Untangling Knots Through Curve Repulsion

Department of Mathematics
University of Oxford
Trinity Term 2022

Untangling Knots Through Curve Repulsion

Abstract

Curves are one of the fundamental objects in geometry and engineering, yet most analysis of curves often disregard their physical characteristics such as their spatial volume or uncrossability. One common situation that such physical characteristics become significant is when one attempts to untangle a knot. An approach to achieve this is to assign an "energy" to a curve such that this energy would increase when two points on "different sides" of a curve are closer, then one continuously deforms the curve to reduce this energy, the expectation being that the curve that achieves minimal energy must be the untangled knot. This dissertation explores numerical methods of achieving this.

Contents

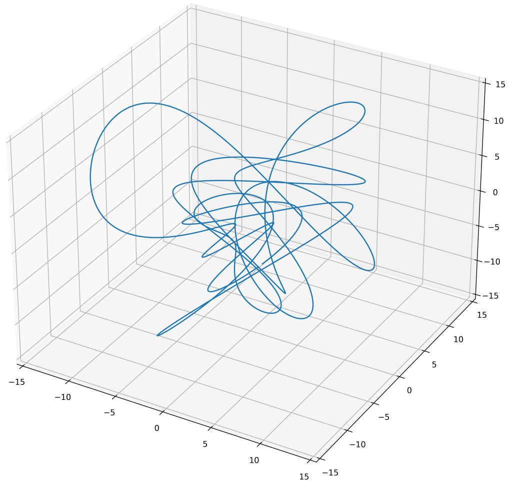


Figure 1.1: A tangled curve in $\mathbb{R}^3$

1 Introduction

Shape optimisation is an important idea in engineering, relevant in anywhere from aircraft designs to packaging ramen noodle. One of the simplest and most fundamental shapes to consider is a curve. While curves are simple objects in theory, they prove to be quite difficult to analyse in practice with realistic physics. Even in absence of other objects, one must consider resiliance to bending, stretching, and especially, impenetrability against itself. With these physical factors in mind, untangling a knot like the one shown (Figure 1.1) becomes a very complicated process, especially in a computer simulation. In this dissertation, we explore numerical methods to achieve this.

The main idea is to *assign energy that penalises "physical entanglement"*. Given a parameterised curve $\boldsymbol{\gamma} : M \to \mathbb{R}^3$ (M being the domain of the parameter, often an interval), one defines some *curve energy* $\mathcal{E}$ of the form:

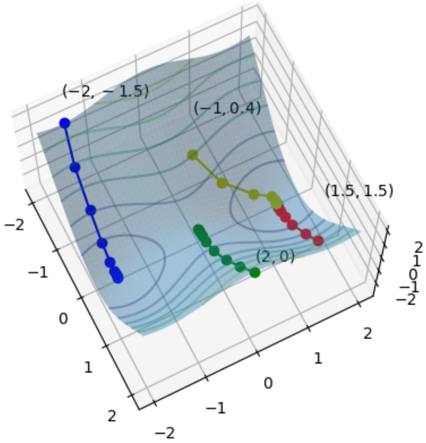
$$\mathcal{E}(\boldsymbol{\gamma}) \coloneqq \iint_{M^2} k\left(\boldsymbol{\gamma}_x, \boldsymbol{\gamma}_y\right) \, \mathrm{d}\gamma_x \, \mathrm{d}\gamma_y \tag{1.1}$$

where $k(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2) \geq 0$ is the *curve energy kernel* such that $k \to +\infty$ as $|\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2| \to 0^+$.

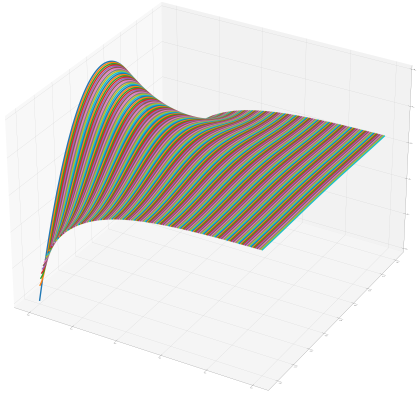
A naïve choice of k satisfying this condition is $k_S(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2) \coloneqq \frac{1}{|\gamma_1 - \gamma_2|}$. However, it turns out that $k_S \sim O\left(\frac{1}{|\gamma_1 - \gamma_2|}\right)$ as $|\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2| \to 0$ (consider neighbouring points), meaning the $\mathcal{E}$ diverges all continuous curves $\boldsymbol{\gamma}$ of nonzero measure.

A more analytically sensible choice of k would be the tangent-point kernel introduced in a paper by Buck and Orloff[2] and later generalised by Yu, Schumacher, and Crane[4].

The next part of the idea is to *reduce $\mathcal{E}$ by continuously deforming the curve* based on a descent method until it reaches a stationary curve, at which, we expect



(a) SDM applied to $f(x, y) = -3\cos x + \cos^2 y$ at different initial points.



(b) Solution to L^2 gradient flow equation with 1D energy functional $\mathcal{E}(f) := \int_{-1}^{1} |\nabla f(x)|^2 \, \mathrm{d}x$, which penealises variation in function. Note that the solution converges to a function with no variation.

Figure 2.1: Gradient flow can be understood as a continuous analogue of steepest descent.

it to be the "unknot" of the original curve. Note that by construction of $\mathcal{E}$, if the curve is to self-intersect, $\mathcal{E}$ increases, and the descent method encourages the curve to repel, preventing the self-intersection.

2 Gradient Flow Equation

Since we pose the problem as continuous reduction of some functional, we need an applicable framework which fits our intention. In our case, **gradient flow equation** seems to be appropriate.

2.1 Motivation of Gradient Flow Equation

For minimising a differentiable function $f : E \subset \mathbb{R}^n \to \mathbb{R}$, there is a well-known method known as **steepest descent method** (SDM)[1].

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha_k \nabla f\left(\mathbf{x}^k\right) \tag{2.1}$$

Starting from the initial input point $\mathbf{x}^0$, at each iteration, input points $\left\{\mathbf{x}^k\right\}$ move in the direction of "steepest" decrease, with specified step size $\alpha_k > 0$, reducing the value at evaluation of f. Note that in general, this method is not guaranteed to find the minimiser as shown in Figure (2.1(a)). On the other hand, convergence is guaranteed under certain assumptions, for example, convexity and L-smoothness[1] with a certain choice of step size α_k.

Analogously, differential equation known describing the reduction process of a functional $F : \mathcal{X} \to \mathbb{R}$ (where $\mathcal{X}$ is an inner product function space) can be motivated. Starting from (2.1), replacing $\mathbf{x}^k$ by f_k and $\nabla f\left(\mathbf{x}^k\right)$ by $\operatorname{grad}_{\mathcal{X}} F\left(f_k\right)$

$$f_{k+1} = f_k - \alpha_k \operatorname{grad}_{\mathcal{X}} F\left(f_k\right) \tag{2.2}$$

Now think of f_k as "snapshots" at certain time $t = t_k$. Without loss of generality, let $\alpha_k \equiv 1$.[2] Dividing (2.2) by time step $\Delta t \coloneqq t_{k+1} - t_k$, and taking the limit as $\Delta t \to 0$, we acquire the **gradient flow equation**[4].

$$\frac{\partial f}{\partial t} = - \operatorname{grad}_{\mathcal{X}} F(f) \tag{2.3}$$

where index k transforms to "time" variable t.

Note that grad of a functional is not defined yet. This depends on the inner product function space $\mathcal{X}$ (eg. L^2, H^1, ...) of interest.

2.2 Gradient of Functional

In order to understand the gradient of a functional, it helps to recall the gradient of a function. Gradient of a function $f : \mathbb{R}^n \to \mathbb{R}$ characterises the first-order variation in a certain direction, that is,

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \quad \nabla f\left(\mathbf{x}\right) \cdot \mathbf{y} = \frac{\partial}{\partial \epsilon} f\left(\mathbf{x} + \epsilon \mathbf{y}\right) |_{\epsilon=0} \tag{2.4}$$

While this is not a conventional definition of function gradient, it is still an equivalent definition.

One could analogously construct the definition of gradient of a functional.

Definition (Gradient of Functional)**.** For a functional $F : \mathcal{X} \to \mathbb{R}$, define functional gradient $\operatorname{grad}_{\mathcal{X}} F(f)$ as:

$$\forall f, g \in \mathcal{X} \;\; \langle \operatorname{grad}_{\mathcal{X}} F(f), g \rangle_{\mathcal{X}} = \frac{\partial}{\partial \epsilon} F\left(f + \epsilon g\right) |_{\epsilon=0} \tag{2.5}$$

where $\langle \cdot, \cdot \rangle_{\mathcal{X}}$ is the inner product defined over the inner product function space $\mathcal{X}$.

Common inner product spaces include L^2, H^1, which are defined in Appendix A.

Remark. By assuming time-independence of f in the gradient flow equation (2.3), we get a stationary state equation $\operatorname{grad}_{\mathcal{X}} F(f) = 0$, which implies $\frac{\partial}{\partial \epsilon} F\left(f + \epsilon g\right) |_{\epsilon=0} = 0$ for all $g \in \mathcal{X}$. This shows that the stationary state of the gradient flow equation is precisely the solution to the Euler-Lagrange equation.

[1] $||\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})|| \leq L||\mathbf{x} - \mathbf{y}||$ for some constant $L > 0$

[2] This is justified by taking a differrent time scale; essentially nondimensionalisation. Also $\alpha_k \sim O(1)$ as $k \to \infty$ (eg. $\alpha = L^{-1}$ for L-smooth optimisation) is in fact a realistic choice.

2.2.1 Gradient on Integer Sobolev Spaces

Assume Ω is a boundary-free domain. Given $L^2(\Omega) = H^0(\Omega)$ gradient (Ω will be omitted in this section when unambiguous), one could express Sobolev gradients of other orders. For functional F, write $h \coloneqq \operatorname{grad}_{L^2} F$ for gradient of F in L^2, that is, by definition

$$\frac{\partial}{\partial \epsilon} F\left(f + \epsilon g\right)|_{\epsilon=0} = \langle h, g\rangle_{L^2} \eqqcolon \mathcal{D} \tag{2.6}$$

By integration by parts (IBP) on boundary-free domain Ω, [3]

$$\begin{aligned}\mathcal{D} &= \langle \Delta\Delta^{-1} h, g\rangle_{L^2} \\ &\overset{\text{IBP}}{=} \langle -\nabla\Delta^{-1} h, \nabla g\rangle_{L^2} \\ &= \langle -\Delta^{-1} h, g\rangle_{H^1}\end{aligned}$$

So one concludes that

$$\operatorname{grad}_{H^1} F = -\Delta^{-1} h \tag{2.7}$$

Similarly, one could acquire gradient in H^2 by

$$\begin{aligned}\mathcal{D} &= \langle \Delta^{-1} h, \Delta g\rangle_{L^2} \\ &= \langle \Delta\left(\Delta^{-2} h\right), \Delta g\rangle_{L^2} \\ &= \langle \Delta^{-2} h, g\rangle_{H^2}\end{aligned}$$

Hence,

$$\operatorname{grad}_{H^2} F = \Delta^{-2} h \tag{2.8}$$

One could take it even further and define gradient in H^{-1},

$$\begin{aligned}\mathcal{D} &= \langle h, \Delta\Delta^{-1} g\rangle_{L^2} \\ &= \langle -\nabla h, \nabla\left(\Delta^{-1} g\right)\rangle_{L^2} \\ &= \langle -\Delta h, g\rangle_{H^{-1}}\end{aligned}$$

So,

$$\operatorname{grad}_{H^{-1}} F = -\Delta h \tag{2.9}$$

Remark. From (2.7), (2.8), and (2.9), one may deduce that choosing the right Sobolev space for a functional may make the gradient flow equation much easier to solve. For example, given functional $F(f) \coloneqq \int_\Omega |\nabla f(x)|^2 \,\mathrm{d}V$ over an open set Ω, L^2 gradient turns out to be $\operatorname{grad}_{L^2} F(f) = h = -\Delta f$. Instead of solving L^2 gradient flow equation $\frac{\partial}{\partial t} f = \Delta f$ (Refer to Figure 2.1(b) for the case $\Omega = (-1, 1)$.), solving H^1 gradient flow equation $\frac{\partial}{\partial t} f = -f$ is more trivial as it practically collapses down to solving an ordinary differential equation.

[3] A more careful treatment of IBP is outlined in appendix B.

3 Tangent-Point Energy

Given a curve $\boldsymbol{\gamma} : M \to \mathbb{R}^3$, a sensible choice of curve energy is the tangent-point energy[4].

Definition (Tangent-Point Energy)**.** For a continuously differentiable parameterised curve $\boldsymbol{\gamma} : M \to \mathbb{R}^3$, define **tangent-point energy** as:

$$\mathcal{E}_\beta^\alpha(\boldsymbol{\gamma}) \coloneqq \iint_{M^2} k_\beta^\alpha(\boldsymbol{\gamma}_x, \boldsymbol{\gamma}_y, \mathbf{T}_x)\, \mathrm{d}\gamma_x\, \mathrm{d}\gamma_y \tag{3.1}$$

where $\mathbf{T}_x \coloneqq \frac{\mathrm{d}\boldsymbol{\gamma}_x}{\mathrm{d}t} \Big/ \left|\frac{\mathrm{d}\boldsymbol{\gamma}_x}{\mathrm{d}t}\right|$ is the unit tangent vector at $\boldsymbol{\gamma}_x$ along the curve, and **tangent-point kernel** is given as:

$$k_\beta^\alpha(\mathbf{p}, \mathbf{q}, \mathbf{T}) \coloneqq \frac{|\mathbf{T} \wedge (\mathbf{p} - \mathbf{q})|^\alpha}{|\mathbf{p} - \mathbf{q}|^\beta} \tag{3.2}$$

α and β are parameters one could choose, but for tangent-point energy to be well-defined, one may choose them to satisfy $\alpha > 1$ and $\beta \in [\alpha + 2, 2\alpha + 1)$.

Note that choosing $\alpha = 2$ and $\beta = 4$ results in the scaled version of the original tangent-point energy by Buck and Orloff[2].

The geometric intuition of $k_4^2(\boldsymbol{\gamma}_x, \boldsymbol{\gamma}_y, \mathbf{T}_x)$ (kernel of "Buck-Orloff tangent-point energy") is that the kernel evaluates to $\frac{1}{4r^2}$ where r is the radius of the smallest circle drawn that is tangent at $\boldsymbol{\gamma}_x$ and crosses through $\boldsymbol{\gamma}_y$ as shown in Figure 3.1.

The choice of parameters α and β changes the behaviour of the tangent-point energy as stated in the following lemma.

Lemma 3.1. *Tangent-point energy $\mathcal{E}_\beta^\alpha$ defined as (3.1) is scale invariant with respect to the curve if and only if $\beta = \alpha + 2$. Moreover, if $\beta > \alpha + 2$, then $\mathcal{E}_\beta^\alpha$ scales inversely with the curve.*

Proof. Take a parameterised curve $\boldsymbol{\gamma} : M \to \mathbb{R}^3$ and $\boldsymbol{\Gamma} \coloneqq c\boldsymbol{\gamma}$, a curve scaled by factor $c > 0$ of $\boldsymbol{\gamma}$. Note that the unit tangent vector is identical for $\boldsymbol{\gamma}$ and $\boldsymbol{\Gamma}$, that is, $\mathbf{T}_x \coloneqq \frac{\mathrm{d}\boldsymbol{\gamma}_x}{\mathrm{d}t} \Big/ \left|\frac{\mathrm{d}\boldsymbol{\gamma}_x}{\mathrm{d}t}\right| = \frac{\mathrm{d}\boldsymbol{\Gamma}_x}{\mathrm{d}t} / \left|\frac{\mathrm{d}\boldsymbol{\Gamma}_x}{\mathrm{d}t}\right|$

Then,

$$\frac{k_\beta^\alpha(\boldsymbol{\gamma}_x, \boldsymbol{\gamma}_y, \mathbf{T}_x)}{k_\beta^\alpha(\boldsymbol{\Gamma}_x, \boldsymbol{\Gamma}_y, \mathbf{T}_x)} = \frac{|\boldsymbol{\gamma}_x - \boldsymbol{\gamma}_y|^\alpha}{|\boldsymbol{\Gamma}_x - \boldsymbol{\Gamma}_y|^\beta} = c^{\beta - \alpha} \tag{3.3}$$

Also note that

$$\mathrm{d}\Gamma = c\, \mathrm{d}\gamma \tag{3.4}$$

So, we deduce from (3.1),

$$\mathcal{E}_\beta^\alpha(\boldsymbol{\Gamma}) = c^{\alpha - \beta + 2} \mathcal{E}_\beta^\alpha(\boldsymbol{\gamma}) \tag{3.5}$$

Hence, $\mathcal{E}_\beta^\alpha$ is scale invariant with respect to the curve if and only if $\alpha - \beta + 2 = 0$, and if $\beta > \alpha + 2$, $\mathcal{E}_\beta^\alpha$ scales as $O\left(\frac{1}{c^{\beta - (\alpha + 2)}}\right)$ as $c \to \infty$. ∎

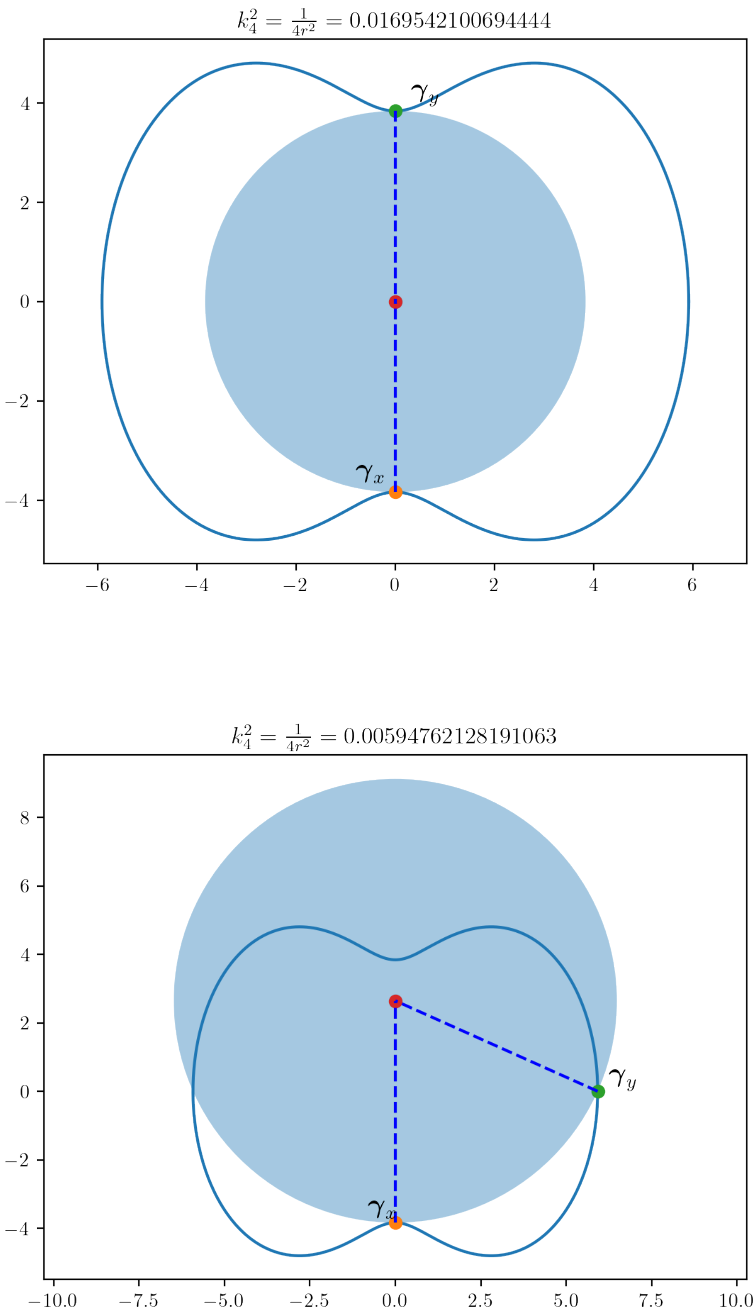


Figure 3.1: Intuition behind k_4^2

Remark. If $\beta < \alpha + 2$, then the energy scales with the size of the curve, meaning that the energy is trivially minimised by scaling down the curve to a singularity, which is not desirable in our context.

One could also justify the condition on α and β by the following lemma.

Lemma 3.2. *Given α and β, the singularity of the kernel at a point and another point of which is arc-length $\epsilon > 0$ away from it is of order $O\left(\epsilon^{2\alpha-\beta}\right)$, that is, $k^{\alpha}_{\beta}\left(\boldsymbol{\gamma}(s), \boldsymbol{\gamma}(s+\epsilon), \mathbf{T}(s)\right) = O\left(\epsilon^{2\alpha-\beta}\right)$ as $\epsilon \to 0$. Moreover, if $2\alpha = \beta$, then the kernel converges to $\left(\frac{\kappa}{2}\right)^{\alpha}$ as the two points get closer, where κ is the curvature of the curve at the point.*

Proof. For $\boldsymbol{\gamma}(s) = (x(s), y(s), z(s))$ parameterised by arc-length, one recognises that the tangent vector at this point is $\mathbf{T} = \boldsymbol{\gamma}'(s)$. Note that $||\boldsymbol{\gamma}'(s)|| = ||\mathbf{T}|| = 1$.

By Taylor expansion:

$$\boldsymbol{\gamma}\left(s+\epsilon\right) = \boldsymbol{\gamma}(s) + \epsilon\boldsymbol{\gamma}'(s) + \frac{1}{2}\epsilon^2\boldsymbol{\gamma}''(s) + O\left(\epsilon^3\right) \tag{3.6}$$

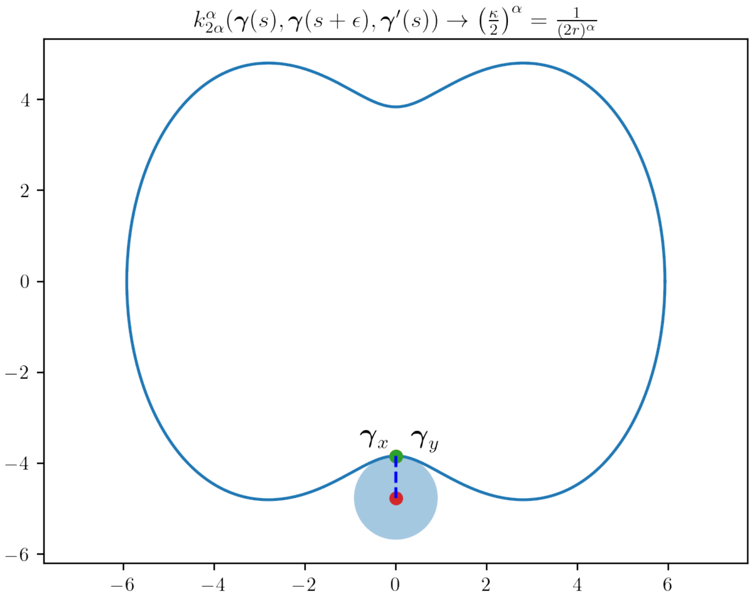
Then around $\epsilon = 0$,

$$\begin{aligned}
&k^{\alpha}_{\beta}\left(\boldsymbol{\gamma}(s), \boldsymbol{\gamma}(s+\epsilon), \boldsymbol{\gamma}'(s)\right) \\
&= \frac{||\boldsymbol{\gamma}'(s) \wedge \left(\boldsymbol{\gamma}(s+\epsilon) - \boldsymbol{\gamma}(s)\right)||^{\alpha}}{||\boldsymbol{\gamma}(s+\epsilon) - \boldsymbol{\gamma}(s)||^{\beta}} \\
&= \frac{||\boldsymbol{\gamma}'(s) \wedge \left(\epsilon\boldsymbol{\gamma}'(s) + \frac{1}{2}\epsilon^2\boldsymbol{\gamma}''(s) + O\left(\epsilon^3\right)\right)||^{\alpha}}{||\epsilon\boldsymbol{\gamma}'(s) + O\left(\epsilon^2\right)||^{\beta}} \\
&= \frac{\epsilon^{2\alpha-\beta}}{2^{\alpha}} \frac{||\boldsymbol{\gamma}'(s) \wedge \boldsymbol{\gamma}''(s) + O\left(\epsilon\right)||^{\alpha}}{||\boldsymbol{\gamma}'(s) + O\left(\epsilon\right)||^{\beta}} && \because \boldsymbol{\gamma}'(s) \wedge \boldsymbol{\gamma}'(s) = \mathbf{0} \\
&= \frac{\epsilon^{2\alpha-\beta}}{2^{\alpha}} \left(\frac{||\boldsymbol{\gamma}'(s)||^{\alpha}||\boldsymbol{\gamma}''(s)||^{\alpha}}{||\boldsymbol{\gamma}'(s)||^{\beta}} + O\left(\epsilon\right)\right) && \because \boldsymbol{\gamma}'(s) \perp \boldsymbol{\gamma}''(s) \\
&= \frac{\epsilon^{2\alpha-\beta}}{2^{\alpha}} \left(\kappa^{\alpha} + O\left(\epsilon\right)\right) && \because ||\boldsymbol{\gamma}'(s)|| = ||\mathbf{T}|| = 1 \\
& && \text{and } ||\boldsymbol{\gamma}''(s)|| = \kappa
\end{aligned}$$

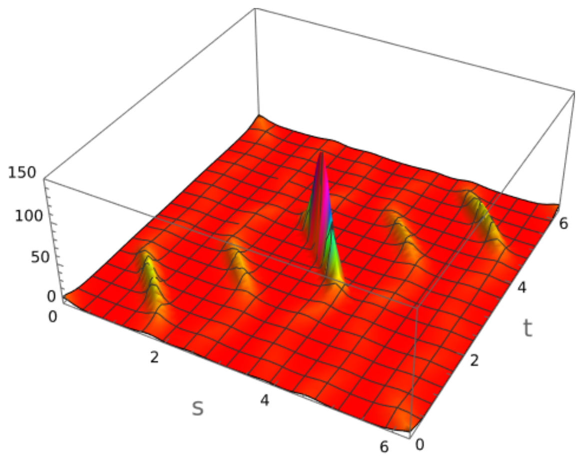
So with arc-length perturbation of ϵ, the order of singularity of the kernel is $O\left(\epsilon^{2\alpha-\beta}\right)$. In particular, if $2\alpha - \beta = 0$, the kernel converges to $\left(\frac{\kappa}{2}\right)^{\alpha}$, as demonstrated in Figure 3.2(a). ■

From lemma 3.1 and lemma 3.2, the well-definedness condition of tangent-point energy immediately follows from arguing by considering integrability of functions with isolated poles of order lower than 1. (See Figure 3.3)

Corollary 3.3. *If $\alpha > 1$ and $\beta \in [\alpha + 2, 2\alpha + 1)$, tangent-point energy $\mathcal{E}^{\alpha}_{\beta}$ is well-defined.*



(a) $k^{\alpha}_{2\alpha}\left(\boldsymbol{\gamma}(s), \boldsymbol{\gamma}(s+\epsilon), \boldsymbol{\gamma}'(s)\right)$ converges to $\left(\frac{\kappa}{2}\right)^{\alpha}$.



(b) $k^{\alpha}_{2\alpha}$ heat map

Figure 3.2: Behaviours of $\beta = 2\alpha$ kernel

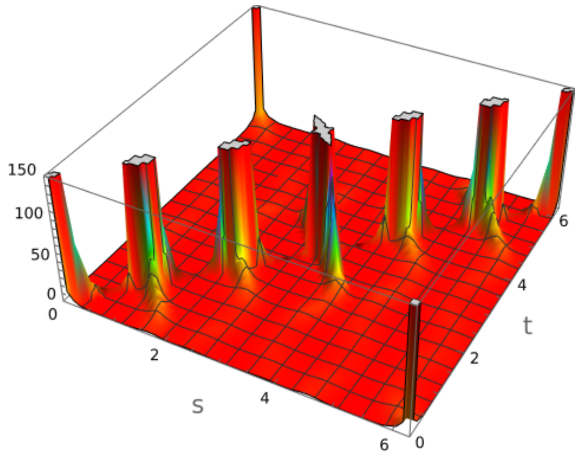


Figure 3.3: Heat map of $k^2_{4.5}$: Even with singularities, because their order is less than 1, the kernel is integrable, hence $\mathcal{E}^2_{4.5}$ is well-defined.

4 Unknotting Curves via Gradient Flow Equation

Now that gradient flow equation and tangent-point energy are introduced, one can formalise the process of untangling a tangled curve:

Definition (Curve Untangling Process)**.** Given a parameterised curve $\boldsymbol{\gamma} : M \times T \to \mathbb{R}^3$ over an interval M and time domain T, denote the following initial value problem as **curve untangling process**:

$$\frac{\partial \boldsymbol{\gamma}}{\partial t} = -\operatorname{grad}_{\mathcal{X}} \mathcal{E}^{\alpha}_{\beta}(\boldsymbol{\gamma}) - \operatorname{grad}_{\mathcal{X}} \mathcal{C}(\boldsymbol{\gamma}) \tag{4.1}$$

$$\boldsymbol{\gamma}(s; 0) = \boldsymbol{\gamma}_0(s) \tag{4.2}$$

where

- $\gamma_0(s)$ is the parameterisation of the initial (tangled) curve (prescribed at $t = 0$)
- $\mathcal{E}^{\alpha}_{\beta}$ is the tangent-point energy (See (3.1))
- $\mathcal{C}$ is additional constraint energy to control behaviour of curve untangling process.

4.1 Discretisation for Numerical Computation

Solving (4.1), (4.2) analytically is challenging. Rather, we aim to acquire a numerical solution. Assume for simplicity that the curve of interest is simple closed.[4]

[4]We may assume that the function definition of $\boldsymbol{\gamma} : M \to \mathbb{R}^3$ extends to $\boldsymbol{\gamma} : \mathbb{R} \to \mathbb{R}^3$ by periodicity.

4.1.1 Discretisation of Curve

We start by discretising the initial curve $\boldsymbol{\gamma}_0$ by taking N points on a curve as shown in Figure 4.1. Represent the initially discretised curve in the tensor form: $\boldsymbol{\Gamma}^0 = \left(\mathbf{x}_0^0, \mathbf{x}_1^0, \cdots, \mathbf{x}_{N-1}^0\right)$, and the discretised curve at subsequent time step $k \in \mathbb{N} \cup \{0\}$ as $\boldsymbol{\Gamma}^k = \left(\mathbf{x}_0^k, \mathbf{x}_1^k, \cdots, \mathbf{x}_{N-1}^k\right)$. Since we restrict our attention to a simple closed curve, it is convenient to extend the indexing rule by:

$$\mathbf{x}_i^k = \mathbf{x}_{r(i,N)}^k \qquad \text{where } r\,(i,N) = (\text{remainder of } i \div N) \tag{4.3}$$

so that $\mathbf{x}_N^k = \mathbf{x}_0^k$, $\mathbf{x}_{N+1}^k = \mathbf{x}_1^k$, etc.

Define (right) operator $[\cdot] : \mathbb{R}^{3\times N} \to \mathbb{R}$ such that for tensor $T \in \mathbb{R}^{3\times N}$,

$$T[i] \coloneqq T\mathbf{e}_{r(i,N)}$$

where $\mathbf{e}_i$ is the canonical vector with 1 as its only nontrivial i^{th} coordinate. With this operator, one could write

$$\boldsymbol{\Gamma}^k[i] = \mathbf{x}_{r(i,N)}^k = \mathbf{x}_i^k \tag{4.4}$$

anaogous to $\boldsymbol{\gamma} = \boldsymbol{\gamma}(s)$ being a parameterised curve, which is a vector-valued function.

Finally, denote by e_i^k for the (undirected) edge with vertex pair $\left(\mathbf{x}_i^k, \mathbf{x}_{i+1}^k\right)$.

Remark. $\boldsymbol{\Gamma}^k$ is a tensor in $\mathbb{R}^{3\times N}$.

4.1.2 Discretisation of Tangent-Point Energy

In order to acquire numerical solution, one must also be able to numerically compute the tangent-point energy. For this, we pose the energy quadrature $E_\beta^\alpha\,(\boldsymbol{\Gamma})$ of the following form:

$$\mathcal{E}_\beta^\alpha\,(\boldsymbol{\gamma}\,(\cdot,t)) \coloneqq \iint_{M^2} k_\beta^\alpha(\boldsymbol{\gamma}_x, \boldsymbol{\gamma}_y)\, \mathrm{d}\gamma_x\, \mathrm{d}\gamma_y \approx E_\beta^\alpha\left(\boldsymbol{\Gamma}^k\right) \coloneqq \sum_{i,j\in\{0,\cdots,N-1\}} K_\beta^\alpha\,(i,j)\, ||e_i^k||\, ||e_j^k|| \tag{4.5}$$

where K_β^α is an approximation of tangent-point kernel k_β^α (which is specified at (4.6)), and $||e_i^k||$ is the length of edge e_i^k, that is, $||e_i^k|| = ||\mathbf{x}_i^{k+1} - \mathbf{x}_i^k||$. Note that $\boldsymbol{\Gamma}^k$ is a polygonal curve, for which tangent-point energy (3.1) is not well-defined due to locally non-integrable contributions from vertices:

One way to resolve the issue is to "ignore" the adjacent edge contribution[4] in the energy quadrature E_β^α as shown in Figure 4.2(a). The justification is that as we take a finer mesh (N sufficiently large), the product of edge lengths ($||e_i||||e_j||$) should tend to zero sufficiently fast, resulting in approximation of the energy for the smooth curve, which did not have vertices resulting in local non-integrability in the first place.

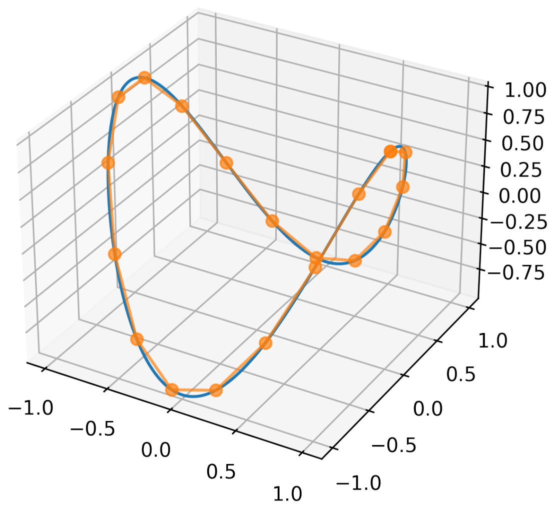


Figure 4.1: Discretisation of a simple closed curve by sampling the points along the curve.

It still remains to sensibly approximate the kernel $k^\alpha_\beta(\boldsymbol{\gamma}_x, \boldsymbol{\gamma}_y) \approx K^\alpha_\beta(i,j)$. One sensible approximation is to use the following 4-point quadrature[4].

$$K^\alpha_\beta(i,j) := \frac{1}{4}\Big(k^\alpha_\beta\left(\mathbf{x}^k_i, \mathbf{x}^k_j, \mathbf{T}^k_i\right) + k^\alpha_\beta\left(\mathbf{x}^k_i, \mathbf{x}^k_{j+1}, \mathbf{T}^k_i\right) + k^\alpha_\beta\left(\mathbf{x}^k_{i+1}, \mathbf{x}^k_j, \mathbf{T}^k_i\right) + k^\alpha_\beta\left(\mathbf{x}^k_{i+1}, \mathbf{x}^k_{j+1}, \mathbf{T}^k_i\right)\Big) \tag{4.6}$$

where $\mathbf{T}^k_i := \frac{\mathbf{x}^k_{i+1}-\mathbf{x}^k_i}{||\mathbf{x}^k_{i+1}-\mathbf{x}^k_i||}$ approximates the tangent vector to the curve at $\boldsymbol{\gamma}_x$. (See Figure 4.2(b).)

Putting (4.5) and (4.6) together, one can write the **tangent-point energy quadrature** as:

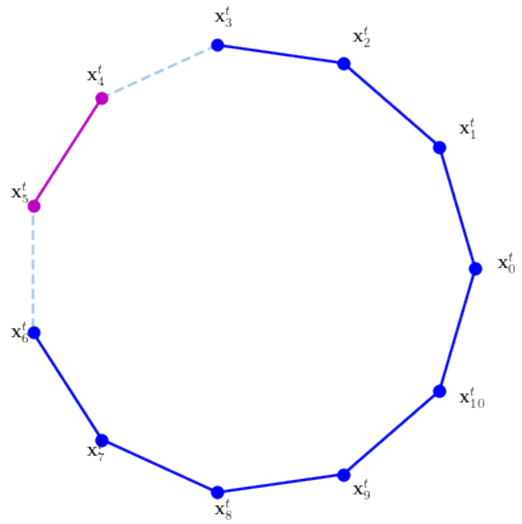
$$E^\alpha_\beta := \sum_{\substack{i,j\in\{0,\cdots,N-1\}\\ r(i-j,N)>1}} K^\alpha_\beta(i,j)||e^k_i||\,||e^k_j|| \tag{4.7}$$

where $r\,(i-j,N)$ is the geodesic distance between i and j in modulo N, characterising the avoidance of adjacent edges on simple closed polygonal curve.

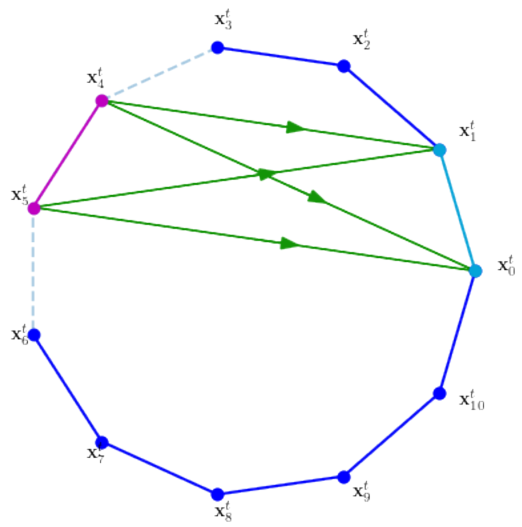
4.1.3 Finite Difference Scheme of Curve Untangling Process

Based on (4.1), one writes the following finite difference scheme:

$$\mathcal{D}_t\boldsymbol{\Gamma}^k = -\operatorname{Grad}_X E^\alpha_\beta(\boldsymbol{\Gamma}^k) - \operatorname{Grad}_X C\left(\boldsymbol{\Gamma}^k\right) \qquad \text{for } k = 0, 1, \cdots \tag{4.8}$$



(a) For a chosen edge e_i^k, ignore the two adjacent edges e_{i-1}^k, e_{i+1}^k. In the limit as $N \to 0$, because the edge lengths tend to zero, the discrepancy between the quadrature and the analytical value of the energy is expected to tend to zero.



(b) Tangent-point kernel is approximated by 4-point quadrature defined as (4.6).

Figure 4.2: Quadrature for approximation of tangent-point energy.

where $\mathcal{D}_t$ is the finite difference operator over time, Grad_X is discrete equivalent of $\mathrm{grad}_{\mathcal{X}}$ on discrete inner product space X, E^{α}_{β} is the tangent-point energy quadrature defined as (4.7), and C is the discretised version of the constraint energy $\mathcal{C}$. For the simplest scheme, one could take the forward difference operator as $\mathcal{D}_t\mathbf{\Gamma}^k[i] := \frac{\mathbf{\Gamma}^{k+1}[i]-\mathbf{\Gamma}^k[i]}{\Delta T}$.

4.2 Example: L^2 Explicit Euler Scheme for Buck-Orloff Energy

Now we visit the simplest concrete numerical scheme for curve untangling process. We will demonstrate with Buck-Orloff energy ($\alpha = 2$, $\beta = 4$), as one need not take into account of constraint energy $\mathcal{C}$ due to its scale invariance (lemma 3.1).

In L^2, $\mathrm{grad}_{L^2}\,\mathcal{E}^2_4(\boldsymbol{\gamma})$ is simply the "first-order perturbation" as in (2.6). Discrete equivalent is the ℓ^2 space, and $\mathrm{Grad}_{\ell^2} = \nabla_{\mathbf{\Gamma}}$. Taking $\mathcal{D}_t$ from (4.8) to be forward difference operator,

$$\frac{\mathbf{\Gamma}^{k+1}-\mathbf{\Gamma}^k}{\Delta T} = -\,\mathrm{Grad}_{\ell^2}\,E^2_4\left(\mathbf{\Gamma}^k\right) \tag{4.9}$$

where one could explicitly write $\mathrm{Grad}_{\ell^2}\,E^2_4\left(\mathbf{\Gamma}^k\right)$ as

$$\mathrm{Grad}_{\ell^2}\,E^2_4\left(\mathbf{\Gamma}^k\right) = \begin{pmatrix} \frac{\partial}{\partial x_{1,1}} & \frac{\partial}{\partial x_{1,2}} & \cdots & \frac{\partial}{\partial x_{1,N-1}} \\ \frac{\partial}{\partial x_{2,1}} & \frac{\partial}{\partial x_{2,2}} & \cdots & \frac{\partial}{\partial x_{1,N-1}} \\ \frac{\partial}{\partial x_{3,1}} & \frac{\partial}{\partial x_{3,2}} & \cdots & \frac{\partial}{\partial x_{3,N-1}} \end{pmatrix} E^2_4\left(\mathbf{\Gamma}^k\right)$$

where $x_{j,i}$ refers to the (j,i) coordinate variable for $3\times N$ tensor representation of $\mathbf{\Gamma}^k$. Note that $\mathrm{Grad}_{\ell^2}\,E^2_4\left(\mathbf{\Gamma}^k\right) \in \mathbb{R}^{3\times N}$ is also a tensor of the same shape as $\mathbf{\Gamma}^k$, so naturally, most arithmetics[5] needed for (4.9) is well-defined.

Remark. Grad_{ℓ^2} is in fact a linear operator with respect to its input, energy.

One could use this exact form of $\mathrm{Grad}_{\ell^2}\,E^2_4\left(\mathbf{\Gamma}^k\right)$ as given in appendix D, but it could be considered cumbersome (even though there are benefits to implementing this). One could alternatively approximate $\mathrm{Grad}_{\ell^2}\,E^2_4\left(\mathbf{\Gamma}^k\right)$ by central difference scheme, for example: for $i = 0, 1, \cdots, N-1$ and $j = 1, 2, 3$,

$$\mathbf{e}_j \cdot \left(\mathrm{Grad}_{\ell^2}\,E^2_4\left(\mathbf{\Gamma}^k\right)[i]\right) = \frac{\partial E^2_4(\mathbf{\Gamma})}{\partial x_{j,i}} \approx \frac{1}{2\Delta X}\left(\bar{E}^2_4(i) - \underline{E}^2_4(i)\right) \tag{4.10}$$

where

$$\bar{E}^2_4(i) := E^2_4\left((\mathbf{x}_0, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_i + \Delta X\mathbf{e}_j, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_{N-1})\right)$$
$$\underline{E}^2_4(i) := E^2_4\left((\mathbf{x}_0, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_i - \Delta X\mathbf{e}_j, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_{N-1})\right)$$

Here $\mathbf{e}_j$ is the canonical vector of which has entry 1 at j^{th} component.

Notice that (4.9) can be interpreted as SDM over all the coordinates on the curve.

[5] Such as addition, subtraction, and scalar multiplication.

Appendices

A Definitions of Important Inner Product Spaces

(Adapted based on definitions given in lecture note by Endre Süli[3]. For our purposes, we focus our attention to real domain.) Here are some of the notable inner product spaces.

A.1 L^2 Space

Definition (L^2 Space)**.** For interval $I \subset \mathbb{R}$, L^2 space over is defined as

$$L^2 = \left\{ f : I \to \mathbb{R} \mid \left(\int_I |f|^2 \, \mathrm{d}x \right)^{1/2} < \infty \right\} \tag{A.1}$$

L^2 inner product is defined as

$$\forall f, g \in L^2 \;\; \langle f, g \rangle_{L^2} = \int_I f g \, \mathrm{d}x \tag{A.2}$$

A.2 H^k Space

To define Sobolev (inner product) spaces (denoted H^k where $k \in \mathbb{N} \cup \{0\}$) one must define weak derivative operator D:

Definition (Weak Derivative)**.** For u locally integrable on I, if there exists w such that for all infinitely smooth $v : I \to \mathbb{R}$ with compact support,

$$\int_I wv \, \mathrm{d}x = (-1)^\alpha \int_I u \frac{\mathrm{d}^\alpha v}{\mathrm{d}x^\alpha} \, \mathrm{d}x \tag{A.3}$$

then w is said to be **weak derivative** of order α of u, and one writes $D^\alpha u = w$.

Weak derivative extends the definition of conventional derivative, and is equivalent to the conventional derivative for smooth functions. With weak derivatives introduced, one may now define Sobolev inner product spaces.

Definition (H^k Space)**.** Sobolev inner product space of order $k \in \mathbb{N} \cup \{0\}$ (denoted H^k) is defined as

$$H^k = \left\{ f \in L^2 \mid D^\alpha f \in L^2, \alpha \leq k \right\} \tag{A.4}$$

H^k inner product is defined as:

$$\forall f, g \in H^k \;\; \langle f, g \rangle_{H^2} = \sum_{\alpha \leq k} \langle D^\alpha f, D^\alpha g \rangle_{L^2} \tag{A.5}$$

Remark. Note that $H^0 = L^2$ by definition. One could say that Sobolev inner product spaces extend L^2 space. It also turns out that H^k are Hilbert spaces.

B Gradient Operator in Sobolev Spaces

For acquiring gradient on integer Sobolev spaces from L^2 gradient over boundary-free Ω, there are two main "rules" one could use.

Lemma B.1 (Shifting Gradient Operator)**.**

$$\langle \nabla f, \nabla g \rangle_{L^2} = -\langle \Delta f, g \rangle_{L^2} = -\langle f, \Delta g \rangle_{L^2} \tag{B.1}$$

Proof.

$$\begin{aligned}
\langle \nabla f, \nabla g \rangle_{L^2} &= \int_\Omega \nabla f \cdot \nabla g \, \mathrm{d}V \\
&= \int_\Omega \left(\nabla \cdot (g \nabla f) - g \Delta f \right) \mathrm{d}V \\
&\overset{\text{IBP}}{=} \underbrace{\oint_{\partial\Omega} g \nabla f \cdot \mathbf{n} \, \mathrm{d}s}_{\text{Boundary Term}} - \int_\Omega g \Delta f \, \mathrm{d}V \\
&= -\langle \Delta f, g \rangle_{L^2} + \text{Boundary Term}
\end{aligned}$$

For boundary-free Ω, boundary terms can be taken to be zero. ■

Lemma B.2 (Shifting Laplacian Operator)**.**

$$\langle \Delta f, g \rangle_{L^2} = \langle f, \Delta g \rangle \tag{B.2}$$

Proof.

$$\begin{aligned}
\langle \Delta f, g \rangle_{L^2} &= \int_\Omega g \Delta f \, \mathrm{d}V \\
&= \int_\Omega \left(\nabla \cdot (g \nabla f) - \nabla f \cdot \nabla g \right) \mathrm{d}V \\
&\overset{\text{IBP}}{=} \underbrace{\oint_{\partial\Omega} g \nabla f \cdot \mathbf{n} \, \mathrm{d}s}_{\text{Boundary Term}} - \int_\Omega \nabla f \cdot \nabla g \, \mathrm{d}V \\
&= \text{Boundary Term} - \langle \nabla f, \nabla g \rangle_{L^2}
\end{aligned}$$

Now, use lemma B.1 and take boundary terms to be zero. ■

C Tangent-Point Energy Quadrature: Other Homeomorphism Classes

We have defined tangent-point energy quadrature for simple closed curve (any curve homeomorphic to circle) $\boldsymbol{\gamma}$ at (4.7).

For curves in different homeomorphism class from S^1, one may make minor changes to (4.7) to get a sensible quadrature.

C.1 Curves Homeomorphic to a Line Segment

D Exact ℓ^2 Gradient of E_β^α

Recall the discretisation of tangent-point energy for simple closed curve $\mathbf{\Gamma}$ from (4.7) with K_β^α defined by (4.6).

To compute $\mathrm{Grad}_{\ell^2} E_\beta^\alpha(\mathbf{\Gamma})$, one can consider subproblem of computing $\nabla_{\mathbf{x}_k} E_\beta^\alpha(\mathbf{\Gamma})$ for each $k = 0, 1, \cdots, N-1$.

Fix k. There are $4(N-3)$ pairs of indices (i, j) where the summand $K_\beta^\alpha(i,j)||\mathbf{e}_i||\,||\mathbf{e}_j||$ involves $\mathbf{x}_k$ are:

- $(k, 0), \cdots, (k, k-2), (k, k+2), \cdots, (k, N-1)$
- $(k-1, 0), \cdots, (k-1, k-3), (k-1, k+1), \cdots, (k-1, N-1)$
- $(0, k), \cdots, (k-2, k), (k+2, k), \cdots, (N-1, k)$
- $(0, k-1), \cdots, (k-3, k-1), (k+1, k-1), \cdots, (N-1, k-1)$

We now attempt to construct explicit derivative in a "modular fashion". If we take gradient of the summand directly,[6]

$$\nabla_{\mathbf{x}_k}\left(K_\beta^\alpha(i,j)||\mathbf{e}_i||\,||\mathbf{e}_j||\right) = ||\mathbf{e}_i||\,||\mathbf{x}_j||\nabla_{\mathbf{x}_k} K_\beta^\alpha(i,j) + K_\beta^\alpha(i,j)\,\nabla_{\mathbf{x}_k}(||\mathbf{e}_i||\,||\mathbf{e}_j||) \tag{D.1}$$

Due to restriction $r(i-j, N) > 1$, at most one of $||\mathbf{e}_i||$ and $||\mathbf{x}_j||$ may involve $\mathbf{x}_k$ at a time.

First note that, if $m \neq k$,

$$\nabla_{\mathbf{x}_k}||\mathbf{x}_k - \mathbf{x}_m|| = \frac{\mathbf{x}_k - \mathbf{x}_m}{||\mathbf{x}_k - \mathbf{x}_m||} \tag{D.2}$$

Now, the demanding part is to compute $\nabla_{\mathbf{x}_k} k_\beta^\alpha$, since it is needed for computing $\nabla_{\mathbf{x}_k} K_\beta^\alpha$. Note that

$$\begin{aligned} k_\beta^\alpha(\mathbf{x}_p, \mathbf{x}_q, \mathbf{T}_r) &= k_\beta^\alpha\left(\mathbf{x}_p, \mathbf{x}_q, \frac{\mathbf{x}_{r+1} - \mathbf{x}_r}{||\mathbf{x}_{r+1} - \mathbf{x}_r||}\right) \\ &= \frac{\sqrt{||\mathbf{x}_{r+1} - \mathbf{x}_r||^2||\mathbf{x}_p - \mathbf{x}_q||^2 - ((\mathbf{x}_{r+1} - \mathbf{x}_r)\cdot(\mathbf{x}_p - \mathbf{x}_q))^2}}{||\mathbf{x}_p - \mathbf{x}_q||^\beta||\mathbf{x}_{r+1} - \mathbf{x}_r||^\alpha} \\ &= \frac{(\xi_{p,q,r})^{\alpha/2}}{\eta_{p,q,r}} \end{aligned} \tag{D.3}$$

where we have defined

$$\begin{aligned} \xi_{p,q,r} &= ||\mathbf{x}_{r+1} - \mathbf{x}_r||^2||\mathbf{x}_p - \mathbf{x}_q||^2 - ((\mathbf{x}_{r+1} - \mathbf{x}_r)\cdot(\mathbf{x}_p - \mathbf{x}_q))^2 \\ \eta_{p,q,r} &= ||\mathbf{x}_p - \mathbf{x}_q||^\beta||\mathbf{x}_{r+1} - \mathbf{x}_r||^\alpha \end{aligned}$$

[6] recall $||\mathbf{e}_i|| = ||\mathbf{x}_i - \mathbf{x}_{i+1}||$ and $||\mathbf{e}_j|| = ||\mathbf{x}_j - \mathbf{x}_{j+1}||$.

Then, we may express $\nabla_{\mathbf{x}_k} k_\beta^\alpha \left(\mathbf{x}_p, \mathbf{x}_q, \mathbf{T}_r\right)$ as:

$$\begin{aligned}\nabla_{\mathbf{x}_k} k_\beta^\alpha \left(\mathbf{x}_p, \mathbf{x}_q, \mathbf{T}_r\right) &= \nabla_{\mathbf{x}_k} \left(\frac{\left(\xi_{p,q,r}\right)^{\alpha/2}}{\eta_{p,q,r}} \right) \\ &= \frac{1}{\left(\eta_{p,q,r}\right)^2} \left(\frac{\alpha}{2} \left(\xi_{p,q,r}\right)^{\alpha/2-1} \eta_{p,q,r} \nabla_{\mathbf{x}_k} \xi_{p,q,r} - \left(\xi_{p,q,r}\right)^{\alpha/2} \nabla_{\mathbf{x}_k} \eta_{p,q,r} \right)\end{aligned} \tag{D.4}$$

It now remains to compute $\nabla_{\mathbf{x}_k} \xi_{p,q,r}$ and $\nabla_{\mathbf{x}_k} \eta_{p,q,r}$ for relevant (p, q, r) tuples.

There are five classes of (p, q, r) relevant tuples.

If $(p, q, r) = (k, j, k)$,

$$\begin{aligned}\xi_{k,j,k} &= ||\mathbf{x}_{k+1} - \mathbf{x}_k||^2 ||\mathbf{x}_k - \mathbf{x}_j||^2 - \left(\left(\mathbf{x}_{k+1} - \mathbf{x}_k\right) \cdot \left(\mathbf{x}_k - \mathbf{x}_j\right)\right)^2 \\ \eta_{k,j,k} &= ||\mathbf{x}_k - \mathbf{x}_j||^\beta ||\mathbf{x}_k - \mathbf{x}_{k+1}||^\alpha \\ \nabla_{\mathbf{x}_k} \xi_{k,j,k} &= 2\left(\mathbf{x}_k - \mathbf{x}_{k+1}\right) ||\mathbf{x}_k - \mathbf{x}_j||^2 + 2||\mathbf{x}_{k+1} - \mathbf{x}_k||^2 \left(\mathbf{x}_k - \mathbf{x}_j\right) \\ &\quad - 2\left(\left(\mathbf{x}_{k+1} - \mathbf{x}_k\right) \cdot \left(\mathbf{x}_k - \mathbf{x}_j\right)\right) \left(\mathbf{x}_j + \mathbf{x}_{k+1} - 2\mathbf{x}_k\right) \\ \nabla_{\mathbf{x}_k} \eta_{k,j,k} &= \beta ||\mathbf{x}_k - \mathbf{x}_j||^{\beta-2} ||\mathbf{x}_k - \mathbf{x}_{k+1}||^\alpha \left(\mathbf{x}_k - \mathbf{x}_j\right) \\ &\quad + \alpha ||\mathbf{x}_k - \mathbf{x}_j||^\beta ||\mathbf{x}_k - \mathbf{x}_{k+1}||^{\alpha-2} \left(\mathbf{x}_k - \mathbf{x}_{k+1}\right)\end{aligned}$$

If $(p, q, r) = (i, j, k)$ where $i \neq k$ and $i \neq k-1$

$$\begin{aligned}\xi_{i,j,k} &= ||\mathbf{x}_{k+1} - \mathbf{x}_k||^2 ||\mathbf{x}_i - \mathbf{x}_j||^2 - \left(\left(\mathbf{x}_{k+1} - \mathbf{x}_k\right) \cdot \left(\mathbf{x}_i - \mathbf{x}_j\right)\right)^2 \\ \eta_{i,j,k} &= ||\mathbf{x}_i - \mathbf{x}_j||^\beta ||\mathbf{x}_k - \mathbf{x}_{k+1}||^\alpha \\ \nabla_{\mathbf{x}_k} \xi_{i,j,k} &= 2||\mathbf{x}_i - \mathbf{x}_j||^2 \left(\mathbf{x}_k - \mathbf{x}_{k+1}\right) \\ &\quad - 2\left(\left(\mathbf{x}_{k+1} - \mathbf{x}_k\right) \cdot \left(\mathbf{x}_i - \mathbf{x}_j\right)\right) \left(\mathbf{x}_j - \mathbf{x}_i\right) \\ \nabla_{\mathbf{x}_k} \eta_{i,j,k} &= \alpha ||\mathbf{x}_i - \mathbf{x}_j||^\beta ||\mathbf{x}_k - \mathbf{x}_{k+1}||^{\alpha-2} \left(\mathbf{x}_k - \mathbf{x}_{k+1}\right)\end{aligned}$$

If $(p, q, r) = (k-1, j, k-1)$,

$$\begin{aligned}\xi_{k-1,j,k-1} &= ||\mathbf{x}_k - \mathbf{x}_{k-1}||^2 ||\mathbf{x}_{k-1} - \mathbf{x}_j||^2 - \left(\left(\mathbf{x}_k - \mathbf{x}_{k-1}\right) \cdot \left(\mathbf{x}_{k-1} - \mathbf{x}_j\right)\right)^2 \\ \eta_{k-1,j,k-1} &= ||\mathbf{x}_{k-1} - \mathbf{x}_j||^\beta ||\mathbf{x}_{k-1} - \mathbf{x}_k||^\alpha \\ \nabla_{\mathbf{x}_k} \xi_{k-1,j,k-1} &= 2||\mathbf{x}_{k-1} - \mathbf{x}_j||^2 \left(\mathbf{x}_k - \mathbf{x}_{k-1}\right) \\ &\quad - 2\left(\left(\mathbf{x}_k - \mathbf{x}_{k-1}\right) \cdot \left(\mathbf{x}_{k-1} - \mathbf{x}_j\right)\right) \left(\mathbf{x}_{k-1} - \mathbf{x}_j\right) \\ \nabla_{\mathbf{x}_k} \eta_{k-1,j,k-1} &= \alpha ||\mathbf{x}_{k-1} - \mathbf{x}_j||^\beta ||\mathbf{x}_k - \mathbf{x}_{k-1}||^{\alpha-2} \left(\mathbf{x}_k - \mathbf{x}_{k-1}\right)\end{aligned}$$

If $(p, q, r) = (k, j, k-1)$,

$$\begin{aligned}\xi_{k,j,k-1} &= ||\mathbf{x}_k - \mathbf{x}_{k-1}||^2 ||\mathbf{x}_k - \mathbf{x}_j||^2 - \left(\left(\mathbf{x}_k - \mathbf{x}_{k-1}\right) \cdot \left(\mathbf{x}_k - \mathbf{x}_j\right)\right)^2 \\ \eta_{k,j,k-1} &= ||\mathbf{x}_k - \mathbf{x}_j||^\beta ||\mathbf{x}_{k-1} - \mathbf{x}_k||^\alpha \\ \nabla_{\mathbf{x}_k} \xi_{k,j,k-1} &= 2||\mathbf{x}_k - \mathbf{x}_j||^2 \left(\mathbf{x}_k - \mathbf{x}_{k-1}\right) \\ &\quad + 2||\mathbf{x}_k - \mathbf{x}_{k-1}||^2 \left(\mathbf{x}_k - \mathbf{x}_j\right) \\ &\quad - 2\left(\left(\mathbf{x}_k - \mathbf{x}_{k-1}\right) \cdot \left(\mathbf{x}_k - \mathbf{x}_j\right)\right) \left(2\mathbf{x}_k - \mathbf{x}_j - \mathbf{x}_{k-1}\right) \\ \nabla_{\mathbf{x}_k} \eta_{k,j,k-1} &= \beta ||\mathbf{x}_{k-1} - \mathbf{x}_k||^\alpha ||\mathbf{x}_k - \mathbf{x}_j||^{\beta-2} \left(\mathbf{x}_k - \mathbf{x}_j\right) \\ &\quad + \alpha ||\mathbf{x}_k - \mathbf{x}_j||^\beta ||\mathbf{x}_k - \mathbf{x}_{k-1}||^{\alpha-2} \left(\mathbf{x}_k - \mathbf{x}_{k-1}\right)\end{aligned}$$

If $(p, q, r) = (i, k, j)$,

$$
\begin{aligned}
\xi_{i,k,j} &= ||\mathbf{x}_{j+1} - \mathbf{x}_j||^2 ||\mathbf{x}_k - \mathbf{x}_i||^2 - \left((\mathbf{x}_{j+1} - \mathbf{x}_j) \cdot (\mathbf{x}_k - \mathbf{x}_i)\right)^2 \\
\eta_{i,k,j} &= ||\mathbf{x}_k - \mathbf{x}_i||^\beta ||\mathbf{x}_j - \mathbf{x}_{j+1}||^\alpha \\
\nabla_{\mathbf{x}_k} \xi_{i,k,j} &= 2||\mathbf{x}_{j+1} - \mathbf{x}_j||^2 \left(\mathbf{x}_k - \mathbf{x}_i\right) \\
&\quad - 2\left((\mathbf{x}_{j+1} - \mathbf{x}_j) \cdot (\mathbf{x}_k - \mathbf{x}_i)\right)\left(\mathbf{x}_{j+1} - \mathbf{x}_j\right) \\
\nabla_{\mathbf{x}_k} \eta_{i,k,j} &= \beta ||\mathbf{x}_j - \mathbf{x}_{j+1}||^\alpha ||\mathbf{x}_k - \mathbf{x}_i||^{\beta-2} \left(\mathbf{x}_k - \mathbf{x}_i\right)
\end{aligned}
$$

With all these cases covered, one may back substitute to (D.1) $\sim$ (D.4) to acquire the exact gradient by summing over the $4\,(N-3)$ pairs of indices.

References

[1] Amir Beck. "Chapter 8: Primal and Dual Projected Subgradient Methods". In: *First-Order Methods in Optimization*, pp. 195–245. DOI: 10.1137/1.9781611974997.ch8. eprint: https://epubs.siam.org/doi/pdf/10.1137/1.9781611974997.ch8. URL: https://epubs.siam.org/doi/abs/10.1137/1.9781611974997.ch8.

[2] Gregory Buck and Jeremey Orloff. "A simple energy function for knots". In: *Topology and its Applications* 61.3 (Feb. 1995), pp. 205–214. DOI: 10.1016/0166-8641(94)00024-w.

[3] Endre Süli. *B6.1 Numerical Solution of Partial Differential Equations.* Lecture Note. 2021.

[4] Chris Yu, Henrik Schumacher, and Keenan Crane. "Repulsive Curves". In: *ACM Transactions on Graphics* 40.2 (Apr. 2021), pp. 1–21. DOI: 10.1145/3439429.