Dimentionless Policies based on the Buckingham π Theorem: Is it a good way to Generalize Numerical Results?

Alexandre Girard[1]

***Abstract*— Yes if the context, the list of variable defining the control problem, are dimentionnaly similar. Here we show that by modifying the problem formulation using dimentionless variables, we can re-use the optimal policy generated numerically for a specific system to a sub-space of dimentionnaly similar problem. This is demonstrated, with numerical results of optimal policies, for the classic motion control problem of swinging-up a torque-limited inverted pendulum. We also demonstrate that by leveraging this scheme when using reinforcement learning, multiple systems of various dimentions can share a data-base during the learning phase, which can be a big advantage for data efficiency [TODO!]. It remains to be seen if this approach can also help generalizing policies for more complex high-dimentional problems.**

I. Introduction

a) Many numerical algorithms = black box mapping: -Trajectory optimization -Reinforcement learning -etc.

b) With numerical results, unlike analytical solutions, system and problem parameters are not explicitly in the solution :

c) This makes the results "context specific" and makes its harder to generalize :

II. Dimentionless Policy: Concept

In the following section, the Buckingham pi theorem is use to develop the concept of dimentionless policies, and it is shown that multiple systems should have the same dimentionless policy if they have what we will call a similar dimentionless context.

A. Context variables in the policy mapping

A state feedback law is defined here as a mapping f, specific to a given system, from an vector space representing the state x of the dynamic system, to a vector space representing the control inputs u of the system:

$$u = f(x) \tag{1}$$

Under some assumptions, mainly a fully observable systems, an additive cost and an infinite time horizon, the optimal policy is also guarantee to be of this form [Cite Bertsekas]. Only this case is considered is considered for the following analysis.

To consider the question of how can this feedback law (a form of system specific knowledge) can be transfered in a different context, it is usefull to think about a higher

[1]Alexandre Girard is with the Department of Mechanical Engineering, Universite de Sherbrooke, Qc, Canada alex.girard@usherbrooke.ca

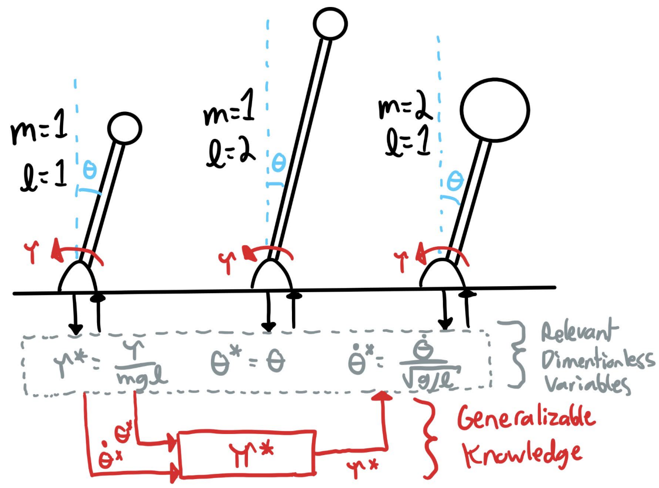


Fig. 1: Big picture

dimension mapping, that we will note π, also having input arguments a vector of variables c describing the context:

$$u = \pi(x, c) \tag{2}$$

The context c is the vector of all variables that would affect what is the feedback law solution to a motion control problem. For instance, in section III, a case study is conducted considering the optimal feedback law for swinging-up an inverted pendulum. For this exemple, the context variables are the pendulum mass m, the gravitational constant g, the lenght l, but also what we will call task parameters: a parameter in the cost function q and a constraint τ_{max} on the maximum input torque, see Fig. 2. For a given state of the system, the torque solution might be different if any of the context variable is modify, for instance if the pendulum is heavier, more torque limited, etc.

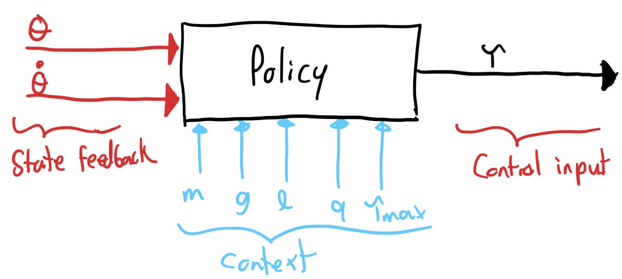


Fig. 2: For the pendulum swing-up task, the context is 5 variables: the parameter of the system: m, g, and l, a parameter of the cost function: q and a parameter defining the constraints: τ_{max}. The feedback law solution i

The context must include all the variables, system parameters and task parameters, that would affect the solution to the control problem:

$$\underbrace{\begin{bmatrix}u_1\\ \vdots\\ u_k\end{bmatrix}}_{\text{inputs}} = \pi\left(\underbrace{\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}}_{\text{states}}, \underbrace{\underbrace{\begin{bmatrix}c_1\\ \vdots\\ \vdots\\ c_m\end{bmatrix}}_{\text{system}}, \underbrace{\begin{bmatrix}c_{m+1}\\ \vdots\\ c_{m+l}\end{bmatrix}}_{\text{task}}}_{\text{Context } c}\right) \tag{3}$$

Then we can formalize the goal of generalizing a policy to a different context: if a good feedback policy f_a is found for a system in a context a described by variables c_a, can this knowledge help finding an equivalent good feedback policy in a different context c_b?

$$\pi\left(x, c = c_a\right) = f_a\left(x\right) \quad \Rightarrow \quad \pi\left(x, c = c_b\right) = ? \tag{4}$$

Using the Buckinham Pi theorem [Cite something], we will show that if a dimentionless version of the context is equal, then a dimentionless version of the policy mapping must be also equivalent.

B. Dimentional analysis of the augmented policy mapping

For a system with k control inputs, we can treat the augmented policy as k mapping from states and context variables to each scalar control input u_j:

$$u_j = \pi_j\left(x_1, \ldots, x_n, c_1, \ldots \ldots, c_{m+l}\right) \tag{5}$$

where eq. (5) is the jth line of the policy in vector form described by eq. (3). Then, if the state vector is define by n variables, and the context is defined by m system parameters plus l tasks parameters, then each mapping π_j involves $1 + n + m + l$ (usually dimentionnal) variables. Here, we will assume that the policy fonction is physically meaningful, in the sense of the requirement for applying the Buckingham Pi theorem [Cite]. This means for exemple, that a policy that compute a force based on position and velocity measurement would be in this framework, but not of policy for playing chess for instance.

Applying the Buckingham Pi theorem to the relashionship, tell us that if d dimensions are involved in all thoses variables, then 5 can be restated into an equivalent dimentionless relationship between p dimentionless Π groups where $p \geq (1 + n + m + l) - d$ [cite bukhingham pi]. Assuming d dimentions are involved in the m system parameter, and that we are in a situation where the maximum reduction is possible $p = (1 + n + m + l) - d$, we can pick d context variable $\{c_1, c_2, \ldots, c_d\}$ as the basis (the repeated variables) to scale all other variables in a dimentionless form (we will note dimentionless Π group as variables with an * subscript):

$$u_j^* = u_j\,[c_1]^{e^u_{1j}}\,[c_2]^{e^u_{2j}} \ldots [c_d]^{e^u_{dj}} \qquad j=\{1,\ldots,k\} \tag{6}$$

$$x_i^* = x_i\,[c_1]^{e^x_{1i}}\,[c_2]^{e^x_{2i}} \ldots [c_d]^{e^x_{di}} \qquad i=\{1,\ldots,n\} \tag{7}$$

$$c_i^* = c_i\,[c_1]^{e^c_{1i}}\,[c_2]^{e^c_{2i}} \ldots [c_d]^{e^c_{di}} \qquad i=\{d+1,\ldots,m+l\} \tag{8}$$

where exposant e are rational numbers selected to make all equations dimentionless. Then, the buckingham theorem tell us that the relationship described by eq. (5) can be restated as the following relationship between dimentionsless variables:

$$u_j^* = \pi_j^*\left(x_1^*, \ldots, x_n^*, c_{d+1}^*, \ldots, c_{m+l}^*, \right) \tag{9}$$

involving d less dimentionless variables in eq. (5). If we apply the same procedure to all control inputs, when can then assemble the k mapping back into a vector form:

$$\underbrace{\begin{bmatrix}u_1^*\\ \vdots\\ u_k^*\end{bmatrix} = \pi^*\left(\begin{bmatrix}x_1^*\\ \vdots\\ x_n^*\end{bmatrix}\right.}_{\text{Dimentionless feedback law } f^*} , \underbrace{\left.\begin{bmatrix}c_{d+1}^*\\ \vdots\\ c_m^*\end{bmatrix}, \begin{bmatrix}c_{m+1}^*\\ \vdots\\ c_{m+l}^*\end{bmatrix}\right)}_{\text{Dimentionless context } c^*} \tag{10}$$

that we will sometime write in compact form as:

$$u^* = \pi^*(x^*, c^*) \tag{11}$$

One interesting perk of this dimentional analysis, is that we can remove p variable form the context (typically p would be 2 or 3 for controlling a physical system involving time, force and length). The global problem of learning $\pi(x, c)$, i.e. the good feeback policy for all possible context is thus simplified in a dimentionless form. Also, an even more interesting feature to help transfering feedback law between systems, is that the dimentional analysis procedure shows that a global policy $\pi(x, c)$, will have an equivalent dimentionless form for multiple context c. As illustrated at Fig. 3, the dimentionless context c^* is a lower dimentional space $(m + l - d)$ and multiple context vector c will correspond to the same dimentionless vector c^*.

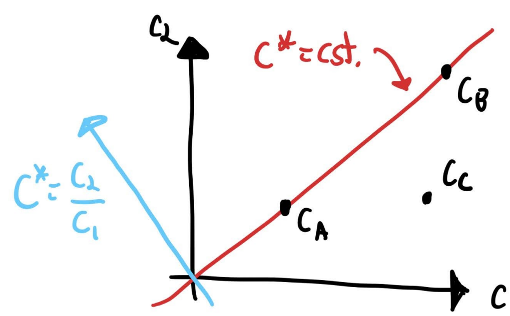


Fig. 3: Dimentionlaly similar context

For a given control problem, if the dimentionless context are equal, then the dimentionless feedback law should be eactly equivalent:

$$\text{if} \quad c_a^* = c_b^* \quad \text{then} \quad f_a^*(x^*) = f_b^*(x^*)\ \forall x^* \tag{12}$$

where

$$f_a^*(x^*) = \pi^*(x^*, c^* = c_a^*) \tag{13}$$

$$f_b^*(x^*) = \pi^*(x^*, c^* = c_b^*) \tag{14}$$

This is simply based on the fact that the dimentionless policy (eq. (11)) give the same output for the same inputs. This results means that the knowledge of a policy for a specific context c can actually be generalized to a sub-space of all context for which the dimentionless context c^* is equal.

Lets illustrate this with an example of a spherical submarine for witch we have an optimal policy of trust as a function of two context variables: the sphere radius and its velocity. The dimentionless policy would only be a function of a single dimentionless context variable: the Reynolds number, and would be the same for all submarines with the same Reynolds number.

C. Transfering policies between contexts

In order to exploit this property, it is usefull to define transformation matrices based on scalar equations (6), (7) and (8):

$$u^* = [T_u(c)]\ u \tag{15}$$

$$x^* = [T_x(c)]\ x \tag{16}$$

$$c^* = [T_c(c)]\ c \tag{17}$$

where matrices T_u and T_x are square diagonal matrix, where each diagonal term is a multiplication of the first d context variables ($\{c_1, c_2, \ldots, c_d\}$) up to a rational power (found by applying the buckinghan pi theorem). Equations (15) and (16) are inversible (unless a context variable is equal to zero) and can be used to go back-and-forth between dimentional and dimentionless state and input variables. The matrix T_c however have d less row than columns and equation (17) is not inversible, for a given context c there is only one dimentionless context c^*, however a dimentionless context c^* correspond to multiple dimentional context c.

To summarize the dimentional analysis procedure, using the transformation matrices, if a dimentional feedback law f_a for a context c_a is known:

$$f_a(x) = \pi\left(x, c = c_a\right) \tag{18}$$

its representation in dimentionless form:

$$f_a^*(x^*) = \pi\left(x^*, c^* = c_a^*)\right) \tag{19}$$

can be found by scaling the input and output of f_a with T_u and T_x:

$$f_a^*(x^*) = \underbrace{T_u(c_a)\, f_a\left(\underbrace{T_x^{-1}(c_a)\, x^*}_{x}\right)}_{u} \tag{20}$$

and the dimentionless context values c_a^* can be found using eq. (17). Inversely, if we know a dimentionless feedback law f_b^*, matrices T_u and T_x can be used to scale it back to a specific context c_b:

$$f_b(x) = \underbrace{T_u^{-1}(c_b)\, f_b^*\left(\underbrace{T_x(c_b)\, x}_{x^*}\right)}_{u^*} \tag{21}$$

Thus, eq. (20) and eq. (20) can be used to take any context specific feedback law, finding its dimentionless form, and

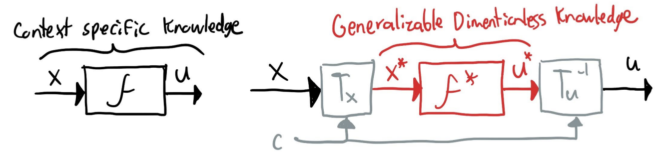


Fig. 4: Isolating the dimentionless knowledge in a policy

scale it back to a new context. In general, there is no guarantee that the behavior of the scaled feedback law in the new context will be similar to the behavior of the feedback law in the original context. However, if the dimentionless context are equal, then the behavior should be eactly equivalent, because the motion control problem was actually the same problem but scaled. For instance, lets suppose an optimal policy f_a is known for a specific context c_a, then applying the scaled policy:

$$f_b(x) = \left[T_u^{-1}(c_b) T_u(c_a)\right]\ f_a\left(\left[T_x^{-1}(c_a) T_x(c_b)\right]\ x\right) \tag{22}$$

on a system described by the context c_b will be equaly equivalent (up to scaling factors) if

$$c_b^* = T_c(c_b)\ c_b = T_c(c_a)\ c_a = c_a^* \tag{23}$$

In section III, we show exemples of this results with numerical solution to dimentionnaly similar pendulum swing-up problems.

III. Optimal pendulum swing-up task

In this paper, we will use a version of the pendulum swing-up task as a prototype problem to test the proposed ideas of dimentionless policies. The motion control problem is defined as finding a feedback law for controling the dynamic system is described by differential equations:

$$ml^2\ddot{\theta} - mgl\sin\theta = \tau \tag{24}$$

that minimize the cost function given by:

$$J = \int (q^2\theta^2 + 0\,\dot{\theta}^2 + 1\,\tau^2)dt \tag{25}$$

subject to input constraints given by:

$$-\tau_{max} \leq \tau \leq \tau_{max} \tag{26}$$

Note that here, the cost function parameter q is include in way to have units of torque, it was chosen to fixe the weight on velocity to zero for simplicity, and the weight on torque to one without loss of generality as only the relative value of weight with impact the solution.

Thus, assuming there is no hidden variables and that equations (51), (52) and (26) fully describe the problem. The solution, i.e. the optimal policy for all context, should be of the form given by:

$$\underbrace{\tau}_{\text{inputs}} = \pi\left(\underbrace{\theta, \dot{\theta}}_{\text{states}},\ \underbrace{\underbrace{m, g, l}_{\text{system parameters}},\ \underbrace{q, \tau_{max}}_{\text{task parameters}}}_{\text{Context } c}\right) \tag{27}$$

TABLE I: Pendulum swing-up optimal policy variables

Variable	Description	Units	Dimensions
	Control inputs		
τ	Actuator torque	Nm	$[ML^2T^{-2}]$
	State variables		
θ	Joint angle	rad	[]
$\dot{\theta}$	Joint angular velocity	rad/sec	$[T^{-1}]$
	System parameters		
m	Pendulum mass	kg	$[M]$
g	Gravity	m/s^2	$[LT^{-2}]$
l	Pendulum lenght	m	$[L]$
	Problem parameters		
q	Weight parameter	Nm	$[ML^2T^{-2}]$
τ_{max}	Maximum torque	Nm	$[ML^2T^{-2}]$

and involving variables listed in table I.

Before, conducting the dimentional analysis, it is interesting to note that while there are 3 system paramters m, g and l, they only appear indepedently in two groups in the dynamic equation. We can thus consider only two system parameters, and for convenience mgl coresponding (the maximum static gravitationnal torque) and ω are selected, as listed at table II

TABLE II: Pendulum swing-up minimal system variables

	System parameters		
mgl	Maximum gravitational torque	Nm	$[ML^2T^{-2}]$
$\omega = \sqrt{\frac{g}{l}}$	Natural frequency	sec^{-1}	$[T^{-1}]$

A. Dimentional analysis

Here we have one control input, two state, two system parameter and two task parameter, for a total of $1 + (n = 2) + (m = 2) + (l = 2) = 7$ variables are involved. In those variables, only $d = 2$ independants dimensions (ML^2T^{-2} and T^{-1}) are present. Using $c_1 = mgl$ and $c_2 = \omega$ as the repeating variables leads to the following dimentionless groups:

$$\Pi_1 = \tau^* = \frac{\tau}{mgl} \qquad \frac{[ML^2T^{-2}]}{[M][LT^{-2}][L]} \tag{28}$$

$$\Pi_2 = \theta^* = \theta \qquad [] \tag{29}$$

$$\Pi_3 = \dot{\theta}^* = \frac{\dot{\theta}}{\omega} \qquad \frac{[T^{-1}]}{[T^{-1}]} \tag{30}$$

$$\Pi_4 = \tau^*_{max} = \frac{\tau_{max}}{mgl} \qquad \frac{[ML^2T^{-2}]}{[M][LT^{-2}][L]} \tag{31}$$

$$\Pi_5 = q^* = \frac{q}{mgl} \qquad \frac{[ML^2T^{-2}]}{[M][LT^{-2}][L]} \tag{32}$$

Here, all 3 torque variables (τ, q and τ_{max}) are scaled by the maximum gravitationnal torque, and the pendulum velocity variable is scaled by the pendulum natural frequency. The transformation matrices are then

$$\tau^* = \underbrace{[1/mgl]}_{T_u} \tau \tag{33}$$

$$\begin{bmatrix} \theta^* \\ \dot{\theta}^* \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1/\omega \end{bmatrix}}_{T_x} \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} \tag{34}$$

$$\begin{bmatrix} q^* \\ \tau^*_{max} \end{bmatrix} = \underbrace{\begin{bmatrix} 1/mgl & 0 \\ 0 & 1/mgl \end{bmatrix}}_{T_c} \begin{bmatrix} q \\ \tau_{max} \end{bmatrix} \tag{35}$$

According to the theorem, any policy that is only based on the variable included in our analysis can be expressed as a relationship between the 5 dimentionless pi groups in the form:

$$\tau^* = \pi^* \left(\theta, \dot{\theta}^*, q^*, \tau^*_{max} \right) \tag{36}$$

Here, the dimentional analysis shows that, for dimentionnally similar swing-up problem (which means here equal ratios q^* and τ^*_{max}) the optimal feedback laws should be equivalent in their dimentionless form. In other words, if we have an optimal policy f_a found in a specific context $c_a = [m_a, l_a, g_a, q_a, \tau_{max,a}]$, and an optimal policy f_b for a second context $c_b = [m_b, l_b, g_b, q_b, \tau_{max,b}]$. Then, both dimentionless form will be equal $f^*_a = f^*_b$ if $q^*_1 = q^*_2$ and $\tau^*_{max,1} = \tau^*_{max,2}$, what we call a *dimentionnally similar context*. Furthermore, we can thus find f_b using f_a or vice-versa unsing the scaling given by eq. (22). However, if $q^*_a \neq q^*_b$ or $\tau^*_{max,a} \neq \tau^*_{max,b}$ then f_a doesn't give us information on f_b wihtout additionnal assumptions.

B. Numerical results

Here, a numerical algorithm (details of the methodology is presented at ...) is use to compute numerical solution to the prototype motion control problem defined by eq (51), (52) and (26). The used numerical recipe produce feedback law solution in the form of look-up table, based on the discretize grid of the state-space.

The optimal (up to discretization error) feedback laws are computed for 9 contexts listed at table III. In those 9 contexts, there is 3 sub-group of 3 cases with dimentionless similar context. Also each sub-group inlcude the same 3 pendulums, illustrated at Fig. 1, a regular, a twice longer and a twice heavier. Contexts 1, 2 and 3 describe a task where the torque is limited to half the satic maximum torque. Contexts 4, 5 and 6 describe a task where the cost highly penalize applying large forces relatively to position errors. Contexts 7, 8 and 9 describe a task where the cost highly penalize position errors, relatively to applying large forces.

TABLE III: Pendulum swing-up problems parameters

	m	g	l	q	τ_{max}
Problems with $\tau^*_{max} = 0.5$ **and** $q^* = 0.1$					
Context no 1 :	1.0	10.0	1.0	1.0	5.0
Context no 2 :	1.0	10.0	2.0	2.0	10.0
Context no 3 :	2.0	10.0	1.0	2.0	10.0
Problems with $\tau^*_{max} = 1.0$ **and** $q^* = 0.05$					
Context no 4 :	1.0	10.0	1.0	0.5	10.0
Context no 5 :	1.0	10.0	2.0	1.0	20.0
Context no 6 :	2.0	10.0	1.0	1.0	20.0
Problems with $\tau^*_{max} = 1.0$ **and** $q^* = 10$					
Context no 7 :	1.0	10.0	1.0	100.0	10.0
Context no 8 :	1.0	10.0	2.0	200.0	20.0
Context no 9 :	2.0	10.0	1.0	200.0	20.0

Figure XXXX illustrate that for each sub-group with equal dimentionless context, the dimentional feedback law generated numerically, and also the exemple trajectory of the system starting from rest at the bottom position, looks very similar. They are similar up to a scaling of the axis, if we neglect slight differences due to discretization errors. Furthermore, when we compute the dimentionless verion of the feedback laws f^*, using eq. (20), the dimentionless version is actually equal within the dimentionaly similar context sub-group. This was the expected results given by the dimentional analysis of section II.

C. Methodology

We used the value-iteration [cite] algoithm on a discretized version of the continuous system...

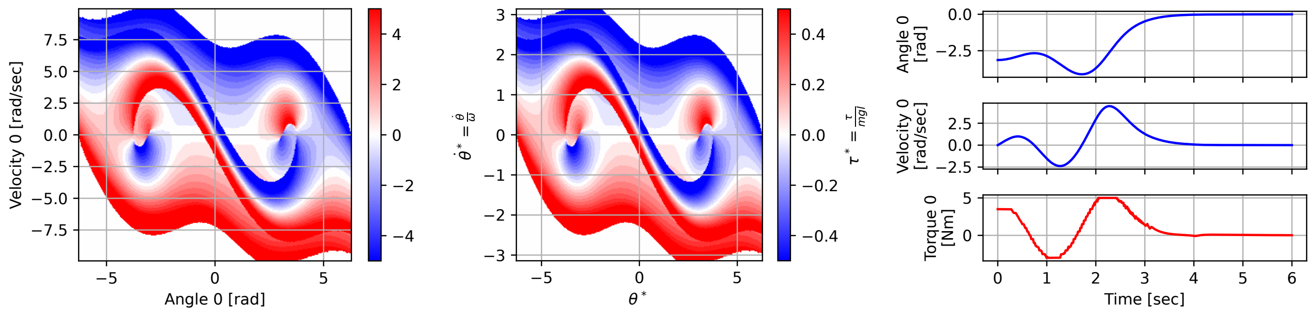
1) Additionnal dimentionless parameters for the solver: Using dynamic programming for solving the optimal policy numerically require setting additionnal parameter that define the domain. Altough those parameter should not affect the optimal policy far away from the boundaries, here a dimensionless version of those parameters was kept fixed in all the experiments:

$$\theta^*_{max} = \theta_{max} = 2\pi \tag{37}$$

$$\dot{\theta}^*_{max} = \frac{\dot{\theta}_{max}}{\omega} = 2 \tag{38}$$

$$t^*_f = t_f \, \omega = 10 \times 2\pi \tag{39}$$

θ_{max} is the range of angle for witch the optimal policy is solved, here set at one full revolution. $\dot{\theta}_{max}$ is the range of angular velocity for witch the optimal policy is solved, here the dimentionless ratio scaled with the natural frequency is set at 2. t_f is the time horizon, here its associated dimentionless ratio is fixed to always corespond to 10 periods of the pendulum using the natural frequency.

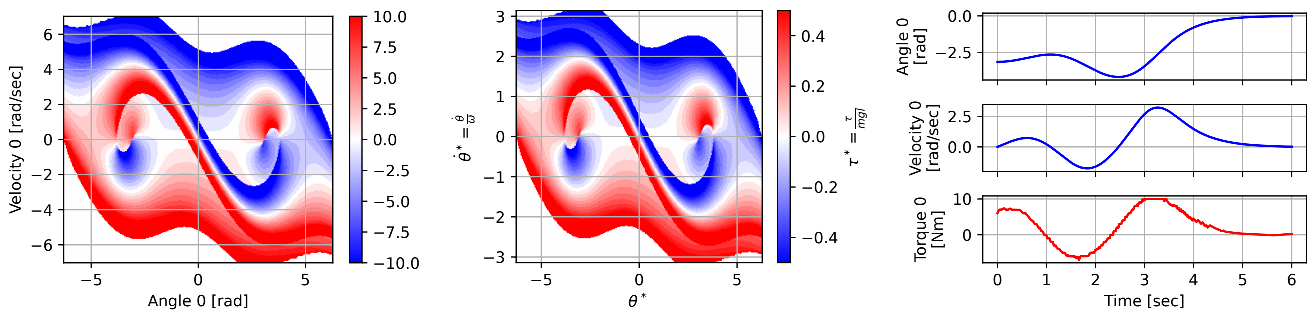


(a) Feedback law f

(b) Dimentionless feedback law f^*

(c) Exemple trajectory starting at $\theta = -\pi$

Fig. 5: Numerical results for context no 1

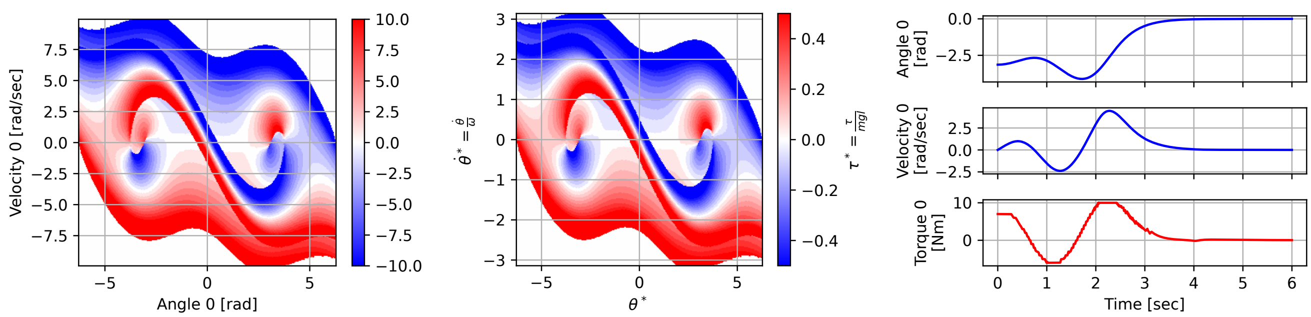


(a) Feedback law f

(b) Dimentionless feedback law f^*

(c) Exemple trajectory starting at $\theta = -\pi$

Fig. 6: Numerical results for context no 2

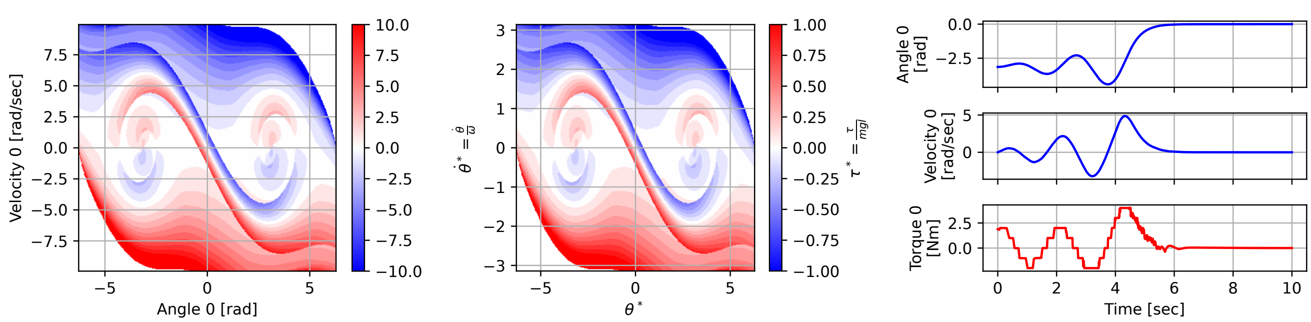


(a) Feedback law f

(b) Dimentionless feedback law f^*

(c) Exemple trajectory starting at $\theta = -\pi$

Fig. 7: Numerical results for context no 3

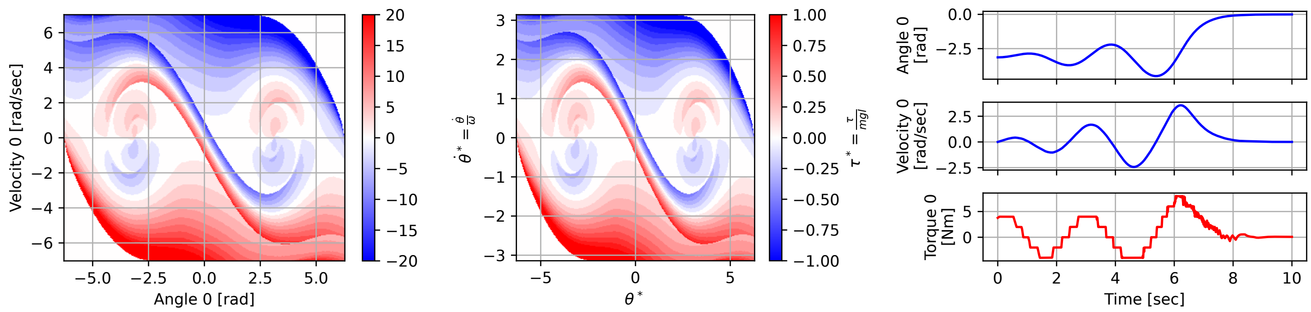


(a) Feedback law f

(b) Dimentionless feedback law f^*

(c) Exemple trajectory starting at $\theta = -\pi$

Fig. 8: Numerical results for context no 4

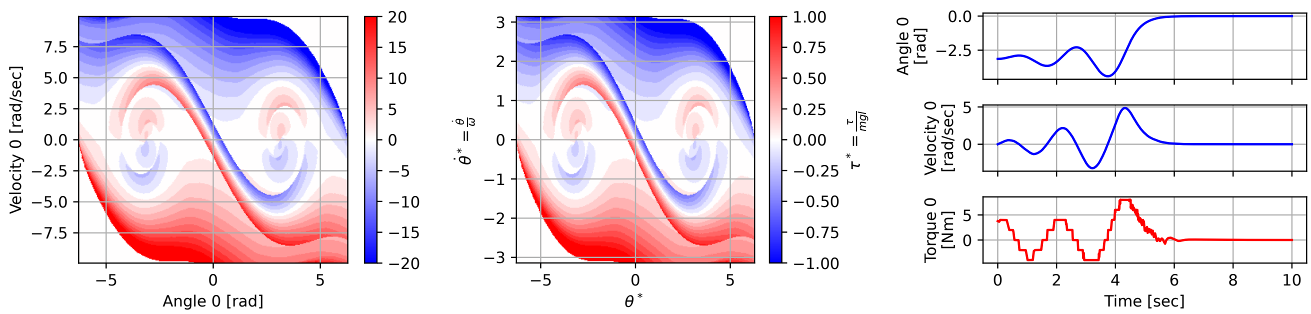


(a) Feedback law f

(b) Dimentionless feedback law f^*

(c) Exemple trajectory starting at $\theta = -\pi$

Fig. 9: Numerical results for context no 5

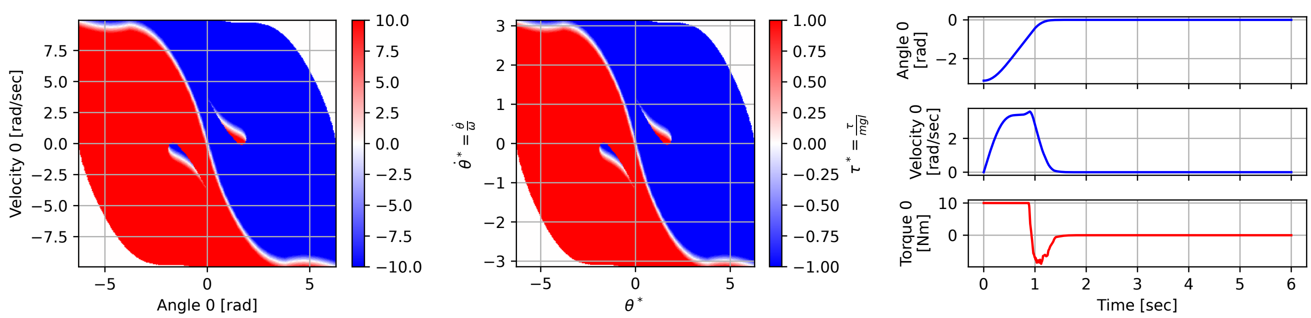


(a) Feedback law f

(b) Dimentionless feedback law f^*

(c) Exemple trajectory starting at $\theta = -\pi$

Fig. 10: Numerical results for context no 6

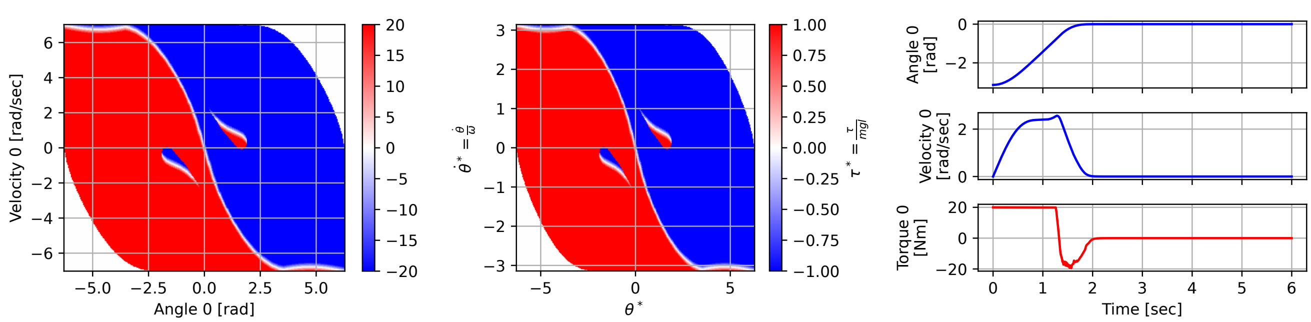


(a) Feedback law f

(b) Dimentionless feedback law f^*

(c) Exemple trajectory starting at $\theta = -\pi$

Fig. 11: Numerical results for context no 7

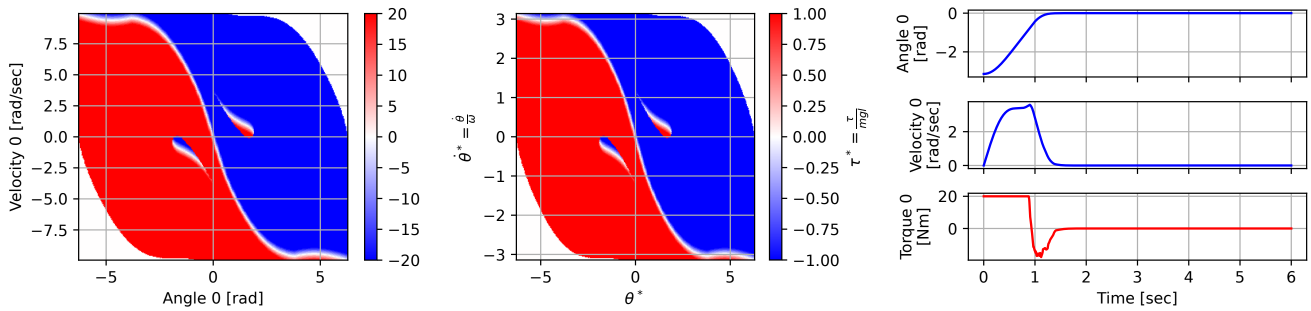


(a) Feedback law f

(b) Dimentionless feedback law f^*

(c) Exemple trajectory starting at $\theta = -\pi$

Fig. 12: Numerical results for context no 8



(a) Feedback law f (b) Dimentionless feedback law f^* (c) Exemple trajectory starting at $\theta = -\pi$

Fig. 13: Numerical results for context no 9

IV. Closed-form parametric policies

To better understand the concept of a dimentionless policy, here we apply the buckingham pi theorem on well-known closed form solution.

A. Computed torque

The computed torque control law provide is a model-based policy that force the system (assuming no torque limits here) on a 2nd order exponential convergence on the desired trajectory:

$$0 = (\ddot{\theta}_d - \ddot{\theta}) + 2\omega_d \zeta (\dot{\theta}_d - \dot{\theta}) + \omega_d^2 (\theta - \theta) \quad (40)$$

For the specific case of the pendulum-swing up, the desired trajectory is simply the up-right position ($\ddot{\theta}_d = \dot{\theta}_d = \theta_d = 0$), and the control law takes this form:

$$\tau = mgl \sin\theta - 2ml^2 \omega_d \zeta \dot{\theta} - ml^2 \omega_d^2 \theta \quad (41)$$

where the only parameters are the system parameters and two variables caracterizing the convergence speed. Hence, the torque policy is a function of those variables:

$$\underbrace{\tau}_{\text{inputs}} = \pi_{ct} \left(\underbrace{\theta, \dot{\theta}}_{\text{states}}, \underbrace{m, g, l}_{\text{system parameters}}, \underbrace{\omega_d, \zeta}_{\text{task parameters}} \right) \quad (42)$$

having the dimension presented at table XXX.

Here $n = 7$ variables are involved and only $p = 2$ independants dimensions (ML^2T^{-2} and T^{-1})

$$m = (n = 7) - (p = 2) = 5 \quad (43)$$

Using mgl and ω, the system parameters, as the repeating variables lead to the following dimentionless groups:

$$\Pi_1 = \tau^* = \frac{\tau}{mgl} \qquad \frac{[ML^2T^{-2}]}{[M][LT^{-2}][L]} \quad (44)$$

$$\Pi_2 = \theta^* = \theta \qquad [-] \quad (45)$$

$$\Pi_3 = \dot{\theta}^* = \frac{\dot{\theta}}{\omega} \qquad \frac{[T^{-1}]}{[T^{-1}]} \quad (46)$$

$$\Pi_4 = \omega_d^* = \frac{\omega_d}{\omega} \qquad \frac{[T^{-1}]}{[T^{-1}]} \quad (47)$$

$$\Pi_5 = \zeta^* = \zeta \qquad [] \quad (48)$$

TABLE IV: Computed torque variables

Variable	Description	Units	Dimensions
Control inputs			
τ	Actuator torque	Nm	$[ML^2T^{-2}]$
State variables			
θ	Joint angle	rad	[]
$\dot{\theta}$	Joint angular velocity	rad/sec	$[T^{-1}]$
System parameters			
mgl	Maximum gravitational torque	Nm	$[ML^2T^{-2}]$
$\omega = \sqrt{\frac{g}{l}}$	Natural frequency	sec^{-1}	$[T^{-1}]$
Policy parameters			
ω_d	Desired closed-loop frequency	sec^{-1}	$[T^{-1}]$
ζ	Desired closed-loop damping	$-$	[ ]

$$\tau^* = \pi_{ct}^* \left(\theta, \dot{\theta}^*, \omega_d^*, \zeta^* \right) \quad (49)$$

Here we can confirme directly, dividing eq (41) by mgl leads to :

$$\tau^* = \sin\theta - 2\omega_d^* \zeta \dot{\theta}^* - (\omega_d^*)^2 \theta \quad (50)$$

B. Linear Quatratic Reglator (LQR) solution

Here we analyse the simplified control problem with the LQR framework. A linearized verison of the equation of motion is used:

$$ml^2\ddot{\theta} - mgl\theta = \tau \tag{51}$$

Also, the same cost function that was used in section III is used to formulate the optimal control problem:

$$J = \int (q^2\theta^2 + 0\,\dot{\theta}^2 + 1\,\tau^2)dt \tag{52}$$

However, here no constraints on the torque are included in the problem. With this problem definition, the same variable as in section III are presents, except the torque limit, see table V. The global policy solution should then have the form:

$$\underbrace{\tau}_{\text{inputs}} = \pi \left(\underbrace{\theta, \dot{\theta}}_{\text{states}}, \underbrace{m, g, l}_{\text{system parameters}}, \underbrace{q}_{\text{task parameters}} \right) \tag{53}$$

We can thus select the same dimentionless group as before,

TABLE V: Pendulum swing-up optimal policy variables

Variable	Description	Units	Dimensions
Control inputs			
τ	Actuator torque	Nm	$[ML^2T^{-2}]$
State variables			
θ	Joint angle	rad	[]
$\dot{\theta}$	Joint angular velocity	rad/sec	$[T^{-1}]$
System parameters			
mgl	Maximum gravitational torque	Nm	$[ML^2T^{-2}]$
$\omega = \sqrt{\frac{g}{l}}$	Natural frequency	sec^{-1}	$[T^{-1}]$
Task parameters			
q	Weight parameter	Nm	$[ML^2T^{-2}]$

and conclude that eq. (53) can be restated under this form:

$$\tau^* = \pi^* \left(\theta, \dot{\theta}^*, q^* \right) \tag{54}$$

For this motion control problem, an analytical solution exist (for the development, see alex screen shot 4 july 2023), and the policy is

$$\tau = [mgl]\,\theta + \left[\sqrt{(mgl)^2 + q^2} \right] \theta + \tag{55}$$

$$\left[\sqrt{2ml^2)} \sqrt{mgl + \sqrt{(mgl)^2 + q^2}} \right] \dot{\theta}^* \tag{56}$$

It is possible to, by this and that.., arrive to this dimentionless form:

$$\tau^* = \left[1 + \sqrt{1 + (q^*)^2} \right] \theta + \left[\sqrt{2} \sqrt{1 + \sqrt{1 + (q^*)^2}} \right] \dot{\theta}^* \tag{57}$$

That correspond to the form predicted by the dimentional analysis of eq. (56).

I am not sure where I am going with this..

V. Conclusion

The concept of dimentionless context is powerful, in the sense that it shows how to transfer a control policy to diferent context where the results should be exactly equivalent. However, it is limited because if the dimentionless context c^* is not exactly equal, then nothing can be deduced regarding if a policy is transferable. Furthermore, the challenge of leveraging this idea is to include all meaningful context variables. If a meaningful variable (in the sense that the policy would be different if its value is changed) is omitted from the context vector c in the dimentional analysis, then the dimentional analysis results might be wrong. On the other hand, if we include too many variables to fully describe a context, then dimentionnaly similar context space will probably be so specific it won't be pratical to use for transfering policy between systems. Henseforth, finding the appropriate parametrization of the context will be critical in order to leverage this principle for sharing policy between similar system,