TamedPUMA: safe and stable imitation learning with geometric fabrics

Saray Bakker[1], **Rodrigo Pérez-Dattari**[1], **Cosimo Della Santina**[1],
Wendelin Böhmer[2], **and Javier Alonso-Mora**[1]
1. Department of Mechanical Engineering
2. Department of Electrical Engineering, Mathematics & Computer Science
Delft University of Technology
The Netherlands
s.bakker-7@tudelft.nl

1 Theoretical details on TamedPUMA

1.1 Energization and Finsler energies

In fabrics [1], the concept of energization is used to transform the system $\ddot{\boldsymbol{q}} = \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ into an energy-conserving system. To ensure path-consistency after energization, the system $\ddot{\boldsymbol{q}} = \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ is designed to be *Homogenous of Degree 2 (HD2)* which holds if $\boldsymbol{h}(\boldsymbol{q}, \alpha\dot{\boldsymbol{q}}) = \alpha^2\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ for $\alpha \geq 0$. An energized system that is HD2 follows the same path as the original system and differs only by an acceleration along the direction of motion.

It is common to use a Finsler energy to energize the system, although any Lagrangian can be used. A Finsler energy extends the concept of kinetic energy by enabling the modeling of directionally dependence metric tensors, e.g. the Finsler energy is HD2 in the velocity [2]. Finsler energies have the property that the Hamiltonian $\mathcal{H}_{\mathcal{L}_e}$ associated with the Finsler energy $\mathcal{L}_e$, is the Finsler energy itself. An energized system that conserves a Finsler energy and is path consistent is called a *geometric fabric*.

To transform the original system $\ddot{\boldsymbol{q}} = \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ into a geometric fabric, we find $\bar{\alpha}$ for which the system, $\ddot{\boldsymbol{q}} = energize_{\mathcal{H}_{\mathcal{L}_e}}[\boldsymbol{h}] = \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) + \bar{\alpha}\dot{\boldsymbol{q}}$, conserves the Finsler energy $\mathcal{L}_e$. This *energization* is performed by setting the time-derivative of the Hamilonian to zero,

$$\dot{\mathcal{H}}_{\mathcal{L}_e} = \dot{\boldsymbol{q}}^\top[\boldsymbol{M}_{\mathcal{L}_e}\ddot{\boldsymbol{q}} + \boldsymbol{\xi}_{\mathcal{L}_e}] = 0, \tag{1a}$$

$$= \dot{\boldsymbol{q}}^\top[\boldsymbol{M}_{\mathcal{L}_e}(\boldsymbol{h} + \bar{\alpha}\dot{\boldsymbol{q}}) + \boldsymbol{\xi}_{\mathcal{L}_e}] = 0, \tag{1b}$$

$$\bar{\alpha} = -\frac{\dot{\boldsymbol{q}}^\top\left(\boldsymbol{M}_{\mathcal{L}_e}\boldsymbol{h} + \boldsymbol{\xi}_{\mathcal{L}_e}\right)}{\dot{\boldsymbol{q}}^\top\mathbf{M}_{\mathcal{L}_\mathbf{e}}\dot{\mathbf{q}}}, \tag{1c}$$

where $(\boldsymbol{M}_{\mathcal{L}_e}\ddot{\boldsymbol{q}} + \boldsymbol{\xi}_{\mathcal{L}_e})$ are the Euler-Lagrange equations of $\mathcal{L}_e$ with $\boldsymbol{M}_{\mathcal{L}_e} = \partial^2_{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}}\mathcal{L}_e$ and $\boldsymbol{\xi}_{\mathcal{L}_e} = \partial_{\dot{\boldsymbol{q}}\boldsymbol{q}}\mathcal{L}_e\dot{\boldsymbol{q}} - \partial_{\boldsymbol{q}}\mathcal{L}_e$ [1].

1.2 Theoretical details on the Compatible Potential Method

In the following, a theoretical analysis is provided of the stability and convergence properties related to the Compatible Potential Method (CPM). Firstly, more details are provided regarding the stability and convergence proof for a navigation policy $\boldsymbol{f} = \boldsymbol{f}^{\mathcal{C}}_\theta$ with a compatible potential $\psi(\boldsymbol{q})$ and the required assumptions. This is followed by Section 1.3 providing Figure 1, an illustration of the CPM with its corresponding latent space, task spaces and configuration space, and the compatible potential.

The stability and convergence properties of the CPM are based on Theorem III.5 in [1]. The Theorem reads as follows, where we correct for two typos, replacing $\mathcal{H}$ with $\mathcal{H}_{\mathcal{L}_e}$ and $\gamma\dot{\boldsymbol{q}}$ with γ in the original

description,

$$\ddot{\boldsymbol{q}} = energize_{\mathcal{H}_{\mathcal{L}_e}}[\boldsymbol{h}+\boldsymbol{f}] + \gamma\dot{\boldsymbol{q}} \text{ is replaced by: } \ddot{\boldsymbol{q}} = energize_{\mathcal{H}}[\boldsymbol{h}+\boldsymbol{f}] + \gamma. \tag{2}$$

Theorem III.5 adapted from [1]: *Let* $\text{energize}_{\mathcal{L}_e}[\boldsymbol{h}(\boldsymbol{q},\dot{\boldsymbol{q}})]$ *be a fabric with generator* $\boldsymbol{h}$ *and Finsler energy* $\mathcal{L}_e$*, and let* $\boldsymbol{f}(\boldsymbol{q},\dot{\boldsymbol{q}})$ *be a navigation policy with compatible potential* $\psi(\boldsymbol{q})$*. Denote the total energy by* $\mathcal{H} = \mathcal{L}_e + \psi$*. The system*

$$\ddot{\boldsymbol{q}} = energize_{\mathcal{H}}[\boldsymbol{h}+\boldsymbol{f}] - \beta\dot{\boldsymbol{q}} \tag{3a}$$

$$= energize_{\mathcal{H}_{\mathcal{L}_e}}[\boldsymbol{h}+\boldsymbol{f}] + \gamma \tag{3b}$$

with energy regulator,

$$\gamma(\boldsymbol{q},\dot{\boldsymbol{q}}) = -\left(\frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top M_{\mathcal{L}_e}\dot{\boldsymbol{q}}}\right)\delta\psi - \beta\dot{\boldsymbol{q}}, \tag{4}$$

converges to the zero set of $\boldsymbol{f}$ *for* $\beta > 0$*.*

The proof of Theorem III.5 consists of two parts: (1) Showing that the system in Eq. (3) is energy-decreasing and therefore it results in $\dot{\boldsymbol{q}} \to \boldsymbol{0}$ and $\ddot{\boldsymbol{q}} \to \boldsymbol{0}$ as time goes to infinity. (2) Assuring that when the system converges, it converges to the zero set of the navigation policy $\boldsymbol{f}$.

Step 1: To ensure that the damped system decreases in energy resulting in $\dot{\boldsymbol{q}} = \boldsymbol{0}$ and $\ddot{\boldsymbol{q}} = \boldsymbol{0}$, we first analyze the energy-conservative system, $energize_{\mathcal{H}}[\boldsymbol{h}+\boldsymbol{f}]$.

Step 1a: Let's start with finding α for which the system $\ddot{\boldsymbol{q}} = energize_{\mathcal{H}}[\boldsymbol{h}+\boldsymbol{f}] = \boldsymbol{h}+\boldsymbol{f}+\alpha\dot{\boldsymbol{q}}$ is energy-conservative, i.e. the derivative of the Hamiltonian is zero, $\dot{\mathcal{H}} = 0$. The total energy is a summation of the Hamiltonian associated with the Finsler energy, $\mathcal{H}_{\mathcal{L}_e}$, and the potential energy ψ. The total energy and its derivative therefore become,

$$\mathcal{H} = \mathcal{H}_{\mathcal{L}_e} + \psi, \qquad \dot{\mathcal{H}} = \frac{\partial \boldsymbol{q}}{\partial t}^\top \frac{\partial \mathcal{H}}{\partial \boldsymbol{q}} = \dot{\boldsymbol{q}}^\top[\boldsymbol{M}_{\mathcal{L}_e}\ddot{\boldsymbol{q}} + \boldsymbol{\xi}_{\mathcal{L}_e} + \delta\psi]. \tag{5}$$

To find α for which $\dot{\mathcal{H}} = 0$, we substitute $\ddot{\boldsymbol{q}} = \boldsymbol{h}+\boldsymbol{f}+\alpha\dot{\boldsymbol{q}}$ into the derivative of the Hamiltonian and set this equal to zero,

$$\dot{\mathcal{H}} = \dot{\boldsymbol{q}}^\top[\boldsymbol{M}_{\mathcal{L}_e}\ddot{\boldsymbol{q}} + \boldsymbol{\xi}_{\mathcal{L}_e} + \delta\psi] = 0, \tag{6a}$$

$$= \dot{\boldsymbol{q}}^\top[\boldsymbol{M}_{\mathcal{L}_e}(\boldsymbol{h}+\boldsymbol{f}+\alpha\dot{\boldsymbol{q}}) + \boldsymbol{\xi}_{\mathcal{L}_e} + \delta\psi] = 0, \tag{6b}$$

$$= \alpha(\dot{\boldsymbol{q}}^\top\boldsymbol{M}_{\mathcal{L}_e}\dot{\boldsymbol{q}}) + \dot{\boldsymbol{q}}^\top(\boldsymbol{M}_{\mathcal{L}_e}\boldsymbol{h} + \boldsymbol{\xi}_{\mathcal{L}_e}) + \dot{\boldsymbol{q}}^\top(\boldsymbol{M}_{\mathcal{L}_e}\boldsymbol{f} + \partial\psi) = 0. \tag{6c}$$

$$\alpha = -\left(\frac{\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top\boldsymbol{M}_{\mathcal{L}_e}\dot{\boldsymbol{q}}}\right)[\boldsymbol{M}_{\mathcal{L}_e}(\boldsymbol{h}+\boldsymbol{f}) + \boldsymbol{\xi}_{\mathcal{L}_e}] - \left(\frac{\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top\boldsymbol{M}_{\mathcal{L}_e}\dot{\boldsymbol{q}}}\right)\partial\psi. \tag{6d}$$

The expression for α is substituted into the system $\ddot{\boldsymbol{q}} = \boldsymbol{h}+\boldsymbol{f}+\alpha\dot{\boldsymbol{q}}$, to obtain,

$$\ddot{\boldsymbol{q}} = energize_{\mathcal{H}}[\boldsymbol{h}+\boldsymbol{f}], \tag{7a}$$

$$= \boldsymbol{h}+\boldsymbol{f}+\alpha\dot{\boldsymbol{q}}, \tag{7b}$$

$$= \underbrace{\boldsymbol{h}+\boldsymbol{f}\underbrace{-\left(\frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top\boldsymbol{M}_{\mathcal{L}_e}\dot{\boldsymbol{q}}}\right)[\boldsymbol{M}_{\mathcal{L}_e}(\boldsymbol{h}+\boldsymbol{f}) + \boldsymbol{\xi}_{\mathcal{L}_e}]}_{\bar{\alpha}}}_{energize_{\mathcal{H}_{\mathcal{L}_e}}[\boldsymbol{h}+\boldsymbol{f}]} \underbrace{-\left(\frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top\boldsymbol{M}_{\mathcal{L}_e}\dot{\boldsymbol{q}}}\right)\partial\psi}_{\gamma \text{ with } \beta=0}. \tag{7c}$$

Analyzing Eq. 7, the system can be split into the energized system of $\ddot{\boldsymbol{q}} = \boldsymbol{h}+\boldsymbol{f}$ with the Finsler energy $\mathcal{L}_e$ and corresponding energization vector $\bar{\alpha}$, and the term γ where $\beta = 0$. By adding damping to Eq. (7), we obtain the damped system as represented in Eq. (3). In Step 1b, it is proven that the damped system decreases energy and converges to $\dot{\mathcal{H}} \to 0$, since $\mathcal{H}$ is decreasing and lower bounded, leading to $\dot{\boldsymbol{q}} \to \boldsymbol{0}$ and $\ddot{\boldsymbol{q}} \to \boldsymbol{0}$ as time goes to infinity.

Step 1b: Damping is added to Eq. (7) via $-\beta\dot{\boldsymbol{q}}$ with $\beta > 0$,

$$\ddot{\boldsymbol{q}} = \boldsymbol{h}+\boldsymbol{f}+\alpha\dot{\boldsymbol{q}} - \beta\dot{\boldsymbol{q}}. \tag{8}$$

As the derivative of the Hamiltonian is zero for $\beta = 0$, this leads to the following derivative of the Hamiltonian for the damped system,

$$\dot{\mathcal{H}} = \dot{\boldsymbol{q}}^\top [\boldsymbol{M}_{\mathcal{L}_e} \ddot{\boldsymbol{q}} + \boldsymbol{\xi}_{\mathcal{L}_e} + \delta\psi] - \beta \dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}}, \tag{9a}$$

$$= -\beta \dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}}. \tag{9b}$$

As $\boldsymbol{M}_{\mathcal{L}_e}$ is strictly positive, Eq. (9) is less than zero for all $\dot{\boldsymbol{q}} \neq \boldsymbol{0}$ and zero for $\dot{\boldsymbol{q}} = \boldsymbol{0}$. Since the total energy $\mathcal{H}$ is always decreasing and lower bounded by zero, the rate of the decrease must converge to zero, $\dot{\mathcal{H}} \to 0$, which means that $\dot{\mathcal{H}} = -\beta \dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}} \to 0$ converges to zero, and therefore $\dot{\boldsymbol{q}} \to \boldsymbol{0}$ leading to $\ddot{\boldsymbol{q}} \to \boldsymbol{0}$.

Step 2: The second step is to ensure convergence of the system in Eq. (3) to the zero set of the navigation policy $\boldsymbol{f}$. For the CPM, this ensures that the system in Eq. (2) converges to the zero set of the pulled dynamical system of PUMA ($\ddot{\boldsymbol{q}} = \boldsymbol{f}_\theta^{\mathcal{C}}(\boldsymbol{q}, \dot{\boldsymbol{q}})$, Eq. (6) in Section 3.2) which contains the desired goal. To explore convergence of Eq. (3) over infinite time to the zero set of $\boldsymbol{f}$, we take the limit with $\dot{\boldsymbol{q}}, \ddot{\boldsymbol{q}} \to \boldsymbol{0}$,

$$\ddot{\boldsymbol{q}} = \boldsymbol{h} + \boldsymbol{f} + \alpha \dot{\boldsymbol{q}} - \beta \dot{\boldsymbol{q}}, \tag{10a}$$

$$= \boldsymbol{h} + \boldsymbol{f} - \frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}}} [\boldsymbol{M}_{\mathcal{L}_e}(\boldsymbol{h} + \boldsymbol{f}) + \boldsymbol{\xi}_{\mathcal{L}_e}] - \frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}}} \partial\psi - \beta \dot{\boldsymbol{q}}, \tag{10b}$$

$$= \boldsymbol{h} - \frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}}} [\boldsymbol{M}_{\mathcal{L}_e} \boldsymbol{h} + \boldsymbol{\xi}_{\mathcal{L}_e}] - \beta \dot{\boldsymbol{q}} + \boldsymbol{f} - \frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}}} \left(\boldsymbol{M}_{\mathcal{L}_e} \boldsymbol{f} + \partial\psi\right), \tag{10c}$$

$$\xrightarrow[t\to\infty]{} \boldsymbol{0} = \underbrace{energize_{\mathcal{H}_{\mathcal{L}_e}}[\boldsymbol{h}] - \beta \dot{\boldsymbol{q}}}_{\xrightarrow[t\to\infty]{} \boldsymbol{0}} + \boldsymbol{f} - \frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}}} \left(\boldsymbol{M}_{\mathcal{L}_e} \boldsymbol{f} + \partial\psi\right), \tag{10d}$$

$$\xrightarrow[t\to\infty]{} \boldsymbol{0} = \boldsymbol{f} - \frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}}} \left(\boldsymbol{M}_{\mathcal{L}_e} \boldsymbol{f} + \partial\psi\right). \tag{10e}$$

In the limit, both $\beta \dot{\boldsymbol{q}} \to \boldsymbol{0}$ and $energize_{\mathcal{H}_{\mathcal{L}_e}}[\boldsymbol{h}]$ converge to zero in Eq. (10d). In the following, we will elaborate why the equality in Eq. (10e) requires $\boldsymbol{f} = \boldsymbol{0}$ in the limit.

The fraction $\frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}}}$ in Eq. (10e) has two occurances of $\dot{\boldsymbol{q}}$ both in the numerator and denominator. As $\boldsymbol{M}_{\mathcal{L}_e}$ is positive definite and bounded, the fraction becomes a projection operator in the limit,

$$\lim_{t\to\infty} \frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}}} = \boldsymbol{A} = \frac{\boldsymbol{v}\boldsymbol{v}^\top}{\boldsymbol{v}^\top \boldsymbol{M}_{\mathcal{L}_e} \boldsymbol{v}} \quad \text{where} \quad \boldsymbol{v} = \lim_{t\to\infty} \frac{\dot{\boldsymbol{q}}}{\|\dot{\boldsymbol{q}}\|}. \tag{11}$$

Using the definition for $\boldsymbol{A}$ in Eq. (11), Eq. 10e can be rewritten as,

$$\xrightarrow[t\to\infty]{} \boldsymbol{0} = \underbrace{[\boldsymbol{I} - \boldsymbol{A}\boldsymbol{M}_{\mathcal{L}_e}]\boldsymbol{f}}_{\lambda_1} + \underbrace{\boldsymbol{A}(\delta\psi)}_{\lambda_2}, \tag{12}$$

In the limit, the term $[\boldsymbol{I} - \boldsymbol{A}\boldsymbol{M}_{\mathcal{L}_e}]$ has nullspace $\boldsymbol{v}$ as $[\boldsymbol{I} - \boldsymbol{A}\boldsymbol{M}_{\mathcal{L}_e}]\boldsymbol{v} = \boldsymbol{0}$,

$$[\boldsymbol{I} - \boldsymbol{A}\boldsymbol{M}_{\mathcal{L}_e}]\boldsymbol{v} = \boldsymbol{v} - \frac{\boldsymbol{v}\boldsymbol{v}^\top}{\boldsymbol{v}^\top \boldsymbol{M}_{\mathcal{L}_e} \boldsymbol{v}} \boldsymbol{M}_{\mathcal{L}_e} \boldsymbol{v} = \boldsymbol{v} - \boldsymbol{v} \frac{\boldsymbol{v}\boldsymbol{M}_{\mathcal{L}_e}\boldsymbol{v}^\top}{\boldsymbol{v}^\top \boldsymbol{M}_{\mathcal{L}_e} \boldsymbol{v}} = \boldsymbol{v} - \boldsymbol{v} = \boldsymbol{0}, \tag{13}$$

which implies that λ_1 and λ_2 in Eq. (12), are orthogonal, $\lambda_1 \perp \lambda_2$. Both terms must be zero, $\lambda_1 = \boldsymbol{0},\ \lambda_2 = \boldsymbol{0}$ for Eq. (12) to hold. By contradiction, it is proven that $\boldsymbol{f}$ is equal to zero as time goes to infinity.

Proof by contradiction: Let us assume that $\boldsymbol{f} \neq \boldsymbol{0}$. First, note that if $\boldsymbol{f} \neq \boldsymbol{0}$, for λ_1 to be zero, $\boldsymbol{f}$ must be in the nullspace of $[\boldsymbol{I} - \boldsymbol{A}\boldsymbol{M}_{\mathcal{L}_e}]$, which we have already noted is equal to $\boldsymbol{v}$. As a consequence, $\boldsymbol{f} \in \text{span}(\boldsymbol{v})$. In contrast, for λ_2 to be equal to zero, i.e., $\boldsymbol{A}(\partial\psi) = \boldsymbol{0}$, two cases exist:

1. **Case 1**: The gradient of the potential is equal to zero, $\partial\psi = \boldsymbol{0}$. For $\boldsymbol{f} \neq \boldsymbol{0}$, this case cannot hold, as the potential is a compatible potential of $\boldsymbol{f}$, which indicates that $\partial\psi = \boldsymbol{0}$ if only if $\boldsymbol{f} = 0$, which would lead to a contradiction.

2. **Case 2**: The other possibility is that $\partial\psi$ is in the nullspace of $\boldsymbol{A}$, which implies that $\partial\psi \perp \boldsymbol{v}$. Consequently, since $\boldsymbol{f} \in \text{span}(\boldsymbol{v})$, $\partial\psi \perp \boldsymbol{v}$ implies that $\partial\psi \perp \boldsymbol{f}$; hence, $\partial\psi^\top \boldsymbol{f} = 0$. However, a compatible potential also has the property $-\partial\psi^\top \boldsymbol{f} > \boldsymbol{0}$ wherever $\boldsymbol{f} \neq \boldsymbol{0}$, which once again leads to a contradiciton.

As a result, since both possible cases lead to a contradiction for $\boldsymbol{f} \neq 0$, $\boldsymbol{f}$ must be zero in the limit.

1.2.1 Discussion

Let's take a closer look at the implications of the assumptions on $\boldsymbol{M}_{\mathcal{L}_e}$. For a fabric describing collision avoidance, two cases exist as the Spec describing the fabric must be boundary conforming [3]: (1) The metric $\boldsymbol{M}_{\mathcal{L}_e}$ is finite along the Eigen-directions parallel with the boundary's tangent space but goes to infinite along directions orthogonal to the tangent space. (2) The metric $\boldsymbol{M}_{\mathcal{L}_e}$ is a finite matrix along all trajectories, implying that $\boldsymbol{M}_{\mathcal{L}_e}$ is also finite in the limit when $t \to \infty$. Observing the assumption on $\boldsymbol{M}_{\mathcal{L}_e}$ in the Compatible Potential Method (CPM) that the metric is bounded in a finite region and strictly positive definite everywhere, only fabrics can be designed following the second case. This implies that in the limit we ensure convergence to the zero set of the forcing policy $\boldsymbol{f}_\theta^{\mathcal{C}}(\boldsymbol{q}, \dot{\boldsymbol{q}})$, but collision avoidance is not guaranteed as barrier-like functions going to infinity on the boundary, cannot be used to construct $\boldsymbol{M}_{\mathcal{L}_e}$

The metric $\boldsymbol{M}_{\mathcal{L}_e}$ is additionally assumed to not vanish in the limit $\dot{\boldsymbol{q}} \to \boldsymbol{0}$. In practice, numerical instability of the fraction $\frac{\dot{\boldsymbol{q}}\dot{\boldsymbol{q}}^\top}{\dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}}}$ is avoided by replacing the denominator by $\dot{\boldsymbol{q}}^\top \boldsymbol{M}_{\mathcal{L}_e} \dot{\boldsymbol{q}} + \epsilon$ with $\epsilon > 0$.

1.3 Compatible potentials for TamedPUMA

In Section 3.3, the Compatible Potential Method (CPM) is discussed which is extended with a detailed theoretical analysis of the stability and convergence properties in Appendix 1.2. In Fig. 1, we provide an illustration of the different components of the CPM and its compatible potential function. The potential in the latent space, $\psi = \|\boldsymbol{y}_g - \boldsymbol{y}\|^2$ can be expressed in the task space variables $\boldsymbol{x}$ by the encoder ρ_θ. The energy regulator γ of the CPM, Eq. (2), is defined in the configuration space and requires the partial derivative of the potential function with respect to $\boldsymbol{q}$, $\partial_{\boldsymbol{q}}\psi$, as explored in Eq. 10 (Section 3.3 in the paper).

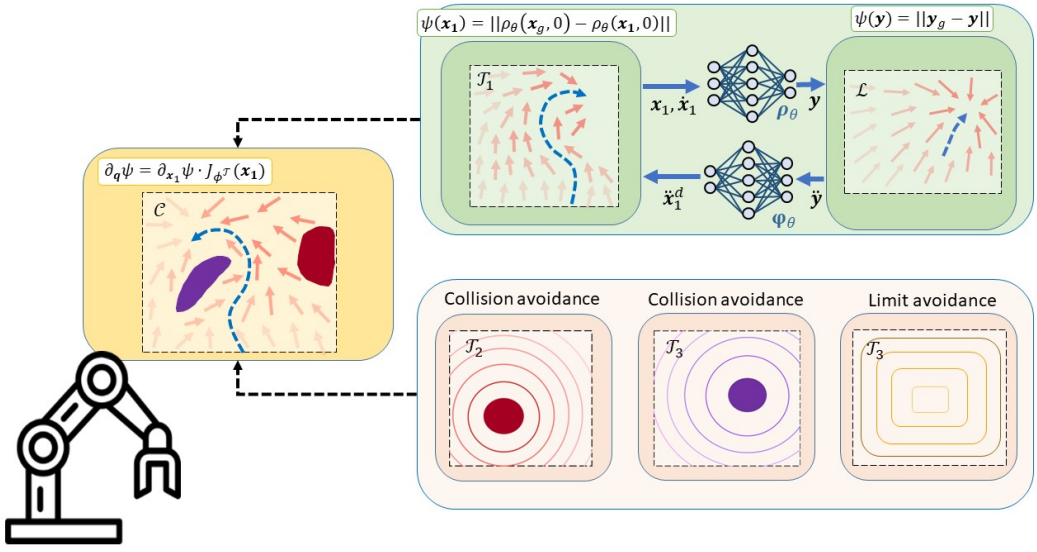

Figure 1: Illustration of TamedPUMA with the compatible potential for the latent space $\mathcal{L}$ towards the task space $\mathcal{T}$ and configuration space $\mathcal{C}$.

2 Differential mappings

2.1 A differential map towards the end-effector pose

The differentiable map $\phi_{\text{ee}}(\boldsymbol{q})$ connecting the joint configuration $\boldsymbol{q}$ and end-effector pose $\boldsymbol{x}_{\text{ee}}$, is obtained via the forward kinematics,

$$\boldsymbol{x}_{\text{ee}}(\boldsymbol{q}) = [\mathbf{x}_{\text{ee}}, \boldsymbol{\zeta}_{\text{ee}}]^\top = \text{FK}(\boldsymbol{q}). \tag{14}$$

given the end-effector position $\mathbf{x}_{\text{ee}}$ and orientation $\boldsymbol{\zeta}_{\text{ee}} = [\zeta_0, \check{\boldsymbol{\zeta}}]$ in quaternions. The end-effector Jacobian provides the first-order derivative of this mapping,

$$\dot{\boldsymbol{x}}_{\text{ee}} = \boldsymbol{J}_{\text{ee}}\dot{\boldsymbol{q}} = \begin{bmatrix} \boldsymbol{I}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \boldsymbol{E}^{-1}(\boldsymbol{\zeta}_{\text{ee}}) \end{bmatrix} \boldsymbol{J}_{\text{ee,p}}\dot{\boldsymbol{q}}, \qquad \boldsymbol{E} = 2\begin{bmatrix} -\check{\boldsymbol{\zeta}} & [\check{\boldsymbol{\zeta}}]_\times + \zeta_0 \boldsymbol{I}_{3\times 3} \end{bmatrix}, \tag{15}$$

where $\boldsymbol{J}_{\text{ee,p}}$ is the positional part of the Jacobian and $\boldsymbol{I}$ the identity matrix. The second-order derivative of the differential map is given by $\ddot{\boldsymbol{x}}_{\text{ee}} = \boldsymbol{J}_{\text{ee}}\ddot{\boldsymbol{q}} + \dot{\boldsymbol{J}}_{\text{ee}}\dot{\boldsymbol{q}}$.

2.2 A configuration space mapping of PUMA

A deep neural network (DNN) is trained on a small set of demonstrations, resulting in the second-order dynamical system $\ddot{\boldsymbol{x}} = \boldsymbol{f}_\theta^{\mathcal{T}}(\boldsymbol{x}, \dot{\boldsymbol{x}})$. Since the inputs and outputs of the DNN are normalized and transformed, some operations are required to obtain the correct inputs to the DNN as well as to transform the output of the DNN to a desired end-effector acceleration of the robotic system. First, the end-effector pose is normalized given the limits of the workspace, $\mathbf{x}_{min}$ and $\mathbf{x}_{max}$, followed by a transformation with respect to the goal position, $\mathbf{x}_{\text{offset}}$, and orientation, $\boldsymbol{\zeta}_{\text{offset}}$, to obtain the input $\boldsymbol{x}$ to the DNN,

$$\boldsymbol{x}_{\text{NN}} = \sigma_{\boldsymbol{x}}(\boldsymbol{q}) = \begin{bmatrix} 2\left(\frac{\mathbf{x}_{\text{ee}}(\boldsymbol{q}) - \mathbf{x}_{\text{min}}}{\mathbf{x}_{\text{max}} - \mathbf{x}_{\text{min}}} - \frac{1}{2}\right) - \mathbf{x}_{\text{offset}} \\ \boldsymbol{\zeta}_{\text{ee}}(\boldsymbol{q}) \otimes \boldsymbol{\zeta}_{\text{offset}}^{-1} \end{bmatrix} \tag{16}$$

A similar reasoning is done for the end-effector velocities using time-step Δt and the Jacobian $\boldsymbol{J}_{\text{ee}}$ as in Eq. (15),

$$\dot{\boldsymbol{x}}_{\text{NN}} = \sigma_{\dot{\boldsymbol{x}}}(\boldsymbol{q}) = \begin{bmatrix} 2\left(\frac{\Delta t \cdot \dot{\mathbf{x}}_{\text{ee}}(\boldsymbol{q}) - \dot{\mathbf{x}}_{\text{min}}}{\dot{\mathbf{x}}_{\text{max}} - \dot{\mathbf{x}}_{\text{min}}} - \frac{1}{2}\right) \\ \dot{\boldsymbol{\zeta}}_{\text{ee}}(\boldsymbol{q}) \otimes \boldsymbol{\zeta}_{\text{offset}}^{-1} \end{bmatrix} \tag{17}$$

The reverse operations are performed to obtain the desired velocity and acceleration of the end-effector from the output of the DNN,

$$\dot{\boldsymbol{x}}_{\text{ee}}^{\text{d}} = \sigma_{\dot{\boldsymbol{x}}}^{-1}(\dot{\boldsymbol{x}}_{\text{NN}}) \approx \begin{bmatrix} \frac{1}{\Delta t} \cdot \dot{\mathbf{x}}_{\text{NN}} \\ \left(\frac{1}{\Delta t} \cdot \boldsymbol{\zeta}_{\text{NN}}\right) \otimes \boldsymbol{\zeta}_{\text{offset}} \end{bmatrix}, \quad \ddot{\boldsymbol{x}}_{\text{ee}}^{d} = \sigma_{\ddot{\boldsymbol{x}}}^{-1}(\ddot{\boldsymbol{x}}_{\text{NN}}) \approx \begin{bmatrix} \frac{1}{\Delta t^2} \cdot \ddot{\mathbf{x}}_{\text{NN}} \\ \left(\frac{1}{\Delta t^2} \cdot \boldsymbol{\zeta}_{\text{NN}}\right) \otimes \boldsymbol{\zeta}_{\text{offset}} \end{bmatrix} \tag{18}$$

Using the pseudo-inverse of the Jacobian $\boldsymbol{J}_{\text{ee}}^\dagger$ as stated in Eq. 15, the desired velocity and acceleration in configuration space becomes,

$$\dot{\boldsymbol{q}}^{\text{d}} = \boldsymbol{J}_{\text{ee}}^\dagger \dot{\boldsymbol{x}}_{\text{ee}}^{\text{d}}, \qquad \ddot{\boldsymbol{q}}^{\text{d}} = \boldsymbol{f}_\theta^{\mathcal{C}}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \boldsymbol{J}_{\text{ee}}^\dagger \left(\ddot{\boldsymbol{x}}_{\text{ee}}^{\text{d}} - \dot{\boldsymbol{J}}_{\text{ee}}\boldsymbol{J}_{\text{ee}}^\dagger \dot{\boldsymbol{x}}_{\text{ee}}^{\text{d}}\right). \tag{19}$$

To the function $\ddot{\boldsymbol{q}}^{\text{d}} = \boldsymbol{f}_\theta^{\mathcal{C}}(\boldsymbol{q}, \dot{\boldsymbol{q}})$, we will refer as the configuration space mapping of Policy via neUral Metric leArning (PUMA) obtained via pullback operation $\text{pull}_{\phi^{\mathcal{T}}}\left(\boldsymbol{f}_\theta^{\mathcal{T}}(\boldsymbol{x}, \dot{\boldsymbol{x}})\right)$ as described in this section. As the KUKA LBR iiwa has seven degrees of freedom while a pose specifies six degrees of movement, a pseudo-inverse is required to approximate the Jacobian. To resolve this issue, a null-space controller is proposed with as its primary objective, end-effector pose reaching and as a secondary objective, the placing the elbow joint in an upwards position.

3 Algorithm description

Algorith 1 provides a short description the TamedPUMA at time-step t. The parameters of the joint impedance controller can be found on the website. Torques, $\boldsymbol{\tau}$, are send to the robot at 1000Hz.

Algorithm 1: TamedPUMA

Input : $(\boldsymbol{q}, \dot{\boldsymbol{q}})$, ▷ *Current joint-state of robot*
$\boldsymbol{x}_{\text{obsts}}, \boldsymbol{x}_g$ ▷ *Obstacles' and target poses*

$\boldsymbol{x}_j = \phi_j(\boldsymbol{q}),\ \forall j \in [M]$, and $\boldsymbol{x}_{\text{ee}} = \text{FK}(\boldsymbol{q})$ ▷ *Forward mappings, Sec. 2*

$\dot{\boldsymbol{x}}_j = \boldsymbol{J}_{\phi_j}(\boldsymbol{q})\dot{\boldsymbol{q}},\ \forall j \in [M]$, and $\dot{\boldsymbol{x}}_{\text{ee}} = \boldsymbol{J}_{\text{ee}}(\boldsymbol{q})\dot{\boldsymbol{q}}$

$\ddot{\boldsymbol{x}}_j = \boldsymbol{h}(\boldsymbol{x}_{\text{j}}, \dot{\boldsymbol{x}}_{\text{j}}),\ \forall j \in [M]$ ▷ *Fabrics in $\mathcal{X}_j$*

$\ddot{\boldsymbol{x}}_{\text{ee}} = \boldsymbol{f}_\theta^{\mathcal{T}}(\boldsymbol{x}_{\text{ee}}, \dot{\boldsymbol{x}}_{\text{ee}})$ ▷ *DNN action*

$\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \sum_{j=1}^{M} \boldsymbol{h}_j(\boldsymbol{q}, \dot{\boldsymbol{q}}),\ \tilde{\boldsymbol{h}}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \text{energize}_{\mathcal{L}_e}[\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})]$ ▷ *Combined Fabric*

$\boldsymbol{f}_\theta^{\mathcal{C}}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \text{pull}_{\phi_{\mathcal{T}}}(\ddot{\boldsymbol{x}}_{\text{ee}})$ ▷ *DNN pullback, Sec. 2*

$\ddot{\boldsymbol{q}}^{\text{d}} = \tilde{\boldsymbol{h}}(\boldsymbol{q}, \dot{\boldsymbol{q}}) + \boldsymbol{f}_\theta^{\mathcal{C}}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ ▷ *FPM*

or

$\ddot{\boldsymbol{q}}^{\text{d}} = \text{energize}_{\mathcal{H}}[\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) + \boldsymbol{f}_\theta^{\mathcal{C}}(\boldsymbol{q}, \dot{\boldsymbol{q}})] + \gamma(\boldsymbol{q}, \dot{\boldsymbol{q}})$ ▷ *CPM*

$\boldsymbol{\tau} = joint\_impedance\_control(\boldsymbol{q}, \dot{\boldsymbol{q}}, \ddot{\boldsymbol{q}}^d)$

References

[1] N. Ratliff and K. Van Wyk. Fabrics: A foundationally stable medium for encoding prior experience. *arXiv preprint:2309.07368*, 2023.

[2] N. D. Ratliff, K. Van Wyk, M. Xie, A. Li, and M. A. Rana. Optimization fabrics. *arXiv preprint arXiv:2008.02399*, 2020.

[3] N. D. Ratliff, J. Issac, D. Kappler, S. Birchfield, and D. Fox. Riemannian motion policies. *arXiv preprint arXiv:1801.02854*, 2018.