Zero Dynamics

March 16, 2023

Consider system dynamics of the form

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})\boldsymbol{u}, \tag{1}$$

where $\boldsymbol{x} = [\boldsymbol{q}, \dot{\boldsymbol{q}}]^T$, and an output of the form

$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{q}), \tag{2}$$

which depends only on configuration variables. Because setting $\boldsymbol{h}(\boldsymbol{q})$ can only provide constraints on the M actuated degrees of freedom, there exists a smooth real-valued function $\boldsymbol{\theta}(\boldsymbol{q})$ for the $(N-M)$ unactuated degrees of freedom such that

$$[\boldsymbol{h}(\boldsymbol{q}); \boldsymbol{\theta}(\boldsymbol{q})] : \tilde{\mathcal{Q}} \to \mathcal{R}^N \tag{3}$$

is a diffeomorphism onto its image. Then the corresponding zero dynamics manifold $T\tilde{\mathcal{Q}}$ is a smooth $2(N-M)$-dimensional embedded submanifold of $T\mathcal{Q}$, defined as

$$\mathcal{Z} = \{\boldsymbol{x} \in T\tilde{\mathcal{Q}} | \boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) = 0, \dot{\boldsymbol{y}} = L_{\boldsymbol{f}}\boldsymbol{h}(\boldsymbol{x}) = 0\} \tag{4}$$

Note that because

$$\begin{aligned} \frac{d\boldsymbol{y}}{dt} &= \frac{d\boldsymbol{h}}{dt} = \frac{d\boldsymbol{h}}{d\boldsymbol{x}}\frac{d\boldsymbol{x}}{dt} = \frac{d\boldsymbol{h}}{d\boldsymbol{x}}\dot{\boldsymbol{x}} = \begin{bmatrix} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}} & \frac{\partial \boldsymbol{h}}{\partial \dot{\boldsymbol{q}}} \end{bmatrix} \dot{\boldsymbol{x}} = \begin{bmatrix} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}} & \boldsymbol{0} \end{bmatrix} \dot{\boldsymbol{x}} \\ &= L_{\boldsymbol{f}}\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}}\dot{\boldsymbol{q}}, \end{aligned} \tag{5}$$

and

$$\begin{aligned} \frac{d^2\boldsymbol{y}}{dt^2} &= \frac{d}{dt}(\frac{d\boldsymbol{y}}{dt}) = \frac{d}{dt}(\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}}\dot{\boldsymbol{q}}) = \frac{d}{d\boldsymbol{x}}(\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}}\dot{\boldsymbol{q}})\dot{\boldsymbol{x}} = \begin{bmatrix} \frac{\partial}{\partial \boldsymbol{q}}(\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}}\dot{\boldsymbol{q}}) & \frac{\partial}{\partial \dot{\boldsymbol{q}}}(\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}}\dot{\boldsymbol{q}}) \end{bmatrix} \dot{\boldsymbol{x}} = \begin{bmatrix} \frac{\partial}{\partial \boldsymbol{q}}(\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}}\dot{\boldsymbol{q}}) & \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}} \end{bmatrix} \dot{\boldsymbol{x}} \\ &= L_{\boldsymbol{f}}^2\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) + L_{\boldsymbol{g}}L_{\boldsymbol{f}}\boldsymbol{h}(\boldsymbol{q})\boldsymbol{u}, \end{aligned} \tag{6}$$

where L stands for the Lie derivative, the relative degree of the output is at least two and thus $\dot{\boldsymbol{y}} = 0$ is needed in the definition of zero dynamics. The corresponding feedback control $\boldsymbol{u}$ can be decomposed into two parts as

$$\boldsymbol{u} = \boldsymbol{u}^* + \boldsymbol{v}, \tag{7}$$

where $\boldsymbol{u}^*$ is the term that renders $\mathcal{Z}$ invariant, *i.e.* $\ddot{\boldsymbol{y}} = 0$, defined as

$$\boldsymbol{u}^*(\boldsymbol{x}) = -(L_{\boldsymbol{g}}L_{\boldsymbol{f}}\boldsymbol{h}(\boldsymbol{q}))^{-1}L_{\boldsymbol{f}}^2\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) \tag{8}$$

and $\boldsymbol{v}$ is the term that ensures the output $\boldsymbol{y}$ and its derivative vanish to zero under the designed control law. A simple example of the control law would be a linear (PD) controller designed in the form

$$\boldsymbol{v} = -(L_{\boldsymbol{g}}L_{\boldsymbol{f}}\boldsymbol{h}(\boldsymbol{q}))^{-1}[\frac{1}{\epsilon}\boldsymbol{K}_D L_{\boldsymbol{f}}\boldsymbol{h}(\boldsymbol{q}) + \frac{1}{\epsilon^2}\boldsymbol{K}_P\boldsymbol{h}(\boldsymbol{q})]. \tag{9}$$

To simplify the zero dynamics, further coordinate transformation is performed. Set

$$\boldsymbol{\Phi}(\boldsymbol{q}) = \begin{bmatrix} \boldsymbol{h}(\boldsymbol{q}) \\ \boldsymbol{\theta}(\boldsymbol{q}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\xi}_1 \end{bmatrix}, \tag{10}$$

which is a coordinate transformation on $\tilde{\mathcal{Q}}$, thus

$$\begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\xi}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\xi}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\xi}_1 \\ \dot{\boldsymbol{\eta}}_1 \\ \dot{\boldsymbol{\xi}}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}(\boldsymbol{q}) \\ \boldsymbol{\theta}(\boldsymbol{q}) \\ L_{\boldsymbol{f}}\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) \\ L_{\boldsymbol{f}}\boldsymbol{\theta}(\boldsymbol{q}, \dot{\boldsymbol{q}}) \end{bmatrix} \tag{11}$$

is a coordinate transformation on $T\tilde{\mathcal{Q}}$, the inverse transformation of which is

$$\begin{bmatrix} \boldsymbol{q} \\ \dot{\boldsymbol{q}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}^{-1}(\boldsymbol{\eta}_1, \boldsymbol{\xi}_1) \\ (\frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{q}})^{-1} \begin{bmatrix} \boldsymbol{\eta}_2 \\ \boldsymbol{\xi}_2 \end{bmatrix} \end{bmatrix}. \tag{12}$$

Note that the above relationship has used the following derivation,

$$\begin{bmatrix} \boldsymbol{\eta}_2 \\ \boldsymbol{\xi}_2 \end{bmatrix} = \begin{bmatrix} \dot{\boldsymbol{\eta}}_1 \\ \dot{\boldsymbol{\xi}}_1 \end{bmatrix} = \frac{d\boldsymbol{\Phi}}{dt} = \frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{q}} \dot{\boldsymbol{q}} \Rightarrow \dot{\boldsymbol{q}} = (\frac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{q}})^{-1} \begin{bmatrix} \boldsymbol{\eta}_2 \\ \boldsymbol{\xi}_2 \end{bmatrix}. \tag{13}$$

Using the new coordinates $[\boldsymbol{\eta}_1; \boldsymbol{\xi}_1; \boldsymbol{\eta}_2; \boldsymbol{\xi}_2]$, the system dynamics become

$$\begin{bmatrix} \dot{\boldsymbol{\eta}}_1 \\ \dot{\boldsymbol{\xi}}_1 \\ \dot{\boldsymbol{\eta}}_2 \\ \dot{\boldsymbol{\xi}}_2 \end{bmatrix} = \begin{bmatrix} L_{\boldsymbol{f}}\boldsymbol{h} \\ L_{\boldsymbol{f}}\boldsymbol{\theta} \\ L_{\boldsymbol{f}}^2\boldsymbol{h} + L_{\boldsymbol{g}}L_{\boldsymbol{f}}\boldsymbol{h}\boldsymbol{u} \\ L_{\boldsymbol{f}}^2\boldsymbol{\theta} + L_{\boldsymbol{g}}L_{\boldsymbol{f}}\boldsymbol{\theta}\boldsymbol{u} \end{bmatrix}, \tag{14}$$

and the zero dynamics become

$$\begin{cases} \boldsymbol{y} = \boldsymbol{\eta}_1 = \boldsymbol{0}, \dot{\boldsymbol{y}} = \dot{\boldsymbol{\eta}}_1 = \boldsymbol{\eta}_2 = \boldsymbol{0} \\ \dot{\boldsymbol{\xi}}_1 = L_{\boldsymbol{f}}\boldsymbol{\theta} \\ \ddot{\boldsymbol{y}} = \dot{\boldsymbol{\eta}}_2 = L_{\boldsymbol{f}}^2\boldsymbol{h} + L_{\boldsymbol{g}}L_{\boldsymbol{f}}\boldsymbol{h}\boldsymbol{u}^* = \boldsymbol{0} \\ \dot{\boldsymbol{\xi}}_2 = L_{\boldsymbol{f}}^2\boldsymbol{\theta} + L_{\boldsymbol{g}}L_{\boldsymbol{f}}\boldsymbol{\theta}\boldsymbol{u} \end{cases} \Rightarrow \begin{cases} \dot{\boldsymbol{\xi}}_1 = L_{\boldsymbol{f}}\boldsymbol{\theta} \\ \dot{\boldsymbol{\xi}}_2 = L_{\boldsymbol{f}}^2\boldsymbol{\theta} + L_{\boldsymbol{g}}L_{\boldsymbol{f}}\boldsymbol{\theta}\boldsymbol{u} \end{cases}. \tag{15}$$

Because the columns of $\boldsymbol{g}(\boldsymbol{x})$ are involutive, the inversion of the decoupling matrix in the zero dynamics can be avoided by using a smooth scalar function $\boldsymbol{\gamma}$, such that

$$\begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\xi}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\xi}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}(\boldsymbol{q}) \\ \boldsymbol{\theta}(\boldsymbol{q}) \\ L_{\boldsymbol{f}}\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) \\ \boldsymbol{\gamma}(\boldsymbol{q}, \dot{\boldsymbol{q}}) \end{bmatrix} \tag{16}$$

is a valid coordinate transformation and

$$L_{\boldsymbol{g}}\boldsymbol{\gamma} = \boldsymbol{0}. \tag{17}$$

Additionally, it can be proved that $\boldsymbol{\gamma}$ can be explicitly computed to be the last $(N - M)$ entry of $\boldsymbol{D}(\boldsymbol{q})\dot{\boldsymbol{q}}$, hence it can be assumed that

$$\boldsymbol{\gamma}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \boldsymbol{\gamma}_0(\boldsymbol{q})\dot{\boldsymbol{q}}. \tag{18}$$

The corresponding inverse transformation thus becomes

$$\begin{bmatrix} \boldsymbol{q} \\ \dot{\boldsymbol{q}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}^{-1}(\boldsymbol{\eta}_1, \boldsymbol{\xi}_1) \\ \begin{bmatrix} \dfrac{\partial \boldsymbol{\Phi}}{\partial \boldsymbol{q}} \\ \boldsymbol{\gamma}_0 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{\eta}_2 \\ \boldsymbol{\xi}_2 \end{bmatrix} \end{bmatrix}. \tag{19}$$

Consequently, the system dynamics become

$$\begin{bmatrix} \dot{\boldsymbol{\eta}}_1 \\ \dot{\boldsymbol{\xi}}_1 \\ \dot{\boldsymbol{\eta}}_2 \\ \dot{\boldsymbol{\xi}}_2 \end{bmatrix} = \begin{bmatrix} L_{\boldsymbol{f}} \boldsymbol{h} \\ \boldsymbol{\gamma} \\ L_{\boldsymbol{f}}^2 \boldsymbol{h} + L_{\boldsymbol{g}} L_{\boldsymbol{f}} \boldsymbol{h} \boldsymbol{u} \\ L_{\boldsymbol{f}} \boldsymbol{\gamma} \end{bmatrix}, \tag{20}$$

and the zero dynamics become

$$\begin{cases} \dot{\boldsymbol{\xi}}_1 = \boldsymbol{\gamma} \\ \dot{\boldsymbol{\xi}}_2 = L_{\boldsymbol{f}} \boldsymbol{\gamma} \end{cases}, \tag{21}$$

where the corresponding inverse transformation becomes

$$\begin{bmatrix} \boldsymbol{q} \\ \dot{\boldsymbol{q}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}^{-1}(\boldsymbol{0}, \boldsymbol{\xi}_1) \\ \begin{bmatrix} \dfrac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}} \\ \boldsymbol{\gamma}_0 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{\xi}_2 \end{bmatrix} \end{bmatrix}. \tag{22}$$

To relate the zero dynamics to the original system dynamics, further derivation can be conducted. Set the last $(N-M)$ rows of matrices and vectors $\boldsymbol{D}$, $\boldsymbol{C}$, $\boldsymbol{G}$ and $\boldsymbol{q}$ as $\boldsymbol{D}_M$, $\boldsymbol{C}_M$, $\boldsymbol{G}_M$ and $\boldsymbol{q}_M$. Then we can have

$$\boldsymbol{\gamma}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \boldsymbol{\gamma}_0(\boldsymbol{q})\dot{\boldsymbol{q}} = \boldsymbol{D}_M \dot{\boldsymbol{q}} = (\frac{\partial K}{\partial \dot{\boldsymbol{q}}_M})^T. \tag{23}$$

Therefore, the zero dynamics become

$$\begin{cases} \boldsymbol{\xi}_1 = \boldsymbol{\theta} \\ \boldsymbol{\xi}_2 = (\dfrac{\partial K}{\partial \dot{\boldsymbol{q}}_M})^T \end{cases} \tag{24}$$

and

$$\begin{cases} \dot{\boldsymbol{\xi}}_1 = \dfrac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{q}} \dot{\boldsymbol{q}} = \dfrac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{q}} \begin{bmatrix} \dfrac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}} \\ \boldsymbol{\gamma}_0 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{\xi}_2 \end{bmatrix} \\ \dot{\boldsymbol{\xi}}_2 = \begin{bmatrix} \dfrac{\partial (\boldsymbol{D}_M \dot{\boldsymbol{q}})}{\partial \boldsymbol{q}} & \dfrac{\partial (\boldsymbol{D}_M \dot{\boldsymbol{q}})}{\partial \dot{\boldsymbol{q}}} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{q}} \\ -\boldsymbol{D}^{-1}[\boldsymbol{C}\dot{\boldsymbol{q}} + \boldsymbol{G}] \end{bmatrix} \\ \quad = \begin{bmatrix} \dot{\boldsymbol{q}}^T \dfrac{\partial \boldsymbol{D}_M^T}{\partial \boldsymbol{q}} & \boldsymbol{D}_M \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{q}} \\ -\boldsymbol{D}^{-1}[\boldsymbol{C}\dot{\boldsymbol{q}} + \boldsymbol{G}] \end{bmatrix} \\ \quad = \dot{\boldsymbol{q}}^T \dfrac{\partial \boldsymbol{D}_M^T}{\partial \boldsymbol{q}} \dot{\boldsymbol{q}} - \boldsymbol{C}_M \dot{\boldsymbol{q}} - \boldsymbol{G}_M \\ \quad = \dot{\boldsymbol{q}}^T \dfrac{\partial \boldsymbol{D}_M^T}{\partial \boldsymbol{q}} \dot{\boldsymbol{q}} - (\dot{\boldsymbol{q}}^T \dfrac{\partial \boldsymbol{D}_M^T}{\partial \boldsymbol{q}} - \dfrac{1}{2} \dot{\boldsymbol{q}}^T \dfrac{\partial \boldsymbol{D}}{\partial \boldsymbol{q}_M}) \dot{\boldsymbol{q}} - \boldsymbol{G}_M \\ \quad = \dfrac{1}{2} \dot{\boldsymbol{q}}^T \dfrac{\partial \boldsymbol{D}}{\partial \boldsymbol{q}_M} \dot{\boldsymbol{q}} - \boldsymbol{G}_M \\ \quad = -\boldsymbol{G}_M \quad \text{(Because } \boldsymbol{D} \text{ is related to the kinetic energy which has nothing to do with } \boldsymbol{q}_M) \\ \quad = -\dfrac{\partial V}{\partial \boldsymbol{q}_M} \end{cases} \tag{25}$$