Dimensionless 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 variables defining the motion control problem, is dimensionally similar. Here we show that by modifying the problem formulation using dimensionless variables, we can re-use the optimal control law generated numerically for a specific system to a sub-space of dimensionally similar systems. This is demonstrated, with numerically generated optimal controllers, for the classic motion control problem of swinging-up a torque-limited inverted pendulum. We also discuss the concept of regime, a region in the space of context variables, that can help relax the condition on dimensional similarity. It remains to be seen if this approach can also help generalizing policies for more complex high-dimensional problems.**

I. Introduction

The state-of-the-art toolbox of control engineers include many numerical algorithms and data-driven schemes. Many control approaches now include a type of mathematical optimization that has no closed-form solution and that is thus solved numerically. More specifically, reinforcement learning (RL), which we can be seen as a data-driven offline optimization, is starting to be a viable option for solving some motion control problems [1]. All in all, numerical tool are very useful and have been used to solve many hard control problem. However, they have a major drawback compared to simpler analytical approaches: parameters of the problem are not appearing explicitly in the solutions, which makes it much harder to generalize and reuse them.

Analytical solutions to control problems have the useful property of allowing the solution to be adjusted to different system parameters by simply substituting the new values in the equation. Most numerical solutions, including RL, usually act has black-boxes with respect to the parameters. For instance, if we use a reinforcement learning (RL) approach to find a good feedback law for a given task with a robot, the solution will be specific to this system. If the RL-generated feedback law is transferred to a robot with a longer arm, there is a good chance it will not behave as intended. With an analytical feedback law solution, we would simply update the value of the length variable in the equation to adjust it, while with a RL solution we would have to re-conduct all the training implying generally multiple hours of data collection and/or computation. It would be a great asset to have the ability to adjust black-box numerical solutions with respect to some problem parameters.

In this paper, we explore the concept of dimensionless policies, a more generic form of knowledge, has a mean to generalize numerical solutions to motion control problems. First, in section II, we use dimensional analysis (i.e. the Buckingham π theorem) to show that motion control problem with dimensionally similar context variables must share the same feedback law solution when expressed in a dimensionless form, and discuss the implications. Then in section III, this is demonstrated for the classical motion control problem of swinging-up an inverted pendulum, using numerically generated optimal feedback laws with a dynamic programming algorithm. Also, in section IV, we illustrate with two examples that the proposed dimensional scaling is equivalent to changing parameters in an analytical solution.

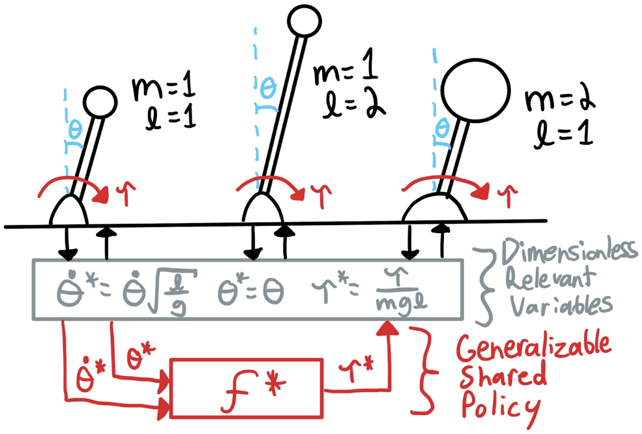

m=1
l=1
θ
m=1
l=2
θ
m=2
l=1
θ
τ
τ
τ
$\dot{\theta}^* = \dot{\theta}\sqrt{\frac{l}{g}}$
$\theta^* = \theta$
$\tau^* = \frac{\tau}{mgl}$
Dimensionless Relevant Variables
$\dot{\theta}^*$
θ^*
f^*
τ^*
Generalizable Shared Policy

Fig. 1: Shared dimensionless policy for various systems

II. Dimensionless Policies

In the following section, we develop the concept of dimensionless policies based on the Buckingham π and show that multiple motion control problems will have the same policy solution when restated in a dimensionless form, if they have a dimensionally similar context.

A. Context variables in the policy mapping

Here, we will call a feedback law a mapping noted 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 feedback law is guarantee to be in this state feedback form [2]. We will only consider this case in the following analysis.

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

To consider the question of how can this system-specific feedback law be transferred in a different context, it is useful to think about a higher dimension mapping, that we will call a policy and note π, also having as additional input arguments, a vector of variables c describing the context:

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

where the context c is the vector of all variables that would affect what is the feedback law solution. We can think about the policy π as the solution to a motion control problem, and c as a vector of parameters in the problem definition. The policy π outputs the control action as a function of the state x, but also of problem parameters. The policy π thus contains the feedback laws for all possible contexts. For example, in section III a case study is conducted considering the optimal feedback law for swinging-up a torque-limited inverted pendulum. For this example, the context variables are the pendulum mass m, the gravitational constant g, the length 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 policy might be different if any of the context variable change value, for instance if the pendulum is heavier, more torque limited, etc.

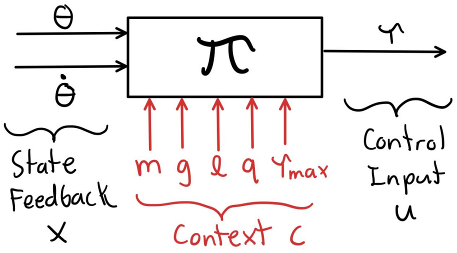


Fig. 2: For the pendulum swing-up control problem, the context includes 5 variables: the parameter of the system: m, g, and l, a parameter of the cost function: q and a parameter defining the constraints: τ_{max}.

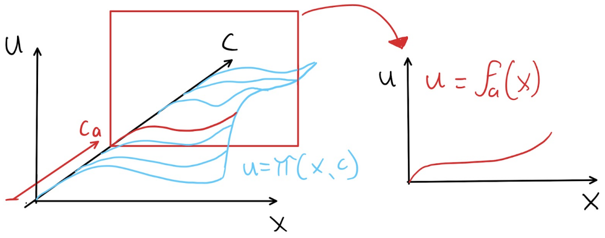


Fig. 3: A feedback law f, is a slice of the higher dimensional policy mapping π for a specific context.

We will use a subscript letter to refer to a specific context, for instance we will note f_a the feedback law solution to a motion problem defined by an instance of context variables c_a. The feedback law f_a is thus a slice of the global policy when the context variables are fixed at values c_a:

$$f_a\left(x\right) = \pi\left(x, c = c_a\right) \tag{3}$$

as illustrated at Fig. 3. Then we can formalize the goal of generalizing a feedback law to a different context: if a feedback law f_a is known for 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 Buckingham π theorem [3], we will show that if the context is dimensionnally similar, then both feedback laws must be equal when restated in dimensionless form. It is important to note that for the following dimensional analysis to hold, we must include all the variables, system parameters and task parameters, that would affect the policy solution π to the control problem in the context vector c:

$$\underbrace{\begin{bmatrix} u_1 \\ \vdots \\ u_k \end{bmatrix}}_{\text{inputs}} = \pi \left(\underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{\text{states}}, \underbrace{\underbrace{\begin{bmatrix} c_1 \\ \vdots \\ \vdots \\ \vdots \end{bmatrix}}_{\text{system}}, \underbrace{\begin{bmatrix} \vdots \\ c_m \end{bmatrix}}_{\text{task}}}_{\text{Context } c} \right) \tag{5}$$

B. Dimensional analysis of the policy mapping

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

$$u_j = \pi_j\left(x_1, \ldots, x_n, c_1, \ldots\ldots, c_m\right) \tag{6}$$

where eq. (6) is the jth line of the policy in vector form described by eq. (5). Then, if the state vector is defined by n variables, and the context is defined by m system plus tasks parameters, then each mapping π_j involves $1 + n + m$ variables. Here, we will assume that the policy is physically meaningful, in the sense of the requirement for applying the Buckingham π theorem [3]. This means for example, that a policy that computes a force based on position and velocity measurements would be in this framework, but not a policy for playing chess for instance.

Applying the Buckingham π theorem to this relationship, tell us that if d dimensions are involved in all those variables, then eq. (6) can be restated into an equivalent relationship between p dimensionless Π groups where $p \geq (1+n+m) - d$. Assuming d dimensions are involved in the m context variables, and that we are in the usual situation where the maximum reduction is possible, i.e. $p = (1 + n + m) - d$, then we can pick d context variables $\{c_1, c_2, \ldots, c_d\}$ as the basis (the repeated variables) to scale all other variables in dimensionless Π groups. We will note dimensionless Π group as the base variable with an $*$:

$$u_j^* = u_j \left[c_1\right]^{e^u_{1j}} \left[c_2\right]^{e^u_{2j}} \ldots \left[c_d\right]^{e^u_{dj}} \quad {\scriptstyle j=\{1,\ldots,k\}} \tag{7}$$

$$x_i^* = x_i \left[c_1\right]^{e^x_{1i}} \left[c_2\right]^{e^x_{2i}} \ldots \left[c_d\right]^{e^x_{di}} \quad {\scriptstyle i=\{1,\ldots,n\}} \tag{8}$$

$$c_i^* = c_i \left[c_1\right]^{e^c_{1i}} \left[c_2\right]^{e^c_{2i}} \ldots \left[c_d\right]^{e^c_{di}} \quad {\scriptstyle i=\{d+1,\ldots,m\}} \tag{9}$$

where exponents e_{ij} are rational numbers selected to make all equations dimensionless. Then, the Buckingham π theorem tell us that the relationship described by eq. (6) can be restated as the following relationship between dimensionless variables:

$$u_j^* = \pi_j^* \left(x_1^*, \ldots, x_n^*, c_{d+1}^*, \ldots, c_{m+l}^*, \right) \tag{10}$$

involving d less dimensionless variables. If we apply the same procedure to all control inputs, when can then assemble the k mappings back into a vector form:

$$\underbrace{\begin{bmatrix} u_1^* \\ \vdots \\ u_k^* \end{bmatrix} = \pi^* \left(\begin{bmatrix} x_1^* \\ \vdots \\ x_n^* \end{bmatrix} \right.}_{\text{Dimensionless feedback law } f^*} , \underbrace{\left. \begin{bmatrix} c_{d+1}^* \\ \vdots \\ c_m^* \end{bmatrix} \right)}_{\text{context } c^*} \tag{11}$$

that we will sometimes write in compact form as:

$$u^* = \pi^*(x^*, c^*) \tag{12}$$

One interesting perk of this dimensional analysis, is that we can remove p variables from 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 feedback policy for all possible contexts is thus simplified in a dimensionless form. Also, an even more interesting feature, that we can use for transferring feedback laws between systems, is that a global policy $\pi(x, c)$, will have an equivalent dimensionless form for multiple context c. As illustrated at Fig. 4, the dimensionless context c^* is in a lower dimensional space $(m - d)$, thus multiple context vector c will correspond to the same dimensionless vector c^*.

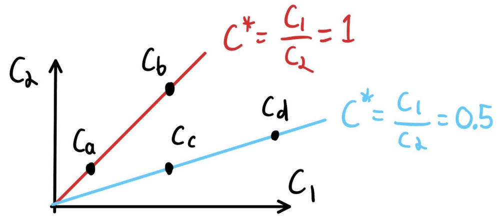


Fig. 4: Dimensionally similar contexts example with dimensions: $m = 2$ and $d = 1$: c_a is dimensionally similar to c_b but not to c_c or c_d.

For a given motion control problem, if the dimensionless context are equal, then the dimensionless feedback law should be exactly equivalent:

$$\text{if} \quad c_a^* = c_b^* \quad \text{then} \quad f_a^*(x^*) = f_b^*(x^*) \; \forall x^* \tag{13}$$

where dimensionless feedback laws are defined as slices of the dimensionless policy for specific contexts:

$$f_a^*(x^*) = \pi^*(x^*, c^* = c_a^*) \tag{14}$$

$$f_b^*(x^*) = \pi^*(x^*, c^* = c_b^*) \tag{15}$$

This is simply based on the fact that the dimensionless policy, i.e. eq. (12), gives the same outputs for the same inputs. This results means that the knowledge of the policy for a specific context c can actually be generalized to a sub-space of all context for which the dimensionless context c^* is equal. For instance, lets imagine we have a global policy for a spherical submarine that depends only on the velocity and the radius. In dimensionless form we would find the policy depends only on the Reynolds number, thus would be equivalent for all pair of velocity and radius that correspond to the same Reynolds number.

C. Transferring policies between contexts

In order to exploit this property, it is useful to define transformation matrices based on scalar equations (7), (8) and (9):

$$u^* = [T_u(c)] \; u \tag{16}$$

$$x^* = [T_x(c)] \; x \tag{17}$$

$$c^* = [T_c(c)] \; c \tag{18}$$

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 Buckingham π theorem). Equations (16) and (17) are inversible (unless a context variable is equal to zero) and can be used to go back-and-forth between dimensional and dimensionless states and inputs variables. The matrix T_c however have d less row than columns and eq. (18) is not inversible: for a given context c there is only one dimensionless context c^*, however a dimensionless context c^* correspond to multiple dimensional context c.

Using the transformation matrices, if a dimensional feedback law f_a for a context c_a is known:

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

its representation in dimensionless form:

$$f_a^*(x^*) = \pi\left(x^*, c^* = c_a^*\right)) \tag{20}$$

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

$$f_a^*(x^*) = T_u(c_a) \underbrace{f_a \left(\underbrace{T_x^{-1}(c_a) \; x^*}_{x} \right)}_{u} \tag{21}$$

Inversely, if we know a dimensionless feedback law f_b^*, matrices T_u and T_x can be used to scale it back to a specific context c_b:

$$f_b(x) = T_u^{-1}(c_b) \underbrace{f_b^* \left(\underbrace{T_x(c_b) \; x}_{x^*} \right)}_{u^*} \tag{22}$$

Thus, eq. (21) and eq. (22) can be used to take any context specific feedback law, finding its dimensionless form, and scale it back to a new context, as illustrated at Fig. 5. In general, there is no guarantee that the behaviour of the scaled feedback law in the new context will be similar to the behaviour of the feedback law in the original context, only if

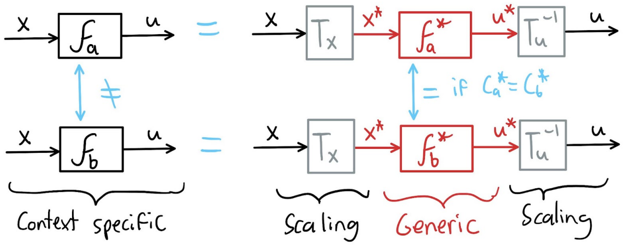


Fig. 5: Isolating the dimensionless knowledge in a policy

the context are dimensionally similar, i.e. the dimensionless context c^* are equal.

Lets suppose an optimal policy solution f_a is known for a specific context c_a, then the scaled policy:

$$f_b(x) = \left[T_u^{-1}(c_b)T_u(c_a)\right] \; f_a\left(\left[T_x^{-1}(c_a)T_x(c_b)\right] \; x\right) \tag{23}$$

will be the optimal solution to the same motion control problem for a context c_b if

$$c_b^* = T_c(c_b)\; c_b = T_c(c_a)\; c_a = c_a^* \tag{24}$$

In some sense, this similar context condition means that the motion problem with parameters c_a and the motion problem with parameter c_b were actually the exact same problem up to scaling factors. It thus make sense that the two solutions should thus also be equivalent up to scaling factors.

In section III, we show examples of this result with numerical solution to dimensionally similar pendulum swing-up problems. Furthermore, we demonstrate that in some situations, eq. (24) equality conditions can be relaxed into inequality conditions, using the concept of regimes.

III. Optimal pendulum swing-up task

In this paper, we will use the classical pendulum swing-up task to test the ideas of dimensionless policies. The motion control problem is formally defined here as finding a feedback law for controlling the dynamic system described by the differential equation:

$$ml^2\ddot{\theta} - mgl\sin\theta = \tau \tag{25}$$

that minimize the quadratic cost function given by:

$$J = \int (q^2\theta^2 + 0\,\dot{\theta}^2 + 1\,\tau^2)dt \tag{26}$$

subject to input constraints given by:

$$-\tau_{max} \leq \tau \leq \tau_{max} \tag{27}$$

Note that here, 1) the cost function parameter q is included with a power of two to have units of torque, 2) it was chosen to set to zero the weight on velocity for simplicity, and 3) the weight multiplying the torque is set to one without loss of generality as only the relative values of weights will impact the optimal solution.

Thus, assuming there is no hidden variables and that equations (25), (26) and (27) fully describe the problem. The solution, i.e. the optimal policy for all context, should be of the form given by:

$$\underbrace{\tau}_{\text{inputs}} = \pi\left(\underbrace{\theta,\dot{\theta}}_{\text{states}}, \underbrace{\underbrace{m,g,l}_{\text{system parameters}} \;,\; \underbrace{q,\tau_{max}}_{\text{task parameters}}}_{\text{Context } c} \right) \tag{28}$$

involving variables listed in table I.

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	$[ML^2T^{-2}]$

Before, conducting the dimensional analysis, it is interesting to note that while there are 3 system parameters m, g and l, they only appear independently in two groups in the dynamic equation. We can thus consider only two system parameters. For convenience we selected mgl, corresponding to the maximum static gravitational torque (i.e. when the pendulum is horizontal) and ω, as listed at table II

TABLE II: Pendulum system parameters

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

A. Dimensional analysis

Here we have one control input, two states, two system parameters and two task parameters, for a total of $1 + (n = 2) + (m = 4) = 7$ variables are involved. In those variables, only $d = 2$ independents 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 dimensionless groups:

$$\Pi_1 = \tau^* = \frac{\tau}{mgl} \qquad \frac{[ML^2T^{-2}]}{[M][LT^{-2}][L]} \tag{29}$$

$$\Pi_2 = \theta^* = \theta \qquad [] \tag{30}$$

$$\Pi_3 = \dot{\theta}^* = \frac{\dot{\theta}}{\omega} \qquad \frac{[T^{-1}]}{[T^{-1}]} \tag{31}$$

$$\Pi_4 = \tau^*_{max} = \frac{\tau_{max}}{mgl} \qquad \frac{[ML^2T^{-2}]}{[M][LT^{-2}][L]} \tag{32}$$

$$\Pi_5 = q^* = \frac{q}{mgl} \qquad \frac{[ML^2T^{-2}]}{[M][LT^{-2}][L]} \tag{33}$$

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

$$\tau^* = \underbrace{[1/mgl]}_{T_u}\, \tau \tag{34}$$

$$\begin{bmatrix} \theta^* \\ \dot{\theta}^* \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1/\omega \end{bmatrix}}_{T_x} \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} \tag{35}$$

$$\underbrace{\begin{bmatrix} q^* \\ \tau^*_{max} \end{bmatrix}}_{c^*} = \underbrace{\begin{bmatrix} 0 & 0 & 1/mgl & 0 \\ 0 & 0 & 0 & 1/mgl \end{bmatrix}}_{T_c} \underbrace{\begin{bmatrix} mgl \\ \omega \\ q \\ \tau_{max} \end{bmatrix}}_{c} \tag{36}$$

According to the theorem, any policy that is only based on the variable included in our analysis can be restated as a relationship between the 5 dimensionless Π groups:

$$\tau^* = \pi^* \left(\theta, \dot{\theta}^*, q^*, \tau^*_{max} \right) \tag{37}$$

The dimensional analysis conducted at sec. II) told us that, for dimensionally similar swing-up problem (which means here equal ratios q^* and τ^*_{max}) the optimal feedback laws should be equivalent in their dimensionless 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 dimensionless form will be equal $f^*_a = f^*_b$ if $q^*_a = q^*_b$ and $\tau^*_{max,a} = \tau^*_{max,b}$. Furthermore, we can thus find f_b using f_a or vice-versa using the scaling formula given by eq. (23) if this condition is met. 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 without additional assumptions.

B. Numerical results

Here, we use a numerical algorithm (we give the details of the methodology at section III-E) to compute numerical solutions to the motion control problem defined by eq (25), (26) and (27). The used numerical recipe produce feedback laws in the form of look-up tables, based on a discretized grid of the state-space. The optimal (up to discretization errors) feedback laws are computed for 9 contexts listed at table III. In those 9 contexts, there are 3 sub-groups of 3 with dimensionally similar contexts. Also each sub-group inlcudes 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 static maximum torque. Contexts 4, 5 and 6 describe a task where the cost highly penalize applying large torques. Contexts 7, 8 and 9 describe a task where the cost highly penalize position errors.

TABLE III: Pendulum swing-up problems parameters

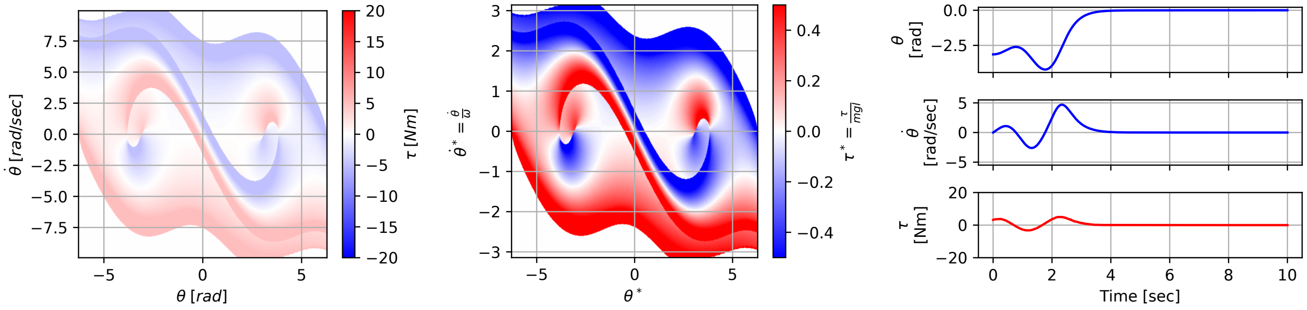
	m	g	l	q	τ_{max}
Problems with $\tau^*_{max} = 0.5$ and $q^* = 0.1$					
Context no 1 :	1.0	10.0	1.0	1.0	5.0
Context no 2 :	1.0	10.0	2.0	2.0	10.0
Context no 3 :	2.0	10.0	1.0	2.0	10.0
Problems with $\tau^*_{max} = 1.0$ and $q^* = 0.05$					
Context no 4 :	1.0	10.0	1.0	0.5	10.0
Context no 5 :	1.0	10.0	2.0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

Figures 6 to 14 illustrate that for each sub-group with equal dimensionless context, the dimensional feedback law generated numerically looks very similar. They are similar up to a scaling of their axis, if we neglect slight differences due to discretization errors. Furthermore, when we compute the dimensionless version of the feedback laws f^*, using eq. (21), the dimensionless version is actually equal within each similar sub-group. This was the expected results predicted by the dimensional analysis of section II.

In terms of how to use this in a practical scenario, we see that if we computed the feedback law given by Fig. 6(a), we can get the feedback law given by Fig. 7(a) or Fig. 8(a) directly by scaling the original policy with eq. (23), using the appropriate context variables, without having to recompute. In some sense, we got back the ability to adjust the feedback law on the fly to new system parameters mgl or ω, as it would be the case with an analytical solution. But this only works within the dimensionally similar context sub-group. The feedback law given by Fig. 6(a) cannot be scaled into the feedback law given by Fig. 9(a) or Fig. 12(a) for instance, since τ^*_{max} and q^* are not equals.

C. Trajectory solutions

Trajectory solutions are also generalizable in their dimensionless forms. Figures 6(c) to 14(c) show optimal trajectory solutions starting from the bottom position at rest. Those trajectories were computed by executing the computed optimal policies in a simulation. We can also see that within the 3 dimensionally similar sub-groups, the optimal trajectories are actually the same but scaled. Lets assume we were trying to

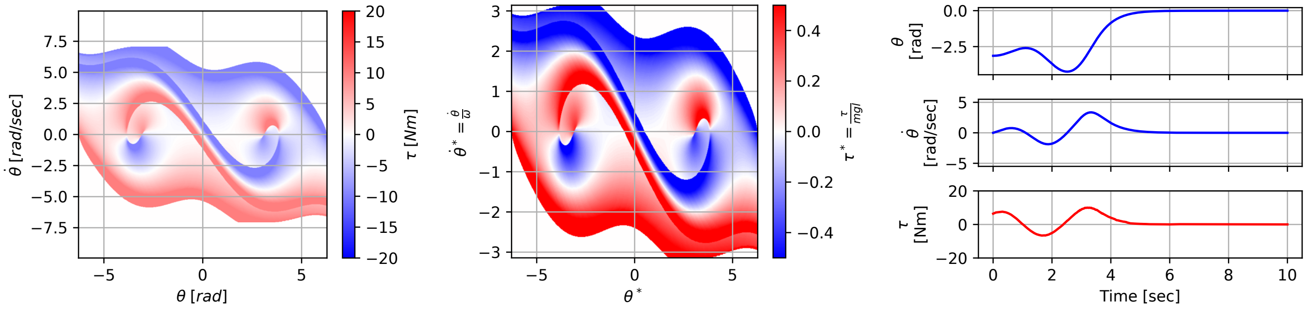


(a) Feedback law f

(b) Dimensionless feedback law f^*

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

Fig. 6: Numerical results for context no 1

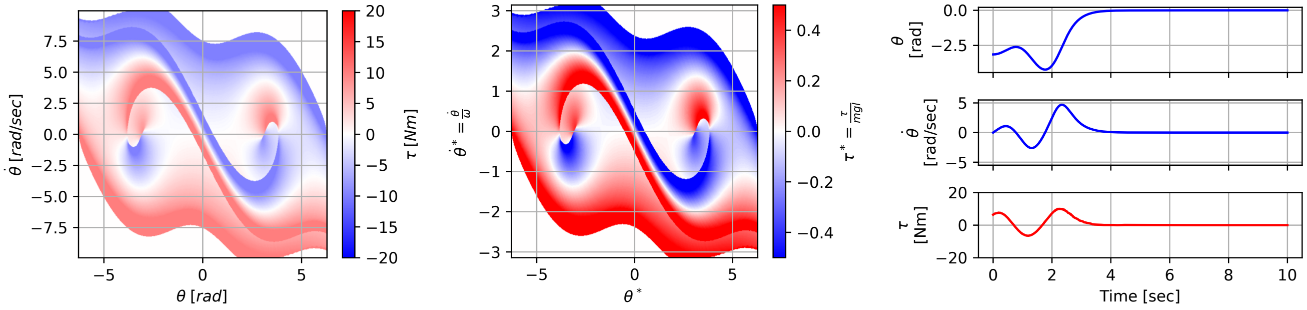


(a) Feedback law f

(b) Dimensionless feedback law f^*

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

Fig. 7: Numerical results for context no 2

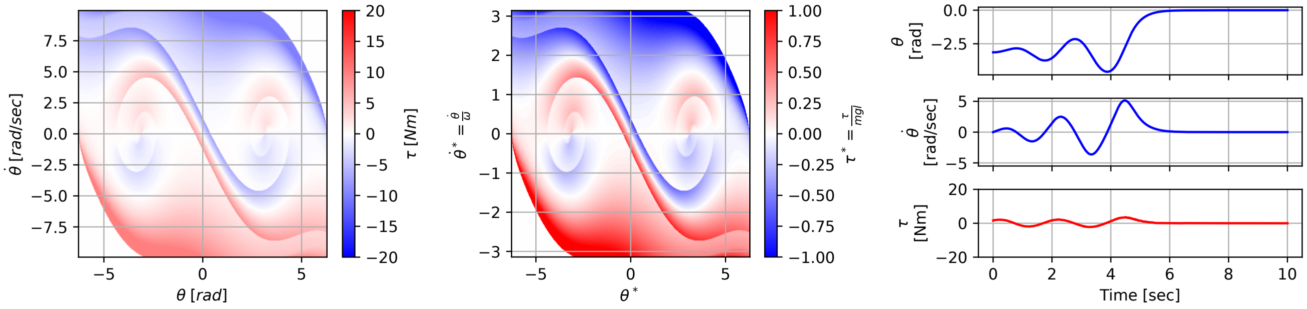


(a) Feedback law f

(b) Dimensionless feedback law f^*

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

Fig. 8: Numerical results for context no 3

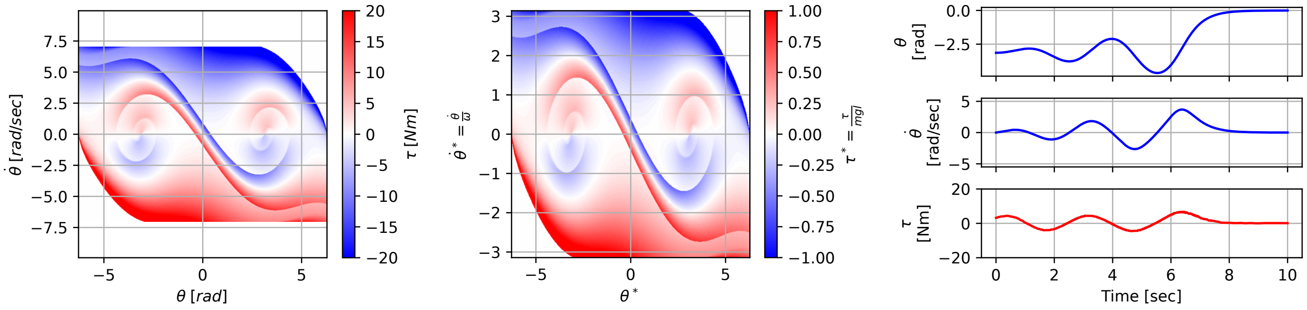


(a) Feedback law f

(b) Dimensionless feedback law f^*

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

Fig. 9: Numerical results for context no 4

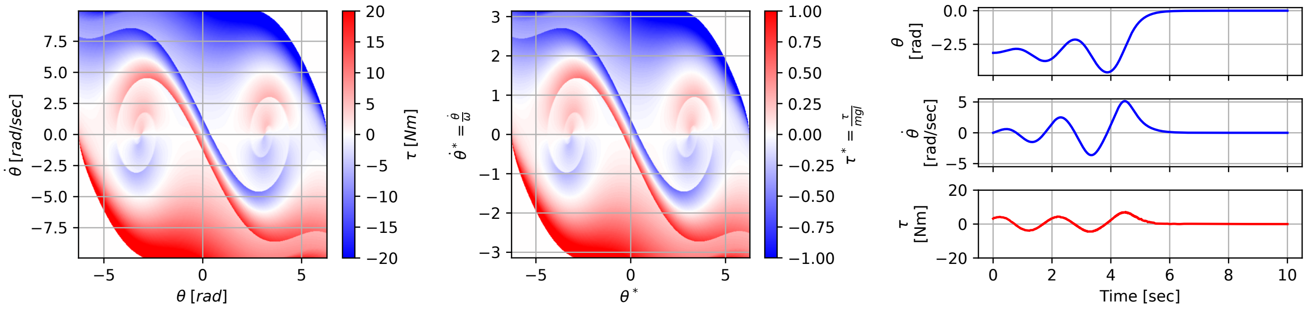


(a) Feedback law f

(b) Dimensionless feedback law f^*

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

Fig. 10: Numerical results for context no 5

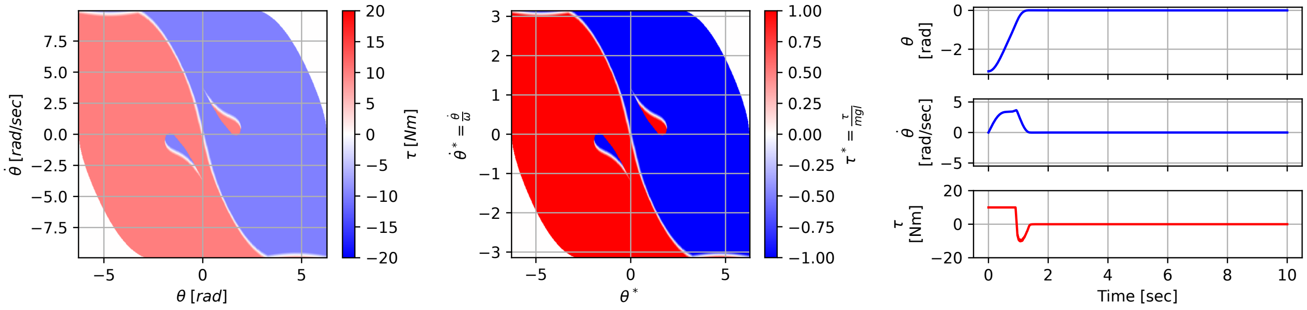


(a) Feedback law f

(b) Dimensionless feedback law f^*

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

Fig. 11: Numerical results for context no 6

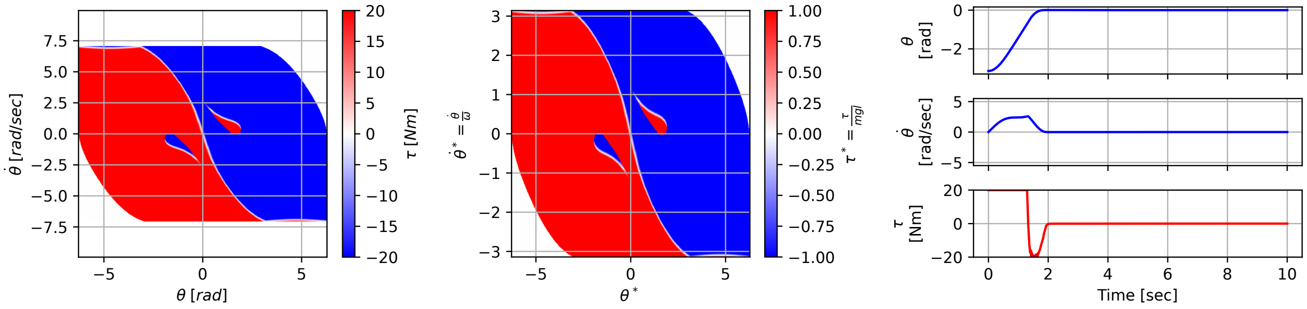


(a) Feedback law f

(b) Dimensionless feedback law f^*

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

Fig. 12: Numerical results for context no 7

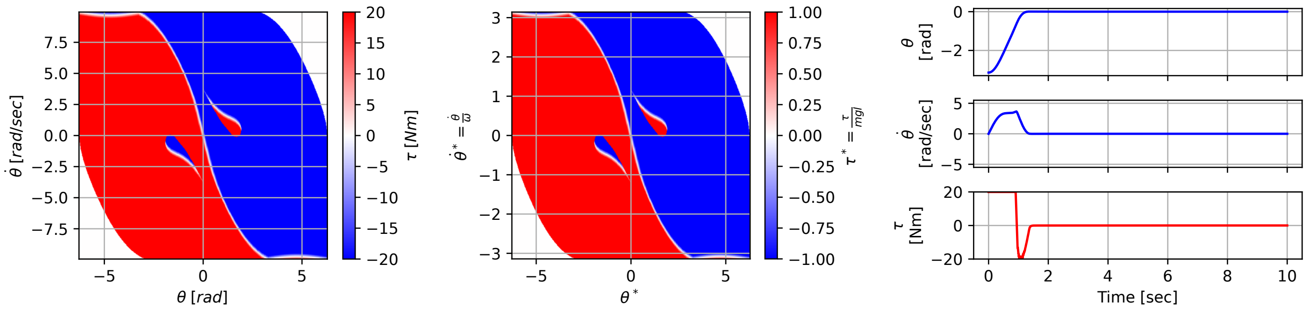


(a) Feedback law f

(b) Dimensionless feedback law f^*

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

Fig. 13: Numerical results for context no 8



(a) Feedback law f

(b) Dimensionless feedback law f^*

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

Fig. 14: Numerical results for context no 9

solve a slightly different motion control problem. Instead of looking for feedback laws, lets assume we were looking for optimal trajectories. We must then think about this motion problem as looking for three time-based policies defining trajectories for states and inputs:

$$\tau = \pi_\tau(t, m, g, l, q, \tau_{max}) \tag{38}$$

$$\theta = \pi_\theta(t, m, g, l, q, \tau_{max}) \tag{39}$$

$$\underbrace{\dot{\theta} = \pi_{\dot{\theta}}(t,}_{\text{trajectory}} \underbrace{m, g, l, q, \tau_{max})}_{\text{context}} \tag{40}$$

Here, we can thus use the same dimensionless group as before, but with the addition of a dimensionless time:

$$t^* = t\,\omega \qquad [T][T^{-1}] \tag{41}$$

since the time t is now explicitly an argument of the mappings. Applying the same dimensional analysis, to eq. (38), eq. (39) and eq. (40), we would conclude that it can be restated as:

$$\tau^* = \pi^*_\tau(t^*, q^*, \tau^*_{max}) \tag{42}$$

$$\theta^* = \pi^*_\theta(t^*, q^*, \tau^*_{max}) \tag{43}$$

$$\underbrace{\dot{\theta}^* = \pi^*_{\dot{\theta}}(t^*,}_{\text{dim. trajectory}} \underbrace{q^*, \tau^*_{max})}_{c^*} \tag{44}$$

Hence, as it is the case for the feedback laws, the optimal trajectories should be equivalent in their dimensionless version, if the dimensionless context is equal.

D. Regimes of solutions

In some situation, changing a context variable will not have any effect on the optimal policy. For instance, for the torque-limited optimal pendulum swing-up problem, augmenting τ_{max} or q while keeping the other value fixed will have little effect pass a given threshold. If we look at the solutions for context no 4, 5 and 6, using a lot of torque is so highly penalized by the cost function that the saturation limit is not really impacting the solution (except edges cases on the boundary), hence we would expect that augmenting τ_{max} should not change the solution.

Fig. 15 and 16 show a slice (to allow visualization) of the optimal policy solution, for various contexts. Fig. 15 illustrates changing τ^*_{max} while keeping q^* fixed. We can see that when $\tau^*_{max} < 0.3$ the policy is almost always on the min-max allowable values, this behaviour is often called *bang-bang*. At the other extreme when $\tau^*_{max} > 2.5$ the policy solution is continuous and almost never affected by the saturation. Fig. 16 illustrates changing q^* while keeping τ^*_{max} fixed. We can see that when $q^* < 0.1$ the optimal policy solution does not reach the min-max saturation, while when $q^* > 1.0$, the policy is almost always on the min-max allowable values.

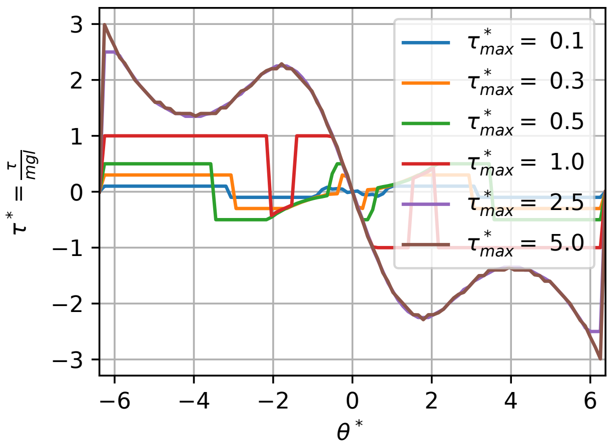


Fig. 15: Optimal dimensionless policy for various contexts $\tau^* = \pi^*(\theta^*, \dot{\theta}^* = 0, q^* = 0.5, \tau^*_{max} = [0.1, ..., 5.0])$

We can see that for extreme context values, we have two type of behaviour, illustrated as region in the dimensionless context space at Fig. 17. Those regions are best caracterized by a ratio of q^* and τ^*_{max}, a new dimensionless value that we will define as the ratio of the maximum torque saturation τ_{max} over the weight in the cost function q:

$$R^* = \frac{\tau^*_{max}}{q^*} = \frac{\tau_{max}}{q} \tag{45}$$

When the value of $R^* \approx 1$, the policy solution is partially continuous, and on the min-max value in some other region of the state-space, a behaviour we will call the transition regime. When the value of $R^* \ll 1$, the constraint on torque

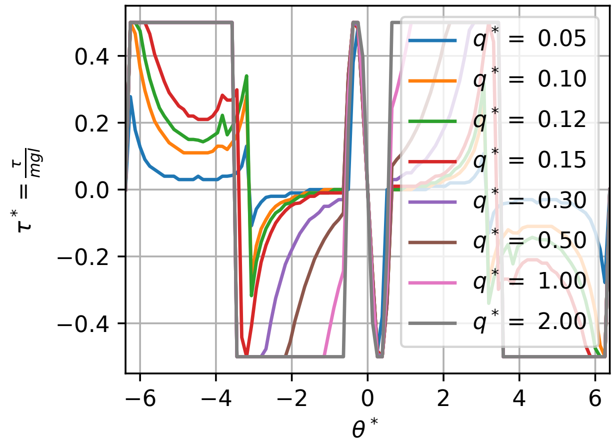


Fig. 16: Optimal dimensionless policy for various contexts $\tau^* = \pi^*(\theta^*, \dot{\theta}^* = 0, q^* = [0.05, ..., 2.0], \tau^*_{max} = 0.5)$

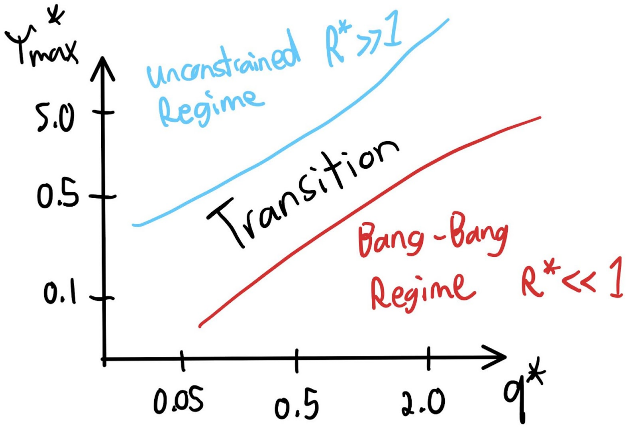


Fig. 17: Regime zones

drives the solution to have have a bang-bang type behaviour. In this region, that we would approximate here based on our sensitivity analysis to $R^* \ll 0.1$, the global policy is only a function of τ^*_{max}:

$$\pi^*(\theta^*, \dot{\theta}^*, q^*, \tau^*_{max}) \approx \pi^*(\theta^*, \dot{\theta}^*, \tau^*_{max}) \text{ if } R^* \ll 1 \quad (46)$$

i.e. the value of q^* is not affecting the solution. On the other hang, when the value of $R^* \gg 1$, the policy is unconstrained. In this region, that we would approximate here based on our sensitivity analysis to $R^* \gg 10$, the global policy is only a function of q^* since the constraint is so far away:

$$\pi^*(\theta^*, \dot{\theta}^*, q^*, \tau^*_{max}) \approx \pi^*(\theta^*, \dot{\theta}^*, q^*) \text{ if } R^* \gg 1 \quad (47)$$

The concept of regime is often leveraged in fluid mechanics, it allow to generalize results between situation where the relevant dimensionless number are not exactly matched. For instance, when the Mach number is small $Ma < 0.3$, we can generally assume to be in an incompressible regime where various speed of sound (for instance having $Ma = 0.1$ or $Ma = 0.3$) would not really change the behaviour. Here, for the purpose of transferring policy solution between contexts, it means that the condition of having the same exact dimensionless context variables can be relaxed with an inequality that correspond to a regime. For instance, for two motion control problems, if both context are in the unconstrained regime, it is sufficient to match only q^* to have equivalent dimensionless policies. More formally, from eq. (47), we can say that:

$$f^*_a(x^*) = f^*_b(x^*) \; \forall \; x^* \quad (48)$$
$$\text{if} \quad q^*_a = q^*_b \; , \quad R^*_a \gg 1 \quad \text{and} \quad R^*_b \gg 1 \quad (49)$$

Also, for two motion control problems, if both context are in the bang-bang regime, it is sufficient to match only τ^*_{max} to have equivalent dimensionless policies. More formally, from eq. (46), we can say that:

$$f^*_a(x^*) = f^*_b(x^*) \; \forall \; x^* \quad (50)$$
$$\text{if} \quad \tau^*_{max,a} = \tau^*_{max,b} \; , \; R^*_a \ll 1 \text{ and } R^*_b \ll 1 \quad (51)$$

Another point of view, is that assuming we are in one of those regime means we could have removed one variable from the context from the start of the dimensional analysis. All in all, the impact of having such regimes identified is that we can increase the sub-space of contexts for which the dimensionless version of the policy should be equivalent, leading to a potentially larger pool of systems that can share learned policy or numerical results.

E. Methodology

We obtained the optimal feedback law by using the basic dynamic programming algorithm [2] on a discretized version of the continuous system. The approach is almost equivalent to the value iteration algorithm [4], sometime refer to as model-based reinforcement learning, with the exception that here the total number of iteration steps was fixed (corresponding to a very long time horizon), instead of stopping the iteration after reaching a convergence criterion. This approach was chosen to have consistent results across all contexts, that lead to wide range of order-of-magnitude cost-to-go solution J. The selected discretization parameters are as follow: the time step is 0.025 sec, the state space is discretized into an even 501 x 501 grid and the continuous torque is discretized into 101 discrete control options. Special out-of-bound and on-target termination states are included to guarantee convergence [2]. The source code is available here: `https://github.com/SherbyRobotics/pyro/blob/dimensionless/dev/dimensionless/cases_master.py` and this google colab page allow reproducing the results: `https://colab.research.google.com/drive/1kf3apyHlf5t7XzJ3uVM8mgDsneVK_63r?usp=sharing`.

1) Additionnal dimensionless parameters for the solver: Using dynamic programming for solving the optimal policy numerically required setting additional parameters that define the domain. Although 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:

$$\theta^*_{max} = \theta_{max} = 2\pi \tag{52}$$

$$\dot{\theta}^*_{max} = \frac{\dot{\theta}_{max}}{\omega} = \pi \tag{53}$$

$$t^*_f = t_f\ \omega = 20 \times 2\pi \tag{54}$$

where θ_{max} is the range of angles 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, and t_f is the time horizon, set to 20 periods of the pendulum using the natural frequency.

IV. Closed-form parametric policies

To better understand the concept of a dimensionless policy, here we apply the Buckingham π theorem on well-known closed form solution to classical motion control problems.

A. Dimensionless Linear Quatratic Reglator (LQR) solution

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

$$ml^2\ddot{\theta} - mgl\theta = \tau \tag{55}$$

and we keep the same quadratic cost function:

$$J = \int (q^2\theta^2 + 0\,\dot{\theta}^2 + 1\,\tau^2)dt \tag{56}$$

However, here no constraints on the torque are included in the problem. With this problem definition, the same variable as before except the torque limit τ_{max} are presents. The global policy solution should then have the form:

$$\underbrace{\tau}_{\text{inputs}} = \pi \left(\underbrace{\theta, \dot{\theta}}_{\text{states}}, \underbrace{\underbrace{m, g, l}_{\text{system parameters}} \quad , \quad \underbrace{q}_{\text{task parameters}}}_{\text{context } c} \right) \tag{57}$$

We can thus select the same dimensionless Π group as before, and conclude that eq. (57) can be restated under this form:

$$\tau^* = \pi^* \left(\theta, \dot{\theta}^*, q^* \right) \tag{58}$$

For this motion control problem, an analytical solution exist (see appendix A), and the optimal policy is:

$$\begin{aligned} \tau = & \left[mgl + \sqrt{(mgl)^2 + q^2} \right] \theta \\ & + \left[\sqrt{2ml^2)} \sqrt{mgl + \sqrt{(mgl)^2 + q^2}} \right] \dot{\theta} \end{aligned} \tag{59}$$

Applying eq. (21) to this feedback law given leads to the dimensionless form, using $G = mgl$ and $H = ml^2$ for shortness:

$$\begin{aligned} \tau^* = & \left[\frac{1}{G} \right] \left[G + \sqrt{G^2 + q^2} \right] \theta \\ & + \left[\frac{1}{G} \right] \left[\sqrt{2H(G + \sqrt{G^2 + q^2})} \right] \left[\omega \dot{\theta}^* \right] \end{aligned} \tag{60}$$

$$\tau^* = \left[1 + \sqrt{\frac{G^2 + q^2}{G^2}} \right] \theta + \left[\sqrt{\frac{2H\omega^2}{G} \frac{G + \sqrt{G^2 + q^2}}{G}} \right] \dot{\theta}^* \tag{61}$$

$$\tau^* = \left[1 + \sqrt{1 + (q^*)^2} \right] \theta + \left[\sqrt{2} \sqrt{1 + \sqrt{1 + (q^*)^2}} \right] \dot{\theta}^* \tag{62}$$

The dimensionless form is only a function of the states and the dimensionless cost parameter, as predicted by eq. (58) based on the dimensional analysis.

We can also use this analytical policy solution to show that scaling the policy with eq. (23) is equivalent to substituting new context variables, when the context are dimentionally similar. Lets say we have two contexts, labelled a and b, and that we use the global policy solution of eq. (59) to have two versions of the context-specific feedback laws:

$$f_a = \left[G_a + \sqrt{G_a^2 + q_a^2} \right] \theta + \left[\sqrt{2H_a(G_a + \sqrt{G_a^2 + q_a^2})} \right] \dot{\theta} \tag{63}$$

$$f_b = \left[G_b + \sqrt{G_b^2 + q_b^2} \right] \theta + \left[\sqrt{2H_b(G_b + \sqrt{G_b^2 + q_b^2})} \right] \dot{\theta} \tag{64}$$

where variables:

$$G_a = m_a g_a l_a \quad H_a = m_a l_a^2 \quad \omega_a = \sqrt{G_a / H_a} \tag{65}$$

$$G_b = m_b g_b l_b \quad H_b = m_b l_b^2 \quad \omega_b = \sqrt{G_b / H_b} \tag{66}$$

represent the values of the parameter groups for each contexts a and b. We can then try to find f_b by scaling f_a using eq. (23):

$$f_b = \left[\frac{G_b}{G_a} \right] f_a \left(\theta, \left[\frac{\omega_a}{\omega_b} \right] \dot{\theta} \right) \tag{67}$$

$$\begin{aligned} f_b = & G_b \left[1 + \sqrt{1 + (q_a^*)^2} \right] \theta \\ & + G_b \left[\sqrt{2} \sqrt{1 + \sqrt{1 + (q_a^*)^2}} \right] \frac{\dot{\theta}}{\omega_b} \end{aligned} \tag{68}$$

where $q_a^* = q_a / G_a$ is the dimensionless version of the cost parameter. Then, distributing G_b lead to:

$$\begin{aligned} f_b = & \left[G_b + \sqrt{G_b^2 + (G_b q_a^*)^2} \right] \theta \\ & + \left[\sqrt{\frac{2G_b}{\omega_b^2}} \sqrt{G_b + \sqrt{G_b^2 + (G_b q_a^*)^2}} \right] \dot{\theta} \end{aligned} \tag{69}$$

$$\begin{aligned} f_b = & \left[G_b + \sqrt{G_b^2 + (G_b q_a^*)^2} \right] \theta \\ & + \left[\sqrt{2H_b} \sqrt{G_b + \sqrt{G_b^2 + (G_b q_a^*)^2}} \right] \dot{\theta} \end{aligned} \tag{70}$$

which is equivalent to eq. (64), if

$$G_b q_a^* = q_b \quad \text{or equivalently} \quad q_a^* = q_b^* \tag{71}$$

which is the condition of having equal dimensionless context $c_a^* = c_b^*$ for this motion control problem. This example illustrates that applying the scaling of eq. (23) based on the dimensional analysis framework, is equivalent to changing the context variables in an analytical solution, when dimensionless context variables are equal.

B. Dimensionless Computed torque

The computed torque feedback law is a model-based policy (assuming no torque limits here), that is the solution to the motion control problem of making a mechanical system converging on a desired trajectory, with a specified 2nd order exponential time profile defined by

$$0 = (\ddot{\theta}_d - \ddot{\theta}) + 2\omega_d\zeta(\dot{\theta}_d - \dot{\theta}) + \omega_d^2(\theta - \theta) \tag{72}$$

For the specific case of the pendulum-swing up, the desired trajectory is simply the up-right position ($\ddot{\theta}_d = \dot{\theta}_d = \theta_d = 0$), leaving only two parameter defining the tasks: ω_d and ζ. Then, the computed torque policy takes this form:

$$\tau = mgl\sin\theta - 2ml^2\omega_d\zeta\dot{\theta} - ml^2\omega_d^2\theta \tag{73}$$

where the context includes the system parameters and two variables caracterizing the convergence speed. Hence, the dimensional global policy is a function of those variables:

$$\underbrace{\tau}_{\text{inputs}} = \pi_{ct}\left(\underbrace{\theta, \dot{\theta}}_{\text{states}}, \underbrace{\underbrace{m, g, l}_{\text{system parameters}} \quad , \quad \underbrace{\omega_d, \zeta}_{\text{task parameters}}}_{\text{context } c} \right) \tag{74}$$

Note that here, the task parameters define directly the desired behaviour as opposed to the previous examples where they were defining the behaviour indirectly thought a cost function. The states, control inputs and system parameters are the same as before, only the task parameter differ, having the dimension presented at table 74.

TABLE IV: Computed torque task variables

Variable	Description	Units	Dimensions
	Task parameters		
ω_d	Desired closed-loop frequency	sec^{-1}	$[T^{-1}]$
ζ	Desired closed-loop damping	$-$	[ ]

Here, 7 variables are involved and only $p = 2$ independents dimensions (ML^2T^{-2} and T^{-1}), and 5 dimensionless groups can be formed:

$$p = (1 + (n = 2) + (m = 4)) - (p = 2) = 5 \tag{75}$$

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

$$\Pi_1 = \tau^* = \frac{\tau}{mgl} \qquad \frac{[ML^2T^{-2}]}{[M][LT^{-2}][L]} \tag{76}$$

$$\Pi_2 = \theta^* = \theta \qquad [-] \tag{77}$$

$$\Pi_3 = \dot{\theta}^* = \frac{\dot{\theta}}{\omega} \qquad \frac{[T^{-1}]}{[T^{-1}]} \tag{78}$$

$$\Pi_4 = \omega_d^* = \frac{\omega_d}{\omega} \qquad \frac{[T^{-1}]}{[T^{-1}]} \tag{79}$$

$$\Pi_5 = \zeta^* = \zeta \qquad [] \tag{80}$$

Then, applying the Buckingham π theorem tell us that the computed torque policy can be restated as a relation between the dimensionless variable:

$$\tau^* = \pi_{ct}^* \left(\theta, \dot{\theta}^*, \omega_d^*, \zeta^* \right) \tag{81}$$

Here we can confirm directly since we have an analytical solution: applying eq. (21) to the computed torque feedback law given by eq. (73), leads to the dimensionless form following form:

$$\tau^* = \left[\frac{1}{mgl}\right] \left(mgl\sin\theta - 2ml^2\omega_d\zeta\left(\omega\dot{\theta}^*\right) - ml^2\omega_d^2\theta \right) \tag{82}$$

$$\tau^* = \sin\theta^* - 2\omega_d^*\zeta\dot{\theta}^* - (\omega_d^*)^2\theta^* \tag{83}$$

confirming the structure predicted by eq. (81).

Also, we can again use this example to show that, when the dimensionless context is equal, scaling a policy using eq. (23), is equivalent to substituting new values of the system parameters in the analytical equation. By substituting the variable in the analytical policy solution given by eq (73), we can get two instances of the feedback law:

$$f_a = G_a\sin\theta - 2H_a(\omega_d\zeta)_a\dot{\theta} - H_a(\omega_d^2)_a\theta \tag{84}$$

$$f_b = G_b\sin\theta - 2H_b(\omega_d\zeta)_b\dot{\theta} - H_b(\omega_d^2)_b\theta \tag{85}$$

Now, lets try to get to f_b from f_a using eq. (23):

$$f_b = \left[\frac{G_b}{G_a}\right]\left[G_a\sin\theta - 2H_a(\omega_d\zeta)_a\left[\frac{\omega_a}{\omega_b}\right]\dot{\theta} - H_a(\omega_d^2)_a\theta\right] \tag{86}$$

Then, distributing $1/G_a$ leads to:

$$f_b = G_b\left[\sin\theta - 2\frac{(\omega_d)_a}{\omega_a}\zeta_a\frac{\dot{\theta}}{\omega_b} - \left(\frac{(\omega_d)_a}{\omega_a}\right)^2\theta\right] \tag{87}$$

where the this intermediate step illustrate the dimensionless form in the brackets. Finally, by distributing the G_b factor we can get:

$$f_b = G_b\sin\theta - 2H_b\left(\frac{\omega_b}{\omega_a}(\omega_d)_a\right)\zeta_a\dot{\theta} - H_b\left(\frac{\omega_b}{\omega_a}(\omega_d)_a\right)^2\theta \tag{88}$$

which will be exactly equivalent to eq. (85) (i.e. equivalent to changing the value of context variable from the a to the b version) if:

$$\frac{\omega_b}{\omega_a}(\omega_d)_a = (\omega_d)_b \quad \text{and} \quad \zeta_a = \zeta_b \tag{89}$$

which is the condition of having equal dimensionless context $c_a^* = c_b^*$ for this motion control problem:

$$\frac{(\omega_d)_a}{\omega_a} = \omega_a^* = \omega_b^* = \frac{(\omega_d)_b}{\omega_b} \quad \text{and} \quad \zeta_a^* = \zeta_b^* \tag{90}$$

V. CONCLUSION

The concept of dimensionless 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 dimensionless 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,

APPENDIX

A. LQR analytic solution

In this section, we show that the policy given by eq. (59), is optimal with respect to the LQR problem defined in section IV-A.

We can write the equation of motion given by eq. (55) in state-space form, using $G = mgl$ and $H = ml^2$, as :

$$\frac{d}{dt}\begin{bmatrix}\theta\\ \dot{\theta}\end{bmatrix} = \underbrace{\begin{bmatrix}0 & 1\\ G/H & 0\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}\theta\\ \dot{\theta}\end{bmatrix}}_{x} + \underbrace{\begin{bmatrix}0\\ 1/H\end{bmatrix}}_{B}\underbrace{\begin{bmatrix}\tau\end{bmatrix}}_{u} \tag{91}$$

Then, adapting a solution from [5], if we parameterize the weight matrix of the cost function this way:

$$J = \int_0^\infty \left(x^T \underbrace{\begin{bmatrix}a(a-2G) & 0\\ 0 & b^2 - 2aH\end{bmatrix}}_{Q} x + u^T \underbrace{\begin{bmatrix}1\end{bmatrix}}_{R} u \right) dt \tag{92}$$

the optimal cost-to-go is given by

$$J = x^T \underbrace{\begin{bmatrix}b(a-G) & aH\\ aH & bH\end{bmatrix}}_{S} x \tag{93}$$

and the optimal feedback policy is given by

$$u = -\underbrace{\begin{bmatrix}R^{-1}B^T S\end{bmatrix}}_{K} x = -\underbrace{\begin{bmatrix}a & b\end{bmatrix}}_{K} x \tag{94}$$

This solution can by verifyied by subtituting matrices into the algebraic Riccati equation given by:

$$0 = SA + A^T S - SBR^{-1}B^T S + Q \tag{95}$$

since the problem fit in the framework of the classical infinite horizon LQR result [2]. Then, we can see that the cost function defined in section IV-A, is a special case where $Q_{11} = q^2$ and $Q_{22} = 0$, leading to:

$$q^2 = a(a - 2G) \tag{96}$$

$$0 = b^2 - 2aH \tag{97}$$

Solving for a and b, and keeping the positive solution, leads to

$$a = G + \sqrt{G^2 + q^2} \tag{98}$$

$$b = \sqrt{2aH} = \sqrt{2H(G + \sqrt{G^2 + q^2})} \tag{99}$$

that when substituted into eq. (94) is equal to the policy given by eq. (59) in section IV-A.

REFERENCES

[1] N. Rudin, D. Hoeller, P. Reist, and M. Hutter, "Learning to Walk in Minutes Using Massively Parallel Deep Reinforcement Learning," in *Proceedings of the 5th Conference on Robot Learning*. PMLR, Jan. 2022, pp. 91–100, iSSN: 2640-3498. [Online]. Available: https://proceedings.mlr.press/v164/rudin22a.html

[2] D. P. Bertsekas, *Dynamic Programming and Optimal Control: Approximate Dynamic Programming*, Nashua, NH, 2012.

[3] M. E. Buckingham, "On Physically Similar Systems; Illustrations of the Use of Dimensional Equations," *Physical Review*, Oct. 1914, publisher: American Physical Society (APS). [Online]. Available: https://www.scienceopen.com/document?vid=805fe995-1849-413a-b228-3fe616732290

[4] R. S. Sutton and A. G. Barto, *Reinforcement Learning, second edition: An Introduction*, second edition ed. Cambridge, Massachusetts: Bradford Books, Nov. 2018.

[5] B. HANKS and R. SKELTON, "Closed-form solutions for linear regulator-design of mechanical systems including optimal weighting matrix selection," National Aeronautics and Space Administration, NASA Technical Memorandum 104052, Jan. 1991, _eprint: https://arc.aiaa.org/doi/pdf/10.2514/6.1991-1117. [Online]. Available: https://arc.aiaa.org/doi/abs/10.2514/6.1991-1117