}_A(x) V_k(x) + \nabla V_k^T \mathbb{E}_{b \sim N}(f(x) + g(x)u + b)\Delta t \quad (16)$$

$$\rightarrow \quad (17)$$

$$\frac{V_{k-1} - V_k(x)}{\Delta t} = \nabla V_k^T \mathbb{E}_{b \sim N}(f(x) + g(x)u + b) \quad (18)$$

$$(19)$$

$$\Delta t \rightarrow 0 \quad (20)$$

$$-\frac{\delta V}{\delta t} = \nabla V_k^T \mathbb{E}_{b \sim N}(f(x) + g(x)u + b) \quad (21)$$

$$-\frac{\delta V}{\delta t} = \nabla V_k^T(f(x) + g(x)u + \mu) \quad (22)$$

### I.4 Mitchell Magic

$$dx = F(x, u)dt + \sigma(x)dB \quad (23)$$

$$-\frac{\delta V}{\delta t} = \nabla V_k^T \mathbb{E}_{b \sim N}(f(x) + g(x)u + b) + 0.5 \text{trace}[\sigma(x)\sigma(x)^T \nabla_x^2 V_k] \quad (24)$$

$$dx = (f(x) + g(x)u + \mu(x))dt + \sigma(x)dB \quad (25)$$