2021

Steepest Descent, Elastic Energy, and Stability in infinite Dimensional Hilbert Manifolds

Scott W. Rexford
scottevil_421@yahoo.com

Follow this and additional works at: https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations

Part of the Mathematics Commons

Recommended Citation
https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations/7592

This Dissertation/Thesis is brought to you for free and open access by the Graduate Research & Artistry at Huskie Commons. It has been accepted for inclusion in Graduate Research Theses & Dissertations by an authorized administrator of Huskie Commons. For more information, please contact jschumacher@niu.edu.
Consider constant-speed planar curves \( \gamma = (x, y) : [0, 1] \to \mathbb{R}^2 \) subject to \( \gamma(0) = (0, 0) \) and a prescribed value of \( x(1) \), but with \( y(1) \) unconstrained. We analyze the existence and the stability of critical curves of the elastic energy. Elastic curves are here thought of as points in an infinite-dimensional Sobolev manifold where the intrinsic gradient of the elastic energy vanishes. The admissible curves subject to constraints are members of a Riemannian submanifold using the ambient metric. A curve is stable with respect to the negative gradient flow if it is global or local minimum, and unstable otherwise. We analyze the stability of the classical Euler-Bernoulli elastica with vanishing curvature at the endpoints, and show that all are unstable. Our geometric treatment takes full advantage of the Riemannian structure. There is a smooth scalar field defined on the whole manifold that corresponds to the classical Lagrange multiplier values when evaluating the scalar field at the critical points. We illustrate the delicate nature of the stability issues here with natural examples in the presence of non-convex isoperimetric constraints. In the case of the constraint \( x(1) = 0 \), a half of the Euler figure eight is a candidate to be a critical point. It is of course possible to rotate the curve and still satisfy the constraints given in this case. Only the two vertical
positions turn out to be critical, and this has implications concerning mountain-pass issues when examining traversals between the upwards and the downwards globally minimizing straight-line segments, while insisting that the maximal elastic energy of the curves during the transition is as small as possible. The numerically delicate nature of stability involving isoperimetric constraints in the infinite dimensional Sobolev space setting is illustrated.
STEPEST DESCENT, ELASTIC ENERGY, AND STABILITY IN INFINITE DIMENSIONAL HILBERT MANIFOLDS

BY

SCOTT REXFORD
© 2021 Scott Rexford

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICAL SCIENCES

Dissertation Director:
Anders Linnér
ACKNOWLEDGEMENTS

I would first acknowledge the direction and contributions herein from Anders Linnér, and the support from everyone at NIU who put up with me over the years, including Doug Bowman, Sien Deng, Zhuan Ye, Gleb Sirotkin, Hamid Bellout, Dan Grubb, Jeff Thunder, Bernie Harris, John Wolfskill, Qingkai Kong, Deepak Naidu, Alastair Fletcher, Nathan Krislock, Joseph Stephen, Ilya Krishtal, Michael Geline, Yunier Bello-Cruz, Harvey Blau, Alon Regev, James Benson, Peter Waterman, Richard Blecksmith and Bela Erdelyi. Secondly to the people at Kishwaukee college where my education began who helped guide me down this path, including Laurie Hoecherl, Georgina Traush, Neil McKenna and Matt Gillespie, Chris Cullop, Roger Layton, Steve Dunn and George Overby. And finally, to people who have built up the subject to make this work possible, who aside from Anders Linnér I have never met, including, but definitely not limited to, Johan and Daniel Bernoulli, Leonhard Euler, Carl Jacobi, John Milnor, Joel Langer, David Singer, and Joe Jerome. Also, a special acknowledgement for Michael Johnson for contributions to elasticity theory, and also many helpful suggestions.
DEDICATION

To all of the committee members, who actually have to read this,
and Joseph “Buck” Stephen who unfortunately cannot.

Also, in memory of Pretty Kitty (who was given many names over the years) who succame
to old age on March 22, 2021 — she was in her early twenties. Believe it or not, she made
significant contributions to this work, and was a very clever cat.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
<th>vi</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF APPENDICES</td>
<td>xi</td>
</tr>
</tbody>
</table>

1.1 General Ideas and Motivating Examples in Euclidean 3-Space .......................... 1
1.2 Notation and general structures ................................................................. 12
1.3 Manifold Structure .................................................................................... 16
1.4 The Curve Straightening Flow ....................................................................... 20

## Chapter 2: Paths in $\mathcal{M}_L$ and Analysis of Critical Points
2.1 The Notion of Stability and Critical Point Classification ............................... 24
2.2 The Existence of Critical (and ‘Nearly’ Critical) Points ................................. 25
2.3 The Relationship Between Length and Elliptic Modulus .................................... 41
2.4 Analysis of the Critical Points in $\mathcal{M}_\infty$ ...................................... 42
2.5 Analysis of the Critical Points in $\mathcal{M}_L$ for $L < \infty$ ......................... 50
2.6 Regarding Critical Points with $y(1) \neq 0$ ................................................... 61
2.7 Searching for Special Mountain Pass Paths in $\mathcal{M}_L$ ................................. 63
2.8 Geodesic Lines in $\mathcal{M}_L$ and the Geodesic Spray .................................... 66

## Chapter 3: Additional Constraints and The Corresponding Submanifolds of $\mathcal{M}_L$
3.1 Submanifolds of Intersection ............................................................................. 68
3.2 Affine Subspaces Induced by Evaluation Functionals .......................................... 71
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3</td>
<td>73</td>
</tr>
<tr>
<td>3.4</td>
<td>84</td>
</tr>
<tr>
<td>4</td>
<td>88</td>
</tr>
<tr>
<td>4.1</td>
<td>88</td>
</tr>
<tr>
<td>4.2</td>
<td>91</td>
</tr>
<tr>
<td>5</td>
<td>96</td>
</tr>
<tr>
<td>5.1</td>
<td>96</td>
</tr>
<tr>
<td>5.2</td>
<td>99</td>
</tr>
<tr>
<td>5.3</td>
<td>105</td>
</tr>
<tr>
<td>5.4</td>
<td>107</td>
</tr>
<tr>
<td>5.5</td>
<td>107</td>
</tr>
<tr>
<td>5.6</td>
<td>108</td>
</tr>
<tr>
<td>5.6.1</td>
<td>109</td>
</tr>
<tr>
<td>5.6.2</td>
<td>110</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>114</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>118</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>The figure from the opening example to Milnor’s classic text, <em>Morse Theory</em>.</td>
</tr>
<tr>
<td>1.2</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>A geodesic meridian on a surface of revolution, and a non-meridian winding geodesic.</td>
</tr>
<tr>
<td>1.3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>From left to right, some members of the projected gradient field, some typical descent trajectories of (1.3) on the Milnor torus, and geodesic descents in the planes $x_1, x_2 = 0$.</td>
</tr>
<tr>
<td>1.4</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>A gradient descent (black) and two geodesic paths (gray) joining a point to the global minimizer.</td>
</tr>
<tr>
<td>1.5</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Gradient descents for $\alpha &lt; \frac{3\sqrt{3}}{2}$ on the left, and $\alpha &gt; \frac{3\sqrt{3}}{2}$ on the right. For alpha above the bifurcation, some planar sections height=const. cut the surface in two components, but below the bifurcation, all such sections only cut the surface in a single component.</td>
</tr>
<tr>
<td>1.6</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>A plot of the graph of a quartic surface and the plane at the mountain pass level and a contour plot. In the contour plot, the dotted line is the intersection with the plane plot, which is a ‘plateau,’ and the dashed line is the mountain pass path.</td>
</tr>
<tr>
<td>1.7</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>In the figure typical planar curves $\gamma$ (black) with turning angles $\theta$ (gray) satisfying constraint (1.17) are plotted for each of the two distinguishable cases, with $C = 0$ above and $C = 1$ below.</td>
</tr>
<tr>
<td>1.8</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>A depiction of the transported frame used to recover a planar curve $\gamma$ from its indicatrix, via equation (1.20).</td>
</tr>
<tr>
<td>2.1</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Plotted are elastic loops with indicatrices $\xi_1, \xi_2, \xi_{10}$, with $n - 1$ inflections, and $n$ self intersections.</td>
</tr>
<tr>
<td>2.2</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Plotted are elastic arches, with indicatrices $\xi_1, \xi_2, \xi_{10}$, with $n - 1$ inflections, and no self intersections.</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>2.3</td>
<td>Plotted in black are dn curvature loops with indicatrices $\eta_n$ with $n = 1, 2, 5$, having $n - 1$ inflections and $n$ self intersections. Also, plotted in gray are the corresponding cn curvature loops $\xi_n$, for comparison. For $k$ near $1$ these become indistinguishable to the curve straightening flow, as both functionals $F_L$ and $G_L$ are even in $\theta$.</td>
</tr>
<tr>
<td>2.4</td>
<td>Plots comparing cn and dn loops and corresponding squared curvatures $\kappa^2$ for $L = 1.7, 3$, in gray and black, respectively. The dn loop cannot ‘relax’ at the endpoints, and must more evenly distribute the curvature; however, for $k \approx 1$, the endpoint curvature nearly does vanish. While the cn loop attains the largest curvature of the two over the interval $[0, 1]$, occurring at the midpoint, the totality of squared curvature is lesser for the cn loop. As $k$ gets near $1$, these plots become indistinguishable.</td>
</tr>
<tr>
<td>2.5</td>
<td>Plots comparing cn and dn double loops, with darker curves for increasing $k$. Each loop becomes increasingly similar as $k$ increases, while the even number loops, (from left to right,) are opposing.</td>
</tr>
<tr>
<td>2.6</td>
<td>Plots of $\xi$ as per (2.9) with increasing gray level for $1 \leq n \leq 10$, where in this setting, regardless of $p$, the associated planar curve consists of $n/2$ windings of the Euler figure $8$ — this is taken to mean an additional half covering of the Euler figure $8$ when $n$ is odd. Each curve has total length $1$, which is of course the distance to the line $x = 1$, which is included in the figure for scale.</td>
</tr>
<tr>
<td>2.7</td>
<td>Construction of $\theta_t(s)$ for the proof of Theorem 2.5.1 in the leftmost plot, and also plots of discrete sets along the path for $n = 1, 3$.</td>
</tr>
<tr>
<td>2.8</td>
<td>On the left, a discrete set of points in the half of the of the path from $\xi_1$ to $\xi_1^*$ with decreasing darkness. On the right, a plot with two points along the path for even $n$, where the final $y$-coordinate always is zero.</td>
</tr>
<tr>
<td>2.9</td>
<td>Translating the gray section of the curve suggests that in totality $\theta_t(s)$ should have the same energy as the critical elastica — indeed, this is proven analytically in Theorem 2.5.2.</td>
</tr>
<tr>
<td>2.10</td>
<td>Discrete plots pertaining to three mountain pass candidates. From left to right, a variation in elliptic modulus $0 \leq k &lt; k_\infty$ with gaps ranging from $0 \leq g &lt; 1$ with half a rotation of the half figure $8$ (which is similar to the path used in the proof of Theorem 2.5.1), the path used in the proof of Theorem 2.5.2 with rotated ambient segment joining the extremities, and a gradient descent from the half figure $8$ curve with the largest projected gradient among curves on the mountain pass plateau, where via a half rotation this could produce an ascent-descent pair.</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>3.1</td>
<td>69</td>
</tr>
<tr>
<td>3.2</td>
<td>70</td>
</tr>
<tr>
<td>3.3</td>
<td>81</td>
</tr>
<tr>
<td>5.1</td>
<td>98</td>
</tr>
<tr>
<td>5.2</td>
<td>100</td>
</tr>
<tr>
<td>5.3</td>
<td>100</td>
</tr>
<tr>
<td>5.4</td>
<td>101</td>
</tr>
<tr>
<td>5.5</td>
<td>102</td>
</tr>
<tr>
<td>5.6</td>
<td>103</td>
</tr>
<tr>
<td>5.7</td>
<td>103</td>
</tr>
<tr>
<td>5.8</td>
<td>104</td>
</tr>
</tbody>
</table>

3.1 Some 2-surfaces in $\mathbb{R}^3$ where intersection fails to be transversal at certain points, at which the intersection may fail to be a submanifold.

3.2 Some affine sections of 2-surfaces in $\mathbb{R}^3$ which intersect in submanifolds, which may be disconnected.

3.3 Some cn curvature curves and corresponding surfaces defined by the system (3.32); each curve corresponds a lattice point in the intersection of the three families of surfaces. There is a fourth dimension in the parameter space which may enable the construction of cn curvature paths between these curves.

5.1 The unwinding of a critical elastica under the unconstrained flow with decreasing gray level.

5.2 A tightly wound single loop flows near the critical elastica $\xi_1$, and the gradient gets very small. The flow nearly stops when close to the saddle point.

5.3 For more loops, similar behavior is seen if tightly wound.

5.4 With enough length, the loop freely unwinds; the uppermost curve is the final iterate plotted, and this quickly approaches (until nearby) the upper straight line segment which is a global minimizer. Here $\theta_0$, in dashed black, is a smooth function constructed using tangent inverse.

5.5 The behavior is a bit different with an even number of loops, but with enough length, it still freely unwinds. On the left, a discrete set of points flowing toward the saddle point $\xi_2$. On the right another discrete set of points from the same discretization; after slowing near the saddle point, the flow increases speed and the curve unwinds, where a further curve down the flow is plotted in black.

5.6 With a slightly different configuration, tightly wound loops still flow near critical elastica.

5.7 Beginning with winding number two, tightly wound. This will very slowly flow past the 'nearly' critical dn curvature double loop.

5.8 Plotted is the flow of a circular loop with winding number four, which was quite costly to compute. Below the initial curve (dotted black,) final curve in the flow (gray,) and the dn curvature indicatrix $\eta_4$ (black) are plotted. Eventually this will flow near $\eta_4$. 
<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.9</td>
<td>104</td>
</tr>
<tr>
<td>5.10</td>
<td>105</td>
</tr>
<tr>
<td>5.11</td>
<td>106</td>
</tr>
<tr>
<td>5.12</td>
<td>108</td>
</tr>
<tr>
<td>5.13</td>
<td>109</td>
</tr>
<tr>
<td>5.14</td>
<td>111</td>
</tr>
<tr>
<td>5.15</td>
<td>112</td>
</tr>
<tr>
<td>5.16</td>
<td>113</td>
</tr>
<tr>
<td>A.1</td>
<td>120</td>
</tr>
<tr>
<td>A.2</td>
<td>123</td>
</tr>
</tbody>
</table>

5.9 On the left is a plot of the energy $F(\theta_n)$, and on the right the invariant $x(1)$ for initial curve a smooth single non-elastic loop with $L \approx 3$ and step size $h = .01$ with the AB3 discretization. Aside from an initial jump in the invariant, there is just small self correcting oscillation. This level of accuracy in long term invariant preservation is typical for the Adams-Bashforth methods.

5.10 Discretizations for fixed terminal angles converging to cn sections. On the left initial points in the discretizations are plotted, beginning with the dotted triple loop. On the right the same discretization is run to a limiting black curve, a critical cn curvature elastica.

5.11 Discretizations for fixed terminal angles with $\Delta = 2\pi, 5\pi, 16\pi$ the respective angular variations.

5.12 Plots of iterates on the left, and objective value on the right for discretization of the basic problem with $x_0 = \sqrt{t}$.

5.13 A drift in the invariant is typically seen for the attracting flow, (as in the list plot on the left,) unlike the curve straightening flow which enjoys long term invariant preservation. On the right a path joining two points is produced numerically with the attracting flow on a monkey saddle in $H = \mathbb{R}^3$.

5.14 Above the gradient projection is depicted with discrete sets for the linear segment on the left, and the projected path on the right. Below a discrete collection of indicatrices along the resulting path joining the two points. It is noted that one point is a $C^1$ pasting of lines and circles, which is not actually in $H$, (like many of our initial curves for discretization,) but the method still works well.

5.15 Beginning with winding number three and enough slack to freely unwind. Also included in the plots are some associated non-critical dn curvature elastica.

5.16 A gradient ascent run in $M_\infty$, where as in all plots the curves darken with flow time. This endlessly winds with increasing angular variation and energy.

A.1 The unit ellipse.

A.2 The functions $cn$, $sn$, $dn$ plotted over one period for $k = .25$ on the left, and $k = .9$ on the right. These converge to cosine, sine and 1, respectively, as $k \downarrow 0$. 
<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.1 Early figures from the study of elastica, from left to right given by Galileo, Leibniz, J. Bernoulli, and Euler, respectively.</td>
<td>139</td>
</tr>
<tr>
<td>C.2 Sketches of elastica from Euler’s work on the subject.</td>
<td>140</td>
</tr>
<tr>
<td>C.3 A collection of elastic loops for $L = 1.4$ with darker shading for increasing energy. Each has the same elliptic modulus $k \approx .99999$, and containing $n$ self-similar $1/n$ scalings for $n &gt; 1$.</td>
<td>144</td>
</tr>
<tr>
<td>C.4 A collection of elastic arches for $L = 1.1$ with darker shading for increasing energy. Each has the same elliptic modulus $k \approx .08985$, and containing $n$ self-similar $1/n$ scalings for $n &gt; 1$.</td>
<td>144</td>
</tr>
<tr>
<td>C.5 A collection of pseudo-elastic loops for $L = 1.35$ with darker shading for increasing energy. Each has the same elliptic modulus $k \approx .999997$, and containing $n$ self-similar $1/n$ scalings for $n &gt; 1$.</td>
<td>145</td>
</tr>
<tr>
<td>C.6 A variation of elastica in $k$ with $0 &lt; k &lt; 1$, and darker curves plotted for larger $k$. The collection of all such curves is the totality of elastica in the manifolds $\mathcal{M}_L$, up to scaling and reflection for $n &gt; 1$, represented here with unit length.</td>
<td>147</td>
</tr>
<tr>
<td>C.7 A variation of dn curvature elastica in $k$ with $0 \leq k &lt; 1$ and $L = 1$, with darker curves plotted for larger $k$. This collection curves is the totality of dn curvature elastica in the manifolds $\mathcal{M}_L$, up to scaling and periodicity for $n &gt; 1$, represented here with unit length. At $k = 0$ we have a circle, and as $k \uparrow 1$, these curves approach the cn curvature elastica.</td>
<td>148</td>
</tr>
<tr>
<td>C.8 Plotted are $k \mapsto 2E(k)$, $k \mapsto K(k)$ and the auxiliary function $f(k) = 2E(k) - K(k)$ corresponding to $\mathcal{M}<em>\infty$, with the former two graphs intersecting in $k</em>\infty$, and the latter function having this unique real root for $0 &lt; k &lt; 1$.</td>
<td>149</td>
</tr>
</tbody>
</table>
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>JACOBI ELLIPTIC FUNCTIONS, AND ELLIPTIC INTEGRALS</td>
<td>118</td>
</tr>
<tr>
<td>A.1</td>
<td>Trigonometry on an Ellipse</td>
<td>119</td>
</tr>
<tr>
<td>A.2</td>
<td>The Jacobi Elliptic Functions</td>
<td>121</td>
</tr>
<tr>
<td>A.3</td>
<td>Elliptic Integrals and the Jacobi Amplitude</td>
<td>124</td>
</tr>
<tr>
<td>B</td>
<td>SOLVING THE PENDULUM EQUATION</td>
<td>130</td>
</tr>
<tr>
<td>B.1</td>
<td>Transformation to Algebraic Differential Operator</td>
<td>131</td>
</tr>
<tr>
<td>B.2</td>
<td>Curvature Types for Pendulum Equation Solutions</td>
<td>134</td>
</tr>
<tr>
<td>C</td>
<td>ELASTIC CURVES</td>
<td>138</td>
</tr>
<tr>
<td>C.1</td>
<td>A brief overview of the history of elasticity theory</td>
<td>139</td>
</tr>
<tr>
<td>C.2</td>
<td>Elastica in the manifolds $M_L$</td>
<td>142</td>
</tr>
<tr>
<td>C.2.1</td>
<td>General form of the indicatrices</td>
<td>142</td>
</tr>
<tr>
<td>C.2.2</td>
<td>The Elliptic Modulus for a Given Gap</td>
<td>145</td>
</tr>
<tr>
<td>C.2.3</td>
<td>Explicit Energy of the Elastica</td>
<td>149</td>
</tr>
<tr>
<td>D</td>
<td>NUMERICAL DISCRETIZATION OF DIFFERENTIAL EQUATIONS</td>
<td>151</td>
</tr>
<tr>
<td>D.1</td>
<td>Initial Value Problems</td>
<td>152</td>
</tr>
<tr>
<td>D.2</td>
<td>Numerical Discretization of Initial Value Problems</td>
<td>153</td>
</tr>
<tr>
<td>D.3</td>
<td>Linear Multi-Step Methods</td>
<td>154</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION: THE SETTING, THE MANIFOLDS, AND THE FLOW

1.1 General Ideas and Motivating Examples in Euclidean 3-Space

In a Hilbert space $H$ on which a functional $f : H \rightarrow [0, \infty)$ is defined, consider the functional restricted to a manifold in $H$, induced by $g^{-1}(\{c\})$ for $g : H \rightarrow \mathbb{R}$ a smooth functional. A famous instructive example is the height functional on a torus from Milnor’s classic text *Morse Theory* [1], (with borrowed figure below.)

![Diagram 1.](image)

Figure 1.1: The figure from the opening example to Milnor’s classic text, *Morse Theory*.

To be explicit, take $H = \mathbb{R}^3$ and consider the *height functional*

$$f(x_1, x_2, x_3) = x_3$$  \hspace{1cm} (1.1)
acting on the manifold \( T = g^{-1}(\{1\}) \), a torus, induced by the smooth functional

\[
g(x_1, x_2, x_3) = \left(2 - \sqrt{x_2^2 + (x_3 - 3)^2}\right)^2 + x_1^2.
\]  

(1.2)

For the restricted functional \( f : T \rightarrow [0, \infty) \) there are four exceptional points, which are ‘critical’ with respect to the height, labelled in Milnor’s figure (Figure 1.1) as

\[
p = (0, 0, 0), \quad q = (0, 0, 2), \quad r = (0, 0, 4), \quad s = (0, 0, 6),
\]

although Milnor did not give explicit form or coordinates. Of course \( V \) in Milnor’s figure is the plane \( x_3 = 0 \) in this explicit formulation. Milnor was primarily concerned with the topological type of the sublevel sets\(^1\), while our concern here is geometric and analytical.

The critical point types here are rather clear, where \( p \) is a global minimum, \( q \) are \( r \) are saddle points, and \( s \) is a global maximum. Of course \( f \) cannot be lesser than the height of \( f(p) = 0 \), nor exceed the height of \( f(s) = 6 \) on \( T \), where \( p, s \) are global extrema. Also, considering the curves of intersection with the planes \( x_1 = 0 \) and \( x_2 = 0 \), there are four circular geodesic curves. At both \( q, r \) one of these circles attains minimal height, and another maximal height, so it is evident that these are critical points of the saddle type.

To be clear, a geodesic curve in a Riemannian manifold is a generalization of a straight line in a curved space, and perhaps the simplest description is a path which locally minimizes distance. We consider submanifolds of ambient spaces where at each point the bases for the tangent and normal spaces generate the ambient space. For a geodesic curve, the acceleration vector is always in the normal component, so there is no change in speed or turning in the manifold. Typically geodesics are defined to have unit speed, and thus are parametrized by arclength. Moreover, a pregeodesic curve is more general as it need not have constant

\(^{1}\)An \( \alpha \)-sublevel set of \( f \) is given \( \{ x : f(x) \leq \alpha \} \).
speed, and can be defined by the condition $\dot{\gamma}(t) \in \text{span}\{\dot{\gamma}(t), \nu\}$ for manifold normal vector $\nu$. In either case there is no turning in the manifold along the curve except for that in the ambient space which is necessary to stay on the manifold, and in either case the curve describes geodesic path which is taken as a generalization of a line in Euclidean space. The quintessential example seems to be a great circle on a sphere. A pregeodesic differs from a geodesic only in parametrization, and in theory can be reparametrized to have unit speed, and thus no tangential component of acceleration. However, finding a unit speed reparametrization is non-trivial. At every point in a Riemannian manifold there is a geodesic line in every direction, the so called geodesic spray, in its simplest description.

To formally define criticality, and investigate extrema of $f$ over $\mathcal{T}$, a tangent vector field is defined by projecting the gradient

$$\nabla f(x_1, x_2, x_3) = (0, 0, 1) = e_3$$

of the height functional onto the tangent space of the torus. The projected gradient vector

$$\nabla^\pi f = \nabla f - \lambda \nabla g$$

is constructed by choosing a scalar $\lambda$ so that $\nabla^\pi f \in T_p \mathcal{T}$ at any given $p = (x_1, x_2, x_3) \in \mathcal{T}$. This happens if $\lambda$ satisfies

$$0 = \left\langle \nabla^\pi f, \frac{\nabla g}{||\nabla g||} \right\rangle = \left\langle \nabla f - \lambda \nabla g, \frac{\nabla g}{||\nabla g||} \right\rangle,$$

and solving for $\lambda$ gives

$$\lambda = \frac{\langle \nabla f, \nabla g \rangle}{\langle \nabla g, \nabla g \rangle} = \langle e_3, \nu \rangle$$

Footnote: In general $T_p \mathcal{M}$ denotes the tangent space to a manifold $\mathcal{M}$ at a point $p \in \mathcal{M}$. 
for $\nu = \nabla g/||\nabla g||$. As $\nu = \pm e_3$ at exactly four exceptional points, namely $p$, $q$, $r$, $s$; these are the only critical points in the sense that $\nabla^\pi f$ vanishes.

Investigation of the critical points can be done via consideration of the negative projected gradient flow, induced by the solution trajectories of the differential equation

$$\frac{d\gamma}{dt} = -\nabla^\pi f(\gamma) \quad \text{with} \quad \nabla^\pi f = e_3 - \langle e_3, \nu \rangle \nu.$$  

(1.3)

Critical points of $f$ over $\mathcal{T}$ are rest points of the flow equation. In terms of dynamic stability local minima are stable rest points, while saddle points and local maxima are unstable rest points. However, saddle points and local maxima differ in the sense that the latter are repelling for the gradient descent while stable for the gradient ascent, where the flow is in the positive gradient direction. There is of course a duality for considerations of the gradient ascent. However, the ascent and descent are typically much differently behaved, unless the manifold exhibits symmetry with respect to height, such as the Milnor torus example.

Remark 1.1.1. Heuristically, one may imagine the negative projected gradient flow as the paths of water droplets running down the surface with gravity acting downward in the vertical direction, with this force acting to decrease height. Excluding adhesive and cohesive forces of water, this should be a fairly accurate model in terms of the path of the water droplets, but the parametrization would surely differ, as the gradient vector field vanishes at critical points, and the trajectories, when convergent, do not converge in finite time. To the contrary, a droplet of water will quickly reach the bottom of a (physical non-porous) surface, or perhaps get stuck at a saddle point. Of course a given choice of inner product has its bearing on the physical reality of such a model, and we do not set out to model this here.

Regarding the behavior of the flow (1.3) on $\mathcal{T}$, for any initial point $x_0 = (x_1, x_2, x_3)$ with $x_1 \neq 0$ and $x_2 \neq 0$, trajectories under the gradient descent flow to the global minimum $p$, so all points excluding a zero measure set on $\mathcal{T}$ flow to $p$. It is noted that for all planes
containing the $x_3$ axis, only those with $x_1 = 0$ or $x_2 = 0$, cut $\mathcal{T}$ in geodesic curves. For these two vertical planes which cut $\mathcal{T}$ in a geodesic curve, the gradient descent does not leave this path. On the geodesic curve containing $p$ and $s$, with the exception of $s$ all points flow to $p$. These claims are all intuitively clear, and none are difficult to prove.

As seen on the Milnor torus $\mathcal{T}$, it is possible for the geodesic in the negative gradient direction to agree with the steepest descent path. This is atypical, but may occur under symmetry. For instance, on a surface of revolution the meridians have this property. Taking a surface of revolution with profile curve $\eta(t) = (u(t), 0, v(t))$ to be rotated about the $x_3$-axis with chart

$$x(t, \theta) = (u(t) \cos(\theta), u(t) \sin(\theta), v(t)),$$ 

(1.4)
a meridian is a $t$ curvilinear curve $\gamma(t) = x(t, \theta_0)$. This curve is in general a pregeodesic, as the planar section containing the vertical axis and the meridian contains the curve tangent, curve normal, and surface normal vectors, (with the latter two parallel,) and the acceleration
of a space curve is always in the span of the unit tangent and normal vectors. Moreover, a gradient descent with respect to the height (1.1) cannot leave the meridian upon which the initial point resides, as the gradient vector field is everywhere parallel to the tangent vector of the meridian curve. However, the parametrizations of these curves do not agree, and may vastly differ.

Figure 1.3: From left to right, some members of the projected gradient field, some typical descent trajectories of (1.3) on the Milnor torus, and geodesic descents in the planes $x_1, x_2 = 0$.

Often the gradient descent and the geodesic in that direction follow different paths to the same point. Of these one may ask which is a more advantageous route, which would correspond to optimality of some path functional, such as path length. On the Milnor torus there are examples of geodesic paths to the global minimizer which only agree with the steepest descent trajectory at the terminal points, (taking the path to the minimizer as a limit point of the descent, as it does not converge in finite time.) For example, in Figure 1.4 the steepest descent curve (red) and a geodesic (yellow) start at the same point in the projected gradient direction (the projected gradient vector in black) and give a path to the global minimizer. Also, another geodesic (in blue) gives a path between these points. The plot also includes a set of paths in geodesic spray (magenta) plus normal and tangent vectors.
at the initial point. Of the three plotted curves, if competing with respect to length, the geodesic with same initial direction gives the shortest path to the minimizer.

![Figure 1.4](image)

Figure 1.4: A gradient descent (black) and two geodesic paths (gray) joining a point to the global minimizer.

A related geometric concept which has garnered interest in the last century is that of a mountain pass between distinct global minimizers. By a celebrated theorem of Courant, the so called *classical mountain pass theorem*, a *coercive* functional on $\mathbb{R}^n$ with distinct minimizers also has a saddle point at which the functional assumes a greater value. By the construction in the standard proof of this theorem, (for instance, as in [2]) there is an associated mountain pass path which is a minimax navigation between the minimizers which passes thru the saddle point. There are also different notions of a mountain pass path, for instance a *length minimizing mountain pass path* which (as the name suggests) is a curve $\gamma \in \Omega_{\xi_1, \xi_2}$ characterized by

$$\mathcal{L}(\gamma) = \inf_{\eta \in \Omega_{\xi_1, \xi_2}} \mathcal{L}(\eta).$$

(1.5)

---

3Here coercive is taken to mean that $f(x) \to \infty$ as $||x|| \to \infty$. 
A length minimizing mountain pass path does not necessarily pass through the saddle point; to see this, consider a diffeomorphism of the toy surface (1.6) below where the saddle point is shifted along the $x_2$-axis and the parabolic cross section is scaled to have a very small derivative — if sufficiently skewed as such, it may not be advantageous, lengthwise, to pass thru the saddle point. This theorem has generalizations to the infinite dimensional settings, with applications in the calculus of variations and partial differential equations; for instance, in the partial differential equations text [3] existence theorems involving both gradient flows and mountain passes are covered, and an excellent source of history and generalizations of the mountain pass theorem is the book [2].

For a geometrically intuitive example, again take $H = \mathbb{R}^3$ and again consider the height functional (1.1) over a parametrized family of quartic polynomial surfaces

$$g_\alpha(x_1, x_2, x_3) = \frac{x_1^4}{4} - \frac{x_1^2}{2} - \frac{x_1^3}{\alpha} + x_2^2 - x_3$$

for $\alpha > 0$, and consider $\mathcal{M}_\alpha = g_\alpha^{-1}[0]$. Depending on $\alpha$, the height functional $f$ has one, two or three critical points. The possibilities are a distinct global minimum, a distinct global minimum and a saddle point, or two local minimums and a saddle point. When there are three critical points, there are two local minima, and a mountain pass saddle point. The analysis of the critical points here is simple, for by construction, in the plane $x_2 = 0$ this section of the manifold $\mathcal{M}_\alpha$ is the graph of a quartic polynomial, and since quadratic in $x_2$ the critical points of this surface are exactly those critical for the polynomial in this section. For the cubic derivative the Cardano-Tartaglia formula for its roots involves a quantity $\sqrt{\frac{729}{\alpha^2} - 108}$ and the sign under the radicand determines the number of roots, where there is a distinct real root when $0 < \alpha < 3\sqrt{3}/2$, and three distinct real roots when $\alpha > 3\sqrt{3}/2$. When $\alpha > 3\sqrt{3}/2$ the mountain pass path between the local minimizers is clear, provided by the surface curve $\gamma_\alpha(t) = (t, 0, g_\alpha(t, 0, 0))$. This is a pregeodesic on $\mathcal{M}_\alpha$ which furnishes a
length minimizing mountain pass path in the sense of (1.5) since it is a distance minimizing geodesic path. To see that $\gamma_\alpha$ is a pregeodesic, one could verify that $\ddot{\gamma}(t) \in \text{span}\{\dot{\gamma}(t), \nu\}$ by calculating $\ddot{\gamma}_\alpha = \frac{\langle \ddot{\gamma}_\alpha, \dot{\gamma}_\alpha \rangle}{\langle \gamma_\alpha, \gamma_\alpha \rangle} \gamma_\alpha + \frac{\langle \gamma_\alpha, \nu \rangle}{\langle \nu, \nu \rangle} \nu$ for surface normal $\nu(t)$, but this is seen to be a pregeodesic by an argument similar to that for a geodesic meridian on a surface of revolution, for the surface normal and the curve’s velocity and acceleration vectors all clearly reside in the plane $x_2 = 0$. Moreover, in the same way this pregeodesic agrees with the gradient descent, and likewise with the gradient ascent. It may be said that this mountain pass is an ascent-descent pair between the local minima, as neither the descent or ascent ever leave this path — rather both follow it, but in infinite time.

Figure 1.5: Gradient descents for $\alpha < \frac{3\sqrt{3}}{2}$ on the left, and $\alpha > \frac{3\sqrt{3}}{2}$ on the right. For alpha above the bifurcation, some planar sections height=const. cut the surface in two components, but below the bifurcation, all such sections only cut the surface in a single component.

A topological characterization of a mountain pass point is given in terms of connectedness, or lack thereof, of the sublevel sets determined by the function value. In this sense, $\xi$ is a topological mountain pass point of $f : X \to \mathbb{R}$ if for every neighborhood $\mathcal{N}$ of $\xi$, the set

$$\mathcal{N} \cap \{x \in X : f(x) < f(\xi)\}$$

(1.7)
is disconnected. Our toy example makes the general idea clear, where in Figure 1.6 the sublevel sets are disconnected below the saddle point height, and moreover, the components lose convexity above this level.

![Figure 1.6: A plot of the graph of a quartic surface and the plane at the mountain pass level and a contour plot. In the contour plot, the dotted line is the intersection with the plane plot, which is a ‘plateau,’ and the dashed line is the mountain pass path.](image)

In this thesis we consider these notions in an infinite dimensional Hilbert space with an embedded Riemannian manifold. The analysis of the critical points in such a setting is non-trivial. An understanding of this structure sheds light on the qualitative nature of the flow. We consider a particular setting which has implications on elasticity dynamics by providing a model for the evolution of an elastic wire or beam released from a certain configuration while constrained so that one endpoint is fixed, and the other is free to slide on a vertical line. This is done by representing indicatrices for the curvatures in a Sobolev manifold, where the associated functional for the curve straightening flow is the squared curvature functional. The physical interpretation is that the dynamics under the curve straightening flow purport to give some qualitative description of the dynamics of an elastic wire. This makes use of the least action principle applied to the elastica problem by Daniel Bernoulli, also in application of Hooke’s law. No claim is made regarding the physical reality of the
description, particularly with freedom to choose an inner product, but some meaningful qualitative description in elasticity dynamics is sought.

For us the height functional is the elastic energy, (often called the bending energy,) which is taken proportional to the integral of the squared curvature of the curve. The existence of energy-minimizing curves has been known for non-linear splines of fixed length going back to work by Jerome [4]. In the spline case the curves pass through a given sequence of points in a prescribed order. The case of smooth periodic curves has also been of considerable interest as exemplified in Langer-Singer [5]. More recently there is a renewed interest in both existence and stability; see Borbély-Johnson [6] and [7]. Some recent work has taken a different approach, such as [8] where a curve straightening flow is considered, and [9] where interestingly a curve ‘shortening−straightening’ flow is considered, where the associated functional is the sum of the length and energy functionals; there is a stark contrast to our approach, where PDE methods are utilized and ‘gradient flow’ is not used in the technical sense, as the action is not on a proper Hilbert manifold; while nice results have been obtained from PDE driven approaches, our approach is geometric and makes use of a Riemannian structure. In the above notation, for the case of interest we investigate, $H$ is a first order Sobolev space over the unit interval $[0,1]$, with $f$ the squared curvature functional acting on $H$. Here, members of $H$ are proportional to the curvature of associated planar curves, and $g$ is a non-linear functional which is chosen to impose certain constraints on the associated planar curves. The main setting of interest is specified in Section 1.2, and in Section 1.3 we define the associated infinite-dimensional manifold, which has a nontrivial structure with some mysterious properties. The ability for curves to ‘unwind’ towards a global minimum by reducing the elastic energy through steepest descent, or by other means, is not automatic here even with available freedom to unwind. We construct explicit paths, given in terms of Jacobi elliptic functions, which in fact correspond to curves of constant elastic energy.
This helps produce a path which ‘escapes’ from a neighborhood of a critical point, and this precludes the critical point from being a local minimum.

### 1.2 Notation and general structures

We consider the Sobolev space

$$H = \{ \theta | \theta : [0, 1] \rightarrow \mathbb{R} \text{ absolutely continuous, } \dot{\theta} \in L^2[0, 1] \}, \quad (1.8)$$

(square integrable in the sense of Lebesgue,) with inner product

$$\langle u, v \rangle_H = u(0)v(0) + \int_0^1 \dot{u}(s) \dot{v}(s) \, ds. \quad (1.9)$$

The functions $\theta \in H$ are the so called tangent indicatrices (singular tangent indicatrix) that correspond to the turning angles of constant speed planar curves of length $L$ of the form

$$\gamma(s) = (x(s), y(s)), \quad \text{with } \gamma : [0, 1] \rightarrow \mathbb{R}^2, \text{ and } ||\dot{\gamma}(s)||_{\mathbb{R}^2} = L \text{ for all } s. \quad (1.10)$$

We express the tangent vector of $\gamma$ in terms of $\theta$ as

$$\dot{\gamma}(s) = L(\cos(\theta(s)), \sin(\theta(s))) = Le^{i\theta(s)}, \quad (1.11)$$

(with the point of view that $\mathbb{R}^2$ and $\mathbb{C}$ are isomorphic.) By (1.8) and (1.11), one may write the acceleration as

$$\ddot{\gamma}(s) = L^2 ie^{i\theta(s)} \frac{\dot{\theta}(s)}{L} = L^2 i e^{i\theta(s)} \kappa(s), \quad (1.12)$$
where \( \kappa(s) = \frac{\theta'(s)}{L} \) denotes the \textit{signed curvature}, and we note that the second derivative satisfies \( ||\dot{\gamma}(s)||_{\mathbb{R}^2} = L^2|\kappa(s)| \). From this vantage, we consider planar curves with square integrable signed curvature in the sense of Lebesgue, and further impose the constraints

\[
\gamma(0) = (0,0) \quad \text{and} \quad x(1) = C \in \mathbb{R}. \tag{1.13}
\]

Integrating \( \dot{\gamma} \), by (1.11), one has

\[
\gamma(s) = (x(s), y(s)) = L \left( \int_0^s \cos(\theta(u)) du, \int_0^s \sin(\theta(u)) du \right). \tag{1.14}
\]

Using (1.14), the planar curve corresponding to \( \theta \in H \) has terminal coordinates

\[
x(1) = L \int_0^1 \cos(\theta(s)) ds, \quad y(1) = L \int_0^1 \sin(\theta(s)) ds. \tag{1.15}
\]

While (1.15) suggests a constraint for fixing each terminal coordinate, we only fix the first and leave the second free. However, the fixed length condition restricts the \( y \)-coordinate to the interval

\[
y(1) \in \left[ -\sqrt{L^2 - C^2}, \sqrt{L^2 - C^2} \right], \tag{1.16}
\]

and also imposes that \( |C| \leq L \) for such curves to exist. Using (1.15), we express the terminal point constraint in (1.13) as the integral constraint

\[
L \int_0^1 \cos(\theta(s)) ds = C. \tag{1.17}
\]

We thus consider level sets of a family of non-linear functionals on \( H \), indexed by \( L > 0 \). In particular, in (1.13) we are interested in constants \( C = 0 \), or \( C = 1 \), so planar curves with
initial point the origin and terminal point on one of the lines $x = 0$ or $x = 1$, since these are the only essentially distinguishable cases by scaling and rigid motions.

It is worth noting that with the normalization $C = 1$ for $C \neq 0$ in (1.13), the constraint (1.17) may be written as

$$\int_0^1 \cos(\theta(s)) ds = \frac{1}{L},$$

(1.18)

which tends to the case $C = 0$ as $L \to \infty$. This suggests that the case $C = 0$ is a limiting case in $L$ for $C = 1$.

Figure 1.7: In the figure typical planar curves $\gamma$ (black) with turning angles $\theta$ (gray) satisfying constraint (1.17) are plotted for each of the two distinguishable cases, with $C = 0$ above and $C = 1$ below.

As is well known, the problem of finding a unit speed parametrization of a planar curve requires a closed form integral of the curve’s speed, which typically admits no elementary antiderivative. This difficulty is circumvented by defining the constant speed curves in terms
of their curvature from the onset. The drawback is of course not having an explicit representation of the planar curve. It is then natural to ask how the curve can be generated. One means presented is a representation via integration of the form given in (1.14), where of course in most cases this must be done numerically. For an alternate and more general method of generating the planar curve, one can use an ordinary differential equation derived from a transported frame. Given a unit vector \( \mathbf{X} = (x_1, x_2) \) obtain an orthogonal frame \( \{X, Y\} \) with \( Y = (-x_2, x_1) \). Then (as depicted in Figure 1.8) for all \( s \) the tangent vector has orthogonal decomposition

\[
\dot{\gamma}(s) = L \cos(\theta(s)) X + L \sin(\theta(s)) Y.
\] (1.19)

Then by (1.13) and (1.19) the components of the planar curve \( \gamma = (x, y) \) satisfy the initial value problem

\[
\begin{align*}
\dot{x}(s) &= L (x_1 \cos(\theta(s)) - x_2 \sin(\theta(s))) , \quad x(0) = 0, \\
\dot{y}(s) &= L (x_2 \cos(\theta(s)) + x_1 \sin(\theta(s))) , \quad y(0) = 0.
\end{align*}
\] (1.20)

Now the differential equation (1.20) is rather general inasmuch as it takes \( \theta \) as measured from the ray \( tX \) for \( t > 0 \), so from an arbitrary axis. We (naturally) measure from the positive \( x \)-axis, and thus take \( \{X, Y\} = \{(1, 0), (0, 1)\} \), so that (1.20) simplifies to

\[
\begin{align*}
\dot{x}(s) &= L \cos(\theta(s)) , \quad x(0) = 0, \\
\dot{y}(s) &= L \sin(\theta(s)) , \quad y(0) = 0,
\end{align*}
\] (1.21)

which agrees with differentiating (1.14) with respect to \( s \).

Although the planar curve must be generated numerically in most cases, this can typically be done stably with sufficient accuracy. While this is a drawback in some ways, it greatly simplifies complications involved with finding a unit speed parametrization — removing these complications by studying instead tangent indicatrices has enabled the establishing of the
results presented in this thesis. This approach was long ago devised and it has ushered in great advances in elasticity theory.

Figure 1.8: A depiction of the transported frame used to recover a planar curve $\gamma$ from its indicatrix, via equation (1.20).

1.3 Manifold Structure

Define a one-parameter family of non-linear functionals by

$$G_L(\theta) = L \int_0^1 \cos(\theta(s)) ds,$$  \hspace{1cm} (1.22)

and consider the level sets, denoted as

$$\mathcal{M}_L := G_L^{-1}({1}) \text{ for } L > 1 \text{ and } \mathcal{M}_\infty := G_1^{-1}({0}),$$ \hspace{1cm} (1.23)

where the length restriction is needed to guarantee that $\mathcal{M}_L$ is neither empty nor a singleton. We take the point of view that $\mathcal{M}_\infty = G_1^{-1}({0})$ is a limiting case in $L$, (as this is suggested by (1.18), and later observations,) and since the dynamic behavior is replicated across different

\footnote{We use := to denote equal by definition.}
values of $L$ in this case, we set $L = 1$ when $C = 0$. Throughout, when we write $L = \infty$ we refer to the manifold $\mathcal{M}_\infty$, although we normalize the length of the curves to $L = 1$. It is further claimed that the sets $\mathcal{M}_L$ for $L \in (1, \infty]$ are in fact Riemannian submanifolds of $H$. To verify this, we show that at each point $\theta \in \mathcal{M}_L$, there is a tangent space, which is a co-dimension 1 hyperplane defined via a normal (orthogonal) vector, where the normal vector varies smoothly in $\theta$. This normal vector is in fact the gradient $\nabla G_L$, defined to be the function satisfying the equation

$$DG_L(\theta)\eta = \langle \nabla G_L(\theta), \eta \rangle_H,$$

where $DG_L(\theta)\eta$ is the directional derivative at $\theta$ in the direction $\eta$, and this generalizes the familiar finite dimensional gradient. We will show that at each $\theta \in \mathcal{M}_L$ this gradient is indeed well defined and expressed in terms of integrals of smooth functions acting on $\theta$. This produces a smooth gradient tangent vector field in the tangent bundle of $\mathcal{M}_L$. The directional derivative is given by

$$DG_L(\theta)\eta = \frac{d}{d\varepsilon}[G_L(\theta + \varepsilon \eta)]_{\varepsilon=0} = -L \int_0^1 \sin(\theta(s))\eta(s)ds,$$

where there is sufficient regularity to support differentiating under the integral sign in (1.25). From (1.25) one sees that $\forall \eta \in H$, the mapping $\eta \mapsto DG_L(\theta)\eta$ is a bounded linear functional on $H$ so the Riesz representation theorem implies the existence of a unique gradient $\nabla G_L(\theta)$ at each $\theta \in \mathcal{M}_L$. By (1.24) and (1.25), the gradient satisfies the integro-differential equation

$$-L \int_0^1 \sin(\theta(s))\eta(s)ds = \nabla G_L(\theta; 0)\eta(0) + \int_0^1 \frac{d}{ds}[\nabla G_L(\theta; s)]\eta(s)ds.$$
To solve equation (1.26), integrate by parts to rewrite the left hand side as follows:

\[-L \int_0^1 \sin(\theta(u)) \eta(u) \, du = - \int_0^1 \left[ \frac{d}{du} \left( \int_1^u \sin(\theta(v)) \, dv \right) \right] \eta(u) \, du \]

\[= \int_0^1 \left[ \frac{d}{du} \left( \int_u^1 \sin(\theta(v)) \, dv \right) \right] \eta(u) \, du \]

\[= \left[ \eta(u) L \int_u^1 \sin(\theta(v)) \, dv \right]_0^1 - \int_0^1 \left[ L \int_u^1 \sin(\theta(v)) \, dv \right] \dot{\eta}(u) \, du \]

\[= -\eta(0) L \int_0^1 \sin(\theta(v)) \, dv - \int_0^1 \left( L \int_u^1 \sin(\theta(v)) \, dv \right) \dot{\eta}(u) \, du \]

\[= \left( -L \int_0^1 \sin(\theta(v)) \, dv \right) \eta(0) + \int_0^1 \left( -L \int_u^1 \sin(\theta(v)) \, dv \right) \dot{\eta}(u) \, du \]

Substituting the expression on the last line into the left hand side of equation (1.26), by inspection it is clear that the gradient satisfies the initial value problem

\[\frac{d}{ds}[\nabla G_L(\theta; s)] = -L \int_s^1 \sin(\theta(v)) \, dv, \quad \nabla G_L(\theta; 0) = -L \int_0^1 \sin(\theta(v)) \, dv. \quad (1.27)\]

Integrating, one finds that the gradient is given

\[\nabla G_L(\theta; s) = -L \left[ \int_0^s \left( \int_u^1 \sin(\theta(v)) \, dv \right) \, du + \int_0^1 \sin(\theta(u)) \, du \right]. \quad (1.28)\]

Rewriting the iterated integral in (1.28), one obtains the equivalent form

\[\nabla G_L(\theta; s) = -L \left[ \int_0^s u \sin(\theta(u)) \, du + s \int_s^1 \sin(\theta(u)) \, du + \int_0^1 \sin(\theta(u)) \, du \right]. \quad (1.29)\]

Observe that,

\[\nabla G_L(\theta; 0) = -L \int_0^1 \sin(\theta(u)) \, du = -y(1), \quad (1.30)\]

so that the initial value of the gradient of the constraint functional is the negative of the final free second coordinate value of the associated planar curve.
Remark 1.3.1. We have used that the functions in $H$ are absolutely continuous by definition, (for instance, so that the fundamental theorem of calculus is applicable,) and this is a natural assumption, inasmuch as functions in a first order Sobolev space over the real line have absolutely continuous representatives.

Proposition 1.3.2. For all $L \in (1, \infty]$, the set $\mathcal{M}_L$ is an infinite dimensional Riemannian manifold.

Proof. The formula for the gradient given in (1.28) shows its smooth dependence on $\theta$. By way of contradiction, assume that

$$\nabla G_L(\theta) \equiv 0. \quad (1.31)$$

Differentiation of (1.31) produces

$$\frac{d}{ds} \nabla G_L(\theta) = -L \int_s^1 \sin(\theta(u)) du \equiv 0 \quad (1.32)$$

and differentiating once more gives

$$\sin(\theta(s)) \equiv 0, \quad (1.33)$$

that is $\sin(\theta(s))$ must be the zero function. But the sine function is smooth, and $\theta$ is absolutely continuous, which implies that $\theta(s) \equiv \pi n$ for some integer $n$. The contradiction then follows, since the functions $\theta(s) = \pi n$ are not in any manifold $\mathcal{M}_L$ for any $L \in (1, \infty]$, as is readily verified. First if $L \in (1, \infty)$, one has

$$1 = L \int_0^1 \cos(\pi n) ds = L \cos(\pi n) = \pm L, \quad (\iff)$$
which is absurd. Similarly for \( L = \infty \) one has

\[
0 = \int_0^1 \cos(\pi n) ds = \cos(\pi n) = \pm 1, \quad (\implies \iff)
\]

which is also absurd. Thus, for \( L \in (1, \infty] \) the corresponding level sets are Riemannian manifolds.

\[ \square \]

### 1.4 The Curve Straightening Flow

Consider now the squared curvature nonlinear functional

\[
F_L(\theta) = \frac{1}{2L} \int_0^1 \dot{\theta}(s)^2 ds,
\] (1.34)

which corresponds to the Euler elastic energy of the associated planar curve \( \gamma \). For the case \( L = \infty \), the corresponding energy functional is \( F_1 \).

By a similar calculation as for \( \nabla G_L \), one obtains directional derivative

\[
DF_L(\theta)\eta = \frac{d}{d\varepsilon} \left[ \frac{1}{2L} \int_0^1 \left( \dot{\theta}(s) + \varepsilon \dot{\eta}(s) \right)^2 ds \right]_{\varepsilon=0} = \frac{1}{L} \int_0^1 \dot{\theta}(s) \dot{\eta}(s) ds,
\] (1.35)

where, again, there is sufficient regularity to support differentiating under the integral sign. Then setting \( DF_L(\theta)\eta = (\nabla F_L(\theta), \eta)_H \), it is apparent that the gradient of \( F_L \) must satisfy the differential equation

\[
\frac{d}{ds} \nabla F_L(\theta; s) = \frac{\dot{\theta}(s)}{L} \quad \text{with initial value} \quad \nabla F_L(\theta; 0) = 0.
\] (1.36)
It is readily verified that a solution to this initial value problem is

\[ \nabla F_L(\theta) = \frac{\theta(s) - \theta(0)}{L}. \]  

(1.37)

The gradient (1.37) gives the steepest descent direction in \( H \) with respect to the elastic energy. The \textit{curve straightening flow on} \( H \) is generated by the ‘differential equation’

\[ \frac{d\theta_\tau}{d\tau} = -\nabla F_L(\theta_\tau), \]  

(1.38)

and solution trajectories to (1.38) give the steepest descent path. Here \( \theta_\tau \) is a point in \( H \) for each \( \tau \), where \( \tau \) is the flow-parameter, so this may be thought of as an ordinary differential equation in an infinite-dimensional phase space. Here the solution is given explicitly in closed form as

\[ \theta_\tau(s) = \theta_0(0) + e^{-\frac{s}{\tau}} (\theta_0(s) - \theta_0(0)), \]  

(1.39)

and the gradient descent tends to a constant indicatrix, which corresponds to a straight line in the plane for any indicatrix \( \theta_0 \in H \) acting as initial-point. Note that by (1.37), the gradient of \( F_L \) is zero at \( s = 0 \), suggesting that when there are no other constraints the flow (1.38) avoids rotation of the initial tangent angle; this is in fact the case by (1.39). It follows that for any initial indicatrix in \( \theta_0 \in H \), the associated planar curve will flow to the straight line with angle \( \theta_0(0) \).

The flow is of course much more interesting in the presence of constraints. To define the curve straightening flow on the manifolds \( \mathcal{M}_L \) we project the gradient of the curve straightening functional onto the tangent space of the manifold. Denoting the projected gradient as \( \nabla^\pi F_L \), then for some real valued scalar field \( \lambda_L \) one has

\[ \nabla^\pi F_L(\theta) = \nabla F_L(\theta) - \lambda_L(\theta)\nabla G_L(\theta). \]  

(1.40)
The projected gradient is required to be in the tangent space $T_{\theta}M_L$ at $\theta \in M_L$, so one must have
\[
\langle \nabla^\pi F_L, \nabla G_L \rangle_H = 0 \quad \Rightarrow \quad \lambda_L = \frac{\langle \nabla F_L, \nabla G_L \rangle_H}{\langle \nabla G_L, \nabla G_L \rangle_H} = \frac{DG_L \nabla F_L}{DG_L \nabla G_L}. \tag{1.41}
\]
Thus by formula (1.25) the projection gives a scalar field $\lambda_L$ on $M_L$, with explicit form
\[
\lambda_L(\theta) = -\frac{1}{L^2} \frac{\int_0^1 \sin(\theta(u))(\theta(u) - \theta(0)) \, du}{\int_0^1 \sin(\theta(w)) \left( \int_0^w \left( \int_0^1 \sin(\theta(v)) \, dv \right) \, du + \int_0^1 \sin(\theta(v)) \, dv \right) \, dw}. \tag{1.42}
\]
Similarly as in (1.38) this induces the curve straightening flow on $M_L$, which is generated by the differential equation
\[
\frac{d\theta}{d\tau} = -\nabla^\pi F_L(\theta). \tag{1.43}
\]
If the projected gradient (1.40) vanishes at a point $\theta = \xi$, then the value $\lambda_L(\xi)$ is in fact the value of the classical Lagrange multiplier. We may call such a point $\xi$ a critical point of the projected gradient, which from the dynamical point of view is a rest point of the curve straightening flow on $M_L$.

The existence of the flows (1.38) and (1.43) follow from foundational results in global analysis, which do not belong to the theory of differential equations proper. A great source of the foundational theory is [10], specifically chapter 9. It is noted that in [10] the authors also assume the stronger Palais-Smale condition which they also use to establish convergence properties which may not hold in this setting. For existence it is enough to have local completeness, which we have here since we consider a complete Sobolev space, and inverse images of singletons, which are closed. Linnér has given more specific treatment for free length in [11], which for us is the manifold $M_L \times \mathbb{R}^+$; in particular, see Proposition 3.1.

---

5The Palais-Smale condition is an attempt to extend compactness to infinite dimensions, where for a functional $f : M \to \mathbb{R}$, the condition is that for a null sequence $||\nabla f(x_n)|| \to 0$ there is a convergent subsequence.
Regarding the Palais-Smale condition in $\mathcal{M}_L$, it is not known if it holds, but it has been proven by Langer and Singer to hold for positive tension\textsuperscript{6}; see theorem 9 in [12].

We will make use of the following lemma, which is a direct consequence of the existence of the curve straightening flow.

**Lemma 1.4.1.** Let $\theta$ be a non-critical point in $\mathcal{M}_L$. Then for all $\varepsilon > 0$ there exists $\eta \in \mathcal{M}_L$ with $||\theta - \eta||_H < \varepsilon$ and $F_L(\eta) < F_L(\theta)$.

*Proof.* This is a direct consequence of the negative projected gradient flow (1.43) being defined (at least locally) at any point $\theta \in \mathcal{M}_L$. Thus, a sufficiently close point $\eta$ with lesser energy exists on this trajectory. \hfill $\square$

\textsuperscript{6}Tension is discussed in Appendix C.
CHAPTER 2

PATHS IN $\mathcal{M}_L$ AND ANALYSIS OF CRITICAL POINTS.

2.1 The Notion of Stability and Critical Point Classification

We primarily use geometric concepts, but the critical point type may best be described in terms of dynamic stability of the projected gradient flow. We distinguish three types of critical points, global and local minima, and saddle points. We do not consider maximums, as the gradient flow decreases energy, and energy is unbounded in $\mathcal{M}_L$. We take the following standard definition of Lyapunov stability where further details can be found in many texts on the qualitative theory of differential equations, for instance [13]. In the following definition $\phi_t(x)$ denotes the flow from initial point $x$ at parameter value $t$.

**Definition 2.1.1. Lyapunov stability:** A rest point $\xi$ of a differential equation is stable in the sense of Lyapunov if for each $\varepsilon > 0$, there is $\delta > 0$ such that $|\phi_t(x) - \xi| < \varepsilon$ for all $t \geq 0$ whenever $|x - \xi| < \delta$. Also, given initial point $x_0$ a solution $t \mapsto \phi_t(x_0)$ is stable in the sense of Lyapunov if for each $\varepsilon > 0$, there is $\delta > 0$ such that $|\phi_t(x) - \phi_t(x_0)| < \varepsilon$ for all $t \geq 0$ whenever $|x - x_0| < \delta$. Otherwise, the rest point or solution is said to be unstable. Moreover, the rest point or solution is said to be asymptotically stable if $|\phi_t(x) - \phi_t(x_0)| \longrightarrow 0$ as $t \longrightarrow \infty$ whenever $|x - x_0| < \delta$.

By this notion, under the curve straightening flow a local minimum is asymptotically stable locally, and a saddle point is unstable. A critical point is a minimizer if all sufficiently small perturbations from that point will flow back, and otherwise the critical point is a saddle point where there are perturbations from that point which will not flow back. We prove that
all critical points corresponding to classical Euler-Bernoulli elastica with vanishing curvature at the endpoints are unstable, and thus saddle points, while the zero curvature critical points are ‘nearly’ globally asymptotically stable. Moreover the curve straightening flow is ‘nearly’ globally asymptotically stable, that is for any $\delta$ neighborhood at any non-critical $x_0$ nearly all trajectories flow to the global minimizer; it is unclear if certain trajectories can flow to a saddle point such as in examples shown on the Milnor torus in Section 1.1, and hence the description ‘nearly’ as if such trajectories exist, they are rare – these comprise only a set of zero measure on the Milnor torus. However, in Chapter 5 we demonstrate that this convergence can be very slow.

### 2.2 The Existence of Critical (and ‘Nearly’ Critical) Points

We refer to points $\xi \in \mathcal{M}_L$ with zero projected gradient as critical points, and in this section we identify them. To this end, suppose that $\xi$ is a critical point, so that

$$\nabla^\pi F_L(\xi) = 0. \quad (2.1)$$

Now $\xi$ has at least one derivative, so differentiation of both sides of this equation gives

$$0 = \frac{d}{ds} \nabla^\pi F_L(\xi) = \frac{\dot{\xi}(s)}{L} - \lambda_L L \int_1^s \sin(\xi(u))du. \quad (2.2)$$

A subtle point here is that $\lambda_L$ has already been evaluated at $\xi$, and it is thus constant with respect to $s$. It follows that

$$\dot{\xi}(s) = \lambda_L L^2 \int_1^s \sin(\xi(u))du \quad (2.3)$$
which implies that $\xi$ has at least one more derivative, and thus

$$\ddot{\xi}(s) = \lambda(\xi)L^2 \sin(\xi(s)).$$

(2.4)

Satisfying equation (2.4) is a necessary condition for being a critical point, and we seek solutions to this equation. For a fixed $\lambda_L$ equation (2.4) is the pendulum equation, and here, once again, one sees the connection between elasticity and pendulum motion. Non-trivial solutions to (2.4) are in the class of Jacobi elliptic functions.

Before considering a general solution for equation (2.4), it should be noted the equation is satisfied if

$$\xi = \text{const.},$$

(2.5)

i.e., for a constant indicatrix. Indeed, the scalar formula (1.42) evaluates to $\lambda_L(\text{const.}) = 0$, so both sides of equation (2.4) are null for (2.5). Due to constraints, in each manifold $M_L$ with $L < \infty$ there are only two possibilities, namely

$$\xi = \pm \tan^{-1} \left( \sqrt{L^2 - 1} \right).$$

(2.6)

Moreover the solutions (2.6) pass to the limit for $L = \infty$, with $\xi = \pm \frac{\pi}{2}$. (Of course this is not a formal limit, as we take unit length in $M_\infty$, but this reinforces the view that $C = 0$ in (1.13) is the limiting case for $L$.) Further (2.1) holds, as one can verify that the projected gradient vanishes when $\xi = \text{const}$, as clearly both right hand terms in (1.40) are zero. Since the energy is non-negative, these zero energy curves are clearly global minimizers, so the projected gradient must vanish. As global minimizers these points are necessarily stable.

**Remark 2.2.1.** For the global minima (2.6) the planar curve is a straight line, which is a geodesic in that setting. It is also more generally true that geodesics are global minimizers of the squared curvature functional on a given manifold, and here these are geodesics in the flat
manifold $\mathbb{R}^2$, where the associated planar curves reside. Theory on elastica in more general manifolds can be found in [5].

Turning to non-constant solutions of (2.1), it is noted that additional regularity can be assumed for a critical point, since (2.4) implies the existence of another derivative — in fact, it similarly follows inductively that a critical point must be $C^\infty$, despite having no assumption of such regularity in $H$. We may thus consider critical points in terms of the curvature of the associated planar curve. In Appendix B it is explained that via the existence of a second derivative solutions of (2.1) have squared curvature

$$
\kappa^2(s) = \rho(1 - A \, \text{sn}_k^2(LBs + \phi)),
$$

(2.7)

for length $L$ and constants $\rho$, $A$, $B$, $\phi$. As $\dot{\theta} = \kappa/L$. It is noteworthy that the solutions posses a symmetry where if $\theta$ is a solution so is $-\theta$. It is further explained in Appendix B that $0 \leq A \leq 1$, and when $A = 1$ the solution has cn curvature, and when $A < 1$ the solution has dn curvature. We shall rule out dn curvature solutions from criticality in this setting. However, dn curvature curves may be critical with other constraints, for instance with fixed terminal angles as later discussed in Section 3.3. We are primarily interested in critical points which satisfy $y(1) = 0$. In Section 2.6 we address the possibility of non-trivial critical curves with $y(1) \neq 0$, that is aside from the global minimizers (2.5).

We make much use of the gap equation

$$
2E(k) = \left(1 \pm \frac{g}{L}\right) K(k),
$$

(2.8)

where we take $g$ as a non-negative real number, and $E$, $K$ are the standard elliptic integrals discussed in Appendix A. Here $g$ is the magnitude of the gap between endpoints of a corresponding cn curvature elastica, and the elliptic modulus $k$ is taken to be a real root
Figure 2.1: Plotted are elastic loops with indicatrices $\xi_1, \xi_2, \xi_{10}$, with $n - 1$ inflections, and $n$ self intersections.

with $0 < k < 1$. Further details, a derivation, and a discussion on conventions used are included in Appendices A and C, where it is also shown that for $g < L$ there exists a unique root with $0 < k < 1$. As previously discussed, we take gaps $g \in \{0, 1\}$, and will find it helpful to view $L$ as a function $k$, which will be discussed in more detail in Section 2.3.

As discussed in Appendix C, we are primarily interested in the $\text{cn}$ curvature solutions of (2.4) having the form

$$
\xi(s) = \pi(1 - p) + 2 \sin^{-1}\left(\sqrt{k} \, \text{sn}_k \left( (2ns + (-1)^p)K(k) \right) \right), \quad p \in \{0, 1\},
$$

(2.9)

for counting numbers $n \in \mathbb{N} = \{1, 2, 3, ...\}$, where (depending on $p$) the elliptic modulus $k$ is given as a root of the equation

$$
2E(z) = \left( 1 + (-1)^{1-p} \frac{1}{L} \right) K(z).
$$

(2.10)
Equation (2.10) is evidently a special case of equation (2.8), where the gap is \( g = 1 \), and the sign corresponding to \( p \) is taken. For each \( p \) we get a distinct gap equation, and thus a unique real root.

![Graphs showing elastic arches with indicatrices \( \zeta_1, \zeta_2, \zeta_{10} \), with \( n-1 \) inflections, and no self intersections.]

For the two possible powers \( p \) in (2.9), we will denote the indicatrices as \( \xi_n, \zeta_n \), for \( p = 0, 1 \), respectively. These correspond to elastic curves with \( n - 1 \) inflection points, so no inflection in the case \( n = 1 \). We now show that these are in fact critical points.

**Theorem 2.2.2.** The projected gradient vanishes at the points

\[
\zeta_n(s) = 2 \sin^{-1} \left( \sqrt{k} \, sn(k(K(k)(2ns - 1))) \right), \quad n \in \mathbb{N}, \quad 2E(k) = \left( 1 + \frac{1}{L} \right) K(k) \tag{2.11}
\]

and

\[
\xi_n(s) = \pi + 2 \sin^{-1} \left( \sqrt{k} \, sn(k(K(k)(2ns + 1))) \right), \quad n \in \mathbb{N}, \quad 2E(k) = \left( 1 - \frac{1}{L} \right) K(k) \tag{2.12}
\]
for all \( n \in \mathbb{N}, \ L \in (1, \infty], \) where for the case \( L = \infty \) we take the limit as \( L \to \infty \) in these definitions.

Proof. The proof of this theorem is not overly technical, and the methods are not particularly clever, as we simply verify that \( \nabla^p F_L(\xi) = 0 \) for \( \xi \) any of the critical points in (2.11) or (2.12), which is just calculus. The only difficulty may be a lack of familiarity with Jacobi’s elliptic functions, so the fine details are included. Also, for clarity, it is noted that throughout, for the case \( \mathcal{M}_\infty \) the corresponding functionals used in the calculations are \( F_1 \) and \( G_1 \), so with \( L = 1 \), despite taking the limit as \( L \to \infty \) in the definitions (2.11) and (2.12).

For brevity write

\[
f(s) = K(k)(2ns + (-1)^p).
\]

Then for \( p \in \{0, 1\} \) using the sine sum identity and elliptic identity (A.6), one obtains

\[
\sin(\xi(s)) = 2(-1)^{1-p}\sqrt{k} \ sn_k(f(s))dn_k(f(s)).
\]

Then, from the antiderivative \( \int sn_k(u)dn_k(u)du = -cn_k(u) + \text{const.} \) one has

\[
\int_0^1 \sin(\xi(s))ds = 2(-1)^{p+1}\sqrt{k} \int_0^1 sn_k(f(s))dn_k(f(s)) = 0,
\]

The nullity of the integral (2.15) also follows from the final coordinate of the planar curve being \((1, 0)\), while this is proportional to the \( y \)-coordinate functional.

To obtain more succinct expressions for \( \nabla G_L(\xi) \) and \( \lambda_L(\xi) \) define

\[
I_1(s) = \int_0^s u \sin(\xi(u))du, \quad I_2(s) = \int_s^1 \sin(\xi(u))du,
\]

\[
I_3 = \int_0^1 \sin(\xi(s))\xi(s)ds, \quad I_4 = \int_0^1 \sin(\xi(s))(I_1(s) + sI_2(s))ds.
\]
and use formulas (1.29), (1.42), and (2.15) to obtain

\[ \nabla G_L(\xi) = -L (I_1(s) + s I_2(s)) \quad \text{and} \quad \lambda_L(\xi) = -\frac{1}{L^2} \frac{I_3}{I_4}. \]  

(2.18)

The gradient of \( F_L \) is given by the evaluation

\[ \nabla F_L(\xi) = \frac{2}{L} \left( \sin^{-1} \left( \sqrt{k} \, \text{sn}_k(f(s)) \right) - (-1)^p \sin^{-1} \left( \sqrt{k} \right) \right). \]  

(2.19)

Computing the gradient of \( G_L \) involves integration, and we begin writing

\[ I_1(s) = \int_0^s u \sin(\xi(u))du = 2(-1)^p \sqrt{k} \int_0^s u \, \text{sn}_k(f(u))dn_k(f(u))du. \]

To evaluate \( I_1 \), use the integral

\[ J_1(s) := \int_0^s t \, \text{sn}_k(at + b)dn_k(at + b)dt = -\frac{s}{a} \text{cn}_k(as + b) + \frac{1}{a} \int_0^s \text{cn}_k(at + b)dt, \]

having integrated by parts, differentiating \( t \). This gives another integral

\[ J_2(s) := \int_0^s \text{cn}_k(at + b)dt \]

\[ = \frac{1}{a\sqrt{k}} \int_0^s \frac{ak \, \text{cn}_k(at + b)sn_k(at + b)}{\sqrt{1 - dn_k^2(at + b)}} \text{sign} (\text{sn}_k(at + b)) \, dt \]

\[ = \frac{1}{a\sqrt{k}} \left( \cos^{-1} (dn_k(as + b)) \text{sign} (\text{sn}_k(as + b)) - \cos^{-1} (dn_k(b)) \text{sign} (\text{sn}_k(b)) \right) \]

\[ = \frac{1}{a\sqrt{k}} \left( \sin^{-1} \left( \sqrt{k} \, \text{sn}_k(as + b) \right) - \sin^{-1} \left( \sqrt{k} \, \text{sn}_k(b) \right) \right), \]
and thus

\[ J_1(s) = -\frac{s}{a} \text{cn}_k(as + b) + \frac{1}{a} J_2 \]

\[ = -\frac{s}{a} \text{cn}_k(as + b) + \frac{\sin^{-1}\left(\sqrt{k} \text{sn}_k(as + b)\right) - \sin^{-1}\left(\sqrt{k} \text{sn}_k(b)\right)}{a^2 \sqrt{k}}. \]

Then, using the integrals \( J_1(s), J_2(s) \) one obtains

\[ I_1(s) = \int_0^s u \sin(\xi(u))du = 2(-1)^p \sqrt{k} \int_0^s u \text{sn}_k(f(s)) \text{dn}_k(f(s))du \]

\[ = (-1)^p \left( \frac{\sqrt{k} s \text{cn}_k(f(s))}{nK(k)} - \frac{\sin^{-1}\left(\sqrt{k} \text{sn}_k(f(s))\right) - (-1)^p \sin^{-1}\left(\sqrt{k}\right)}{2n^2 K(k)^2} \right). \]

Continuing with computing the gradient of \( G_L \), next one has

\[ I_2(s) = \int_s^1 \sin(\xi(u))du = 2(-1)^p \sqrt{k} \int_s^1 \text{sn}_k(f(u)) \text{dn}_k(f(u))du \]

\[ = \frac{(-1)^p \sqrt{k}}{nK(k)} (\text{cn}_k(f(s)) - \text{cn}_k(f(1))) = \frac{(-1)^p \sqrt{k} \text{cn}_k(f(s))}{nK(k)}. \]

and combining these, with some simplification the gradient of \( G_L \) becomes

\[ \nabla G_L(\xi) = L \frac{(-1)^p \sin^{-1}\left(\sqrt{k} \text{sn}_k(f(s))\right) - \sin^{-1}\left(\sqrt{k}\right)}{2n^2 K(k)^2}. \]  

(2.20)
Next, turning to the scalar \( \lambda_L \), we find that \( I_3 \) is proportional to an additional integral, as

\[
I_3 = 2 \int_0^1 \sin(\xi(s)) \sin^{-1}\left(\sqrt{k} \, \text{sn}_k(f(s))\right) \, ds
\]

\[
= 4\sqrt{k}(-1)^{p+1} \int_0^1 \, \text{sn}_k(f(s)) \, \text{dn}_k(f(s)) \sin^{-1}\left(\sqrt{k} \, \text{sn}_k(f(s))\right) \, ds
\]

\[
= 4\sqrt{k}(-1)^{p+1} I_5
\]

for

\[
I_5 := \int_0^1 \, \text{sn}_k(f(s)) \, \text{dn}_k(f(s)) \sin^{-1}\left(\sqrt{k} \, \text{sn}_k(f(s))\right) \, ds.
\]

For us it suffices to know that \( I_5 \neq 0 \), which is clear since the integrand is non-negative over \([0, 1]\), as \( \text{sn}_k(f(s)), \sin^{-1}\left(\sqrt{k} \, \text{sn}_k(f(s))\right) \) are of the same sign, and \( \text{dn}_k(f(s)) \geq 0 \). However, integrating by parts by differentiating the sine inverse factor, and using (A.6) and (A.27), one can obtain the closed form

\[
I_5 = \frac{E(k) - K(k)(1 - k)}{\sqrt{k}K(k)}, \quad (2.21)
\]

which may be a useful formula for additional constraints; see Chapter 3. Moreover, we also find \( I_4 \) proportional to \( I_5 \), as

\[
I_4 = (-1)^{p+1} \int_0^1 \sin(\xi(s)) \left( \sin^{-1}\left(\sqrt{k} \, \text{sn}_k(f(s))\right) - (-1)^p \sin^{-1}\left(\sqrt{k}\right) \right) \frac{1}{2n^2K(k)^2} \, ds
\]

\[
= \frac{(-1)^{p+1}}{2n^2K(k)^2} \int_0^1 \sin(\xi(s)) \sin^{-1}\left(\sqrt{k} \, \text{sn}_k(f(s))\right) \, ds = \frac{\sqrt{k}}{n^2K(k)^2} I_5.
\]
Thus, one has
\[
\lambda_L(\xi) = -\frac{1}{L^2} I_3 = -\frac{1}{L^2} \frac{4\sqrt{k}(-1)^{1-p}I_5}{n^2 K(k)^2 I_5} = (-1)^p \frac{4n^2 K(k)^2}{L^2}.
\] (2.22)

Finally, one has \( \nabla^s F_L(\xi) = 0 \iff \nabla F_L(\xi) = \lambda_L(\xi) \nabla G_L(\xi) \) where
\[
\nabla F_L(\xi) = \frac{2}{L} \left( \sin^{-1} \left( \sqrt{k} \operatorname{sn}(f(s)) \right) - (-1)^p \sin^{-1} \left( \sqrt{k} \right) \right)
\] (2.23)

and
\[
\nabla G_L(\xi) = L \left( \frac{(-1)^p \sin^{-1} \left( \sqrt{k} \operatorname{sn}(f(s)) \right) - \sin^{-1} \left( \sqrt{k} \right)}{2n^2 K(k)^2} \right)
\] (2.24)

so multiplying the value of \( \nabla G_L(\xi) \) from (2.24) by the value for \( \lambda_L(\xi) \) from (2.22) one sees that the projected gradient vanishes at the indicatrices \( \xi \) defined in (2.9).

As further discussed in finer detail in Appendix C, the curves (2.4) have by construction cn curvature which vanishes at the endpoints. This is easily verified by differentiating (2.9), to obtain
\[
\dot{\xi}(s) = 4nK(k)\sqrt{k} \operatorname{cn}_k((2ns + (-1)^p)K(k)).
\] (2.25)

Now (2.25) is null at \( s = 0, 1 \), as cn vanishes at odd integer multiples of the quarter period \( K(k) \). It is noteworthy that vanishing curvature at the endpoints is equivalent to the natural boundary conditions \(^1\) for free endpoint variations being satisfied.

\(^1\)The natural boundary conditions are well known in the calculus of variations and are also briefly discussed in Chapter 4 Section 4.1 and stated in (4.3).
In fact, the natural boundary conditions must hold for any critical point that satisfies \( y(1) = 0 \). Indeed, \( \dot{\xi}(1) = 0 \) follows by evaluation in (2.3). Moreover, integrating both sides of equation (2.4) gives

\[
\int_0^1 \dddot{\xi}(s) ds = \lambda_L L^2 \int_0^1 \sin(\xi(s)) ds = \lambda_L Ly(1) = 0,
\]

so one then has

\[
0 = \int_0^1 \dddot{\xi}(s) ds = \dot{\xi}(1) - \dot{\xi}(0) = -\dot{\xi}(0).
\]

This implies vanishing curvature at the endpoints, and thus the natural boundary conditions hold if \( y(1) = 0 \).

Regarding dn curvature curves, equation (2.3) automatically precludes them from criticality, as it implies \( \dot{\xi}(1) = 0 \), while the dn function is non-vanishing. However, a special class of dn curvature curves are of interest in this setting. As discussed in the Appendix C, this is the enumerated family of curves with indicatrices

\[
\eta_n(s) = \pi + 2am_k \left((2ns + 1)K(k)\right), \quad (n \in \mathbb{Z}).
\]

Here \( 0 < k < 1 \) is a root of

\[
\frac{2}{k} \left(1 - \frac{E(k)}{K(k)}\right) - 1 = \frac{1}{L},
\]

as derived in Appendix C, where it is also shown that this root \( k \) uniquely exists. Despite not being critical, these curves are vital to understanding the dynamics of the curve straightening flow. We next directly verify that these curves are non-critical, as the resulting expression for the projected gradient has further implications.

**Claim 2.2.3.** The projected gradient of the dn curvature curves is non-zero in \( \mathcal{M}_L \).
Figure 2.3: Plotted in black are dn curvature loops with indicatrices $\eta_n$ with $n = 1, 2, 5$, having $n - 1$ inflections and $n$ self intersections. Also, plotted in gray are the corresponding cn curvature loops $\xi_n$, for comparison. For $k$ near 1 these become indistinguishable to the curve straightening flow, as both functionals $F_L$ and $G_L$ are even in $\theta$.

Proof. To verify the claim, for brevity write

$$f(s) = (2ns + 1)K(k).$$

(2.30)

Then $\eta_n(s) = \pi - 2am_k(f(s))$, and the squared curvature functional has gradient

$$\nabla F_L(\eta_n) = \frac{\eta_n(s) - \eta_n(0)}{L} = \frac{\eta_n(s)}{L} = \frac{\pi - 2am_k(f(s))}{L},$$

(2.31)

where $\eta_n(0) = 0$ via the identity (A.24). The sine of the indicatrix can be written

$$\sin(\eta_n(s)) = 2\sin(am_k(f(s)))\cos(am_k(f(s))) = 2\text{sn}_k(f(s))\text{cn}_k(f(s)).$$

(2.32)
Then, since \( \int \text{sn}_k(as + b)\text{cn}_k(as + b)ds = -\frac{dn_k(as+b)}{ak} + C \) and \( dn_k((2ns + 1)K(k)) = \sqrt{1-k} \) for all integers \( n \), one has \( \int_0^1 \sin(\eta_n(s))ds = 0 \). Now write

\[
\nabla G_L(\eta_n) = -L(I_1(s) + sI_2(s))
\] (2.33)

with \( I_1(s) := \int_0^s u \sin(\eta_n(u))du \), and \( I_2(s) := \int_s^1 \sin(\eta(u))du \). Integrating by parts and using the derivative formula for \( \text{dn} \) yields

\[
\int u \text{sn}_k(au + b)\text{cn}_k(au + b)ds = \frac{am_k(au + b) - audn_k(au + b)}{ak} + C,
\] (2.34)

from which it follows that

\[
\nabla G_L(\eta_n) = L \left( \frac{\pi - 2am_k((2ns + 1)K(k)) + 4\sqrt{1-knsK(k)}}{4kn^2K(k)^2} \right).
\] (2.35)

To compute \( \lambda_L \), for the numerator

\[
\int_0^1 \sin(\eta_n(s)) (\eta_n(s) - \eta_n(0)) ds = -4 \int_0^1 \text{sn}_k(as + b)\text{cn}_k(as + b)am_k(as + b)ds,
\] (2.36)

which simplifies using \( y(1) = 0 \) and (2.32). Integrating by parts, and using the derivative \( \frac{d}{du}am_k(u) = \text{dn}_k(u) \), one obtains the antiderivative

\[
\int \text{sn}_k(au + b)\text{cn}_k(au + b)am_k(au + b)du = \frac{E'(au + b) - am_k(au + b)dn_k(au + b)}{ak} + \text{const.},
\] (2.37)

from which one has

\[
\int_0^1 \sin(\eta_n(s)) (\eta_n(s) - \eta_n(0)) ds = \frac{2(\pi \sqrt{1-k} - 2E(k))}{kK(k)}.
\] (2.38)
Next, for the denominator of $\lambda_L$, using formula (3.37) above with the integral formulas (3.38), (2.37), one obtains

$$\int_0^1 \sin(\eta_n(s)) \left( I_1(s) + sI_2(s) \right) ds = \frac{(1-k)K(k) - \pi \sqrt{1-k} + E(k)}{k^2 n^2 K(k)^3}. \tag{2.39}$$

Thus one has

$$\lambda_L = 2kn^2 K(k)^2 \frac{\pi \sqrt{1-k} - 2E(k)}{\pi \sqrt{1-k} - E(k) - (1-k)K(k)}. \tag{2.40}$$

Thus the projected gradient is given

$$\nabla^\pi F_L(\eta_n) = \frac{1}{2L} \left( 2 \left( \pi - 2am_k(f(s)) \right) - \frac{\left( \pi \sqrt{k'} - 2E(k) \right) \left( 4n s K(k) \sqrt{k'} + \pi - 2am_k(f(s)) \right)}{\pi \sqrt{k'} - E(k) - k'K(k)} \right). \tag{2.41}$$

for $k' = 1 - k$. It can be shown that this is non-vanishing at $s = 1$, which by absolute continuity verifies the claim.

Despite (2.41) being non-zero for $0 < k < 1$, from the limit

$$\lim_{k \uparrow 1} (1-k)^\alpha K(k) = 0, \quad (\alpha \in (0,1])$$

it follows that

$$\lim_{k \uparrow 1} \nabla^\pi F_L(\eta_n) = 0.$$ 

Hence the gradient gets very small for $k \approx 1$ and corresponding $dn$ curvature curves with length $L \approx 1$ and arbitrarily small projected gradients exist, which amounts to very flat regions of the corresponding manifolds which have no critical points.
Figure 2.4: Plots comparing $\text{cn}$ and $\text{dn}$ loops and corresponding squared curvatures $\kappa^2$ for $L = 1.7, 3$, in gray and black, respectively. The $\text{dn}$ loop cannot ‘relax’ at the endpoints, and must more evenly distribute the curvature; however, for $k \approx 1$, the endpoint curvature nearly does vanish. While the $\text{cn}$ loop attains the largest curvature of the two over the interval $[0, 1]$, occurring at the midpoint, the totality of squared curvature is lesser for the $\text{cn}$ loop. As $k$ gets near 1, these plots become indistinguishable.

The connection between $\text{cn}$ and $\text{dn}$ curves is seen in the limit as $k \uparrow 1$, where the single $\text{cn}$ and $\text{dn}$ loop tend to the same curve. This is seen from

$$\lim_{k \uparrow 1} \text{cn}_k(u) = \text{sech}(u) = \lim_{k \uparrow 1} \text{dn}_k(u).$$

In the limit these tend to sech curvature as the period tends to $\infty$, and essentially these curves each degenerate to a line segment of unit length, while the corresponding manifolds
\( \mathcal{M}_L \) degenerate to the singleton \( \mathcal{M}_1 = \{0\} \). Geometrically the loops on the associated planar curves become arbitrarily small for both \( \text{cn} \) and \( \text{dn} \) curvature, with each tending to the same non-rectifiable curve in the limit; of course the angular variation is discontinuous in this limit. Thus for \( k \approx 1 \), the difference becomes negligible, in which case it is apparent that functionals \( F_L \) and \( G_L \), being even in \( \theta \), can hardly distinguish between the \( \text{cn} \) and \( \text{dn} \) curvature curves; see Figures 2.4 and 2.5. Hence for \( k \approx 1 \) the \( \text{dn} \) curvature curves are ‘nearly’ critical in the sense that the gradient descent behaves much like it would if it were in the neighborhood of a critical point. This is later observed in numerical discretization of the flow in Chapter 5.

Figure 2.5: Plots comparing \( \text{cn} \) and \( \text{dn} \) double loops, with darker curves for increasing \( k \). Each loop becomes increasingly similar as \( k \) increases, while the even number loops, (from left to right,) are opposing.
2.3 The Relationship Between Length and Elliptic Modulus

For \( g = 1 \) in the gap equation (2.8), solving for \( L \), one has

\[
L = \pm \frac{K(k)}{2E(k) - K(k)}.
\]  

(2.42)

Differentiating \( L \) with respect to \( k \), using formulas (A.29), one obtains

\[
\frac{dL}{dk} = \pm \frac{E(k)^2 - 2(1 - k)E(k)K(k) + (1 - k)K(k)^2}{k(1 - k)(2E(k) - K(k))^2}.
\]  

(2.43)

The positive sign in (2.42) and (2.43) corresponds to the arches \( \zeta_n \), while the negative sign in (2.43) corresponds to the loops \( \xi_n \). Using that \( \sqrt{1 - k} > 1 - k \), for \( 0 < k < 1 \), one obtains

\[
\frac{E(k)^2 - 2(1 - k)E(k)K(k) + (1 - k)K(k)^2}{k(1 - k)(2E(k) - K(k))^2} > \frac{(E(k) - \sqrt{1 - k}K(k))^2}{k(1 - k)(2E(k) - K(k))^2} \geq 0.
\]  

(2.44)

Thus for arches \( \zeta_n \) one has \( \frac{dL}{dk} > 0 \) and \( L \) is increasing in \( k \), while for the loops \( \xi_n \) one has \( \frac{dL}{dk} < 0 \) and \( L \) is decreasing in \( k \). Moreover, the singular value in (2.42) which occurs when \( 2E(k) = K(k) \) is the value \( k \) corresponding to gap \( g = 0 \) in (2.8). The length \( L \) of the corresponding elastic curve grows without bound when approaching this value \( k \) from below and above, and this is the single elliptic modulus for \( \mathcal{M}_\infty \). We denote this common elliptic modulus as \( k_\infty \approx .826 \), which is the real number defined by the transcendental equation

\[
2K(k_\infty) = E(k_\infty).
\]  

(2.45)

For any manifold \( \mathcal{M}_L \) with \( L < \infty \), one has elliptic modulus \( k < k_\infty \) for the indicatrices \( \zeta_n \), and elliptic modulus \( k > k_\infty \) for the indicatrices \( \xi_n \).
2.4 Analysis of the Critical Points in $\mathcal{M}_\infty$

We have stated that $\mathcal{M}_\infty$ can be considered as a limiting case; let us first clarify this point of view. As discussed in Section 2.3, when approaching $k_\infty$ among elliptic moduli $k$ in corresponding manifolds $\mathcal{M}_L$, the length $L$ grows without bound. Moreover, in the case $L = \infty$, the gap in equation (2.8) is zero, as the planar curve has the same initial and terminal point. But in (2.8), setting $g = 0$ has the same effect as letting $L \to \infty$, (or as letting $g \downarrow 0$ for that matter.) Then letting $L \to \infty$ for a unit gap, coincides with letting the gap go to zero with unit length. In this way we view $\mathcal{M}_\infty$ as the limiting case for large $L$. As the elliptic modulus is independent of length, we find it reasonable to normalize the length to unity, (i.e., take $L = 1$,) and consider a single manifold $\mathcal{M}_\infty$. However, we do consider different length curves with zero gap in the proof of Theorem 2.4.1.

As there is a single elliptic modulus corresponding to the manifold $\mathcal{M}_\infty$, this suggests that the two critical elastica, $\xi_n$ and $\zeta_n$, should coincide here. This is plausible geometrically, (see Figure C.6,) and is easily verified analytically. First of all, taking the limit as $L \to \infty$ in (2.8), the limit is independent of the power $p$, and in each case the elliptic modulus is $k_\infty$.

Regarding the term $\pi^{1-p}$ occurring in the general form (2.9), for $p = 0$ the additional angle $\pi$ gives a reflection about the origin, and a translation of a half period in the argument of $sn_{k_\infty}(\cdot)$, thus coinciding with the case $p = 1$. In each case we get windings of the Euler figure 8; see Figure 2.6.

In the manifold $\mathcal{M}_\infty$ the critical points remain in the manifold if the corresponding planar curve is rotated, which is in contrast to the general case $L < \infty$. This holds since the coordinates of the terminal point of the associated planar curve are always $(0,0)$ under rotation, and thus the constraint continues to hold. However, the projected gradient does not vanish for most rotations of these critical points.
Figure 2.6: Plots of $\xi$ as per (2.9) with increasing gray level for $1 \leq n \leq 10$, where in this setting, regardless of $p$, the associated planar curve consists of $n/2$ windings of the Euler figure 8 — this is taken to mean an additional half covering of the Euler figure 8 when $n$ is odd. Each curve has total length 1, which is of course the distance to the line $x = 1$, which is included in the figure for scale.

**Theorem 2.4.1.** Consider the elastic Euler figure 8 with its two inflection points at the origin and let $n$ be a positive integer. A loop consisting of $n$ halves, as illustrated in Figure 2.6, possibly rotated by some arbitrary amount, is a critical point in $\mathcal{M}_\infty$ if, and only if, the vertical axis is the axis of symmetry of the loop.

**Proof.** Criticality of the Euler figure 8 when the vertical axis is the axis of symmetry follows from Theorem 2.6.2. It is then only necessary to show that curves rotated from the vertical axis are not critical. Since the curves coincide in $\mathcal{M}_\infty$ with $\xi_n = \zeta_n$, it suffices to show this for rotations of the indicatrices $\zeta_n$. Then, using symmetry it suffices to show that $\zeta_n + t$ has non-vanishing projected gradient for $t \in (0, \pi)$.

By absolute continuity it is enough to show that $\nabla^\pi F_1(\zeta_n + t; s_0) \neq 0$ for any $s_0 \in [0, 1]$. We will show that

$$\nabla^\pi F_1\left(\zeta_n + t; \frac{1}{n}\right) \neq 0 \quad (2.46)$$
for $t \in (0, \pi)$. We first prove this for the half figure 8, which is the case $n = 1$. For $\xi = \zeta_1$, we claim that

$$
\theta_t(s) := 2 \sin^{-1} \left( \sqrt{k_\infty} \, \text{sn}_{k_\infty}(K(k_\infty)(2s - 1)) \right) + t = \xi(s) + t
$$

has non-zero projected gradient for $t \in (0, \pi)$. Now it may be inferred from Appendix B that for $t \neq n\pi$, $\theta_t$ cannot be a solution to the necessary differential equation (2.4), and thus automatically has non-zero projected gradient — however, since we are not in space saving mode, we feel that some insight and rigor is added by showing this directly. First, the gradient of $F$ can be written as

$$
\nabla F_1(\theta_t; s) = \theta_t(s) - \theta_t(0) = \nabla F_1(\xi; s).
$$

Using a sine sum identity and defining an auxiliary function

$$
\phi(s) := \int_0^s \left( \int_u^1 \cos(\xi(v))dv \right) du,
$$

the gradient of $G$ can be written as

$$
\nabla G_1(\theta_t; s) = -\left[ \int_0^s \left( \int_u^1 \sin(\xi(v) + t)dv \right) du + \int_0^1 \sin(\xi(u) + t)du \right] \\
= \cos(t) \nabla G(\xi; s) - \sin(t) \phi(s).
$$

Defining auxiliary integrals

$$
I_1 := \int_0^1 \sin(\xi(s)) \nabla F_1(\xi; s)ds, \quad I_2 := \int_0^1 \cos(\xi(s)) \nabla F_1(\xi; s)ds,
$$
\[ I_3 := \int_0^1 \sin(\xi(s))\nabla G_1(\xi; s)ds, \quad I_4 := \int_0^1 \cos(\xi(s))\nabla G_1(\xi; s)ds, \]

\[ I_5 := \int_0^1 \sin(\xi(s))\phi(s)ds, \quad I_6 := \int_0^1 \cos(\xi(s))\phi(s)ds, \]

the scalar \( \lambda_L \) has the form

\[
\lambda_L(\theta_1(s)) = \frac{\cos(t)I_1 + \sin(t)I_2}{\cos^2(t)I_3 + \sin(t)\cos(t)(I_4 - I_5) - \sin^2(t)I_6}. \tag{2.50}
\]

Now expression (2.50) simplifies since

\[ I_2 = I_4 = I_5 = 0. \tag{2.51} \]

The relevant property used in showing that these integrals are null is that \( \xi \) is odd about \( s = \frac{1}{2} \). First one has

\[ I_2 = \int_0^1 \cos(\xi(s))(\xi(s) - \xi(0))ds = \int_0^1 \cos(\xi(s))\xi(s)ds = 0 \tag{2.52} \]

as the cosine integral is zero since proportional to the terminal coordinate \( x(1) = 0 \), and the rightmost integrand is a product of an even and odd function about \( s = \frac{1}{2} \) over the interval \([0, 1]\). Next, turning to \( I_4 \), by similar calculations as in the proof of Theorem 2.6.2, the gradient is given by

\[
\nabla G_1(\xi; s) = \frac{\sin^{-1}\left(\sqrt{k} \ \text{sn}_{k_\infty}((2s - 1)K(k_\infty))\right) - \sin^{-1}\left(\sqrt{k_\infty}\right)}{2K(k_\infty)^2}, \tag{2.53}
\]

from which one has

\[ I_4 = \frac{1}{4K(k_\infty)^2} \int_0^1 \cos(\xi(s))\xi(s)ds - \frac{\sin^{-1}\left(\sqrt{k_\infty}\right)}{2K(k_\infty)^2} \int_0^1 \cos(\xi(s))ds = 0. \tag{2.54} \]
Finally, turning to the integral $I_5$, integration by parts gives

$$ I_5 = \int_0^1 \left( \int_0^s \sin(\xi(s)) du \right) \left( \int_s^1 \cos(\xi(v)) dv \right) ds := \int_0^1 y(s)x^R(s) ds = 0. $$

Here we denote the two integrated integrals as $y(s)$, $x^R(s)$ respectively, as the first is the $y$ coordinate functional for the associated planar curve, and the second is the $x$ coordinate functional for the reverse parametrized associated planar curve. Again the integrand is a product of even and odd functions about $s = \frac{1}{2}$, and thus vanishes. This verifies claim (2.51).

Thus, denoting

$$ \lambda_{L,t} := \lambda_L(\theta_t(s)) $$

the scalar field along $\theta_t$ takes the form

$$ \lambda_{1,t} = \lambda_1(\theta_t(s)) = \frac{\cos(t)I_1}{\cos^2(t)I_3 - \sin^2(t)I_6}. $$

Now to show that the projected gradient is non-vanishing for a fixed $t \in (0, \pi)$, again, it is enough to show it is non-vanishing for a single value $s \in [0, 1]$. Beginning with the case $n = 1$ of the half figure 8, we take $s = 1$. Then, by the calculations above, the evaluation of the gradient of $F_L$ at $s = 1$ is

$$ \nabla F_1(\theta_t; 1) = \nabla F_1(\xi; 1) = \xi(1) - \xi(0) = 4 \sin^{-1} \left( \frac{1}{\sqrt{k_{\infty}}} \right), $$

and the evaluation of the gradient of $G_L$ at $s = 1$ is

$$ \nabla G_1(\theta_t; 1) = \cos(t) \nabla G_1(\xi; 1) - \sin(t) \left( \int_0^1 u \cos(\xi(u)) du + \int_0^1 \cos(\xi(u)) du \right). $$
Moreover, both integrals in the sine term of (2.58) are null. The rightmost is the constraint functional, which we know is zero. The following calculations also show that the leftmost integral is null, where the steps are elementary, but they are included for clarity:

\[
\int_0^1 s \cos(\xi(s))ds = \int_0^{1/2} s \cos(\xi(s))ds + \int_{1/2}^1 s \cos(\xi(s))ds
\]

\[
= \int_0^{1/2} s \cos(\xi(s))ds + \int_0^{1/2} (1 - s) \cos(\xi(1 - s))ds
\]

\[
= \int_0^{1/2} s \cos(\xi(s))ds + \int_0^{1/2} (1 - s) \cos(\xi(s))ds
\]

\[
= \int_0^{1/2} \cos(\xi(s))ds = \frac{1}{2} \int_0^1 \cos(\xi(s))ds = 0.
\]

The last step follows from \(\cos(\xi(s))\) being even about \(s = 1/2\).

Then, by (2.53) and (2.58), one has

\[
\nabla G_1(\theta_t; 1) = \cos(t) \nabla G_1(\xi; 1) = -\frac{\sin^{-1}(\sqrt{k_\infty})}{K(k_\infty)^2} \cos(t),
\]

(2.59)

so the projected gradient \(\nabla^\pi F_1(\theta_t; 1) = \nabla F_1(\theta_t; 1) - \lambda_{1,t} \nabla G_1(\theta_t; 1)\) is given

\[
\nabla^\pi F_1(\theta_t; 1) = 4 \sin^{-1}(\sqrt{k_\infty}) \left( 1 + \frac{\cos^2(t)I_1}{\cos^2(t)I_3 - \sin^2(t)I_6} \frac{1}{4K(k_\infty)^2} \right).
\]

(2.60)

For (2.60) to vanish, one must have

\[
\frac{\cos^2(t)I_1}{\cos^2(t)I_3 - \sin^2(t)I_6} = -4K(k_\infty)^2.
\]

(2.61)

We know this happens when \(t = 0\), since the ratio has the form \(\lambda_{1,t} \cos(t)\), and this reduces to formula (2.22). Also, it is relevant that

\[
\frac{d}{dt} \left[ \frac{\cos^2(t)I_1}{\cos^2(t)I_3 - \sin^2(t)I_6} \right] = \frac{I_1 I_6 \sin(2t)}{(\cos^2(t)I_3 - \sin^2(t)I_6)^2} > 0, \quad \text{for} \quad 0 < t < \frac{\pi}{2},
\]

(2.62)
where the derivative has positive sign since all three factors in the numerator are positive, while the denominator is a square. Indeed, it is trivial that $\sin(2t) > 0$ on $(0, \frac{\pi}{2})$, and moreover both of the integrals $I_1$, $I_6$ are positive. First considering $I_1$, one has that $\nabla F_1(\xi; s)$ is non-negative and increasing on $[0, 1]$, as $\nabla F_1(\xi; 0) = 0$ and

$$\frac{d}{ds} \nabla F_1(\xi; s) = \dot{\xi}(s) = 4K(k_\infty) \sqrt{k_\infty} \cn_{k_\infty}((2s - 1)K(k_\infty)),$$

(2.63)

which is positive for $0 < s < 1$. Also, $\sin(\xi(s)) < 0$ for $0 < s < \frac{1}{2}$, by formula (2.14), as $\sn$ is negative over this interval, so one has

$$\int_0^{1/2} \sin(\xi(s)) \nabla F_1(\xi; s) ds < 0.$$

Moreover, since $\sin(\xi(s))$ is odd about $s = \frac{1}{2}$, and $\nabla F_1(\xi; s)$ is increasing,

$$0 < -\int_0^{1/2} \sin(\xi(s)) \nabla F_1(\xi; s) ds < \int_{1/2}^1 \sin(\xi(s)) \nabla F_1(\xi; s) ds \implies I_1 > 0.$$

Next, turning to $I_6$, integrating by parts one can rewrite the integral as

$$I_6 = -\int_0^1 \left( \int_0^s \cos(\xi(s)) ds \right) \left( \int_s^1 \cos(\xi(s)) ds \right) ds$$

$$:= \int_0^1 x(s)(-x^R(s)) ds = \int_0^1 x^2(s) ds > 0,$$

where $x(s)$ is the $x$ coordinate functional, and by symmetry about the $y$-axis, one has $x^R(s) = -x(s)$ for the reverse parametrization $x$ coordinate functional.

Thus the derivative (2.62) is positive $0 < t < \frac{\pi}{2}$, and when $t = 0$ one has equality in (2.61). It follows that the gradient is non-vanishing for $0 < t < \frac{\pi}{2}$, while there is symmetry for $\frac{\pi}{2} < t < \pi$, where

$$\nabla^\pi F_1(\theta_\pi - t; 1) = \nabla^\pi F_1(\theta_\pi + t; 1)$$

(2.64)
follows from the identities \( \cos \left( \frac{\pi}{2} - t \right) = -\cos \left( \frac{\pi}{2} + t \right) \) and \( \sin \left( \frac{\pi}{2} - t \right) = \sin \left( \frac{\pi}{2} + t \right) \), using expression (2.60). It follows that \( \nabla^s F_1(\zeta_1 + t; 1) \neq 0 \), since the absolutely continuous function is not zero at \( s = 1 \), and this holds for each \( t \) such that \( 0 < t < \pi \).

To see that for arbitrary \( n \) one has

\[
\nabla^s F_1(\zeta_n + t; 1/n) \neq 0, \tag{2.65}
\]

consider curves of arbitrary length \( L \). The corresponding rotated half figure 8 corresponds to a non-zero projected gradient, and the formula for \( \xi_1 \) stays the same since the elliptic modulus \( k_\infty \) is independent of \( L \) as seen in (2.8) when the gap is zero. The gradient and scalar fields satisfy

\[
\nabla F_L(\cdot) = \frac{1}{L} \nabla F_1(\cdot), \quad \nabla G_L(\cdot) = L \nabla G_1(\cdot), \quad \text{and} \quad \lambda_L(\cdot) = \frac{1}{L^2} \lambda_1(\cdot). \tag{2.66}
\]

Thus one has

\[
\nabla^s F_L(\theta) = \nabla F_L(\theta) - \lambda_{L,t} \nabla G_L(\theta) = \frac{1}{L} \nabla^s F_1(\theta), \tag{2.67}
\]

so for arbitrary \( L \), these rotated half figure 8 curves also have non-zero projected gradient. The restriction of \( \zeta_n + t \) to the interval \([0, 1/n]\), can be viewed as a half figure 8 curve of length \( L = 1/n \), where \( \nabla^s F_1(\zeta_n + t) \neq 0 \) holds since the function is not zero at \( s = 1/n \). It now follows that \( \nabla^s F_1(\zeta_n + t) \) at \( s = 1/n \) is equal to \( \nabla^s F_1(\zeta_n + t) \) at \( s = 1 \). Finally we have established (2.46), which proves the theorem.

Using this theorem the critical points in \( \mathcal{M}_\infty \) are easily classified.

**Theorem 2.4.2.** All of the \( cn \) curvature critical points are saddle points in \( \mathcal{M}_\infty \).

**Proof.** By Theorem 2.4.1, the points \( \zeta_n(s) + t \) have non-zero projected gradient and constant energy for the parameter \( t \in (0, \pi) \). Also, as \( t \downarrow 0 \) the rotation \( \zeta_n(s) + t \) converges to
\[ \zeta_n = \xi_n \text{ in } H\text{-norm}. \] Thus, there are points which are arbitrarily close in \( M_{\infty} \) having the same energy, but with non-zero projected gradient. It follows from Lemma 1.4.1 that the indicatrices \( \zeta_n = \xi_n \), (with equality only in \( M_{\infty} \)), are not local minima.

**Remark 2.4.3.** It is interesting to compare the implications of Theorem 2.4.1 with a physical elastic wire configured as such, with one end fixed at the origin, and the other free to slide frictionlessly along the vertical line — although this is an idealization, it could perhaps be nearly realized by configuring a lubricated wire as such between two sheets of glass. It is intuitive that for \( t = 0 \) there should be a rest point (or tipping point) as the wire is balanced between an inclination to slide in either direction, but for small \( t \neq 0 \), the wire should be able to freely unwind. However, we do not claim that our choice of an inner product (or any choice of one for that matter) gives rise to dynamics coincident with those of the physical world. We would hope that it gives rise to a useful model however, which is qualitatively similar.

### 2.5 Analysis of the Critical Points in \( M_L \) for \( L < \infty \)

Analysis of the critical points is more complicated in the manifolds with \( L < \infty \). The construction from the last section is not applicable, as rotated curves leave the manifold and the indicatrix no longer satisfies the constraint. Moreover, in contrast there are distinct families of critical elastica to consider for \( L < \infty \). We demonstrate that the critical elastica are also unstable by utilizing special paths of non-critical \( cn \) curvature curves, which are thus not rest points of the curve straightening flow. Naturally, we should look to non-critical curves with Jacobi elliptic curvature to find paths of minimal energy, as such curves minimize the bending energy for a given configuration. Also, in contrast to the \( M_{\infty} \) case, there are two distinct families of critical elastica to consider. First, consider the elastica
ζₙ, where geometric intuition may lead one to expect that nearby curves will unwind under the gradient descent. A similar construction as in the last section demonstrates that such intuition is correct, and that the critical points are unstable, and thus of the saddle type.

**Theorem 2.5.1.** All critical points ζₙ in (2.11) corresponding to elastic arches with cn curvature are saddle points in ℳ₇ for L ∈ (1, ∞).

**Proof.** Consider the path

\[ \theta_t(s) := \alpha(t) + 2 \sin^{-1} \left( \sqrt{t} \, \text{sn}(K(t)(2ns - 1)) \right), \quad t \in [0, k], \tag{2.68} \]

where k is the elliptic modulus for ζₙ, (with k is independent of n.), and

\[ \alpha(t) = \sec^{-1}(g(t)), \tag{2.69} \]

for 1 ≤ g(t) ≤ L the gap corresponding to t. Here the gaps vary from L to 1 as t varies from 0 to k, and α(t) is a rotation angle; see Figure 2.7. The modulus t and the gap g are related by equation (2.8). We will show that the parametric curve (2.68) provides a path in ℳ₇ from the global minimizer \( \theta_0(s) = \tan^{-1} (\sqrt{L^2 - 1}) \) to the critical point \( \theta_k(s) = \zeta_n(s) \), with increasing energy.

To see that (2.68) is a path in ℳ₇ it must be shown that the constraint is satisfied for any \( t \in [0, k] \). Setting

\[ \phi_t(s) := 2 \sin^{-1} \left( \sqrt{t} \, \text{sn}(K(t)(2ns - 1)) \right) \]

one has

\[ \int_0^1 \cos(\theta_t(s))ds = \cos(\alpha(t)) \int_0^1 \cos(\phi_t(s))ds - \sin(\alpha(t)) \int_0^1 \sin(\phi_t(s))ds. \tag{2.71} \]
Now $\int_0^1 \sin(\phi_t(s))ds = 0$, which can be shown by direct computation as in Theorem 2.6.2, or inferred by noting that this integral is proportional to the $y$-coordinate functional of the associated planar curve, which is zero. This considerably simplifies formula (2.71). Proceeding, one expresses

$$\cos(\phi_t(s)) = 2\text{dn}^2_t(K(t)(2ns - 1)) - 1,$$ (2.72)

by a cosine identity, and the fundamental elliptic identity (A.6). Thus one has

$$\int_0^1 \cos(\phi_t(s))ds = 2\mathcal{E}(2nK(t) - K(t), t) - \mathcal{E}(-K(t), t) - 1 = 2\frac{E(t)}{K(t)} - 1$$

by (A.27), where $\mathcal{E}(\tau, k) = \int_0^\tau \text{dn}_k^2(u)du$ is the Jacobi Epsilon function (A.26). Moreover

$$2E(t) = K(t) \left(1 + \frac{g}{L}\right) \iff 2\frac{E(t)}{K(t)} - 1 = \frac{g}{L},$$

which holds since $t$ is a root of the left hand equation, equivalently (2.8). Whence

$$\int_0^1 \cos(\theta_t(s))ds = \cos(\sec^{-1}(g)) \int_0^1 \cos(\phi_t(s))ds = \frac{1}{gL} = \frac{1}{L},$$

and the constraint is satisfied.

The energy along the path $\theta_t(s)$ is given by

$$F(\theta_t(s)) = \frac{2a}{L}(a(t - 1) + \mathcal{E}(a + b, t) - \mathcal{E}(b, t))$$ (2.73)
Figure 2.7: Construction of $\theta_t(s)$ for the proof of Theorem 2.5.1 in the leftmost plot, and also plots of discrete sets along the path for $n = 1, 3$.

via a formula derived in Appendix C, for $a = 2nK(t)$ and $b = K(t)$. Using identity (A.27), the energy (2.73) simplifies to

$$F(\theta_t(s)) = \frac{8K(t)(E(t) - K(t)(1 - t))}{L}.$$  \hspace{1cm} (2.74)

Via formulas (A.29), the derivative of the energy (2.73) with respect to the parameter $t$ can be written as

$$\frac{d}{dt} F(\theta_t(s)) = \frac{4n^2 E(t)^2 - 2(1 - t)E(t)K(t) + (1 - t)K(t)^2}{t(1 - t)}$$

$$> \frac{4n^2 E(t)^2 - 2\sqrt{1 - t}E(t)K(t) + (1 - t)K(t)^2}{t(1 - t)}$$

$$= \frac{4n^2 (E(t) - \sqrt{1 - t}K(t))^2}{t(1 - t)} \geq 0,$$

where the inequality is valid since $0 < t < k < 1$; this is similar to (2.43) and (2.44). Thus the energy is a strictly increasing quantity as the path approaches the critical point, and the point is not a local minimum. \hfill $\Box$
Finally, we turn to consideration of the self-intersecting elastic curves in $\mathcal{M}_L$ for $L \in (1, \infty)$. Intuition may lead one to suspect that if the loop is wound tightly enough, there may be a bifurcation value $L$ for which these become stable — however, we demonstrate that this is not the case.

**Theorem 2.5.2.** All critical points $\xi_n$ in (2.12) corresponding to elastic loops with $cn$ curvature are saddle points in $\mathcal{M}_L$ for $L \in (1, \infty)$.

**Proof.** We demonstrate this by showing that there exist points arbitrarily close of lesser energy by using an explicit path. For an elastica $\xi_n$ consider the path

\[
\theta_t(s) := \pi + 2 \sin^{-1}\left(\sqrt{k} \, \text{sn}_k(as + b)\right), \quad a := 2nK(k), \quad b := K(k) + t.
\] (2.75)

It is claimed that $\theta_t$ is a periodic in $t$, constant energy path in $\mathcal{M}_L$, (including the case $L = \infty$,) with period $4K(k)$. First of all, defining

\[
\phi_t(s) = 2 \sin^{-1}\left(\sqrt{k} \, \text{sn}_k(as + b)\right),
\] (2.76)

one has, as in the proof of Theorem 2.3,

\[
\int_0^1 \cos(\theta_t(s))ds = \cos(\pi) \int_0^1 \cos(\phi_t(s))ds = 1 - 2 \frac{E(t)}{K(t)} = \frac{1}{L},
\] (2.77)

and thus $\theta_t(s) \in \mathcal{M}_L$ for arbitrary $t$. Also, as in the proof of Theorem 2.5.1, using formula (2.73), one has energy

\[
F(\theta_t(s)) = \frac{8n^2K(k)(E(k) - K(k)(1 - k))}{L},
\] (2.78)

which is independent of $t$, showing that this is a constant energy path. Finally, the claim that $\theta_t$ is periodic in $t$ follows from the $4K(k)$ periodicity of $\text{sn}_k$. 
Now the path $\theta_t$ passes thru both $\xi_n$ and the indicatrix of the reflection of the associated planar curve about the $x$-axis, say

$$\xi_n^*(s) := \pi + 2 \sin^{-1}(\sqrt{k} \, \text{sn}_k(K(k)(2ns - 1))). \tag{2.79}$$

The projected gradient only vanishes along $\theta_t$ at $\xi_n$, $\xi_n^*$. Observe that differentiation with respect to $s$ gives

$$\frac{d\theta_t(s)}{ds} = 4nK(k)\sqrt{k} \, \text{cn}_k((2ns + 1)K(t) + t). \tag{2.80}$$

Here it is noteworthy $\dot{\theta}_t(1) = 0$ only holds when $t = (2m + 1)K(k)$, as $\text{cn}$ only vanishes at even multiples of the quarter period. Hence, equation (2.3) implies that the projected gradient is non-vanishing away from $\xi_n$, $\xi_n^*$. However, we will directly verify the claim that the projected gradient is non-vanishing, as we are not in space saving mode, and feel that it adds additional insight and rigor.

Turning to computation of the projected gradient along $\theta_t$, first the gradient of $F_L$ is given

$$\nabla F_L(\theta_t(s)) = \frac{1}{L}(\theta_t(s) - \theta_t(0)) = \frac{2}{L} \left( \sin^{-1}\left(\sqrt{k} \, \text{sn}_k(as + b)\right) - \sin^{-1}\left(\sqrt{k} \, \text{sn}_k(b)\right) \right). \tag{2.81}$$
Next, for the gradient of $G_L$, defining auxiliary integrals

$$
I_1 := \int_0^1 \sin(\theta_t(u))du, \quad I_2(s) := \int_0^1 \sin(\theta_t(u))du, \quad I_3(s) := \int_s^1 u \sin(\theta_t(u))du, \quad (2.82)
$$

one can write

$$
\nabla G_L(\theta_t(s)) = -L(I_1 + sI_2(s) + I_3(s)). \quad (2.83)
$$

The calculations here are similar as for points $\xi$ as per (2.9) in the proof of Theorem 2.6.2, but more complicated as many simplifications used there are no longer applicable. First of all, the integral corresponding to $I_1$ was null for those elastica, while here from the derivative formula of $cn$ one has

$$
I_1 = 2\sqrt{k}a(cn_k(a+b) - cn_k(b)) = \frac{2\sqrt{k}}{a}((-1)^n - 1)cn_k(K(k) + t). \quad (2.84)
$$

Throughout we use that

$$
\cn_k(a + b) = \cn_k(a)cn_k(b) = (-1)^n\cn_k(b), \quad (2.85)
$$

which follows from elliptic function identities (A.8) and (A.11). The integral $I_1$ does not in general vanish, but does vanish when $n$ is even by equation (2.85). It is clear that $I_1$ vanishing implies that the final $y$ coordinate is zero, as this integral is proportional to the coordinate value $y(1)$.

Via the same antiderivative, the second integral is given

$$
I_2(s) = 2\sqrt{k}\frac{(-1)^n\cn_k(b) - \cn_k(as + b)}{a}. \quad (2.86)
$$
Also, similarly as in the proof of Theorem 2.6.2, the third integral is given in closed form as
\[
\mathcal{I}_3(s) = 2\sqrt{k} \left( \frac{s \, \text{cn}_k(as + b)}{a} - \frac{\sin^{-1} \left( \sqrt{k} \, \text{sn}_k(as + b) \right)}{a^2 \sqrt{k}} - \sin^{-1} \left( \sqrt{k} \, \text{sn}_k(b) \right) \right). \tag{2.87}
\]

Turning to computing the scalar \( \lambda \), first the numerator is
\[
\int_0^1 \sin(\theta_t(s)) L \nabla F_L(\theta_t(s)) ds = \int_0^1 \sin(\theta_t(s)) \theta_t(s) ds - \theta_t(0) \mathcal{I}_1 = 2\mathcal{I}_4 + (\pi - \theta_t(0)) \mathcal{I}_1 \tag{2.88}
\]
for an additional auxiliary integral
\[
\mathcal{I}_4 := \int_0^1 \sin(\theta_t(s)) \sin^{-1} \left( \sqrt{k} \, \text{sn}_k(as + b) \right) ds. \tag{2.89}
\]
Evaluating, similarly as the integral \( I_5 \), in the proof of Theorem 2.6.2, one finds that
\[
\mathcal{I}_4 = 2 \left( 1 - k - \frac{E(k)}{K(k)} \right) = 1 - 2k + \frac{1}{L}, \tag{2.90}
\]
where the last expression follows from formula (2.10). Next, the denominator of \( \lambda \) is
\[
\int_0^1 \sin(\theta_t(s))(\mathcal{I}_1 + s\mathcal{I}_2(s) + \mathcal{I}_3(s)) ds = \mathcal{I}_1^2 + \mathcal{I}_5 + \mathcal{I}_6, \tag{2.91}
\]
in terms of two more auxiliary integrals
\[
\mathcal{I}_5 := \int_0^1 s \sin(\theta_t(s)) \mathcal{I}_2(s) ds \quad \text{and} \quad \mathcal{I}_6 := \int_0^1 \sin(\theta_t(s)) \mathcal{I}_3(s) ds. \tag{2.92}
\]
Now one has
\[
\mathcal{I}_5 = \int_0^1 s \sin(\theta_t(s)) \mathcal{I}_2 ds = \frac{2\sqrt{k}}{a}((-1)^n \text{cn}_k(b) \mathcal{I}_7 - \mathcal{I}_8) \tag{2.93}
\]
for two more auxiliary integrals

\[ I_7 := \int_0^1 s \sin(\theta_t(s))ds \quad \text{and} \quad I_8 := \int_0^1 s \ \cn_k(as + b) \sin(\theta_t(s))ds. \]  

(2.94)

For these integrals, via calculations similar to those in the proof of Theorem 2.6.2, obtain closed form

\[ I_7 = -2\sqrt{k} \left( \frac{\sin^{-1}(\sqrt{k} \cd_k(t)) ((-1)^n - 1)}{a^2 k} + \frac{(-1)^n \sqrt{1 - k} \sd_k(t)}{a} \right), \]  

(2.95)

where \( \cn_k(b) = -\sqrt{1 - k} \ \sd_k(t) \), (with \( \sd = \sn / \dn \)) follows from (A.8), (A.6), (A.11) and (A.12). Next rewrite the second integral as

\[ I_8 = -2\sqrt{k} \int_0^1 s \ \sn_k(as + b) \cn_k(as + b) \dn_k(as + b)ds := -2\sqrt{k} \mathcal{I}, \]  

(2.96)

where we evaluate \( \mathcal{I} \) by integrating by parts and differentiating the factor \( \sn \cn_k(as + b) \), to get

\[ \mathcal{I} = -\frac{\cn_k^2(a + b)}{a} + \frac{1}{a} \int_0^1 \cn_k^2(as + b)ds - \mathcal{I}. \]  

(2.97)

Solving for \( \mathcal{I} \) one obtains

\[ \mathcal{I}_8 = \frac{a(1 - k) + ak(1 - k)\sd_k^2(t) - 2nE(k)}{a^2 \sqrt{k}}. \]  

(2.98)

Continuing, again using similar calculations as before, one obtains

\[ \mathcal{I}_6 = \frac{2n\sqrt{k}K(k)\mathcal{I}_8 - \mathcal{I}_4 + \sin^{-1}(\sqrt{k} \cd_k(t)) \mathcal{I}_1}{2n^2 K(k)^2}. \]  

(2.99)
The scalar $\lambda$ is then given

$$
\lambda_{L,t} = -\frac{1}{L^2} \frac{2I_4 + (\pi - \theta(0))I_1}{I_1^2 + I_5 + I_6} = \frac{a^2}{L^2} \alpha + \frac{\alpha + \beta}{L^2} \alpha + (1 - (-1)^n)\beta + a k'k(3 - 2(-1)^n)sd^2_k(t)
$$

(2.100)

with $\lambda_{L,t} = \lambda_L(\theta_t(s))$ as defined in (2.55), for

$$
k' := 1 - k, \quad \alpha := 2n(E(k) - k'K(k)), \quad \beta := (1 - (-1)^n)\sqrt{k'k}sd_k(t) \sin^{-1} \left( \sqrt{k'k}cd_k(t) \right).
$$

(2.101)

Notice that when $t = 0$, since $sd_k(0) = sn_k(0)/dn_k(0) = 0$, one has

$$
\lambda_{L,0} = \lambda(\theta_0(s)) = \frac{a^2}{L^2} \alpha + 0 + 0 = \frac{a^2}{L^2} \frac{4n^2K(k)^2}{L^2},
$$

(2.102)

which is consistent with what was previously computed in the proof of Theorem 2.6.2.

Now we can easily write the gradient of $F_L$ in terms of $cd=cn/dn$ as

$$
\nabla F_L(\theta_t(s)) = \frac{2}{L} \left( \sin^{-1} \left( \sqrt{k'k}sd_k(t) \right) - \sin^{-1} \left( \sqrt{k'k}cd_k(t) \right) \right)
$$

(2.103)

using elliptic sum identities, and similarly can write the gradient of $G_L$ as

$$
\nabla G_L(\theta_t(s)) = -2L \frac{a\sqrt{k'k}sd_k(t)(1 - (1 + s)(-1)^n)}{a^2}
+ \frac{2L}{a^2} \left( \sin^{-1} \left( \sqrt{k'k}cd_k(t) \right) - \sin^{-1} \left( \sqrt{k'k}cd_k(t) \right) \right).
$$

Then letting

$$
\mu_t := \frac{L^2}{a^2} \lambda_{L,t} = \frac{\alpha + \beta}{\alpha + (1 - (-1)^n)\beta + a k'k(3 - 2(-1)^n)sd^2_k(t)}.
$$

(2.104)
the projected gradient becomes

$$\nabla^\pi F_L(\theta_t(s)) = (1 - \mu_t)\nabla F_L(\theta_t(s)) + \frac{2\mu_t a\sqrt{kk'}sd_k(t)(1 - (1 + s)(-1)^n)}{L}. \quad (2.105)$$

From (2.105) it follows that in the period interval $[0, 4K(k))$, the projected gradient only vanishes along $\theta_t$ at $t = 0, 2K(k)$. Indeed, first suppose that $\mu_t = 1$. Then by (2.105), $\nabla^\pi F_L(\theta_t(s)) = 0$ if, and only if, $sd_k(t) = 0$. This occurs only when $t = 0, 2K(k)$ in the first period. Moreover, if $sd_k(t) = 0$, from (2.104) it follows that $\mu_t = 1$. Next, suppose that $\mu_t \neq 1$, and by way of contradiction, assume that $\nabla^\pi F_L(\theta_t(s))$ vanishes for $t \in (0, 2K(k)) \cup (2K(k), 4K(k))$. Then from (2.105) one has

$$(\mu_t - 1)\nabla F_L(\theta_t(s)) = \frac{2\mu_t a\sqrt{kk'}sd_k(t)(1 - (1 + s)(-1)^n)}{L}. \quad (2.106)$$

Then one can divide by $\mu_t - 1$ to obtain

$$\nabla F_L(\theta_t(s)) = \frac{2\mu_t a\sqrt{kk'}sd_k(t)}{L(\mu_t - 1)}(1 - (1 + s)(-1)^n), \quad (2.107)$$

from which it follows that $\nabla F_L(\theta_t(s))$ is a linear function of $s$, which is absurd, by (2.103). This proves that the $t$-periodic path $\theta_t$ has vanishing projected gradient only at $\xi_n$, $\xi_n^*$. Thus, we can produce a sequence of arbitrarily close points in $\mathcal{M}_L$ with equal energy to the critical points, and non-vanishing projected gradient, namely $\{\theta_{t/n}\}_{n \in \mathbb{N}}$. From this and Lemma 1.4.1 it follows that these elastica are unstable, for in turn there are arbitrarily close points with lesser energy. This proves the theorem.

$$\Box$$

**Remark 2.5.3.** Theorem 2.5.2 is more general than Theorems 2.5.1 and 2.4.2, but it is worth demonstrating that different constructions can be used to analyze the critical points in the latter cases. It is demonstrated that the manifold $\mathcal{M}_\infty$ is a considerably different setting
than the cases $L \in (1, \infty)$. Also for $L < \infty$, the ability to produce a strictly decreasing energy path of $cn$ curvature curves at the elastic arches suggests that there is much more freedom to release the bending energy in an arch than in a loop.

**Remark 2.5.4.** It is also plausible geometrically that the indicatrices $\theta_t$ are of constant energy, as they can be realized as a construction of the critical elastica via rigid motions, by reflection and translation with a cut and paste. Observe that $\theta_t(s)$ consists of sections of the critical elastica as depicted in Figure 2.9.

![Figure 2.9: Translating the gray section of the curve suggests that in totality $\theta_t(s)$ should have the same energy as the critical elastica – indeed, this is proven analytically in Theorem 2.5.2.](image)

**2.6 Regarding Critical Points with $y(1) \neq 0$.**

As discussed, if a critical point has associated planar curve satisfying $y(1) = 0$, then it must correspond to one of the classical Euler-Bernoulli elastica with $cn$ curvature and vanishing curvature at the endpoints – this is inferred from (2.3). We have shown that all of the critical points with $y(1) = 0$ are unstable. However, the condition $y(1) = 0$ cannot be necessary for criticality, as we have identified two critical points with $y(1) \neq 0$. A simple reason why a critical point with positive energy must satisfy $y(1) = 0$ does not seem
apparent, although intuition about the dynamics of the flow seems to suggest \( y(1) = 0 \) must hold. If such a critical point were to exist it must have an indicatrix

\[
\xi(s) = n\pi \pm 2\sin^{-1}\left(\sqrt{k_0} \, \text{sn}_{k_0}(a_0 s + b_0)\right) \quad (n \in \mathbb{Z})
\]

(2.108)

with associated planar curve satisfying \( y(1) \neq 0 \). From the condition \( \dot{\xi}(1) = 0 \), one has \( a + b = (2n + 1)K(k) \), which enables elimination of the parameter \( b \). If such a critical point were to exist, given the methods applied in this Chapter, it would be natural to seek a non-increasing energy path. A candidate would be a path

\[
\phi_t(s) = \alpha(t) \pm 2\sin^{-1}\left(\sqrt{k(t)} \, \text{sn}_{k(t)}(a(t) s + b(t))\right)
\]

(2.109)

with \( \alpha(0) = n\pi \), \( a(0) = a_0 \), \( b(0) = b_0 \), and \( k(0) = k_0 \). The gap typically varies in the parameters \( a \), \( b \), \( k \), as seen from the integrals \( L \int_0^1 \cos(\phi_t(s))ds \) and \( L \int_0^1 \sin(\phi_t(s))ds \). Typically one can vary any of the parameters \( a \), \( b \), \( k \), and choose a rotation angle \( \alpha \) to correct for the change in gap, (so long as \( y(1) \neq 0 \)), and typically at least one of these parameters can be varied to produce a non-increasing energy path thru (2.108). It is noted that for the path (2.75) no rotation angle correction \( \alpha \) was needed to account for the change in gap, but this is atypical, and was due to the special case of \( a \) being an even multiple of the quarter period. As computed in Appendix C the path (2.109) has energy

\[
\Phi(a,b,k) := F_L(\phi_t) = \frac{2a}{L} (\mathcal{E}(a + b, k) - \mathcal{E}(b, k) - a(1 - k)).
\]

(2.110)

If \( \Phi \) has a non-vanishing gradient in these three variables, then locally a non-increasing energy path can be constructed as described, and the only way that such a non-increasing energy path does not exist is if there is simultaneously a minimum in energy in all three parameters at \( a = a_0 \), \( b = b_0 \), and \( k = k_0 \). (Of course lack of existence of such a path does
not necessarily imply stability.) While the alternate possibility seems unlikely, it does not
appear to be trivial to show that it cannot happen. We can thus only speculate the following:

**Conjecture 2.6.1.** The classical Euler-Bernoulli elastica which satisfy the natural boundary
conditions, (that is with indicatrices (2.11) or (2.12),) are the only positive energy critical
points in \( \mathcal{M}_L \).

**Conjecture 2.6.2.** The only stable critical points in \( \mathcal{M}_L \) are the zero energy global mini-
mizers.

### 2.7 Searching for Special Mountain Pass Paths in \( \mathcal{M}_L \)

In Section 1.1 an example of a mountain pass was given on a quartic surface in \( \mathbb{R}^3 \), where
we made use of geodesic paths, and simplicity of the level sets. By construction this was a
simple example to visualize and analyze. In general, when it exists, this path can be difficult
to produce even in \( \mathbb{R}^3 \), and in the infinite-dimensional setting this is of course a much more
difficult problem.

In \( \mathcal{M}_L \) there are two distinct global minimizers, the two straight line segments, say \( p, q \). Then among the set of curves

\[
\Omega_{p,q} = \{ \theta_r | \theta_r \text{ is a rectifiable path in } \mathcal{M}_L \text{ with } \theta_0 = p \text{ and } \theta_T = q \} \tag{2.111}
\]

we define a mountain pass from \( p \) to \( q \) as a solution to

\[
\mathcal{L}(\theta_r) = \inf_{\theta \in \Omega_{p,q}} \mathcal{L}(\theta). \tag{2.112}
\]
To motivate this definition, note that the path length in $\mathcal{M}_L$ can be written

$$L(\theta, \tau) = \int_0^T \left\| \frac{d\theta}{d\tau} \right\|_H \, d\tau$$

$$= \int_0^T \sqrt{\left( \frac{d\theta(0)}{d\tau} \right)^2 + \int_0^1 \left( \frac{d}{ds} \frac{d\theta(s)}{d\tau} \right)^2 \, ds \, d\tau}$$

$$= \int_0^T \sqrt{\left( \frac{d\theta(0)}{d\tau} \right)^2 + \int_0^1 \left( \frac{d}{d\tau} \frac{d\theta(s)}{ds} \right)^2 \, ds \, d\tau}$$

$$= \int_0^T \sqrt{\left( \frac{d\theta(0)}{d\tau} \right)^2 + \int_0^1 \left( \frac{d\theta(s)}{d\tau} \right)^2 \, ds \, d\tau},$$

where interchanging the order of differentiation is valid for sufficient regularity of $\theta, \tau$. As the length is independent of parametrization, the upper bound $T$ of integration is arbitrary. It is unclear how to produce a minimizer of $L(\theta, \tau)$, let alone if one exists, but this does suggest that minimal elastic energy may be favorable.

In $\mathcal{M}_\infty$ Theorem 2.4.1 is related to a mountain pass in the following way. The vertical line-segments from the origin to the points $(0, \pm L)$ are both global minima. A path from one of the two minima to the other reaches a maximum elastic energy at some point along the path. The infimum of the maximum energy of all such paths that connect the two global minima is here some positive number. Since the path must pass through a closed curve, the half loop of the Euler figure eight must be involved since it has minimal energy among closed curves. This mountain-pass plateau consisting of a circles worth of Euler figure 8 loops must be traversed in the mountain pass. Only the vertical Euler figure 8 loops are critical, and any non-vertical Euler figure 8 loop will instantly become a non-closed curve when switched to steepest descent along the negative gradient of the elastic energy. Also, in Section 1.1 the topological mountain pass description was mentioned, and this is seen here where the energy levels below that of the Euler figure 8 have sublevel sets which are disconnected, with each containing a global minimizer.
Figure 2.10: Discrete plots pertaining to three mountain pass candidates. From left to right, a variation in elliptic modulus $0 \leq k < k_\infty$ with gaps ranging from $0 \leq g < 1$ with half a rotation of the half figure 8 (which is similar to the path used in the proof of Theorem 2.5.1), the path used in the proof of Theorem 2.5.2 with rotated ambient segment joining the extremities, and a gradient descent from the half figure 8 curve with the largest projected gradient among curves on the mountain pass plateau, where via a half rotation this could produce an ascent-descent pair.

Several candidates are illustrated in Figure 2.10. In the left-most example, each curve seen is an elastic curve compatible with the gap between the endpoints without restrictions on the tangent angles at the endpoints. When reaching the mountain-pass energy the gap between the endpoint has vanished and the path spins the Euler loop half a turn until the process retraces itself in the opposite direction. This path is not an example of a steepest ascent-descent pair and its length is greater than some of the steepest descent curves from the plateau to the straight curve that gives the global minimum. The second example consists of a half period of the path (2.75) constructed in Theorem 2.5.2 which passes through the lower half of the vertical Euler figure 8, joined with the rotated line segments joining the extremities to the global minimizers, (where the segments are rotated to stay in $\mathcal{M}_\infty$.)
The third plot illustrates the steepest descent from one of the horizontal Euler loops and by numerical methods it appears to be the shortest of all the descent paths from the non-vertical (and thus non-critical) Euler loops.

In $\mathcal{M}_L$ with $L < \infty$ the path (2.68) used in the proof of Theorem 2.5.1 can be extended to include symmetric rotations $\alpha(t) = -\sec^{-1}(g_t)$ to provide a path joining the straight line segments, which provides a mountain pass path candidate. Here the curves along the path no longer must close to traverse the manifold between minimizers, but to obtain a minimal maximum energy along the path, it must pass thru the elastic arch with no inflection, as this is the minimal energy curve joining the terminal points – here at the minimal positive critical energy level there is no plateau, but rather a single elastic curve. Moreover, the path must assume all terminal points $(1, y)$ for $-\sqrt{L^2 - 1} \leq y \leq \sqrt{L^2 - 1}$, and at all such $y$ the rotated elastic arch has minimal energy among competing curves, and thus likely furnishes the mountain pass path between the global minimizers in $\mathcal{M}_L$ for $L < \infty$.

2.8 Geodesic Lines in $\mathcal{M}_L$ and the Geodesic Spray

Given a Riemannian structure, it is natural to be interested in the geodesics, as the lines are among the most basic geometric objects. As mentioned in the introduction, geodesics generalize straight lines in Euclidean space. Some examples were given, such as a great circle on a sphere, certain planar sections of a torus, or the meridians on a surface of revolution. However geodesics are generally difficult to determine, even on two dimensional surfaces – even for a deformation of a sphere to an ellipsoid, only the three geodesics in the coordinate planes are easy to identify. In a general manifold a geodesic can be described as a curve such that the acceleration is always normal to the tangent space, so with an everywhere non-zero gradient field; in other words the acceleration is always parallel to the gradient.
By the same notion, if a curve $\gamma_t = \gamma_t(s)$ parametrized by $t$ is a geodesic in $\mathcal{M}_L$ then

$$\frac{d^2 \gamma_t}{dt^2} = \mu_t \nabla G_L(\gamma_t), \quad (2.113)$$

for some scalar $\mu_t$ for all $t$. Since $\ddot{\gamma}_t \in \text{span}\{\nabla G_L(\gamma_t)\}$ always holds, it must be equal to its own projection onto the gradient. This gives a formula for $\mu_t$ as thus $\gamma_t$ must satisfy the differential equation

$$\frac{d^2 \gamma_t}{dt^2} = \frac{\langle \ddot{\gamma}_t, \nabla G_L(\gamma_t) \rangle_H}{\langle \nabla G_L(\gamma_t), \nabla G_L(\gamma_t) \rangle_H} \nabla G_L(\gamma_t). \quad (2.114)$$

More generally a pre-geodesic may have a component of acceleration in the tangential direction. The parametric curve would trace out a geodesic path with non-constant speed. In this case $\ddot{\gamma}_t \in \text{span}\{d\gamma_t/dt, \nabla G_L(\gamma_t)\}$ which similarly yields a differential equation

$$\frac{d^2 \gamma_t}{dt^2} = \frac{\langle \ddot{\gamma}_t, \frac{d\gamma_t}{dt} \rangle_H}{\langle \frac{d\gamma_t}{dt}, \frac{d\gamma_t}{dt} \rangle_H} \frac{d\gamma_t}{dt} + \frac{\langle \ddot{\gamma}_t, \nabla G_L(\gamma_t) \rangle_H}{\langle \nabla G_L(\gamma_t), \nabla G_L(\gamma_t) \rangle_H} \nabla G_L(\gamma_t). \quad (2.115)$$

Evidently, explicitly solving either of equations (2.114) or (2.115) seems unlikely. Even numerically solving them poses inherent difficulties, where we consider discretizations of first order differential equations in the Sobolev setting in Chapter 5. However, these equations are useful for checking if a given curve is, or nearly is, a geodesic or pre-geodesic. For instance, one may suspect that the descent from the horizontal Euler figure 8 in Section 2.7 is a geodesic, as there is symmetry similar to the planar section on the Milnor torus example in Section 1.1, but numerical testing with these formulas suggests that it differs slightly from a geodesic.
CHAPTER 3
ADDITIONAL CONSTRAINTS AND THE CORRESPONDING
SUBMANIFOLDS OF $\mathcal{M}_L$ 

The manifolds $\mathcal{M}_L$ are induced by level sets of a functional which returns the terminal $x$-coordinate $x(1)$ of the associated planar curve. Other constraints may be imposed by restricting to level sets of other relevant functionals. The geometric interpretation is the intersection with an additional submanifold of the Sobolev space $H$. In this section we consider submanifolds of $\mathcal{M}_L$ corresponding to the additional constraint of either fixing the terminal angles or endpoints of the associated planar curve.

3.1 Submanifolds of Intersection

A surface in Euclidean 3-space $\mathbb{R}^3$ is a two dimensional submanifold. A particular special case is a plane, which is a flat manifold and affine subspace of the ambient space $\mathbb{R}^3$. In this setting it is easy to visualize cases where the intersection is a smooth curve, and thus a submanifold of both. It is also easy to visualize cases where the intersection fails to be a submanifold, so it is not automatic that the intersection of two submanifolds is itself a submanifold. We are interested in higher dimensional settings, where visualization is lost. In general a tool for determining if two manifolds intersect in a submanifold is the condition of transversality, which is sufficient (but not necessary) to ensure that it is. Two submanifolds are said to intersect transversally if the two tangent spaces generate the ambient space at a given point; see [14] for formal definitions and relevant theorems. Some simple examples are
depicted in Figure 3.1; a twisted ribbon surface intersects a plane through its central axis in perpendicular lines and the intersection fails to locally be a submanifold where these lines intersect — it is easy to verify that the tangent spaces coincide at these points, and fail to generate the ambient space \( \mathbb{R}^3 \); for a torus intersecting a cylinder concentric with its axis of revolution the intersection fails to be transversal when the surfaces are tangent, but the intersection is still a submanifold here; for Viviani’s curve which is given by the intersection of a sphere and a cylinder, there is a self intersection where the tangent planes agree and the intersection of the surfaces fails to be transversal, at which point the curve is not locally a submanifold.

Figure 3.1: Some 2-surfaces in \( \mathbb{R}^3 \) where intersection fails to be transversal at certain points, at which the intersection may fail to be a submanifold.

For the gradient descent acting on two transversally intersecting submanifolds of an ambient space, one may define the gradient descent on the submanifold of intersection in the natural way, via projection onto its tangent space. The ‘height’ functional on \( \mathbb{R}^3 \) discussed in Section 1.1, which provides an analogue to the energy in \( \mathcal{M}_L \), is helpful for visualizing the possibilities, where simple examples here demonstrate that the critical points under the gradient descent may have significantly different structure when acting on a submanifold. For instance, consider a paraboloid, say \( \mathcal{S} \) cut in an affine section by a plane, say \( \mathcal{A} \). On
the paraboloid there is a single critical point with respect to the height, which is a global minimizer at the origin, but on the manifold \( \mathcal{M} = S \cap A \) there may be a distinct minimizer, both a minimum and a maximum, or infinitely many of each, (where in the latter case the plane is normal to the third axis;) see Figure 3.2. Simple examples as such demonstrate that the critical point structure may be in constrast to that of the larger manifold.

Figure 3.2: Some affine sections of 2-surfaces in \( \mathbb{R}^3 \) which intersect in submanifolds, which may be disconnected.

It is also clear from examples that the topological properties may be different for submanifolds; for instance consider a plane cutting the Milnor torus \( T \) discussed in Section 1.1. The plane may cut the surface in two components so that the submanifold of the connected manifold \( T \) is disconnected; see Figure 3.2.

These examples have analogies in infinite dimensions. Similarly as in finite dimensions, an affine subspace of an infinite dimensional space is the possibly translated span of a basis of vectors, say

\[
A = \{ x \in H | x = x_0 + \sum a_k \beta_k \} \quad (3.1)
\]

for some point \( x_0 \in H \), scalars \( a_k \) and basis elements \( \beta_k \). An affine subspace is a flat submanifold with constant normal space which induces a tangent space as its orthogonal complement. The notion of transversal intersection of manifolds still makes sense in infinite dimensions, and it likewise follows that the intersection is a submanifold under this condition; an ex-
cellent source of information on the theory of infinite dimensional manifolds is [14]. There are several ways to define codimension, but assuming that tranversality holds we take the codimension of a submanifold of $H$ to be the dimension of the normal space, which is finite in our examples — for instance, the manifolds $\mathcal{M}_L$ have codimension 1.

### 3.2 Affine Subspaces Induced by Evaluation Functionals

A class of affine subspaces of the Sobolev space $H$ of interest (1.8) is that induced by evaluation functionals, as these correspond to point constraints. To be clear, consider a general evaluation functional, say

$$\phi(x) = x(a), \quad (a \in [0, 1], \ x \in H.) \quad (3.2)$$

To find the gradient, first compute the directional derivative

$$D\phi(x)v = \frac{d}{d\varepsilon} [\phi(x + \varepsilon v)]_{\varepsilon=0} = \frac{d}{d\varepsilon} [x(a) + \varepsilon v(a)]_{\varepsilon=0} = v(a). \quad (3.3)$$

Now the inner product can be written as

$$\langle \nabla \phi(x), v \rangle = \nabla \phi(x; 0)v(0) + \int_0^a \frac{d}{dt} [\nabla \phi(x; t)] \dot{v}(t) dt + \int_a^1 \frac{d}{dt} [\nabla \phi(x; t)] \dot{v}(t) dt, \quad (3.4)$$

and from inspection of (3.4) it becomes clear that

$$\nabla \phi(x; t) = (1 + t)\chi_{[0,a]}(t) + (1 + a)\chi_{[a,1]}(t), \quad (3.5)$$
with $\chi_S(t)$ the characteristic function for the set $S$. Indeed, it is simple to verify that

$$\langle \nabla \phi(x), v \rangle = \nabla \phi(x;0)v(0) + \int_0^a \frac{d}{dt} [\nabla \phi(x;t)] \dot{v}(t)dt + \int_a^1 \frac{d}{dt} [\nabla \phi(x;t)] \dot{v}(t)dt$$

$$= v(0) + \int_0^a \dot{v}(t)dt$$

$$= v(0) + v(a) - v(0)$$

$$= D\phi(x)v.$$ 

**Proposition 3.2.1.** For a finite set of point constraints $\phi_k(x) = x(a_k) = x_k$, with $1 \leq k \leq n$, the feasible set, say $\Omega := \bigcap_{1 \leq k \leq n} \phi_k^{-1}(\{x_k\})$, is an $n$-dimensional affine subspace, and thus a codimension $n$ flat submanifold, of $H$.

*Proof.* To show that $\Omega$ is a submanifold of $H$ it suffices to produce a nonvanishing normal vector field which varies smoothly on $\Omega$. By the gradient formula (3.5) the normal vector is of the form

$$\nu(x) = \sum_{1 \leq k \leq n} \phi_k \sigma_k \nabla \phi_k(x),$$

(3.6)

where $\sigma_1, \ldots, \sigma_n$ are scalars, (which could be chosen to give a unit normal.) Hence, the normal vector $\nu(x)$ is a constant on the Hilbert space $H$ (as a function of a variable $t$, which is a point in $H$.) Thus it is trivial that $\nu$ varies smoothly, so $\Omega$ is a manifold. Moreover, $\Omega$ is its own tangent space at every point, and thus a co-dimension $n$ affine subspace of $H$. 

In particular, for the endpoint constraints

$$\phi_1(x) := x(0), \quad \text{and} \quad \phi_2(x) := x(1),$$

(3.7)

the gradients are

$$\nabla \phi_1(x) = 1, \quad \text{and} \quad \nabla \phi_2(x) = 1 + t.$$  

(3.8)

---

1For $S \subset X$ the characteristic function $\chi_S$ takes the value 1 on $S$, and 0 on the complement $X \setminus S$. 


Corollary 3.2.2. The constraint set \( \phi_1^{-1}(\{0\}) \cap \phi_2^{-1}(\{1\}) \) is a codimension 2 flat submanifold of \( H \).

3.3 Fixed Terminal Angles of the Indicatrix

A natural case of interest for elastic curves is the configuration of fixed terminal angles, which physically correspond to ‘clamping.’ Linnér has given thorough investigation from the point of view of the first two chapters, but with free length, that is where the lengths of feasible curves can be any number larger than the gap; in particular see [15] and [16]. The manifold structure is significantly different for free length, and typically the negative gradient flow is better behaved regarding convergence for fixed length, as without tension with free length curves may blow up to infinity in length while releasing bending energy. Here we briefly consider the fixed length problem via consideration of an affine section and submanifold of \( M_L \).

Take the functionals \( F_L, G_L \) as defined in (1.34), (1.22), and additionally take evaluation functionals

\[
\phi_0(\theta) := \theta(0), \quad \text{and} \quad \phi_1(\theta) := \theta(1). \tag{3.9}
\]

The functionals (3.9) were shown in Section 3.2 to have gradients

\[
\nabla \phi_0(\theta) = 1, \quad \text{and} \quad \nabla \phi_1(\theta) = 1 + s, \tag{3.10}
\]

and induce a flat co-dimension 2 submanifold of \( H \).

Given a choice in two terminal submanifolds

\[
\theta(0) = \theta_0 \in [0, 2\pi) \quad \text{and} \quad \theta(1) = \theta_1, \tag{3.11}
\]
(where the initial angle can be chosen modulo $2\pi$,) consider the sets
\[ M_{\theta_0, \theta_1, L} = M_L \cap \phi_0^{-1}(\{\theta_0\}) \cap \phi_1^{-1}(\{\theta_1\}). \] (3.12)

**Remark 3.3.1.** Here we do not reduce the angular variation $\theta(1) - \theta(0)$ modulo $2\pi$, but if one did, different windings would correspond to different components of the manifolds, with the components pathwise disconnected.

**Proposition 3.3.2.** The level sets $M_{\theta_0, \theta_1, L}$ are submanifolds of the Sobolev space $H$.

*Proof.* The set of functions
\[ B = \{\nabla \phi_0(\theta), \nabla \phi_1(\theta), \nabla G_L(\theta)\} \] (3.13)
have Wronksian
\[ W(\nabla \phi_0(\theta), \nabla \phi_1(\theta), \nabla G_L(\theta)) = -L \sin(\theta(s)), \] (3.14)
which vanishes everywhere on $[0, 1]$ for absolutely continuous $\theta$, only if $\theta \equiv 0$, and thus does not vanish for any $\theta \in M_{\theta_0, \theta_1, L}$. Thus at each $\theta \in M_{\theta_0, \theta_1, L}$ this set is linearly independent, and there is a tangent space
\[ T_0M_{\theta_0, \theta_1, L} = \{\nabla \phi_0(\theta), \nabla \phi_1(\theta), \nabla G_L(\theta)\}^\perp, \] (3.15)
of codimension 3. A linear combination of functions in the basis $B$ varies smoothly in $\theta$, and thus $M_{\theta_0, \theta_1, L}$, is locally a Riemannian manifold of codimension 3.

We can similarly induce the curve straightening flow on these manifolds. The projection of the squared curvature gradient $\nabla F(\theta)$ onto the tangent space takes the form
\[ \nabla^\pi F_L(\theta) = \nabla F_L(\theta) - \lambda_0 \nabla \phi_0(\theta) - \lambda_1 \nabla \phi_1(\theta) - \lambda_2 \nabla G_L(\theta) \] (3.16)
where the scalars $\lambda_j (j \in \{0, 1, 2\},)$ are chosen so that

$$\langle \nabla^xF_L(\theta), \nabla \phi_j(\theta) \rangle = 0, \ (j \in \{0, 1\},) \quad \text{and} \quad \langle \nabla^xF_L(\theta), \nabla G_L(\theta) \rangle = 0, \quad (3.17)$$

i.e., so that the projected gradient is in the tangent space. Note that a vanishing projected gradient $\nabla^xF_L(\theta) = 0$ again implies that $\theta$ satisfies a pendulum equation, since the evaluation functionals are linear. However, this is a significantly different setting from the free endpoints case. Here we do not expect that the natural boundary conditions necessarily hold, for physically clamped endpoints make this infeasible by least action, as curvature would vanish at the endpoints. We find three scalars $\lambda_j$ by solving the following system of equations:

$$\begin{pmatrix}
\langle \nabla \phi_0(\theta), \nabla \phi_0(\theta) \rangle & \langle \nabla \phi_0(\theta), \nabla \phi_0(\theta) \rangle & \langle \nabla G_L(\theta), \nabla \phi_0(\theta) \rangle \\
\langle \nabla \phi_0(\theta), \nabla \phi_1(\theta) \rangle & \langle \nabla \phi_1(\theta), \nabla \phi_0(\theta) \rangle & \langle \nabla G_L(\theta), \nabla \phi_1(\theta) \rangle \\
\langle \nabla \phi_0(\theta), \nabla G_L(\theta) \rangle & \langle \nabla \phi_1(\theta), \nabla G_L(\theta) \rangle & \langle \nabla G_L(\theta), \nabla G_L(\theta) \rangle
\end{pmatrix}
\begin{pmatrix}
\lambda_0 \\
\lambda_1 \\
\lambda_2
\end{pmatrix}
= 
\begin{pmatrix}
\langle \nabla F_L(\theta), \nabla \phi_0(\theta) \rangle \\
\langle \nabla F_L(\theta), \nabla \phi_1(\theta) \rangle \\
\langle \nabla F_L(\theta), \nabla G_L(\theta) \rangle
\end{pmatrix}$$

The upper 2 × 2 block of this Grammian matrix is easily computed, and since

$$\langle \nabla F_L(\theta), \eta \rangle = DF_L(\theta)\eta = \frac{1}{L} \int_0^1 \dot{\theta}\dot{\eta}ds$$

it is also clear that $\langle \nabla F_L(\theta), \nabla \phi_0(\theta) \rangle = 0$. Thus the system can be written as

$$\begin{pmatrix}
1 & 1 & a \\
1 & 2 & b \\
a & b & c
\end{pmatrix}
\begin{pmatrix}
\lambda_0 \\
\lambda_1 \\
\lambda_2
\end{pmatrix}
= 
\begin{pmatrix}
0 \\
d \\
e
\end{pmatrix}$$
with $a, b, c, d, e$ denoting the corresponding matrix elements. We can express these inner products as

$$a = DG_L(\theta)\nabla \phi_0(\theta) = -L \int_0^1 \sin(\theta(s)) ds = -y(1), \quad (3.18)$$

$$b = DG_L(\theta)\nabla \phi_1(\theta) = -L \int_0^1 (s + 1) \sin(\theta(s)) ds \quad (3.19)$$

$$c = DG_L(\theta)\nabla G_L(\theta) = L^2 \int_0^1 \sin(\theta(s)) \left( \int_0^s \left( \int_0^1 \sin(\theta(v)) dv \right) du + \int_0^1 \sin(\theta(u)) du \right) ds \quad (3.20)$$

where the integral formula (1.29) may also be useful,

$$d = \langle \nabla F_L(\theta), \nabla \phi_1(\theta) \rangle = \frac{1}{L} \int_0^1 \dot{\theta} ds = \frac{\theta(1) - \theta(0)}{L}, \quad (3.21)$$

and finally

$$e = DG_L(\theta)\nabla F_L(\theta) = - \int_0^1 \sin(\theta(s)) (\theta(s) - \theta(0)) ds. \quad (3.22)$$

Denoting

$$\delta = 2a^2 - 2ab + b^2 - c,$$

which is the negative determinant of the Grammian matrix, the coefficients $\lambda_j$ are given

$$\lambda_0 = \frac{d(c - ab) + e(2a - b)}{\delta}, \quad \lambda_1 = \frac{d(a^2 - c) + e(b - a)}{\delta}, \quad \lambda_2 = \frac{d(b - a) - e}{\delta}. \quad (3.23)$$

A non-vanishing $\delta$ ensures a well defined vector field, but in some instances these may be an indeterminate ratio with defined limit. One observes that when $a = 0$ one has $\lambda_0 = -\lambda_1$, which occurs when $y(1) = 0$, and $\lambda_0 = -\lambda_1$ also occurs if $e = d(b - a)$, and in the latter case one also has $\lambda_2 = 0$.

As mentioned, a zero projected gradient again implies that a critical point $\xi$ is a solution to the pendulum equation, as twice differentiation of the null equation $\nabla^2 F_L(\xi) = 0$ implies the
necessary differential equation (2.4). Thus, it follows in the same way that possible critical points have either cn or dn curvature. However, a stark contrast in the first derivative of this null equation should be pointed out, where one has

\[ \dot{\xi}(s) = \lambda_1 L + \lambda_2 L^2 \int_1^s \sin(\xi(u))du, \quad (3.24) \]

so in contrast to the pivotal equation (2.3), here one has \( \dot{\xi}(1) = \lambda_1 L \). Hence, one cannot automatically preclude dn curvature curves from criticality in these submanifolds, and in fact we show the ‘nearly’ critical dn curvature curves (2.28) are critical here for corresponding constraints. First, regarding cn curvature curves, the critical points in \( M_L \) must also be critical in these submanifolds when feasible, as the tangent space is a strict subspace of that for \( M_L \), and this amounts to projecting the zero vector onto a vector subspace, which is also zero — this may be referred to as the accidental criticality condition. Hence the following theorem is automatic:

**Theorem 3.3.3.** The cn curvature curves with indicatrices \( \xi \) in (2.9) are critical for

\[ \theta_0 = \pi - p - 2(-1)^p \sin^{-1}\left(\sqrt{k}\right) \quad \text{and} \quad \theta_1 = \pi - p - 2(-1)^{n+p} \sin^{-1}\left(\sqrt{k}\right) \quad (3.25) \]

with \( p \in \{0, 1\} \) in the manifolds \( M_{\theta_0, \theta_1, L} \).

Despite the preceding theorem being automatic, many of the necessary calculations are included, as it is of interest to demonstrate what happens to the Lagrange multipliers at these critical points. The calculations are straightforward, and similar to those in the proof of Theorem 2.6.2, so many of the details are omitted. We consider indicatrices of the previous general form

\[ \xi(s) = \pi - p + 2 \sin^{-1}\left(\sqrt{k} \, \text{sn}_k(K(k)(2ns + (-1)^p))\right), \quad p \in \{0, 1\}. \quad (3.26) \]
Calculating the scalar products \( a, b, c, d, e, f \), first of all one has \( a = 0 \) by (3.18). Next by (3.19) since \( y(1) = 0 \) one has
\[
b = -L \int_0^1 s \sin(\xi) ds,
\]
and by similar computations as in Theorem 2.6.2 one obtains
\[
b = L \frac{((-1)^n - 1) \sin^{-1} \left( \sqrt{k} \right)}{2n^2 K(k)^2}.
\]
Again by similar computations and with \( I_5 \) as denoted in Theorem 2.6.2, using (3.20) one has
\[
c = \frac{L^2 \sqrt{k} I_5}{n^2 K(k)^2}.
\]
Continuing, using (3.21) one has
\[
d = \frac{\xi(1) - \xi(0)}{L} = (-1)^p \frac{2((-1)^n - 1) \sin^{-1} \left( \sqrt{k} \right)}{L}.
\]
and finally using (3.22) with \( y(1) = 0 \) one obtains
\[
e = - \int_0^1 \sin(\xi_n) \xi ds = 4\sqrt{k} (-1)^p I_5.
\]

Turning to the scalars \( \lambda_j \), with these formulas one has
\[
\delta = b^2 - c = L^2 \frac{\left(((-1)^n - 1) \sin^{-1} \left( \sqrt{k} \right)\right)^2 - 4n^2 K(k)^2 \sqrt{k} I_5}{4n^4 K(k)^4}.
\]
Using the formula (2.21) for the integral \( I_5 \) and formula (2.42) for \( L \) gives
\[
\delta = \frac{\left((1 - (-1)^n) \sin^{-1} \left( \sqrt{k} \right)\right)^2 - 4n^2 \sqrt{k} K(k)(E(k) - (1 - k)K(k))}{4n^4 (2E(k) - K(k))^2 K(k)^2}.
\]
Then by (3.23) one has \( \delta \lambda_0 = dc - eb \), while \( dc = eb \) for the indicatrices \( \xi_n \). Also by (3.23), when \( a = 0 \), one has \( \lambda_1 = -\lambda_0 \), and thus

\[
\lambda_0 = 0 = \lambda_1. \tag{3.27}
\]

Finally by (3.23), with \( a = 0 \) one has \( \lambda_2 = (e - db)/\delta = (e - db)/(b^2 - c) \), where this is a ratio of proportional quantities, and thus

\[
\lambda_2 = \frac{e - db}{\delta} = (-1)^p \frac{4n^2 K(k)^2}{L^2}. \tag{3.28}
\]

Observe that we obtain the same scalar \( \lambda \) as in the proof of Theorem 2.6.2 with the free endpoint case, that is \( \lambda_2 = \lambda(\xi_n) \). Hence

\[
\nabla^p F_L(\xi_n) = \nabla F_L(\xi_n) - \lambda_0 - \lambda_1(1 + s) - \lambda_2 \nabla G_L(\xi_n)
\]

\[
= \nabla F_L(\xi_n) - \lambda(\xi_n) \nabla G_L(\xi_n) = 0,
\]

which shows directly that the projected gradient also vanishes in this setting for the indicatrices \( \xi_n \). ■

With these additional constraints it is natural to ask if the critical points analyzed in Chapter 2 are stable in these submanifolds. Here also, possible cn curvature critical points

\[
\theta(s) = n\pi \pm 2 \sin^{-1} \left( \sqrt{k} \text{sn}_k(as + b) \right) \quad (n \in \mathbb{Z},)
\]

while a path of cn curvature curves must also be path in the parameter space

\[
\mathcal{J} = \{(\alpha, a, b, k)|\alpha \in \mathbb{R}, \ a \in \mathbb{R}, \ b \in \mathbb{R}, \ k \in [0,1)\}, \tag{3.30}
\]
as feasible points along the path have the form
\[ \theta(s) = \alpha \pm 2 \sin^{-1} \left( \sqrt{k} \ \text{sn}_k(as + b) \right). \]  \hspace{1cm} (3.31)

Candidates for critical points must satisfy the system of equations
\[
\begin{aligned}
\alpha + 2 \sin^{-1} \left( \sqrt{k} \ \text{sn}_k(b) \right) &= \theta_0 \\
\alpha + 2 \sin^{-1} \left( \sqrt{k} \ \text{sn}_k(a + b) \right) &= \theta_1 \\
L \int_0^1 \cos(\theta(s))ds &= 1
\end{aligned}
\]  \hspace{1cm} (3.32)

As this is a system of three equations in four unknowns, the solution set may be one dimensional, so there may be special paths of cn curvature indicatrices as utilized in Chapter 2 in these manifolds as well. It may be tempting to add an additional angular constraint by fixing an interior angle to give a fourth equation \( \alpha + 2 \sin^{-1} \left( \sqrt{k} \ \text{sn}_k(as_0 + b) \right) = \theta_2 \), for \( 0 < s_0 < 1 \), as this may give a zero dimensional solution set, which would be an isolated lattice of points in the parameter space; this in turn would imply that cn curvature curves would be isolated in the resulting manifold — however, Linnér has shown that interior clamping will not induce a manifold in [16]. Some intuition is given by fixing the rotation angle \( \alpha \) in (3.29) and considering the intersection of three families of surfaces in a three dimensional parameter space, which intersect in a discrete lattice of isolated points; see Figure 3.3. However, the following is suspected since freedom to produce a constant energy curve is severely restricted, and the inability to produce such a path of no greater energy appears to preclude the existence of any such path.

**Conjecture 3.3.4.** The cn curvature curves analyzed in \( \mathcal{M}_L \) are stable with at least three angular constraints.
Figure 3.3: Some cn curvature curves and corresponding surfaces defined by the system (3.32); each curve corresponds a lattice point in the intersection of the three families of surfaces. There is a fourth dimension in the parameter space which may enable the construction of cn curvature paths between these curves.

With fixed terminal angles critical curves are not restricted to those which satisfy the natural boundary conditions, and may possibly have non-vanishing curvature at the endpoints. Thus Theorem 3.3.3 may not provide an exhaustive set of critical cn curvature elastica. Any indicatrix \( \theta(s) = \alpha + 2 \sin^{-1}\left(\sqrt{k} \ sn(\alpha s + b)\right) \), which satisfies the length constraint and also has the specified terminal angles, may be critical, where Theorem 3.3.3 pertains to \( \alpha \in \{0, \pi\}, a = 2nK, b = K \).

In addition, free of the natural boundary conditions, in this setting dn curvature curves can be critical. Critical dn curvature curves are the only possibility for larger angular variation, or winding, where defining angular variation

\[
\Delta := |\theta_1 - \theta_0|, \tag{3.33}
\]
we can say critical dn curves are the only type for $\Delta \geq 2\pi$, as cn curvature curves are precluded by periodicity. However, for $\Delta < 2\pi$ critical dn curvature curves may still exist, as the endpoint constraints can be met.

**Theorem 3.3.5.** The dn curvature curves with indicatrices $\eta$ as in Chapter 2 defined in equation (2.28) are critical in the manifolds $M_{0,\pm 2\pi n, L}$, that is for $\theta_0 = 0$ and $\theta_1 = \pm 2\pi n$, with $n \geq 1$ a counting number.

**Proof.** We restrict attention to $\theta_1 = -2\pi n$, as the other case will follow from symmetry, and use the general form

$$\eta(s) = \pi - 2\text{am}_k((2ns + 1)K(k)),$$

where one has

$$\sin(\eta(s)) = 2\text{sn}_k((2ns + 1)K(k))\text{cn}_k((2ns + 1)K(k)).$$

By calculations similar to those in the proof of Claim 2.2.3 for the case of free terminal angles, one has

$$\nabla F_L(\eta) = \frac{\pi - 2\text{am}_k((2ns + 1)K(k))}{L},$$

and

$$\nabla G_L(\eta) = L\left(\frac{\pi - 2\text{am}_k((2ns + 1)K(k)) + 4\sqrt{1 - knsK(k)}}{4kn^2K(k)^2}\right).$$

Turning to the scalar products $a, b, c, d, e, f$, first one has $a = 0$ since $y(1) = 0$. Also from $y(1) = 0$ it follows that $b = -L \int_0^1 s\sin(\eta)ds$, and using the integral formula

$$\int u \ \text{sn}_k(au + b)\text{cn}_k(au + b)du = \frac{\text{am}_k(au + b) - a\text{dn}_k(au + b)}{ak} + \text{const.},$$

one obtains

$$b = \frac{L \left(2\sqrt{1 - kK(k)} - \pi\right)}{2knK(k)^2}.$$
Using the formulas from the free terminal angle case in Section 2.2, one infers that

\[ c = DG_L(\eta)\nabla G_L(\eta) = L^2 E(k) + (1 - k)K(k) - \pi \sqrt{1 - k} \frac{k^2 n^2 K(k)^3}{k^2 n^2 K(k)^3}. \]

Next by (3.21) one has

\[ d = \frac{\eta(1) - \eta(0)}{L} = -\frac{2\pi n}{L}. \]

It is also noteworthy that

\[ d = \text{sign}(\Delta) \frac{\Delta}{L}. \]

Finally, using formula (3.22) and again that \( y(1) = 0 \) and similar calculations, one obtains

\[ e = \frac{2 \left(2E(k) - \pi \sqrt{1 - k}\right)}{kK(k)}. \]

With these formulas, the scalars simplify to

\[ \lambda_0 = \frac{4\sqrt{1 - knK(k)}}{L}, \quad \lambda_1 = -\frac{4\sqrt{1 - knK(k)}}{L} = -\lambda_0, \quad \lambda_2 = \frac{4kn^2 K(k)^2}{L^2}, \]

where \( y(1) = 0 \) so \( \lambda_1 = -\lambda_0 \) holds. Thus the projected gradient is given

\begin{align*}
\nabla^\pi F_L(\eta) &= \nabla F_L(\eta) - \lambda_0 \nabla \phi_0(\eta) - \lambda_1 \nabla \phi_1(\eta) - \lambda_2 \nabla G_L(\eta) \\
&= \nabla F_L(\eta) - \lambda_0 + \lambda_0 (1 + s) - \lambda_2 \nabla G_L(\eta) \\
&= \nabla F_L(\eta) + \lambda_0 s - \lambda_2 \nabla G_L(\eta).
\end{align*}
Moreover, one has

\[ \lambda_2 \nabla G_L(\eta) = \frac{4kn^2K(k)^2}{L^2} L \left( \pi - 2am_k((2ns + 1)K(k)) + 4\sqrt{1 - knsK(k)} \right) \]

\[ = \frac{(\pi - 2am_k((2ns + 1)K(k)) + 4\sqrt{1 - knsK(k)})}{L} \]

\[ = \nabla F_L(\eta) + \lambda_0 s, \]

which shows directly that \( \nabla F_L(\eta) = 0 \).

Under additional endpoint constraints, global minima may have non-zero curvature, which is in contrast to the critical straight lines which are global minima in the manifolds \( \mathcal{M}_L \). The following statement appears to be true for lack of other possibilities:

**Conjecture 3.3.6.** In the manifolds \( \mathcal{M}_{0,-2\pi n,L} \), the \( dn \) curvature indicatrices are global minima.

There are many other properties of these manifolds which are not understood.

### 3.4 Fixed Endpoints of the Associated Planar Curve

In Section 1.2 it was shown that a constant speed planar curve of length \( L \) parametrized over \([0, 1]\), with initial point \( \gamma(1) = (0, 0) \) has terminal point

\[ \gamma(1) = \left( L \int_0^1 \cos(\theta(u))du, L \int_0^1 \sin(\theta(u))du \right). \] (3.39)
This formula motivated defining the constraint functional \( x(1) = G_L(\theta) \), while in our considerations the terminal \( y \) coordinate was free to vary as much as the length constraint would allow. An alternate case of interest is given by defining

\[
X_L(\theta) := G_L(\theta) \quad \text{and} \quad Y_L(\theta) := L \int_0^1 \sin(\theta(u))du,
\]

and constraining both functionals \( X_L, Y_L \). In this way one obtains fixed endpoints for the associated planar curve. In particular, consider the endpoint constraints

\[
X_L(\theta) = 1, \quad \text{and} \quad Y_L(\theta) = 0,
\]

so that the associated planar curve has initial point \((0, 0)\) and terminal point \((1, 0)\). Also, there is an additional distinguishable case of interest which can be obtained setting

\[
X_L(\theta) = 0, \quad \text{and} \quad Y_L(\theta) = 0,
\]

where feasible curves with indicatrix \( \theta \) are closed and pass thru the origin.

Using that \( Y_L(\theta) = 0 \), the gradient for \( \nabla G_L \) from (1.28) simplifies to give

\[
\nabla X_L(\theta; s) = -L \int_0^s \left( \int_u^1 \sin(\theta(v))dv \right) du.
\]

By similar calculations as in Chapter 2, using that \( X_L(\theta) = 1 \), one obtains

\[
\nabla Y_L(\theta; s) = L \left[ \int_0^s \left( \int_u^1 \cos(\theta(v))dv \right) du + 1 \right].
\]
These gradient vectors generate a manifold, for by themselves each generates a manifold of codimension 1, and the normal vectors are non-parallel, so the tangent spaces generate $H$, and intersect transversally.

As in the previous settings, one has projected gradient

$$\nabla^\pi F(\theta) = \nabla F(\theta) - \lambda \nabla X_L(\theta) - \mu \nabla Y_L(\theta)$$

(3.45)

choosing scalars $\lambda, \mu$ such that

$$\langle \nabla^\pi F(\theta), \nabla X_L(\theta) \rangle = 0, \quad \text{and} \quad \langle \nabla^\pi F(\theta), \nabla Y_L(\theta) \rangle = 0,$$

(3.46)

to obtain a tangent vector field on $\mathcal{M}_L$. These conditions give a system of equations

$$\begin{pmatrix}
\langle \nabla X_L(\theta), \nabla X_L(\theta) \rangle & \langle \nabla Y_L(\theta), \nabla X_L(\theta) \rangle \\
\langle \nabla X_L(\theta), \nabla Y_L(\theta) \rangle & \langle \nabla Y_L(\theta), \nabla Y_L(\theta) \rangle
\end{pmatrix}
\begin{pmatrix}
\lambda \\
\mu
\end{pmatrix}
= \begin{pmatrix}
\langle \nabla F(\theta), \nabla X_L(\theta) \rangle \\
\langle \nabla F(\theta), \nabla Y_L(\theta) \rangle
\end{pmatrix},$$

with solution

$$\lambda = \frac{\langle \nabla F(\theta), \nabla X_L(\theta) \rangle \langle \nabla Y_L(\theta), \nabla Y_L(\theta) \rangle - \langle \nabla F(\theta), \nabla Y_L(\theta) \rangle \langle \nabla X_L(\theta), \nabla Y_L(\theta) \rangle}{\langle \nabla X_L(\theta), \nabla X_L(\theta) \rangle \langle \nabla Y_L(\theta), \nabla Y_L(\theta) \rangle - \langle \nabla X_L(\theta), \nabla Y_L(\theta) \rangle^2},$$

$$\mu = \frac{\langle \nabla F(\theta), \nabla Y_L(\theta) \rangle \langle \nabla X_L(\theta), \nabla X_L(\theta) \rangle - \langle \nabla F(\theta), \nabla X_L(\theta) \rangle \langle \nabla X_L(\theta), \nabla Y_L(\theta) \rangle}{\langle \nabla X_L(\theta), \nabla X_L(\theta) \rangle \langle \nabla Y_L(\theta), \nabla Y_L(\theta) \rangle - \langle \nabla X_L(\theta), \nabla Y_L(\theta) \rangle^2}.$$

For a vanishing projected gradient $\nabla^\pi F(\xi) = 0$, twice differentiation with respect to $s$ gives a necessary differential equation

$$\ddot{\xi}(s) = L^2(\lambda \sin(\xi(s)) - \mu \cos(\xi(s))).$$

(3.47)
It is well known this can be put in the form

\[ \ddot{\xi}(s) = L^2 \sqrt{\lambda^2 + \mu^2} \sin \left( \xi(s) - \arctan \left( \frac{\mu}{\lambda} \right) \right), \] (3.48)

which is again the pendulum equation. The \( cn \) curvature curves of interest in \( \mathcal{M}_L \) do satisfy the constraint, and are automatically critical here. Also, the \( dn \) curvature curves of interest in \( \mathcal{M}_L \) satisfy the constraint, and may be critical here. As far as we know, a classification of the critical points and stability in this setting is not known.
CHAPTER 4
A MODERN APPROACH TO THE CALCULUS OF VARIATIONS

The investigation of an isoperimetric problem in the calculus of variations via consideration of a gradient flow on a Hilbert manifold is more generally applicable. In this section we briefly discuss some associated isoperimetric problems to the curve straightening flow on $\mathcal{M}_L$, and also consider a standard fixed endpoint problem in the calculus of variations.

4.1 The Classical Calculus of Variations

It is apparent from the definition of a gradient flow in the first chapter, that our considerations are related to optimization of a functional over a function space. Classically, this is dealt with in the branch of mathematics known as the calculus of variations. We do not work with the classical theory, but rather take a more geometric point of view. However, a brief discussion of these methods is appropriate as some knowledge of this theory is assumed in this thesis. Here an overview of the assumed knowledge of the calculus of variations is given. There are many good sources of information on the calculus of variations, including [17], [18], [19].

A general problem in the calculus of variations is of the form

$$\text{Minimize } J(x) = \int_a^b f(t, x, \dot{x}) \, dt,$$

(4.1)
often subject to some constraints, in particular, fixed endpoints. The Euler-Lagrange differential equation is

\[ f_x - \frac{d}{dt} f_{\dot{x}} = 0, \quad (4.2) \]

and this equation is analogous to a zero derivative \( f'(x) = 0 \) for a single variable function, and is a necessary condition for a solution to (4.1). Solutions to the Euler-Lagrange equation are commonly referred to as extremals, despite that they may not be optimal – the name persists for historical reasons, and it is lamented in [18] that the terminology is too thoroughly ingrained to change, for sufficiency theory greatly outdates necessity theory. For unconstrained endpoints there are additional necessary conditions

\[ f_{\dot{x}}(a, x(a), \dot{x}(a)) = 0 = f_{\dot{x}}(b, x(b), \dot{x}(b)), \quad (4.3) \]

which are commonly referred to as natural boundary conditions.

A calculus of variations problem with an integral constraint of the form

\[
\text{Minimize } F(x) = \int_0^1 f(t, x, \dot{x}) dt, \quad \text{subject to } G(x) = \int_0^1 g(t, x, \dot{x}) dt = L, \quad (4.4)
\]

is called an isoperimetric problem. The main tool for solving isoperimetric problems is the multiplier rule, which asserts that an extremal for (4.4) is an extremal for the unconstrained problem

\[
\text{Minimize } K(x) = \int_0^1 [f(t, x, \dot{x}) + \lambda g(t, x, \dot{x})] dt, \quad (4.5)
\]

for some scalar \( \lambda \).
As the gradient has the interpretation of steepest descent, we can associate an isoperi-
metric calculus of variations problem to the investigation of critical points in a manifold 
\( M_L \), given by

\[
(P_L) \quad \text{Minimize} \quad \frac{1}{2L} \int_0^1 \dot{\theta}^2(s) ds \quad \text{subject to} \quad L \int_0^1 \cos(\theta(s)) ds = 1. \quad (4.6)
\]

The local nature of the rest points of this flow are connected with this parametrized family 
of problems, where a stable rest point of the gradient flow must be a local minimizer. By 
the multiplier rule one considers the accessory Lagrangian

\[
f_\lambda(s, \theta, \dot{\theta}) = \frac{1}{2L} \dot{\theta}^2(s) + L\lambda \cos(\theta(s)) \quad (4.7)
\]

for which the Euler-Lagrange equation can be put in the form

\[
\ddot{\theta} - \lambda L^2 \sin(\theta) = 0. \quad (4.8)
\]

It is noteworthy that up to the constant \( \lambda \) this is the same necessary differential equation as 
(2.4) for a vanishing projected gradient in \( M_L \).

Investigating sufficiency conditions begins by looking at second order information. The 
second variation is

\[
\delta^2 J(x; h) = \frac{1}{2} \int_0^1 \left[ f_{xx} h^2 + 2f_{x} \dot{h} \dot{h} + f_{xx} \dot{h}^2 \right] dt. \quad (4.9)
\]

A positive definite second variation plays the analogue a positive second derivative \( f''(x) > 0 \) 
for a single variable function. However, in constrast, positive definiteness of the second 
variation is only a necessary condition, and is not sufficient! In [19] pathological examples are 
given to demonstrate this. The culprit for such examples is unboundedness of the derivatives
for controlled pointwise variation. Attempts to rectify this have lead to the notions of *weak variations* versus *strong variations*, where weak variations not only pointwise bound variations from extremals, but also pointwise bound difference in derivatives — the former is weak in the sense of excluding nearby functions in the class of admissible functions. While Jacobi and Weierstrass built beautiful theories to determine sufficiency for weak and strong variations, respectively, each has difficulties in application. Second order information is avoided in Chapter 2 primarily due to these difficulties, but it is possible to consider the notion of a second variation in the Sobolev setting, where some developments can be found in [20] and [21].

It is noteworthy that a first order Sobolev space, which has norm involving the derivative, naturally induces weak variations. Moreover a Sobolev space which is also Hilbert (so with power $p = 2$) is an inner product space which induces a geometric structure. The classical choice of admissible functions being a continuous function space devoid of geometric structure is in many ways a complication, so associating a gradient descent on a Hilbert manifold may be a more natural setting for isoperimetric problems.

### 4.2 A Basic Problem in the Calculus of Variations

To apply the approach of viewing calculus of variations problems as optimization in a Hilbert manifold, consider the problem

$$(\mathcal{P}) \quad \text{Minimize } \Phi(x) = \frac{1}{2} \int_0^1 [x^2 + \dot{x}^2] \, dt \quad \text{Subject to } \ x(0) = 0, \ x(1) = 1. \quad (4.10)$$

We again take admissible functions in the same Hilbert space (1.8) endowed with the same inner product (1.9). (The choice of a dummy variable $t$ rather than $s$ signifies we are not viewing these as an arclength parameter of some curve.) One may view this as a modified
energy integral minimization, also taking the spatial square into account, a problem which arises in connection with the Dirichlet principle [3]. This problem may also be viewed as minimizing the norm induced by the (perhaps more standard) inner product

$$\langle u, v \rangle = \int_0^1 [u(t)v(t) + \dot{u}(t)\dot{v}(t)] dt,$$  \hspace{1cm} (4.11)

where the factor of $1/2$ is included in (4.10) to simplify gradient calculations. This is of the most standard type of calculus of variations problems, those simply having endpoint constraints. We can view the admissible class of curves in $H$ which satisfy this endpoint constraint as a flat submanifold, which is an affine subspace of co-dimension 2, as discussed in Section 3.2. Since this problem amounts to minimization of a squared norm over an affine subspace, this is a convex problem, and is thus quite well behaved.

To find the gradient of $\Phi$, compute

$$D\Phi(x)v = \frac{d}{d\varepsilon}[\Phi(x + \varepsilon v)]_{\varepsilon=0} = \int_0^1 [xv + \dot{x}\dot{v}] dt.$$

Then setting $D\Phi(x)v = \langle f, v \rangle$ for $f = \nabla \Phi$, one needs to solve

$$\int_0^1 [xv + \dot{x}\dot{v}] dt = f(0)v(0) + \int_0^1 \dot{f}(t)\dot{v}(t) dt.$$  \hspace{1cm} (4.12)

First, integrating by parts write

$$\int_0^1 x(t)v(t) dt = \int_0^1 \frac{d}{dt} \left[ \int_0^t x(u) du - \int_0^1 x(u) du \right] v(t) dt$$

$$= v(0) \int_0^1 x(u) du - \int_0^1 \left( \int_0^t x(u) du - \int_0^1 x(u) du \right) \dot{v}(t) dt.$$
Plugging this into (4.12) one deduces that $\nabla \Phi$ satisfies the initial value problem

$$
\frac{d}{dt} \Phi(x; t) = \dot{x}(t) - \int_0^t x(u) du - \int_0^1 x(u) du \quad \text{with} \quad \nabla \Phi(x; 0) = \int_0^1 x(u) du.
$$

Integrating one finds that $\Phi$ has gradient

$$
\nabla \Phi(x; t) = x(t) - \int_0^t \left( \int_0^u x(v) dv \right) du + t \int_0^1 x(u) du + \int_1^0 x(u) du. \quad (4.13)
$$

Now we have already computed the gradients of the projection functions in Section 3.2, with $\nabla \phi_1(x) = 1$, $\nabla \phi_2(x) = 1 + t$. We can thus compute a projected gradient

$$
\nabla^\pi \Phi(x) = \nabla \Phi(x) - \lambda_1 \nabla \phi_1(x) - \lambda_2 \nabla \phi_2(x),
$$

where $\lambda_1, \lambda_2$ are found choosing the scalars so that

$$
\langle \nabla \Phi(x) - \lambda_1 \nabla \phi_1(x) - \lambda_2 \nabla \phi_2(x), \nabla \phi_j(x) \rangle = 0 \quad \text{for} \quad j \in \{1, 2\},
$$

whence the projected gradient is in the tangent space of the manifold $\Omega$. These null inner products result in a linear system of equations, which can be written as a matrix equation involving the Grammian for $\{\nabla \phi_1(x), \nabla \phi_2(x)\}$. This is readily solved as follows:

$$
\begin{pmatrix}
\langle \nabla \phi_1(x), \nabla \phi_1(x) \rangle & \langle \nabla \phi_2(x), \nabla \phi_1(x) \rangle \\
\langle \nabla \phi_1(x), \nabla \phi_2(x) \rangle & \langle \nabla \phi_2(x), \nabla \phi_2(x) \rangle
\end{pmatrix}
\begin{pmatrix}
\lambda_1 \\
\lambda_2
\end{pmatrix}
= 
\begin{pmatrix}
\langle \nabla \Phi(x), \nabla \phi_1(x) \rangle \\
\langle \nabla \Phi(x), \nabla \phi_2(x) \rangle
\end{pmatrix}
$$

$$
\implies 
\begin{pmatrix}
1 & 1 \\
1 & 2
\end{pmatrix}
\begin{pmatrix}
\lambda_1 \\
\lambda_2
\end{pmatrix}
= 
\begin{pmatrix}
\langle \nabla \Phi(x), \nabla \phi_1(x) \rangle \\
\langle \nabla \Phi(x), \nabla \phi_2(x) \rangle
\end{pmatrix}
\[
\begin{pmatrix}
\lambda_1 \\
\lambda_2
\end{pmatrix} = \begin{pmatrix}
2\langle \nabla \Phi(x), \nabla \phi_1(x) \rangle - \langle \nabla \Phi(x), \nabla \phi_2(x) \rangle \\
\langle \nabla \Phi(x), \nabla \phi_2(x) \rangle - \langle \nabla \Phi(x), \nabla \phi_1(x) \rangle
\end{pmatrix}
\]

The inner products occurring in the scalars \( \lambda_j \) are computed as

\[
\langle \nabla \Phi(x), \nabla \phi_1(x) \rangle = x(0) + \int_0^1 x(u)du,
\]

and

\[
\langle \nabla \Phi(x), \nabla \phi_2(x) \rangle = x(0) + 2 \int_0^1 x(u)du + \int_0^1 \left( \dot{x}(t) - \int_0^t x(u)du \right) dt.
\]

We can then consider a projected negative gradient flow induced by

\[
\frac{dx}{dt} = -\nabla^\pi \Phi(x),
\]  

(4.14)

so that a solution of \((P)\) must also be a rest point of this flow. Since \( \Phi \) is convex, a rest point of this flow must be a solution of \((P)\). Then suppose that \( \xi \) is a rest point, \( i.e., \) suppose that \( \nabla^\pi \Phi(\xi) = 0 \). Then one has \( \frac{d}{dt} \nabla^\pi \Phi(\xi) = 0 \) also, so computing this gives the equation

\[
\dot{\xi}(t) = \int_0^t \xi(u)du + \int_0^1 \xi(u)du.
\]

This equation implies that \( \xi \) has at least one more derivative, with

\[
\ddot{\xi}(t) = \xi(t).
\]

This differential equation with boundary values \( \xi(0) = 0, \xi(1) = 1 \) admits a unique solution
\[ \xi(t) = \frac{2e}{e^2 - 1} \sinh(t), \]

which is in fact \( C^\infty \).
CHAPTER 5
NUMERICAL METHODS

We do not know how to solve the projected gradient flow equation (1.43) explicitly, and the same is typically true for other differential equations on submanifolds of $H$. Thus, it is natural to seek numerical means of investigating the flow, not only for application, but also for testing conjectures, and even simply to check for consistency with theoretical deduction. A slight modification of the family of Adams-Bashforth methods has been found to perform remarkably well in numerical discretization of the projected gradient flow equations considered in this thesis; for further details on the Adams-Bashforth methods than given here we reference [22]. What is presented in this chapter is largely empirical.

5.1 Discretization of Flows on Submanifolds

For a ‘differential equation’ $\frac{dx}{dt} = X(x)$ with $x \in \mathcal{M} \subset H$, on a submanifold $\mathcal{M}$ of an ambient Hilbert space $H$, we have experimented with a slight modification of the discretizations (D.7) for numerical solution. Despite the differential equation taking place in an infinite dimensional phase space, given the coordinate free representation, one simply has discretization

$$x_{n+1} + \sum_{j=1}^{k} \alpha_j x_{n+1-j} = h \sum_{j=0}^{k} \beta_j X \left( x_{n+1-j} \right), \quad (5.1)$$

and in particular we have used the Adams-Bashforth (ABK) methods briefly discussed in Appendix D. The slight modification is in the sense that there is only a single equation here, and the output is a function — we essentially treat it as discretization of a single variable
autonomous equation. We are primarily interested in the projected gradient flow (1.43) with \( X = -\nabla^g F \), and the bulk of our experimentation has been with this example. We have found satisfactory results by using an Euler method of discretization to obtain the first \( k \) iterates for initialization, with the Euler method implemented in a similar way. Given the existence of Taylor expansions in the setting of a Hilbert manifold, the usual proofs of convergence properties of these methods can likely be extended to establish similar convergence properties in this case, but we have not worked thru the details on this.

In many runs, an adaptive scheme was implemented by choosing the step size relative to the magnitude of the gradient at a given step, to correct for larger curvature in the manifold. For step size \( h \), it was found to be effective to take adaptive size

\[
h_n = \frac{h}{1 + \alpha \int_0^1 X(x_n(s))^2 ds}
\]

for some positive real parameter \( \alpha \). Here we reduce step size relative to the \( L^2 \) norm of the vector field \( X \), but another metric may be appropriate. For a change in step size when implementing a multistep method, many practitioners will recompute the initial \( k \) values — however, we found satisfactory performance without recomputing the previous \( k \) values when changing the step size.

For the examples in this chapter, we primarily use the method with \( k = 3 \), (AB3). The corresponding parameters \( \beta_j \) are given \( \beta_0 = \frac{5}{12} \), \( \beta_1 = -\frac{4}{3} \), \( \beta_2 = \frac{23}{12} \), \( \beta_3 = 0 \). Other methods in this family have been employed, but we have found AB3 to give satisfactory speed and performance. When using an adaptive step size (5.2), often \( \alpha = 50 \) was used on an empirical basis. Optimality among these choices was not investigated in detail, and this was done in a largely experimental manner. Clarification of these methods may be done in future work.

To demonstrate the effectiveness of this discretization, first take \( X = -\nabla F \) in (5.1) and consider the unconstrained flow; as discussed in Section 1.4 it has exact solution \( \theta_t(s) = \)
\[ \theta_0(0) + e^{-\frac{1}{L}} (\theta_0(s) - \theta_0(0)) \]. For instance, discretize with AB3, taking initial value \( \theta_0 = \xi_1 \), the critical elastic loop with \( L = 2 \), which is not a rest point of this flow. Running 1000 iterations with step size \( h = .01 \), the curve nearly straightens, with observed error of the closed form solution of \( \approx 10^{-10} \) in \( L^2 \) norm. Moreover, the error decreases as the curve straightens, which is likely an artifact of global stability. Nevertheless, the discretization of the unconstrained flow is very stable in this example. A plot is in Figure 5.1.\(^1\)

---

\(^1\)In all discretization plots, there is decreasing gray level, so the curves darken along the flow, and the initial curve is in dashed black.
5.2 Discretization in $\mathcal{M}_L$.

To investigate the projected negative gradient flow on $\mathcal{M}_L$ we take $X = -\nabla^\pi F$ in discretization (5.1). There is a slight difficulty in computing the scalar field. Writing

$$\lambda = -\frac{1}{L} \int_0^1 \sin(\theta(u)) (\theta(s) - \theta(0)) du \bigg/ \int_0^1 \sin(\theta(s)) \left( \int_0^s u \sin(\theta(u)) du + s \int_s^1 \sin(\theta(u)) du + \int_0^1 \sin(\theta(u)) du \right) ds,$$

using the non-iterated integral form (1.29) for the gradient $\nabla G(\theta)$, enables numerical integration via an interpolating polynomial

$$p(s) \approx \int_0^s u \sin(\theta(u)) du + s \int_s^1 \sin(\theta(u)) du + \int_0^1 \sin(\theta(u)) du,$$

which is inserted into (5.3). For the interpolating polynomial one can match successive derivatives by setting

$$p'(s_j) = \int_{s_j}^1 \sin(\theta(u)) du, \quad p''(s_j) = \sin(\theta(s_j)), \quad p'''(s_j) = \cos(\theta(s_j)) \dot{\theta}(s_j), ...$$

at mesh points $0 \leq s_j \leq 1$, provided that the derivatives exist; while these do not necessarily exist in the class of functions $H$, for a sufficiently smooth initial indicatrix this can be implemented; we have found that a $C^1$ pasting of lines and circles suffices for an initial value. For $n$ matching derivatives a spline of order $2n + 1$ may be used, and it appears that optimal accuracy occurs with higher degree than cubic splines. However, we have found this discretization to run remarkably stably with splines, and even without matching the first derivatives.

For discretization in $\mathcal{M}_L$ we have investigated both $L < \infty$ and $L = \infty$ in many examples, but as seen in the analysis of the critical points in Chapter 2 the settings are considerably
Figure 5.2: A tightly wound single loop flows near the critical elastica $\xi_1$, and the gradient gets very small. The flow nearly stops when close to the saddle point.

Figure 5.3: For more loops, similar behavior is seen if tightly wound different. We primarily give examples for $L < \infty$ in this section since we found the dynamics near a tightly wound loop in this setting the most interesting. We originally suspected that there may be a bifurcation value $L$ at which the loops became local minimizers if tightly enough wound, but of course in Chapter 2 this was proven to not be the case — this is not clear from the behavior of the discretization, although we have found the discretization to be consistent with the theory. To investigate the behavior near an elastic loop, (of cn or dn curvature type,) perhaps the simplest choice in initial curve is a piecewise pasting of lines and circles which is near the loop. It is simple to set up such an initial curve by pasting lines and circles, as this involves parameters determined by break points and slopes of lines corresponding to circular segments, and one can choose the final parameter to satisfy the constraint $G_L(\theta_0) = 1$. We have set up a program for doing this, but have not found much
Figure 5.4: With enough length, the loop freely unwinds; the uppermost curve is the final iterate plotted, and this quickly approaches (until nearby) the upper straight line segment which is a global minimizer. Here $\theta_0$, in dashed black, is a smooth function constructed using tangent inverse.

depth or theoretical interest in these constructions, so they are not included. Of course smooth candidate curves may also be devised, but the initial curve need only have indicatrix in $H$, so smoothness is not necessarily needed, and in fact it has been observed that for non-smooth initial curves, the flow will quickly smooth them. Also, one need not worry about the constructed initial curve having lesser energy, for this is highly unlikely! Such a pasting will almost certainly have greater energy than an elastic loop, and will thus initially flow toward the loop of interest.

The behavior of a discretization near a critical elastic loop can be rather deceiving when $L$ is near 1, which in early investigations lead to expectations that the critical point structure may be rather different than what is proven in Chapter 2. When $L$ is large, the curve will unwind relatively quickly for large step size. When $L$ is smaller, the flow will come closer to the saddle point, and is surprisingly less stable for larger step sizes, despite the gradient being smaller. Figures 5.2 and 5.3 show discretizations which flow very close to the saddle
Figure 5.5: The behavior is a bit different with an even number of loops, but with enough length, it still freely unwinds. On the left, a discrete set of points flowing toward the saddle point $\xi_2$. On the right another discrete set of points from the same discretization; after slowing near the saddle point, the flow increases speed and the curve unwinds, where a further curve down the flow is plotted in black.

Furthermore, in general when running near a tightly wound curve, the step size must be continually decreased as the gradient tends to zero, which occurs in very flat regions of the manifold — such examples have been run for months essentially stuck near a saddle point, or simply near a ‘nearly’ critical point! However, the discretization remains stable and slowly releases energy if the step size is reduced. In particular this is observed taking initial curves with larger angular variation which produce examples where the gradient can be made arbitrarily small while not being near a critical point by sufficiently increasing the angular variation. Such curves flow near the ‘nearly’ critical dn curvature elastica $\eta_n$. For example, in Figure 5.8 the initial curve has a winding number 4, so $8\pi$ angular variation, and the discrete plot summarizes results of a very long run time.
Figure 5.6: With a slightly different configuration, tightly wound loops still flow near critical elastica.

Figure 5.7: Beginning with winding number two, tightly wound. This will very slowly flow past the 'nearly' critical dn curvature double loop.

It has been claimed that the discretization (5.1) works remarkably well for the negative projected gradient flow on $\mathcal{M}_L$. To clarify, this is evidenced by unusual long term invariant preservation, where the constraint $L \int_0^1 \cos(\theta) ds = 1$ does not show drift, but rather slight oscillation or noise. For instance, the behavior seen in the example of Figure 5.9 is typical, where there is not drift, but rather oscillation as the curve unwinds to a straight line segment. The oscillations stabilize when nearing the global minimizer. In some instances, this was run for several months with such invariant preservation holding until the algorithm was terminated; for instance see Figure 5.8, and also more tightly wound loops were run. This long term accuracy and invariant preservation is not typical behavior for numerical discretization,
Figure 5.8: Plotted is the flow of a circular loop with winding number four, which was quite costly to compute. Below the initial curve (dotted black,) final curve in the flow (gray,) and the \( \eta_4 \) curvature indicatrix (black) are plotted. Eventually this will flow near \( \eta_4 \).

but this property is enjoyed by symplectic integrators of Hamiltonian differential equations which govern mechanical systems in physics, (the monograph [23] is an excellent source of information on such discretizations,) and it is relevant that there is a physical interpretation of the elasticity dynamics of a bent wire or beam. However, this behavior is not understood here.

Figure 5.9: On the left is a plot of the energy \( F(\theta_n) \), and on the right the invariant \( x(1) \) for initial curve a smooth single non-elastic loop with \( L \approx 3 \) and step size \( h = .01 \) with the AB3 discretization. Aside from an initial jump in the invariant, there is just small self correcting oscillation. This level of accuracy in long term invariant preservation is typical for the Adams-Bashforth methods.
We have also run a gradient ascent by taking $X = \nabla^x F$ in discretization (5.1). The behavior is similar, but the dynamics are a bit less predictable; see Figure 5.16.

### 5.3 Discretization With Fixed Terminal Angles

For fixed terminal angles, the discretization (5.1) can be implemented with the corresponding projected gradient using formulas from Section 3.3. For computing the scalar field, numerical integration via interpolation as discussed in Section 5.2 can be employed. We again observe similar preservation in the invariants, three of them now, with no long term drift in the error, and with no error in terminal angles in most discretizations!

![Figure 5.10: Discretizations for fixed terminal angles converging to cn sections. On the left initial points in the discretizations are plotted, beginning with the dotted triple loop. On the right the same discretization is run to a limiting black curve, a critical cn curvature elastica.](image)

For many constraints the dynamics are rather predictable qualitatively, but the exact form of final curve may not be; in this way this discretization is useful for producing an approximation of the minimal energy curve in this configuration, and likewise the shape of a clamped elastic wire when relaxed. For instance, in Figure 5.10 points from a discretization
flowing to sections of a critical cn elastica are plotted. In general it is unclear if or when such an example will flow to this or a nearby dn elastica, as both are critical in this setting.

Figure 5.11: Discretizations for fixed terminal angles with $\Delta = 2\pi$, $5\pi$, $16\pi$ the respective angular variations.

When the angular variation (3.33) satisfies $\Delta \geq 2\pi$ the only possible limiting curve of the flow is a section of a dn curvature elastica, as cn curvature curves are infeasible. For example, given a particular length $L > 1$ consider a family of initial curves with $\theta_0 = \pi$ and $\theta_1 = (n + 1)\pi$ with $n \geq 2$, so that the angular separation is $\Delta = n\pi$. Such initial curves are easily generated by taking a piecewise linear curve with a semi-circle, as straight line, and an additional $n$ semi-circles. As $\Delta \geq 2\pi$ cn curvature curves are infeasible, and it is rather
predictable that suitable initial curves in this manifold will flow to sections of dn curvature curves given

$$\eta = \pi + 2am_k ((ns + 2)K(k)).$$

(5.6)

In Figure 5.11 plots for $n = 2, 5, 16$ are given, and in each case these were run until the projected gradient had $L^2$-norm nearly machine precision zero.

### 5.4 Discretization With Fixed Endpoints

For fixed endpoints as discussed in Section 3.4 we have found the discretization (5.1) to work similarly, and with long term invariant preservation. We do not include any examples here, but it is noteworthy that one should not set the functional evaluations $X_L(\theta_n), Y_L(\theta_n)$ equal to their constraint values when they arise in the gradient formulas, for if this simplification is done, the long term preservation is not seen – apparently the incurred error contributes to this property, and essentially self corrects the drift that is typically seen in discretization.

### 5.5 Discretization of Basic Calculus of Variations Problem

We have mentioned that the unconstrained flow gives a means of testing the numerical discretization employed here, but the setting of the basic problem considered in Section 4.2 perhaps gives a better example since there are constraints. We similarly apply the Adams-Bashforth discretization from Section 5.1 using the gradient formulas from Chapter 4.

Excellent behavior was seen in all examples which were implemented, with fast convergence to the global minimizer found in Section 4.2, but this is not too surprising due to
convexity in that setting with the action taking place on a flat submanifold. In particular, an ill behaved initial function is given as \( x_0 = \sqrt{t} \), where the objective functional evaluates to \( \Phi(x_0) = \frac{1}{2} \int_0^1 \left[ t + \frac{1}{2t} \right] dt \), which is a divergent improper integral — thus, rather ill-behaved in this sense! Nevertheless, running 100 iterations at a rather large step size \( h = .1 \), this took \( \approx 17.06 \) seconds on an average laptop computer, with a final error at machine precision with

\[
\int_0^1 \left( x_{100}(t) - \frac{2e}{e^2 - 1} \sinh(t) \right)^2 dt \approx 1.33 \times 10^{-17}.
\]

Moreover the max invariant error, that is deviation from the fixed endpoints, is just \( \approx 6.66 \times 10^{-16} \), also at machine precision, and this is despite a deliberately ill choice in initial point and a large step size taken! Similar convergence rate and accuracy was observed with every choice in initial curve, even with deliberately ill behaved choices.

### 5.6 Numerically producing paths

We have experimented with many different means of producing paths to the critical points from nearby points in the manifolds \( \mathcal{M}_L \), and here define two such constructions.
5.6.1 The Attracting Flow

The attractor flow with respect to a given point \( \xi \in H \) is induced by projecting the vector from any arbitrary point to \( \xi \) onto the tangent space. For an elastica \( \xi \) at any point \( \theta \in \mathcal{M}_L \), project the vector \( \xi - \theta \) onto the tangent space at \( \theta \) to produce the vector field

\[
V(\theta) = \xi - \theta - \frac{\langle \nabla G_L(\theta), \xi - \theta \rangle}{\langle \nabla G_L(\theta), \nabla G_L(\theta) \rangle} \nabla G_L(\theta).
\]

(5.7)

This induces the attracting flow

\[
\frac{d\theta}{dt} = V(\theta).
\]

(5.8)

We do not know how to solve the differential equation (5.8) but have discretized it with \( X = V \) in (5.1). This was found to be effective in producing an approximate path approaching \( \xi \) in \( \mathcal{M}_L \). However, the attracting flow did not enjoy the long term invariant preservation seen with the curve straightening flow; see Figure 5.13.

Figure 5.13: A drift in the invariant is typically seen for the attracting flow, (as in the list plot on the left,) unlike the curve straightening flow which enjoys long term invariant preservation. On the right a path joining two points is produced numerically with the attracting flow on a monkey saddle in \( H = \mathbb{R}^3 \).
5.6.2 Gradient Projection

Given two points $\theta, \eta \in \mathcal{M}_L$, the segment $[\theta, \eta]$ in the ambient space is given

$$\gamma_0(t) = \theta + (\eta - \theta)t, \quad t \in [0, 1].$$  \hspace{1cm} (5.9)

For $0 < \tau < 1$ it is unlikely that $\gamma_0(\tau) \in \mathcal{M}_L$. However, if $||\theta - \eta||_H$ is not too large, then for some $\tilde{L} \approx L$ one has $\gamma_0(\tau) \in \mathcal{M}_{\tilde{L}}$. It is possible that there exists a scaling of the normal vector at $\gamma_0(\tau)$ which intersects the manifold $\gamma_0(\tau) \in \mathcal{M}_L$, that is there may exist $\rho_\tau$ with

$$\gamma_0(\tau) + \rho_\tau \nabla G_L(\gamma_0(\tau)) \in \mathcal{M}_L. \hspace{1cm} (5.10)$$

If so, since $\tilde{L} \approx L$ one may approximate $\rho_\tau$ as a root of the transcendental equation

$$L \int_0^1 \cos (\gamma_0(t) + \rho_\tau \nabla G_L(\gamma_0(t))) \, ds = 1. \hspace{1cm} (5.11)$$

Using a discrete set of points $0 < \tau < 1$, via interpolation one can obtain an approximate path, or simply estimate a point when needed with formula (5.10). It is not clear when this method will work, but this has typically been found to be an effective way of approximating a path between nearby points in examples; see Figure 5.14.
Figure 5.14: Above the gradient projection is depicted with discrete sets for the linear segment on the left, and the projected path on the right. Below a discrete collection of indicatrices along the resulting path joining the two points. It is noted that one point is a $C^{1}$ pasting of lines and circles, which is not actually in $H$, (like many of our initial curves for discretization,) but the method still works well.
Figure 5.15: Beginning with winding number three and enough slack to freely unwind. Also included in the plots are some associated non-critical dn curvature elastica.
Figure 5.16: A gradient ascent run in $\mathcal{M}_\infty$, where as in all plots the curves darken with flow time. This endlessly winds with increasing angular variation and energy.
REFERENCES


APPENDIX A

JACOBI ELLIPTIC FUNCTIONS, AND ELLIPTIC INTEGRALS
Here, a brief overview of elliptic functions and elliptic integrals is given, which is by no means exhaustive, or completely rigorous; of course the focus is primarily on what is applied in this thesis. There is a wealth of literature on elliptic functions theory, with recommended sources [24], [25], and [26]. However, conventions vary throughout the literature and software libraries, and those used in this thesis are presented here. The Jacobi elliptic functions fall into the taxonomy of doubly periodic meromorphic functions of a complex variable, but we only consider them with real variables here.

A.1 Trigonometry on an Ellipse

The starting point for elliptic functions could be taken as the consideration of a theory of “quasi-trigonometry” per say, considering a more general formulation for the circle replaced with an ellipse. This development is considered in [27], while many mathematicians would find it too rudimentary — yet some others may see aesthetic and heuristic value. For a given ellipse we are only interested in the ratio of semi-axes lengths, say the ratio $a : b$, and can normalize the ellipse by scaling to $b = 1$, to give the unit ellipse. Now for $a \geq b = 1$, take $c \leq a$ as the number such that

$$a^2 = b^2 + c^2 \quad \Longrightarrow \quad c^2 = a^2 - 1.$$ 

The points $(\pm c, 0)$ are well known to be the foci of the ellipse, and geometrically the foci are at a distance $a$ from the vertices $(0, \pm 1)$; see Figure A.1. To a given ellipse we attach a parameter in terms of its normalization, the elliptic modulus may be defined as

$$k = \frac{c^2}{a^2} = \frac{a^2 - 1}{a^2} < 1,$$
Also $k$ is the eccentricity of the ellipse. In the limit as $k \uparrow 1$, the ellipse tends to a parabola.

The unit ellipse is described by the algebraic equation

$$\frac{x^2}{a^2} + y^2 = 1. \quad (A.1)$$

For a point $p = p(x, y)$ on the unit ellipse with parameter $k$ having an angular separation $\theta$ with the positive $x$-axis, the three basic functions are defined as

$$\text{sn}_k(u) = y, \quad \text{cn}_k(u) = \frac{x}{a}, \quad \text{dn}_k(u) = \frac{r}{a},$$

for $r := \sqrt{x^2 + y^2}$, where the argument $u$ is given by the integral

$$u = \int_{(a,0)}^{p(x, y)} rd\theta.$$ 

While $u$ is in general neither an angle nor an arclength, if $a = 1$, (in which case $k = 0$,) one has $u = \theta$, and recovers sine and cosine for sn and cn, and dn becomes 1. From (A.1), one has by construction the identity

$$\text{sn}_k^2(u) + \text{cn}_k^2(u) = 1.$$
While this development is fun, it is rather elementary, and has clear setbacks, so it is not further pursued here, despite having intuitive value.

### A.2 The Jacobi Elliptic Functions

Another development is to consider the functions sn, cn, dn as solutions to the dynamical system
\[
\begin{align*}
\dot{x} &= yz, \\
\dot{y} &= -zx, \\
\dot{z} &= -kxy, \\
(0 \leq k < 1)
\end{align*}
\] (A.2)

with solutions
\[
\begin{align*}
x(t) &= \text{sn}_k(t), \\
y(t) &= \text{cn}_k(t), \\
z(t) &= \text{dn}_k(t),
\end{align*}
\] (A.3)

as in [28]. These functions may be referred to as the elliptic sine, the elliptic cosine, and the elliptic delta, respectively. The parameter $k$ is called the elliptic modulus.

**Remark A.2.1.** It is common to use $k^2$ rather than $k$ in definition (A.2) which would accordingly alter subsequent formulas involving the elliptic modulus. However, throughout this thesis we use the resulting formulas specified in this appendix.

The dynamical system (A.2) generalizes the harmonic oscillator, which may be used as the dynamical system to define sine and cosine, where for $k = 0$ the harmonic oscillator arises with $\text{dn}_k(t) \equiv 1$; by this definition
\[
\begin{align*}
\text{sn}_0(t) &= \sin(t), \\
\text{cn}_0(t) &= \cos(t), \\
\text{dn}_0(t) &\equiv 1.
\end{align*}
\] (A.4)
By Picard’s existence theorem, the system admits a unique solution, so the three functions are well defined as such. From definition (A.3) one automatically has derivative formulas

\[
\frac{d}{dt} sn_k(t)dn_k(t), \quad \frac{d}{dt} cn_k(t)dn_k(t), \quad \frac{d}{dt} dn_k = -k sn_k(t)cn_k(t). \quad (A.5)
\]

There are twelve standard Jacobi elliptic functions which can be expressed in terms of sn, cn, dn, which are all denoted as two letters from \{c, d, s, n\} by a simple rule:

(i) For a ratio, take the first letter in the numerator followed by the first letter from the denominator; e.g. \(sd=sn/dn\), and \(cs=cn/sn\).

(ii) For a reciprocal, reverse the order of the two letters; e.g. \(nd=1/dn\).

In this way we get twelve functions: \(sn, cn, dn, ns, nc, nd, sd, sc, cs, cd, ds, dc\).

There are fundamental identities

\[
\begin{align*}
\text{sn}^2_k(t) + \text{cn}^2_k(t) = 1, \quad \text{and} \quad \text{dn}^2_k(t) + k \text{ sn}^2_k(t) = 1. \\
\end{align*}
\]  

(A.6)

Generalizing trigonometric sum identities, we have elliptic sum identities

\[
\begin{align*}
\text{sn}_k(a + b) &= \frac{\text{sn}_k(a)\text{cn}_k(b)\text{dn}_k(b) + \text{sn}_k(b)\text{cn}_k(a)\text{dn}_k(a)}{1 - k \text{ sn}^2_k(a)\text{sn}^2_k(b)}, \\
\text{cn}_k(a + b) &= \frac{\text{cn}_k(a)\text{cn}_k(b) - \text{sn}_k(a)\text{sn}_k(b)\text{dn}_k(a)\text{dn}_k(b)}{1 - k \text{ sn}^2_k(a)\text{sn}^2_k(b)}, \\
\text{dn}_k(a + b) &= \frac{\text{dn}_k(a)\text{dn}_k(b) - k \text{ sn}_k(a)\text{sn}_k(b)\text{cn}_k(a)\text{cn}_k(b)}{1 - k \text{ sn}^2_k(a)\text{sn}^2_k(b)}. \\
\end{align*}
\]  

(A.7) (A.8) (A.9)
Figure A.2: The functions cn, sn, dn plotted over one period for $k = .25$ on the left, and $k = .9$ on the right. These converge to cosine, sine and 1, respectively, as $k \downarrow 0$.

In the upper limit as $k \uparrow 1$ (which corresponds to the limit of a very stretched out ellipse, that is, as the conic section approaches a parabola,) the elliptic functions tend to hyperbolic trigonometric functions

$$\lim_{k \uparrow 1} sn_k(t) = \tanh(t), \quad \lim_{k \uparrow 1} cn_k(t) = \text{sech}(t), \quad \lim_{k \uparrow 1} dn_k(t) = \text{sech}(t).$$

(A.10)

The functions sn, cn, dn are periodic, where there is a number $K$, which depends on $k$, for which dn has period $2K$, and sn, cn have period $4K$. For all integers $n$, one has

$$sn_k(2nK) = 0 = cn_k((2n + 1)K), \quad \text{and} \quad cn_k(2nK) = (-1)^n = sn_k((2n + 1)K),$$

(A.11)

which is more general than formulas for the zeros of sine and cosine, while for the dn function integer multiples of the half period are given

$$dn_k(2nK) = 1, \quad \text{and} \quad dn_k((2n + 1)K) = \sqrt{1 - k}.$$  

(A.12)

These values are extrema for the dn function, Also, sn is odd, and cn, dn are even.
A.3 Elliptic Integrals and the Jacobi Amplitude

To compute the period, for $x(t) = \text{sn}_k(t)$ use (A.5) and (A.6) to obtain

$$\dot{x}^2 = \text{cn}_k^2(t) \text{dn}_k^2(t) = (1 - x^2)(1 - kx^2). \quad (A.13)$$

This is solvable by quadrature, where one has

$$\frac{dx}{dt} = \sqrt{(1 - x^2)(1 - kx^2)}, \quad (A.14)$$

and integrating gives

$$t = \int_0^{\text{sn}_k(t)} \frac{dx}{\sqrt{1 - x^2} \sqrt{1 - kx^2}}. \quad (A.15)$$

Thus for $y = \text{sn}_k(t)$ one can write

$$\text{sn}_k^{-1}(y) = \int_0^y \frac{dx}{\sqrt{1 - x^2} \sqrt{1 - kx^2}}. \quad (A.16)$$

It is noteworthy that for $k = 0$ this becomes sine inverse. The quarter period $K$ is attained when $\text{sn}_k(t) = 1$, so one has

$$K(k) = \int_0^1 \frac{dx}{\sqrt{1 - x^2} \sqrt{1 - kx^2}} = \int_0^{\frac{\pi}{2}} \frac{d\phi}{\sqrt{1 - k \sin^2(\phi)}}, \quad (A.17)$$

having substituted $x = \sin(\phi)$ in the last integral. The value in (A.17) is known as the elliptic $K$ integral, which may be viewed as a function of the elliptic modulus $k$.

A series representation for $K(k)$ can be obtained from Newton’s binomial theorem, where

$$K(k) = \frac{\pi}{2} + \sum_{n=1}^{\infty} \left( \frac{1 \cdot 3 \cdot 5 \cdots (2n - 1)}{2 \cdot 4 \cdot 6 \cdots (2n)} \right)^2 k^{2n}. \quad (A.18)$$
In particular, one has \( K(0) = \frac{\pi}{2} \) (as expected from (A.4) since this is the quarter period of the sine function,) while \( \lim_{k \uparrow 1} K(k) = \infty \), (as expected from (A.10) since the hyperbolic trigonometric functions are not periodic.)

We make use of limit

\[
\lim_{k \uparrow 1} (1 - k)^\alpha K(k) = 0, \quad \alpha \in (0, 1]. \tag{A.19}
\]

More general than the elliptic \( K \) integral is the **elliptic \( F \) integral**, defined as

\[
F(\phi, k) = \int_0^\phi \frac{du}{\sqrt{1 - k \sin^2(u)}} \tag{A.20}
\]

where one has \( K(k) = F\left(\frac{\pi}{2}, k\right) \). The elliptic \( F \) integral may be referred to as the **Elliptic Integral of the First Kind**, and \( K \) the **Complete Elliptic Integral of the First Kind**. The variable \( \phi \), as denoted, is the **Jacobi amplitude**, which is the inverse of \( F \), with

\[
\phi = \text{am}_k(t) \iff t = \int_0^\phi \frac{du}{\sqrt{1 - k \sin^2(u)}} = F(\phi, k). \tag{A.21}
\]

According to [25], (page 241,) Jacobi wrote this as such, calling it the amplitude, retained from previous use by Legendre. Later Gudermann wrote \( \sin \phi = \text{sn}(t) \), calling it the modular sine. Moreover, since \( \text{cn}(t) = \sqrt{1 - \text{sn}^2(t)} \) one has

\[
\text{sn}_k(t) = \sin(\phi) = \sin(\text{am}_k(t)), \text{ and } \text{cn}_k(t) = \cos(\phi) = \cos(\text{am}_k(t)).
\]

---

\(^1\)Guderman was a student of Gauss and teacher of Weierstrass, who made contributions to elliptic function theory and analysis. It is said that Guderman taught the first university course on elliptic functions in 1839-1840.
Also, the Jacobi amplitude is an antiderivative of \( dn \), as

\[
  t = \int_{0}^{am_k(t)} \frac{du}{\sqrt{1 - k \sin^2(u)}},
\]

and thus differentiation gives

\[
  1 = \frac{1}{\sqrt{1 - k \sin^2(am_k(t))}} \cdot \frac{d}{dt} am_k(t) \quad \Rightarrow \quad \frac{d}{dt} am_k(t) = dn_k(t). \tag{A.22}
\]

Thus, one can also write

\[
  \phi = am_k(t) = \int_{0}^{t} dn_k(u) du = \int_{0}^{t} \sqrt{1 - k \sin^2(u)} du. \tag{A.23}
\]

We make much use of the identity

\[
  am_k(nK(k)) = \frac{n\pi}{2}, \quad (n \in \mathbb{N}). \tag{A.24}
\]

To see this note that

\[
  \frac{n\pi}{2} = am_k(t) \quad \Leftrightarrow \quad t = \int_{0}^{n\pi/2} \frac{du}{\sqrt{1 - k \sin^2(u)}}.
\]

Since the integrand has period \( \pi \), and is even about its half period, one has

\[
  K(k) = \int_{0}^{\pi/2} \frac{du}{\sqrt{1 - k \sin^2(u)}} = \int_{\pi/2}^{\pi} \frac{du}{\sqrt{1 - k \sin^2(u)}} = \cdots = \int_{(n-1)\pi/2}^{n\pi/2} \frac{du}{\sqrt{1 - k \sin^2(u)}},
\]

which implies the claimed formula.

The Elliptic Integral of the Second Kind is the elliptic \( E \) integral

\[
  E(\phi, k) = \int_{0}^{\phi} \sqrt{1 - k \sin^2(u)} du. \tag{A.25}
\]
It is common to denote \( E(\pi/2, k) = E(k) \), which is called the \textit{Complete Elliptic Integral of the Second Kind}.

The elliptic integral \( E(k) \) naturally arises in computing the length an ellipse. A general ellipse can, (up to rigid motions,) be parametrized \( \gamma(t) = (a \sin(t), b \cos(t)) \) for \( a > b > 0 \). The length of this curve is

\[
L = 4 \int_0^{\pi/2} \sqrt{a^2 \cos^2(t) + b^2 \sin^2(t)} dt = 4a^2 \int_0^{\pi/2} \sqrt{1 - \varepsilon^2 \sin^2(t)} dt = 4a^2 E(\varepsilon^2)
\]

with \( \varepsilon = \sqrt{a^2 - b^2}/a \) the eccentricity of the ellipse.

In the literature, it is common to consider elliptic integrals composed with the Jacobi amplitude. Some clarity on what convention is adhered to may help avoid confusion. Let

\[
\phi = \text{am}_k(\tau)
\]

in the integrals for \( F \) and \( E \). For a composition \( u = \text{am}_k(t) \), one has \( du = \text{dn}_k(t) dt \), and \( \sqrt{1 - k \sin^2(u)} = \sqrt{1 - k \sin^2(\text{am}_k(t))} = \text{dn}_k(t) \), so the composition with \( F \) becomes

\[
F(\phi, k) = \int_0^{\text{am}_k(\tau)} \frac{du}{\sqrt{1 - k \sin^2(u)}} = \int_0^\tau \frac{\text{dn}_k(t) dt}{\text{dn}_k(t)} = \int_0^\tau dt = \tau.
\]

Thus \( F(\text{am}_k(t), k) = t \), as expected, since they are defined as inverse functions. Next, this composition in \( E \) gives

\[
E(\phi, k) = \int_0^{\text{am}_k(\tau)} \sqrt{1 - k \sin^2(u)} du = \int_0^\tau \text{dn}_k(t)(\text{dn}_k(t) dt) = \int_0^\tau \text{dn}_k^2(t) dt.
\]

It is common to define the elliptic \( E \) integral as the squared \( \text{dn} \) integral above, such as in [26], but this is under the assumption that the argument is the Jacobi amplitude. If unaware
of this convention, confusion may arise! To avoid confusion and distinguish between these forms, we define the *Jacobi Epsilon function* as

\[ \mathcal{E}(\tau, k) := E(\text{am}_k(\tau), k) = \int_0^\tau \text{dn}_k^2(t)dt. \]  

(A.26)

Calling this the Jacobi epsilon function as we do here is semi-standard, for instance in the tables at the NIST Digital Library of Mathematical Functions (with link https://dlmf.nist.gov/22.16#i). This seems like a good convention to adhere to, particularly when we work with both $E$ and $\mathcal{E}$. We make much use of the fundamental identity

\[ \mathcal{E}(t + 2nK, k) - \mathcal{E}(t, k) = 2nE(k), \quad (n \in \mathbb{N}). \]  

(A.27)

This can be verified directly from the definition.

\[
\begin{align*}
\mathcal{E}(x + 2nK(k), k) - \mathcal{E}(x, k) &= \int_0^{x+2nK(k)} \text{dn}_k^2(t)dt - \int_0^x \text{dn}_k^2(t)dt \\
&= \int_x^{x+2nK(k)} \text{dn}_k^2(t)dt \\
&= \int_0^{2nK(k)} \text{dn}_k^2(t)dt \\
&= \int_0^{2nK(k)} \sqrt{1 - k \text{sn}_k^2(t)} \sqrt{1 - k \text{sn}_k^2(t)} dt.
\end{align*}
\]

Letting $\phi = \text{am}_k(t)$, one has

\[
1 - k \text{sn}_k^2(t) = 1 - k \sin^2(\phi), \quad \text{and} \quad d\phi = \sqrt{1 - k \text{sn}_k^2(t)}dt.
\]

Now $\phi = \text{am}_k(t) \iff t = F(\phi, k)$, while $F(n\pi/2, k) = nK(k)$ for integers $n$, so

\[
\mathcal{E}(x + 2nK(k), k) - \mathcal{E}(x, k) = \int_0^{2n\pi/2} \sqrt{1 - k \sin^2(\phi)} d\phi = 2nE(k).
\]
Newton’s binomial theorem similarly gives a series expansion for \( E(k) \) where

\[
E(k) = \frac{\pi}{2} - \sum_{n=1}^{\infty} \left( \frac{1 \cdot 3 \cdot 5 \cdots (2n - 1)}{2 \cdot 4 \cdot 6 \cdots (2n)} \right)^2 \frac{k^{2n}}{2n - 1}.
\]  
(A.28)

Finally, we make use of the derivative formulas

\[
\frac{d}{dk} E(k) = \frac{E(k) - K(k)}{2k}, \quad \frac{d}{dk} K(k) = \frac{E(k) - (1 - k)K(k)}{2k(1 - k)}.
\]  
(A.29)
APPENDIX B

SOLVING THE PENDULUM EQUATION
We have shown that to identify the critical points in the main setting $\mathcal{M}_L$, and also for certain additional constraints, it is necessary to solve the pendulum equation, which we may write in the form

$$\ddot{\theta}(s) = \lambda \sin(\theta(s)).$$  \hfill (B.1)

In this appendix we derive general solutions, where first via transformations we show that the curvature must be one of the Jacobi elliptic functions $cn$ or $dn$.

### B.1 Transformation to Algebraic Differential Operator

Using an integrating factor $2\dot{\theta}$, one obtains $2\ddot{\theta}\dot{\theta} = 2\lambda \sin(\theta)\dot{\theta}$, which integrates to

$$\dot{\theta}^2 = -2\lambda \cos(\theta) + c_1.$$ \hfill (B.2)

For initial values $\theta(0) = \theta_0$ and $\dot{\theta}(0) = v_0$, the constant of integration is

$$c_1 = v_0^2 + 2\lambda \cos(\theta_0).$$ \hfill (B.3)

Moreover, differentiation of (B.1) gives

$$\ddot{\theta} = \lambda \cos(\theta)\dot{\theta}.$$ \hfill (B.4)
Using (B.4) and the additional integrating factor $\dot{\theta}$ in (B.2) gives

$$
\dot{\theta}^3 = -2\lambda \cos(\theta)\dot{\theta} + c_1 \dot{\theta}
= -2 \left( \lambda \cos(\theta) \dot{\theta} \right) + c_1 \dot{\theta}
= -2 \ddot{\theta} + c_1 \dot{\theta}.
$$

This results in the transformed differential equation

$$
\dddot{\theta} + \frac{\dot{\theta}^3}{2} - \frac{c_1}{2} \dot{\theta} = 0,
$$

which has an algebraic differential operator. It is noteworthy that if $\theta$ is a solution of (B.5), then so is $-\theta$. Next a substitution of $\kappa = \dot{\theta}/L$, for $L \in \mathbb{R}$ a constant, into (B.5) gives

$$
\dddot{\kappa} + \frac{L^2}{2} \kappa^3 - \frac{c_1}{2} \kappa = 0
$$

upon simplification; in regards to the necessary differential equation (2.4) for criticality, $\kappa$ corresponds to the signed curvature of the associated planar curve. We have reduced to a second order differential equation in the curvature, which can tell us the possible curvature types of critical points. To this end, multiply (B.6) by an integrating factor $2\dot{\kappa}$ to obtain

$$
2\dddot{\kappa} + L^2 \kappa^3 \ddot{\kappa} - c_1 \kappa \dot{\kappa} = 0.
$$

This integrates to

$$
\dot{\kappa}^2 + \frac{L^2}{4} \kappa^4 - \frac{c_1}{2} \kappa^2 = c_2.
$$
From here it is apparent that one has curvature given in terms of Jacobi elliptic functions, where equation (B.8) is of the general form

\[ y^2 = ay^4 + by^3 + cy^2 + dy + e, \]  

(B.9)

and via a transformation \( z = z(y) \) this can be reduced to the form

\[ \left( \frac{dz}{dt} \right)^2 = (1 - z^2)(1 - k z^2), \]  

(B.10)

and thus from Appendix A it is clear that \( z = \text{sn}_k(t) \), and hence the desired solution \( y \) can be obtained in terms of Jacobi elliptic functions. More details can be found in the differential equations treatise [29]; in particular see Chapter 7, Section 11, Solution of the General Elliptic Equation.

Now let us focus on the particular case at hand, by making yet another substitution

\[ u = \kappa^2 \]  

(B.11)

in equation (B.8). First, differentiation and squaring of (B.11) gives

\[ \kappa^2 = \frac{\dot{u}^2}{4u}, \]  

(B.12)

which results in a first order differential equation in \( u \), which upon simplification becomes

\[ \dot{u}^2 = -L^2 u^3 + 2c_1 u^2 + 4c_2 u. \]  

(B.13)
In this way we obtain a differential equation

\[ \dot{u}^2 + L^2 P(u) = 0 \quad (B.14) \]

with \( P \) a monic cubic polynomial with roots

\[ \rho_1 = \frac{-\beta - \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}, \quad \rho_2 = 0, \quad \rho_3 = \frac{-\beta + \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}, \quad (B.15) \]

where \( \alpha = 1, \quad \beta = -2c_1/L^2, \quad \gamma = -4c_2/L^2 \). For real roots it is necessary that \( c_1^2 + 4c_2L^2 \geq 0 \).

### B.2 Curvature Types for Pendulum Equation Solutions

A particular case of interest, which occurs when \( c_2 \geq 0 \) is an ordering of roots

\[ \rho_1 \leq 0 \leq \rho_2 \leq \rho_3. \quad (B.16) \]

For constants

\[ k = \frac{\rho_3 - \rho_2}{\rho_3 - \rho_1}, \quad A = \frac{\rho_3 - \rho_2}{\rho_3}, \quad B = \frac{\sqrt{\rho_3 - \rho_1}}{2}, \quad (B.17) \]

as \( u = \kappa^2 \), equation (B.14) has known solution

\[ \kappa^2(s) = \rho_3(1 - A \, \text{sn}^2_k(LBs)), \quad (B.18) \]

where we again reference the treatise [29] for further details on this solution. Also, in [5] a more general derivation involving the sectional curvature of the underlying manifold and torsion\(^1\) of the curve is discussed, but here the underlying manifold is \( \mathbb{R}^2 \) and both of these

---

\(^1\)The torsion is essentially a measure of the ‘twist’ of a space curve, which is similar to the curvature but gives a measure of the rate of change of the binormal vector. A planar space curve has zero torsion.
quantities are zero. Since $0 \leq \rho_2 \leq \rho_3$ it follows that $0 \leq A \leq 1$. Hence, a solution to the pendulum equation has dn derivative when $0 < A < 1$, and cn derivative when $A = 1$, which follows from the identities (A.6). Thus, if we regard the solution as an indicatrix, then the associated planar curve has either cn or dn curvature. Moreover, by periodicity of sn a more general solution is of the form

$$\kappa^2(s) = \rho_3(1 - A \, \text{sn}^2_k(LBs + \phi)), \quad \text{(B.19)}$$

for $\phi$ an appropriate phase shift.

In this thesis we consider indicatrices $\dot{\theta} = L\kappa$, so indicatrices with cn or dn curvature are given by antiderivatives

$$\theta(s) = \alpha + 2 \sin^{-1} \left( \sqrt{k} \, \text{sn}_k(as + b) \right) \quad \text{or} \quad \eta(s) = \alpha + 2 \text{am}_k(as + b). \quad \text{(B.20)}$$

Using formulas from Appendix A one easily verifies that these indicatrices have cn and dn curvature, respectively. Now we seek solutions of (B.1), and one also easily verifies that for $\alpha = 0$ the functions in (B.20) are solutions to equation (B.1) for parameters $\lambda = -a^2$ and $\lambda = -a^2k$, respectively. However, there are restrictions on the parameters $\alpha$ in (B.20). First of all, if $\theta$ is a solution to $\ddot{\theta} = A \sin(\theta)$, then for integers $n$, $\theta + 2n\pi$ is also a solution, and $\theta + n\pi$ is a solution to $\ddot{\theta} = -A \sin(\theta)$. To see if any other parameters $\alpha$ are feasible for cn curvature, setting $\ddot{\theta}(s) = -a^2 \sin(\theta(s))$ for $\theta$ in (B.20), (so with $A = -a^2$), one can rewrite this equation in the form

$$\sqrt{k} (1 - \cos(\alpha)) \text{sn}_k(as + b) \text{dn}_k(as + b) + \frac{\sin(\alpha)}{2} \left(1 - 2k \text{sn}_k^2(as + b)\right) = 0. \quad \text{(B.21)}$$
Moreover, one has Wronskian

\[ W(sn_k(as + b)dn_k(as + b), 1 - 2k sn_k^2(as + b)) = -asn_k(as + b), \]  

(B.22)

which is not everywhere vanishing, so the functions are linearly independent. Hence the only solution to equation (B.21) is that with coefficients

\[ \sqrt{k}(1 - \cos(\alpha)) = 0 = \frac{\sin(\alpha)}{2}, \]  

(B.23)

and it follows that parameters \( \alpha \neq 2n\pi \) will not give a solution to the pendulum equation, which in our context means the planar curves can only vary by full rotations, which is the same planar curve. One similarly arrives at the same conclusion for the equation \( \ddot{\theta} = -A \sin(\theta) \) with cn curvature solutions \( \theta + n\pi \), and also for the dn curvature solutions. Also if \( \theta \) is a solution to equation (B.1), so is \(-\theta\), as easily verified, so we have that for integers \( n \)

\[ \theta(s) = n\pi \pm 2\sin^{-1}\left(\sqrt{k} \ sn_k(as + b)\right) \]  

solves \( \ddot{\theta}(s) = (-1)^{n+1}a^2 \sin(\theta(s)) \)  

(B.24)

and

\[ \eta(s) = n\pi \pm 2am_k(as + b) \]  

solves \( \ddot{\eta}(s) = (-1)^{n+1}a^2k \sin(\eta(s)) \).  

(B.25)

We may further consider limiting cases of these solution types. For the cn curvature solution, in the limit as \( k \downarrow 0 \), one has

\[ \lim_{k \downarrow 0} 2 \sin^{-1}\left(\sqrt{k} \ sn_k(as + b)\right) = 2 \sin^{-1}(0) = 0, \]
and we get the trivial solution in this limit, easily seen to satisfy the pendulum equation. For the dn curvature solution, in the limit as $k \downarrow 0$, one has

$$\lim_{k \downarrow 0} 2\text{am}_k(as + b) = 2s,$$

and this limit does not yield a solution. In the limit as $k \uparrow 1$, for the cn curvature solution one has

$$\lim_{k \uparrow 1} 2\sin^{-1}\left(\sqrt{k} \, \text{sn}_k(as + b)\right) = 2\sin^{-1}(\tanh(as + b)),$$

and it is simple to verify that

$$\theta(s) = 2\sin^{-1}(\tanh(as + b)) \quad (B.26)$$

gives a solution to (B.1), while for the dn curvature solution one has

$$\lim_{k \uparrow 1} 2\text{am}_k(as + b) = -\pi + 4\tan^{-1}(e^{as+b}) = 2\text{gd}(as + b),$$

with gd the Gudermannian\(^2\) function. It is simple to verify that

$$\eta(s) = -\pi + 4\tan^{-1}(e^{as+b}) = 2\text{gd}(as + b), \quad (B.27)$$

gives a solution (B.1). In fact, both solutions (B.26) and (B.27) coincide, which is not difficult to show, while also one has

$$\frac{d\theta}{ds} = 2\text{sech}(as + b) = \frac{d\eta}{ds}, \quad (B.28)$$

with coinciding hyperbolic secant curvature in each case.

\(^2\)Apparently Cayley named this function after Guderman, who was mentioned in Appendix A
APPENDIX C

ELASTIC CURVES
C.1 A brief overview of the history of elasticity theory

The question of what shape a bent beam or wire assumes under stress or clamping is a very natural one to inquire, and one that has been considered since antiquity. Many have posed the problem in some situation, often offering erroneous solutions. Among the earliest to make substantial contributions was Galileo, despite believing it could take the shape of a circular arc — perhaps in vindication Linnér and Jerome have shown in [30] that under certain constraints an elastic curve can contain circular segments. It appears that it was James Bernoulli who first approached it with any level of mathematical rigor, which was in 1691, (despite not being mathematically rigorous by modern standards.) J. Bernoulli considered the problem of finding the shape of a bent lamina with sufficient mass suspended to bend the lamina perpendicular to the hanging mass. J. Bernoulli published a solution in a 1694 treatise, apparently having held on to it for a few years.

Figure C.1: Early figures from the study of elastica, from left to right given by Galileo, Leibniz, J. Bernoulli, and Euler, respectively.

In a 1742 letter from Daniel Bernoulli to Euler, it was suggested that Euler use the emerging theory of the calculus of variations to solve this problem as a minimization of the bending energy. This is an application of the least action principle and Hooke’s law. It appears that D. Bernoulli is the first known to have expressed the problem as a variational
one, that of minimizing the square of the curvature \( \kappa = \frac{\dddot{x}}{(1+\dot{x}^2)^{3/2}} \) with the curve a graph of \( x = f(t) \). In this case with \( ds = \sqrt{1+\dot{x}^2} \, dt \) the problem becomes

\[
\text{Minimize} \quad \int \frac{\dddot{x}^2}{(1+\dot{x}^2)^3} \, ds = \int \frac{\dddot{x}^2}{(1+\dot{x}^2)^{5/2}} \, dt = \int F(t, x, \dot{x}, \dddot{x}) \, dt. \tag{C.1}
\]

Here D. Bernoulli had trouble, as the integrand is a function of the second derivative also; at this time, an early version of the Euler-Lagrange theory for necessity was known which only pertained to integrands involving the first derivative. D. Bernoulli suggested that Euler was the perfect person to make quick work of this analysis. Indeed Euler extended the Euler-Lagrange theory to higher derivatives, and solved the problem, and with remarkable detail. Euler solved (C.1) up to a quadrature as

\[
\frac{dx}{dt} = \frac{a^2 - c^2 + t^2}{\sqrt{(c^2 - t^2)(2a^2 - c^2 + t^2)}}, \tag{C.2}
\]

for parameters \( a, c \). Euler had a very good understanding of the possible solution types, and sketched remarkably accurate figures which are now quite famous.

\[\text{Figure C.2: Sketches of elastica from Euler's work on the subject.}\]

The very natural problem of determining the shape of a beam under stress has attracted the attention of many prominent mathematicians throughout the past two millennia, and there is scant room here to elaborate. Regarding ‘ancient’ history of the elastica problem, of
course historical claims made here are uncertain, but have been pieced together to the best of the author’s ability; for history on elasticity theory, Truesdell and Love have written large volumes [31], [32].

Modern history of elasticity theory is of course more clear, where in the modern approach the application of least action principle as D. Bernoulli did is still standard, but over time it became clear that the squared curvature was difficult to work with, and rather that working with the indicatrix was more convenient. The elastica are typically investigated as rest points of some gradient flow, often on a Riemannian manifold, but many approaches do not take place on a formal manifold, such as the aforementioned PDE driven approaches [8] and [9]. In a given context, elastic curves are defined as critical points of some energy functional. This is typically a functional

\[ F(\gamma) = \frac{1}{2} \int_\gamma [\kappa(s)^2 + 2\nu] ds, \]  

(C.3)

where \( \kappa \) is the signed curvature, and \( \nu \geq 0 \) is a length penalty parameter, which may be thought of (physically) as tension in the context of curve straightening when \( \nu \) is positive. In many treatments, ‘free length’ has been considered, over a manifold \( \mathcal{M} \times \mathbb{R}^+ \), where positive tension prevents the length from increasing without bound to reduce the energy towards zero; for example see [15]. As we take fixed length, we set \( \nu = 0 \) here. There has been recent interest in elastic splines which are energy minimizing curves which pass thru a set of nodes, (which are considered as points along the curve or ‘wire,’) possibly with some angular constraints at the nodes; in this context pinning refers to the curve passing thru a node, and clamping refers to an angular constraint at a node; for example see [16], [33], [6], [7], and [34]. There has also been much interest in convergence of gradient flow descents, often pertaining to the aforementioned Palais-Smale condition; for instance see [35] and [11]. The notion of elastica makes sense in more general manifolds than just the plane considered here,
such as in three dimensional space, in a hyperbolic plane, or on a sphere; for instance see [5] and [36]. Also, Singer’s lectures [12] give a good overview of the modern theory. For general formulas for planar elastic curves with various constraints, see [37]. Here, we will focus on the particular cases relevant to this thesis.

C.2 Elastica in the manifolds $\mathcal{M}_L$

Here attention is restricted to the particular setting $\mathcal{M}_L$ where we work directly with the indicatrix or turning angle. The functional of interest is $F_L$ which has no tension, which we consider restricted to $\mathcal{M}_L$. The elastica are the associated planar curves to the critical points of $F_L$ in the sense that the projected gradient vanishes. It is noteworthy that $A = 1$ in formula (B.18) from Appendix B, since the cubic differential operator has a root $\rho_2 = 0$ the indicatrix has squared curvature

$$\kappa^2(s) = \rho(1 - \text{sn}_k^2(LBs + \phi)) = \rho \text{cn}_k^2(LBs + \phi),$$

so the elastica here have cn curvature. We also consider the dn curvature analogues here, as they help explain the behavior of the gradient flow; we may refer to these as ‘non-critical elastica.’

C.2.1 General form of the indicatrices

It has been shown that in this setting critical points necessarily satisfy the equation

$$\dot{\theta}(s) = \lambda_LL^2\sin(\theta(s)).$$
From Appendix B we know the solutions have the forms

$$\theta(s) = n\pi + 2 \sin^{-1}\left(\sqrt{k}\text{sn}_k(as + b)\right), \quad \text{or} \quad \eta(s) = n\pi + 2 \text{am}_k(as + b).$$

for integers $n$. First of all, we can take $n \in \{0, 1\}$, which is clear from the $2\pi$ periodicity in $\theta$. In the context of indicatrices in these manifolds, the additional angle of $\pi$ merely affects the positioning of the curve, amounting, (via other symmetries,) to a reflection about the $x$ and $y$ axes. The reflectional symmetry of $\theta$ and $-\theta$ gives freedom in choosing one of two representatives, and the choice may be subjective and perhaps aesthetically motivated; dynamically, $\pm \theta$ are the same.

To determine the suitable coefficients $a, b$, we impose the natural boundary conditions (4.3) for the cn curvature form, which become $\dot{\theta}(0) = 0 = \dot{\theta}(1)$ in this setting. Recall that in Chapter 2 we discussed the necessity of the natural boundary conditions for criticality. Since one has

$$\dot{\theta}(s) = 2a\sqrt{k}\text{cn}_k(as + b),$$

and from the discussion in Appendix A we know that the zeros are $\text{cn}_k(2nK(k) \pm K(k))$ for integers $n$, we thus take $a = 2nK(k)$, for $n \in \{1, 2, 3, \ldots\}$ and $b = \pm K(k)$. The two distinguishable cases can be expressed in general form

$$\theta(s) = (1 - p)\pi + 2 \sin^{-1}\left(\sqrt{k}\text{sn}_k(K(k)(2ns + (-1)^p))\right), \quad p \in \{0, 1\}, \quad \text{(C.5)}$$

where for $p = 0$ there is a self intersection, and for $p = 1$, there is not. We refer to the two cases as loops and arches, with plots in Figures C.3 and C.4, respectively. These two critical curves coincide in $\mathcal{M}_\infty$.

Despite not being critical, curves with dn curvature are of interest here as they significantly affect the dynamics of the flow – perhaps pseudo-elastica would be an appropriate
Figure C.3: A collection of elastic loops for $L = 1.4$ with darker shading for increasing energy. Each has the same elliptic modulus $k \approx 0.9999$, and containing $n$ self-similar $1/n$ scalings for $n > 1$.

Figure C.4: A collection of elastic arches for $L = 1.1$ with darker shading for increasing energy. Each has the same elliptic modulus $k \approx 0.08985$, and containing $n$ self-similar $1/n$ scalings for $n > 1$.

name for them. We take the same coefficients $a, b$ for these indicatrices and consider the form

$$\eta(s) = \pi - 2am_k((2ns + 1)K(k)).$$  \hfill (C.6)

This is a pendulum equation solution as discussed in Appendix B, and this form similarly positions $\eta$ with the cn curvature loop, with plots in Figure C.5. The choice in coefficients
\( a = 2nK(k) \) and \( b = K(k) \) serves to make the natural boundary conditions as close as possible to holding in the following sense: one has signed curvature given

\[
\kappa(s) = \frac{\dot{\eta}(s)}{L} = -\frac{4nK(k)}{L} \text{dn}_k((2ns + 1)K(k)),
\]

while \( \text{dn} \) attains a minimum of \( \sqrt{1-k} \) at odd multiples of its half period \( K(k) \). Moreover, one has \( \lim_{k \uparrow 1} K(k) \sqrt{1-k} = 0 \), which may be interpreted as the natural boundary conditions tending to holding in this limit, but never actually holding.

\[
\text{Figure C.5: A collection of pseudo-elastic loops for } L = 1.35 \text{ with darker shading for increasing energy. Each has the same elliptic modulus } k \approx 0.999997, \text{ and containing } n \text{ self-similar } 1/n \text{ scalings for } n > 1.
\]

### C.2.2 The Elliptic Modulus for a Given Gap

In the last section three types of elastica were listed, two of which were critical. For critical elastica, in each manifold \( \mathcal{M}_L \) for \( L < \infty \), there is exactly one of each type, while for \( L = \infty \) this degenerates, where the loop and arch share a common modulus. In this section we derive formulas for the elliptic modulus for each of the three types by choosing \( k \) so that

\[
L \int_0^1 \cos(\theta) ds \in \{0, 1\}.
\]

Here we find the notion of the \textit{gap} helpful, where for a planar curve
the gap $g$ can be taken as the distance between initial and terminal points. For $L < \infty$, the associated planar curve has gap $g = 1$, while for $L = \infty$ it has gap $g = 0$.

For a given $L$, and indicatrix $\theta(s) = 2 \sin^{-1} \left( \sqrt{k} \text{sn}_k(as + b) \right)$ we may choose $k$ so that the associated planar curve has initial point $(0, 0)$ and terminal point $(g, 0)$, for $g < L$. First, using cosine identities, get

$$
\cos(\theta) = 1 - 2 \sin^2 \left( 2 \sin^{-1} \left( \sqrt{k} \text{sn}_k(as + b) \right) \right) = 2 \text{dn}_k^2(as + b) - 1.
$$

Then, the associated integral becomes

$$
\int_0^1 \cos(\theta) ds = 2 \int_0^1 \text{dn}_k^2(as + b) ds - 1 = 2 \frac{E(a + b) - E(b)}{a} - 1.
$$

Note that we get a different sign with an additional rotation angle $(2n + 1)\pi$. Since $a = 2nK(k)$ and $b = K(k)$ for the critical elastica (C.5), by identity (A.27) we get $E(a + b) - E(b) = 2nE(k)$, and thus set

$$
\pm \frac{g}{L} = \int_0^1 \cos(\theta) ds = 2 \frac{E(k)}{K(k)} - 1,
$$

from which we obtain the gap equation

$$
2E(k) = \left(1 \pm \frac{g}{L}\right) K(k). \quad (C.8)
$$

When considering a given manifold $\mathcal{M}_L$, we take $k \in (0, 1)$ as a root of equation (C.8). A visual representation of the totality of these elastica is given by fixing $L = 1$ and letting $k$ vary over $[0, 1)$. For each $L < \infty$ there are two elastica in this collection for each $n$. Also, $n = 1$ gives a suitable representation, as there are generally $n$ self similar curves scaled by $1/n$. A collection of these curves are plotted with $n = 1$ in Figure C.6.
For a given $L$, and indicatrix $\eta(s) = \pi - 2am_k(as + b)$, we may also choose $k$ so that the associated planar curve has a prescribed gap. First, using cosine identities get

$$\cos(\eta) = 2\sin^2(\text{am}_k(as + b)) - 1 = 2\sin^2_k(as + b) - 1 = \frac{2}{k}(1 - \text{dn}^2_k(as + b)) - 1.$$  

Then the associated integral becomes

$$\int_0^1 \cos(\eta)ds = \frac{2}{k} - \frac{2}{k} \int_0^1 \text{dn}^2_k(as + b)ds - 1 = \frac{2}{k} - \frac{2\mathcal{E}(a + b) - \mathcal{E}(b)}{a} - 1.$$  

As $a = 2nK(k)$ and $b = K(k)$, again by identity (A.27) we obtain the gap equation

$$\frac{2}{k} - \frac{2}{k} \frac{E(k)}{K(k)} - 1 = \frac{g}{L}. \tag{C.9}$$  

For a given $L < \infty$ with $g = 1$ we take $k \in (0, 1)$ as this real root of equation (C.9) to get $\eta \in \mathcal{M}_L$, which is the dn curvature loop. For the limiting case of large $L$ this corresponds to a gap $g = 0$, so the manifold $\mathcal{M}_\infty$. Upon clearing $k$ in equation (C.9), this is seen to occur exactly when $k = 0$, and the resulting curve is a circle with indicatrix $\eta(s) = \pi + 2s$, as seen in Figure C.7.
Figure C.7: A variation of dn curvature elastica in $k$ with $0 \leq k < 1$ and $L = 1$, with darker curves plotted for larger $k$. This collection curves is the totality of dn curvature elastica in the manifolds $\mathcal{M}_L$, up to scaling and periodicity for $n > 1$, represented here with unit length. At $k = 0$ we have a circle, and as $k \uparrow 1$, these curves approach the cn curvature elastica.

Regarding solutions of the gap equations (C.8) and (C.9), when $g < L$ there is exactly one solution in all three cases, that is including both signs in (C.8). To see this, first observe that $aE(k) = K(k)$ has unique solution $0 < k < 1$ for $a > 1$. Indeed, defining $f(k) = aE(k) - K(k)$, one has $f(0) = (a - 1)\frac{\pi}{2}$ since $E(0) = K(0) = \frac{\pi}{2}$. Thus $f(0) > 0$ when $a > 1$. Also

$$E(1) = 1 \quad \text{and} \quad \lim_{k \uparrow 1} K(k) = \infty \quad \implies \quad \lim_{k \uparrow 1} f(k) = -\infty.$$ 

The intermediate value theorem thus implies that $f(k^*) = 0$ for some $k^* \in (0, 1)$. Moreover, since $E(k)$ is decreasing while $K(k)$ is increasing in $k$, which is inferred from the derivative formulas (A.29) it follows that there is a unique solution $k \in (0, 1)$. The gap equations can be put in the form $aE(k) = K(k)$ for both cn and dn curvature, and when $g < L$ it follows that there is a unique solution in all three cases. The analysis here is rather simple, and the particular case $L = \infty$ is depicted in Figure C.8. In particular, in each manifold $\mathcal{M}_L$ for $L < \infty$ there are exactly two cn curvature elastica, (which coincide for $L = \infty$,) and one dn curvature ‘elastica.’
Figure C.8: Plotted are $k \mapsto 2E(k), k \mapsto K(k)$ and the auxiliary function $f(k) = 2E(k) - K(k)$ corresponding to $\mathcal{M}_\infty$, with the former two graphs intersecting in $k_\infty$, and the latter function having this unique real root for $0 < k < 1$.

### C.2.3 Explicit Energy of the Elastica

The energy for curves with cn or dn curvature can be exactly calculated up to elliptic integrals, for which we derive formulas in this section.

For a general cn curvature elastica, with $\dot{\theta}(s) = 2\sqrt{k}acn_k(as + b)$, one has

$$F_L(\theta) = \frac{1}{2L} \int_0^1 \dot{\theta}^2(s)ds$$

$$= \frac{4ka^2}{2L} \int_0^1 cn_k^2(as + b)ds$$

$$= \frac{2ka^2}{L} \int_0^1 \left(1 - sn_k^2(as + b)\right)ds$$

$$= \frac{2ka^2}{L} \int_0^1 \left(1 - \frac{1}{k} + \frac{dn_k^2(as + b)}{k}\right)ds$$

$$= \frac{2ka^2}{L} \left(1 - \frac{1}{k} + \frac{\mathcal{E}(a + b, k) - \mathcal{E}(b, k)}{ak}\right)$$

$$= \frac{2a}{L}(\mathcal{E}(a + b, k) - \mathcal{E}(b, k) - a(1 - k)).$$
For a general dn curvature elastica, with $\dot{\eta}(s) = 2a dn_k(as + b)$, one has

$$F_L(\eta) = \frac{1}{2L} \int_0^1 (\dot{\eta})^2(s) ds$$

$$= \frac{2a^2}{L} \int_0^1 dn_k^2(as + b) ds$$

$$= \frac{2a}{L} (E(a + b, k) - E(b, k)).$$

In particular, for $a = 2nK(k)$ and $b = K(k)$ by (A.27) one has energy formulas

$$F_L(\theta) = \frac{8n^2 K(k)}{L} (E(k) - K(k)(1 - k))$$

$$F_L(\eta) = \frac{8n^2 E(k)K(k)}{L}.$$

In each case, the energy tends to $\infty$ as $k \uparrow 1$. From these formulas, it is immediate that for a given $k$, one has $F_L(\theta) < F_L(\eta)$, but this inequality lacks significance since for a given manifold $M_L$, there are different modulus equations which result in different moduli.
APPENDIX D

NUMERICAL DISCRETIZATION OF DIFFERENTIAL EQUATIONS
Here we give a brief overview of numerical discretization which is pertinent to the methods applied in Chapter 5. There is much literature on this topic, and in particular we suggest [38], [22].

## D.1 Initial Value Problems

We of course assume some familiarity with differential equations, and the standard *initial value problem*

\[ x'(t) = f(t, x), \quad x(0) = x_0, \quad x \in \mathbb{R}^n. \quad (D.1) \]

The differential equation is called *autonomous* if \( f \) does not explicitly depend on \( t \). By the standard treatment any \( n^{th} \) order differential equation can be put in this form as a system of \( n \)-equations, and moreover a non-autonomous equation, that is one explicitly involving \( t \) can be put in the form of an autonomous system of \( (n + 1) \)-equations. For our purposes, we need only consider the autonomous single variable form

\[ x'(t) = f(x), \quad x(0) = x_0, \quad x \in \mathbb{R}, \quad (D.2) \]

which is an analogue to the gradient flows considered. Also, most methods from the latter form immediately generalize to the former via vector operations.
D.2 Numerical Discretization of Initial Value Problems

It is well known that there typically little hope of explicitly solving a given differential equation, even simple examples which arise in practice. Thus there is much interest in numerical means for generating approximate solutions, that is a solution

\[ X(t) \approx x(t) \quad \text{where} \quad \frac{dx}{dt} = f(x) \quad \text{and} \quad x(0) = x_0. \]  

(D.3)

The standard approach is to approximate via a finite difference equation to obtain a discrete set of approximate values

\[ x_0, x_1, ..., x_n \quad \text{and} \quad h_1, ..., h_n \quad \text{with} \quad x_j \approx x\left(\sum_{i=1}^{j} h_i\right) \quad \text{for} \quad 1 \leq j \leq n. \]  

(D.4)

The approximate solution \( X(t) \) may be taken as an interpolant of the points \( \left(\sum_{i=1}^{j} h_i, x_j\right) \), (for \( 0 \leq i \leq n \), and the empty sum null.) The values \( h_j \) are referred to as the step sizes, and the method is said to be adaptive if the step sizes vary. Perhaps the simplest example is the well known Euler method

\[ x_{n+1} = x_n + hf(x), \]  

(D.5)

which may be obtained from the difference quotient approximation

\[ \frac{dx}{dt} \approx \frac{x(t+h) - x(t)}{h}. \]  

(D.6)

For a given discretization method, one is interested in the truncation error \( T_n \), which is the error incurred at the \( n^{th} \) iterate.\(^1\) A method is of order \( p \) if \( p \) is the greatest integer such that

\(^1\)There are varying definitions of truncation error, but this simple description suffices here, as we do not give a thorough treatment.
\[ |T_n| \leq Ch_n^p, \text{ for some constant } C. \] A method is said to be \textit{consistent} if \( T_n \rightarrow 0 \) as \( h_n \rightarrow 0 \), and \textit{stable} if the error \( |x(t) - X(t)| \) goes to zero as the truncation error goes to zero; there is a subtlety in the difference, and this is demonstrated in a simple example in \cite{22}. Of course it is desirable to seek methods which are both consistent and stable.

\subsection*{D.3 Linear Multi-Step Methods}

The discretizations used in this thesis are based on \textit{linear multi-step methods}, which have the form

\[ x_{n+1} + \sum_{j=1}^{k} \alpha_j x_{n+1-j} = h \sum_{j=0}^{k} \beta_j f(x_{n+1-j}), \quad (D.7) \]

for parameters \( \alpha_j, \beta_j, \) and \( k \). Here, each iterate is determined from the previous \( k \) iterates via this rule. There are many choices in parameters \( \alpha_j, \beta_j, \) and \( k \) which can give a stable and consistent method, and the parameters \( \alpha_j \) and \( \beta_j \) are typically chosen so that the most possible lower order terms the Taylor expansion of \( x \) cancel. A multistep method (D.7) is stable if the roots of the polynomial

\[ \lambda^k + \alpha_1 \lambda^{k-1} + \cdots + \alpha_{k-1} \lambda + \alpha_k \quad (D.8) \]

are, in complex modulus, less than or equal to 1 for single roots, and strictly less than 1 for multiple roots. In particular, the \textit{Adams-Bashforth} methods are consistent, stable methods of order \( k \) obtained by setting \( \alpha_k = -1, \alpha_j = 0 \) for \( 1 \leq j \leq k - 1, \) and \( \beta_0 = 0, \) and then choosing coefficients \( \beta_j \) to give maximal order; the \( k \) step version may be referred to as ABK. Proofs of these claims can be found in \cite{22}. An inherent difficulty in applying these methods is obtaining the previous iterates to initialize the method, and this is typically done by using some choice in lower order methods to obtain the first \( k \) iterates.