Essays on Numerical Integration in Hamiltonian Monte Carlo

James Brofos
Yale University Graduate School of Arts and Sciences, james@brofos.org

Follow this and additional works at: https://elischolar.library.yale.edu/gsas_dissertations

Recommended Citation
Brofos, James, "Essays on Numerical Integration in Hamiltonian Monte Carlo" (2022). Yale Graduate School of Arts and Sciences Dissertations. 567.
https://elischolar.library.yale.edu/gsas_dissertations/567

This Dissertation is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Graduate School of Arts and Sciences Dissertations by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.
Abstract

Essays on Numerical Integration in Hamiltonian Monte Carlo

James Alexander Brofos

2022

This thesis considers a variety of topics broadly unified under the theme of geometric integration for Riemannian manifold Hamiltonian Monte Carlo. In chapter 2, we review fundamental topics in numerical computing (section 2.1), classical mechanics (section 2.2), integration on manifolds (section 2.3), Riemannian geometry (section 2.5), stochastic differential equations (section 2.4), information geometry (section 2.6), and Markov chain Monte Carlo (section 2.7). The purpose of these sections is to present the topics discussed in the thesis within a broader context. The subsequent chapters are self-contained to an extent, but contain references back to this foundational material where appropriate. Chapter 3 gives a formal means of conceptualizing the Markov chains corresponding to Riemannian manifold Hamiltonian Monte Carlo and related methods; this formalism is useful for understanding the significance of reversibility and volume-preservation for maintaining detailed balance in Markov chain Monte Carlo. Throughout the remainder of the thesis, we investigate alternative methods of geometric numerical integration for use in Riemannian manifold Hamiltonian Monte Carlo, discuss numerical issues involving violations of reversibility and detailed balance, and propose new algorithms with superior theoretical foundations. In chapter 4, we evaluate the implicit midpoint integrator for Riemannian manifold Hamiltonian Monte Carlo, presenting the first time that this integrator has been deployed and assessed within this context. We discuss attributes of the implicit midpoint integrator that make it preferable, and inferior, to alternative methods of geometric integration such as the generalized leapfrog procedure. In chapter 5, we treat an empirical question as to what extent convergence thresholds play a role in geometric
numerical integration in Riemannian manifold Hamiltonian Monte Carlo. If the convergence threshold is too large, then the Markov chain transition kernel will fail to maintain detailed balance, whereas a convergence threshold that is very small will incur computational penalties. We investigate these phenomena and suggest two mechanisms, based on stochastic approximation and higher-order solvers for non-linear equations, which can aid in identifying convergence thresholds or suppress its significance. In chapter 6, we consider a numerical integrator for Markov chain Monte Carlo based on the Lagrangian, rather than Hamiltonian, formalism in classical mechanics. Our contributions include clarifying the order of accuracy of this numerical integrator, which has been underestimated in the literature, and evaluating a simple change that can accelerate the implementation of the method, but which comes at the cost of producing more serially auto-correlated samples. We also discuss robustness properties of the Lagrangian numerical method that do not materialize in the Hamiltonian setting. Chapter 7 examines theories of geometric ergodicity for Riemannian manifold Hamiltonian Monte Carlo and Lagrangian Monte Carlo, and proposes a simple modification to these Markov chain methods that enables geometric ergodicity to be inherited from the manifold Metropolis-adjusted Langevin algorithm. In chapter 8, we show how to revise an explicit integration using a theory of Lagrange multipliers so that the resulting numerical method satisfies the properties of reversibility and volume-preservation. Supplementary content in chapter F investigates topics in the theory of shadow Hamiltonians of the implicit midpoint method in the case of non-canonical Hamiltonian mechanics and chapter G, which treats the continual adaptation of a parameterized proposal distribution in the independent Metropolis-Hastings sampler.
Essays on Numerical Integration in Hamiltonian Monte Carlo

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
James Alexander Brofos

Dissertation Director: Roy R. Lederman

May 2022
I am extremely grateful to my advisor, Roy R. Lederman, for his endless support of my research. It has been a pleasure to work with him over the course of this program, and I will always be impressed by the standards of rigor and clarity that he sets.

I thank my several collaborators who have made this journey through graduate school that much more exciting. Among these are Roy Lederman, Marcus Brubaker, Marylou Gabrié, Pramit Chaudhuri, Joseph Dexter, and Patrick Burns, who have all provided me with valuable advice and mentorship over the years. I also thank my friends Michael, Frank, Seamus, Brett, John, Gurkaran, Ajay, Matthew, and Rui; spending time with them is always a pleasure.

I thank Roy, Marcus, and Joe Chang for graciously volunteering their time to serve on my committee. I thank all the faculty of the Department of Statistics and Data Science, who have provided me with instruction both inside and outside the classroom. I have particularly enjoyed my numerous conversations with David Brinda and Jay Emerson, who have both been generous with their time.

I would like to thank all the staff members of the Statistics and Data Science Department. Among these, I extend special thanks to Joann DelVecchio and Karen Kavanaugh, who, in addition to providing invaluable assistance, have been a pleasure to know.

I am grateful for the generous financial support provided by the National Science Foundation under the auspices of their Graduate Research Fellowship Program. Their support has allowed me to commit myself to my research. I thank the Yale Center for Research Computing for facilitating my use of Yale’s excellent computing resources.

I thank my mother Lynn, my father Alex, and my grandmother Sylvia for their love and support over nearly thirty years and for their insistence that I work hard and continually seek to improve myself. Finally, I would like to thank my wife, Samantha, who has been a source of endless encouragement. It is no understatement to say that I, and this thesis, owe her a deep debt of gratitude.
This dissertation is dedicated to my wife Samantha.
Contents

1 Introduction ............................................. 1

2 Preliminaries ........................................... 8

2.1 Numerical Errors ....................................... 8

2.1.1 Truncation Errors .................................... 8

2.1.2 Finite Precision Number Systems ................... 10

2.1.3 Human Errors ........................................ 15

2.2 Vector Fields, Mechanics, and Numerical Methods .......... 16

2.2.1 Vector Fields and Flow Maps ....................... 18

2.2.2 Numerical Integrators ............................... 24

2.2.3 Hamiltonian and Lagrangian Mechanics ............. 31

2.2.4 Numerical Integrators for Hamiltonian Mechanics .... 39

2.2.5 Conserved Quantities of Numerical Integrators ....... 59

2.2.6 Higher-Order Numerical Methods ................... 62

2.2.7 Linear Stability Analysis ............................ 65

2.3 Numerics and Integration on Manifolds .................. 68

2.3.1 Embedded Manifolds ................................ 68

2.3.2 Differential Forms ................................... 71

2.3.3 Manifold Hamiltonian Mechanics ................... 80

2.3.4 Numerical Integration ............................... 86

2.3.5 Uniqueness of the Lagrange Multipliers ............. 88
<table>
<thead>
<tr>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.6 Properties of Flows on Embedded Manifolds</td>
</tr>
<tr>
<td>2.3.7 Properties of Numerical Integrators on Embedded Manifolds</td>
</tr>
<tr>
<td>2.3.8 Conservation of Volume</td>
</tr>
<tr>
<td><strong>2.4 Stochastic Differential Equations</strong></td>
</tr>
<tr>
<td>2.4.1 Brownian Motion</td>
</tr>
<tr>
<td>2.4.2 Itô’s Lemma</td>
</tr>
<tr>
<td>2.4.3 Infinitesimal Generators</td>
</tr>
<tr>
<td>2.4.4 The Langevin Equation</td>
</tr>
<tr>
<td>2.4.5 The Euler-Maruyama Discretization</td>
</tr>
<tr>
<td><strong>2.5 Basics of Deterministic and Stochastic Motion on Riemannian Manifolds</strong></td>
</tr>
<tr>
<td>2.5.1 Fundamental Objects in Riemannian Geometry</td>
</tr>
<tr>
<td>2.5.2 Riemannian Measures and Probability Densities</td>
</tr>
<tr>
<td>2.5.3 Riemannian Continuity and Fokker-Planck Equations</td>
</tr>
<tr>
<td>2.5.4 Geodesic Motion</td>
</tr>
<tr>
<td>2.5.5 Correspondence Between Stochastic and Deterministic Dynamics on a Manifold</td>
</tr>
<tr>
<td><strong>2.6 Information Geometry</strong></td>
</tr>
<tr>
<td>2.6.1 The Fisher Information Metric</td>
</tr>
<tr>
<td>2.6.2 The Geometric Properties of Fisher Information</td>
</tr>
<tr>
<td>2.6.3 The Natural Gradient Method</td>
</tr>
<tr>
<td>2.6.4 The SoftAbs Metric</td>
</tr>
<tr>
<td><strong>2.7 Markov Chain Monte Carlo</strong></td>
</tr>
<tr>
<td>2.7.1 Markov Chain Transition Kernels</td>
</tr>
<tr>
<td>2.7.2 Stationarity and Detailed Balance</td>
</tr>
<tr>
<td>2.7.3 The Total Variation Distance</td>
</tr>
<tr>
<td>2.7.4 Irreducibility and Small Sets</td>
</tr>
</tbody>
</table>
## 3 Diffeomorphism Monte Carlo

3.1 Construction of the Diffeomorphism Monte Carlo Transition Kernel 165
3.2 Variations of Diffeomorphism Monte Carlo 178

## 4 Evaluating the Implicit Midpoint Integrator for Riemannian Manifold Hamiltonian Monte Carlo

4.1 Introduction 182
4.2 Background 185
4.3 Related Work 189
4.4 Experimental Results 190
  4.4.1 Summary of Integrators 191
  4.4.2 Banana-Shaped Distribution 192
  4.4.3 Hierarchical Neal’s Funnel Distribution 195
  4.4.4 Stochastic Volatility Model 196
  4.4.5 Fitzhugh-Nagumo ODE Model 197
4.5 Conclusion 200

## 4.A Appendices to Chapter 4

4.A.1 Momentum Negation Symmetry of Implicit Midpoint 201
4.A.2 Experimentation with a Quadratic Hamiltonian 204
4.A.3 Bayesian Logistic Regression 205
4.A.4 Volume Preservation and Symmetry Metrics 206
4.A.5 Implementation of Integrators 209
  4.A.5.1 Implementations of the Generalized Leapfrog Integrator 209
  4.A.5.2 Implementations of the Implicit Midpoint Integrator 210
5.A.2 Proofs of Reversibility and Volume Preservation of the Generalized Leapfrog Integrator ........................................ 261
5.A.3 Establishing Detailed Balance in Hamiltonian Monte Carlo ........... 265
5.A.4 Computational Complexity of Implicit Updates ................................. 268
5.A.5 Algorithms ........................................................................ 270
5.A.6 Posterior Summary ............................................................... 272
5.A.7 Detailed Measures and Metrics ................................................... 274
  5.A.7.1 Measuring Reversibility ....................................................... 274
  5.A.7.2 Measuring Volume Preservation ............................................ 275
  5.A.7.3 Measuring Ergodicity ......................................................... 276
  5.A.7.4 Measuring Sample Independence ........................................... 281
  5.A.7.5 Measuring Transition Kernel Similarity ..................................... 282
  5.A.7.6 Measuring Computational Effort ............................................. 283
5.A.8 Additional Details on Experimentation .......................................... 285
  5.A.8.1 Banana-Shaped Distribution ............................................... 285
  5.A.8.2 Bayesian Logistic Regression ............................................... 288
  5.A.8.3 Stochastic Volatility Model .................................................. 290
  5.A.8.4 Log-Gaussian Cox-Poisson Model ......................................... 294
  5.A.8.5 Details for Neal’s Funnel Distribution ..................................... 297
  5.A.8.6 Details for Fitzhugh-Nagumo ................................................. 298
  5.A.8.7 Multi-Scale Phenomena in a Multivariate Student-t Distribution 300
  5.A.8.8 Details for Newton’s Method ................................................ 304
5.A.9 Stability Checks in Newton Methods ........................................... 310
5.A.10 Non-Volume-Preserving Errors in the Fitzhugh-Nagumo Posterior ........ 313
5.A.11 Sensitivity of Jacobian Determinant Estimates to Perturbation Size ........ 316
5.A.12 Relative Error of Reversibility .................................................. 318
5.A.13 The Stationary Distribution of a Reversible Proposal with Non-Unit Jacobian Determinant 320

5.A.14 Ergodicity Issues in Fixed Point Iteration 324

6 Several Remarks on the Numerical Integrator in Lagrangian Monte Carlo 327

6.1 Introduction 328

6.2 Preliminaries 330

6.2.1 Notation 331

6.2.2 Lagrangian Monte Carlo: The Essentials 331

6.3 Related Work 334

6.4 Analytical Apparatus 334

6.4.1 Inversion of the Integration Sequence 335

6.4.2 Aversions to Inverting the Integration Sequence 337

6.4.3 The Order of the Explicit Integrators 339

6.4.4 Efficient Computation of the Jacobian Determinant and the Update to Velocity 340

6.4.5 Built-In Robustness of the Lagrangian Integrator 340

6.5 Experimentation 342

6.5.1 Demonstration of Second-Order Error 343

6.5.2 Banana-Shaped Posterior Distribution 344

6.5.3 Bayesian Logistic Regression 346

6.5.4 Multiscale Student’s t-Distribution 348

6.5.5 Stochastic Volatility Model 351

6.5.6 Fitzhugh-Nagumo Model 352

6.6 Conclusion 354

6.A Appendices to Chapter 6 356

6.A.1 Algorithms 356
8.A.1.1 Proof of Proposition 8.2.3 . . . . . . . . . . . . . . . . . . . . 467
8.A.1.2 Proof of Lemma 8.2.4 . . . . . . . . . . . . . . . . . . . . . 467
8.A.1.3 Proof of Proposition 8.2.6 . . . . . . . . . . . . . . . . . . . . 468
8.A.1.4 Proof of Lemma 8.2.19 . . . . . . . . . . . . . . . . . . . . . 468
8.A.1.5 Proof of Lemma 8.2.20 . . . . . . . . . . . . . . . . . . . . . 469
8.A.1.6 Proof of Theorem 8.2.30 . . . . . . . . . . . . . . . . . . . . 469
8.A.1.7 Proof of Lemma 8.3.2 . . . . . . . . . . . . . . . . . . . . . . 469
8.A.1.8 Proof of Proposition 8.3.4 . . . . . . . . . . . . . . . . . . . . 470
8.A.1.9 Proof of Lemma 8.3.6 . . . . . . . . . . . . . . . . . . . . . . 470
8.A.1.10 Proof of Corollary 8.3.7 . . . . . . . . . . . . . . . . . . . . 470
8.A.1.11 Proof of Proposition 8.3.8 . . . . . . . . . . . . . . . . . . . . 471
8.A.1.12 Proof of Theorem 8.3.9 . . . . . . . . . . . . . . . . . . . . . 471
8.A.1.13 Proof of Proposition 8.3.13 . . . . . . . . . . . . . . . . . . . 472
8.A.1.14 Proof of Proposition 8.3.14 . . . . . . . . . . . . . . . . . . . 472
8.A.1.15 Proof of Corollary 8.3.15 . . . . . . . . . . . . . . . . . . . . 473
8.A.1.16 Proof of Lemma 8.3.16 . . . . . . . . . . . . . . . . . . . . . 473
8.A.1.17 Proof of Proposition 8.3.17 . . . . . . . . . . . . . . . . . . . 474

9 Conclusion 475

F The Shadow Hamiltonian of the Implicit Midpoint Integrator 478
F.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 479
F.2 Mathematical Background . . . . . . . . . . . . . . . . . . . . . . 481
F.3 Identifying the Fourth-Order Shadow Hamiltonian . . . . . . . . . . 487
F.4 Computation of the Shadow Hamiltonian . . . . . . . . . . . . . . . 490
F.4.1 Direct Calculation . . . . . . . . . . . . . . . . . . . . . . . . . 490
F.4.2 Automatic Differentiation . . . . . . . . . . . . . . . . . . . . . 491
F.4.3 The Holomorphic Semi-Shadow . . . . . . . . . . . . . . . . . . 492
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>A summary of the relationship between Riemannian deterministic and stochastic dynamics and (preconditioned) Euclidean dynamics.</td>
<td>113</td>
</tr>
<tr>
<td>4.1</td>
<td>Trajectories computed by the leapfrog integrator and the implicit midpoint integrator for a quadratic, separable Hamiltonian.</td>
<td>182</td>
</tr>
<tr>
<td>4.2</td>
<td>Comparison between the error in symmetry and error in volume preservation properties of the implicit midpoint integrator and the generalized leapfrog integrator on the banana-shaped distribution.</td>
<td>193</td>
</tr>
<tr>
<td>4.3</td>
<td>Comparison of the degree to which the implicit midpoint and generalized leapfrog integrators violate reversibility and volume preservation on Neal’s funnel distribution.</td>
<td>195</td>
</tr>
<tr>
<td>4.4</td>
<td>Comparison of the degree to which the implicit midpoint and generalized leapfrog integrators violate reversibility and volume preservation on the FitzHugh-Nagumo posterior.</td>
<td>198</td>
</tr>
<tr>
<td>4.A.1</td>
<td>Comparison between the energy conservation of the implicit midpoint integrator and the generalized leapfrog integrator on a quadratic Hamiltonian.</td>
<td>204</td>
</tr>
<tr>
<td>5.1</td>
<td>Visualization of the error in reversibility, error in volume-preservation, and the number of decimal digits of similarity in transition kernels for variable thresholds in Neal’s funnel distribution with $10 + 1$ dimensions.</td>
<td>243</td>
</tr>
</tbody>
</table>
5.2

Visualization of the sample quality of the variables in Neal’s funnel distribution as measured by the ESS per second and the distribution of KolmogorovSmirnov (KS) statistics over random one-dimensional subspaces. . . . . . 244

5.3

Additional measures of ergodicity in Neal’s funnel distribution. . . . . . 245

5.4

The use of Ruppert averaging in Neal’s funnel distribution to adaptively
set the convergence threshold to achieve six decimal digits of similarity
compared to a transition kernel with a threshold of 1 × 10−10 . . . . . . . 245

5.5

Visualization of the posterior mean of vt (in blue) and rt (in orange) in the
Fitzhugh-Nagumo posterior distribution. . . . . . . . . . . . . . . . . . . 246

5.6

Visualization of the errors in reversibility and volume preservation in the
Fitzhugh-Nagumo posterior distribution, and the ergodicity of the RMHMC
Markov chain as measured by the Kolmogorov-Smirnov statistics over
random sub-spaces. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 247

5.7

Visualization of the effective sample sizes (ESS) of the variables in the
Fitzhugh-Nagumo posterior. . . . . . . . . . . . . . . . . . . . . . . . . 248

5.8

Additional measures of ergodicity in the Fitzhugh-Nagumo differential
equation model. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 248

5.9

The use of Ruppert averaging in the Fitzhugh-Nagumo posterior distribution to adaptively set the convergence threshold to achieve four decimal
digits of similarity compared to a transition kernel with a threshold of
1 × 10−10 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 249

5.10

Dependency of the reversibility metric on the procedure used to resolve
the implicit updates to momentum and position. . . . . . . . . . . . . . . 251

5.11

The dependency of the Jacobian determinant metric on the procedure used
to resolve the implicit updates to momentum and position. . . . . . . . . 252

xvi


5.A.1 Visualization of the error in reversibility, error in volume-preservation,
and the number of decimal digits of similarity in transition kernels for
variable thresholds in the banana posterior distribution . . . . . . . . . 284

5.A.2 Visualization of the computational effort required to sample with RMHMC
from the banana posterior. . . . . . . . . . . . . . . . . . . . . . . . . . 284

5.A.3 Visualization of the distribution of Kolmogorov-Smirnov statistics over
variable thresholds in the banana-shaped posterior. . . . . . . . . . . . 285

5.A.4 Ergodicity as measured through Biau, MMD, and sliced Wasserstein cri-
teria in the banana-shaped distribution. . . . . . . . . . . . . . . . . . . 285

5.A.5 Visualization of the effective sample size per second of RMHMC com-
pared to HMC on the banana-shaped distribution. . . . . . . . . . . . . 286

5.A.6 The use of Ruppert averaging in the banana-shaped distribution to adap-
tively set the convergence threshold to achieve eight decimal digits of sim-
ilarity compared to a transition kernel with a threshold of $1 \times 10^{-10}$. . . 286

5.A.7 Visualization of the time-normalized effective sample sizes of the vari-
ables in the logistic regression posterior. . . . . . . . . . . . . . . . . 289

5.A.8 Visualization of the error in reversibility, error in volume-preservation,
and the number of decimal digits of similarity in transition kernels for
variable thresholds in the logistic posterior distribution. . . . . . . . . . 289

5.A.9 Visualization of the posterior distribution over the stochastic volatilities
wherein the process hyperparameters $(\sigma^2, \phi, \beta)$ are sampled by RMHMC
with variable threshold. . . . . . . . . . . . . . . . . . . . . . . . . . . . 291

5.A.10 Visualization of the error in reversibility, error in volume-preservation,
and the number of decimal digits of similarity in transition kernels for
variable thresholds in the stochastic volatility posterior. . . . . . . . . . 291

5.A.11 Visualization of the ESS per second for the three latent parameters of the
stochastic volatility model. . . . . . . . . . . . . . . . . . . . . . . . . . 291
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.A.12</td>
<td>Visualization of the number of fixed point iterations required to compute the implicit updates to position and momentum required by the generalized leapfrog integrator in the stochastic volatility model.</td>
</tr>
<tr>
<td>5.A.13</td>
<td>Visualization of the posterior mean (in the first row) and posterior standard deviation (in the second row) of the spatial latent process wherein the kernel parameters ((\sigma^2, \beta)) are sampled by RMHMC with variable thresholds.</td>
</tr>
<tr>
<td>5.A.14</td>
<td>Visualization of the error in reversibility and error in volume-preservation in the log-Gaussian Cox-Poisson posterior.</td>
</tr>
<tr>
<td>5.A.15</td>
<td>Visualization of the number of fixed point iterations required to compute the implicit updates to position and momentum required by the generalized leapfrog integrator in the log-Gaussian Cox-Poisson model.</td>
</tr>
<tr>
<td>5.A.16</td>
<td>Visualization of the computational effort required to sample with RMHMC from Neal’s funnel distribution.</td>
</tr>
<tr>
<td>5.A.17</td>
<td>Visualization of the error in reversibility and error in volume-preservation for variable thresholds in the Fitzhugh-Nagumo posterior distribution.</td>
</tr>
<tr>
<td>5.A.18</td>
<td>Visualization of how multiscale properties of a target multivariate Student-(t) distribution affect the ergodicity and computational efficiency of the Euclidean HMC and RMHMC algorithms.</td>
</tr>
<tr>
<td>5.A.19</td>
<td>Visualization of the error in reversibility, error in volume-preservation, and the number of decimal digits of similarity in transition kernels for variable thresholds in a multivariate Student-(t) distribution.</td>
</tr>
<tr>
<td>5.A.20</td>
<td>Visualization of the sample quality of the variables in a multivariate Student-(t) distribution as measured by the ESS per second and the distribution of Kolmogorov-Smirnov (KS) statistics over random one-dimensional subspaces.</td>
</tr>
<tr>
<td>5.A.21</td>
<td>Measures of ergodicity in the multivariate Student-(t) distribution.</td>
</tr>
</tbody>
</table>
5.A.22 The use of Ruppert averaging in the multiscale Student-$t$ distribution to adaptively set the convergence threshold to achieve eight decimal digits of similarity compared to a transition kernel with a threshold of $1 \times 10^{-10}$. 303

5.A.23 The number of iterations required to resolve the implicit updates to the momentum variable when using either fixed point iteration or Newton’s method. 305

5.A.24 Three circumstances are shown wherein the use of Newton’s method to resolve the implicit update to the momentum resulted in degraded sampling efficiency due to the higher computational burden of Newton’s method compared to fixed point iteration. 307

5.A.25 Three circumstances in which Newton’s method produced performance comparable to fixed point iteration when resolving the implicit update to the momentum. 308

5.A.26 We visualize the ergodicity properties for the three implementations of RMHMC. 309

5.A.27 Comparison of fixed point iterations, Newton’s method (to update both position and momentum), and Newton’s method with a stability check in the banana-shaped distribution. 310

5.A.28 Visualization of how a failure to respect the symmetry of partial derivatives yields severely degraded sample quality. 313

5.A.29 Measures of the error in volume preservation of the generalized leapfrog integrator for variable thresholds. 316

5.A.30 Measures of relative reversibility of the generalized leapfrog integrator for variable thresholds. 318
6.1 Visualization of the difference in expected distance when integrating a trajectory of a separable, quadratic Hamiltonian for a varying number of steps and integration step-sizes when using the standard leapfrog and inverted leapfrog methods. 337

6.2 Visualization of the second-order accuracy (i.e. third order local error) of the Lagrangian integrator on Lagrangian dynamics derived from a non-separable Hamiltonian. 344

6.3 The ESJD, the minimum ESS per second, and the distribution of Kolmogorov-Smirnov statistics for sampling from the banana-shaped distribution using RMHMC, LMC, and ILMC. 345

6.4 The ESJD and the minimum and mean ESS per second for the breast cancer dataset using RMHMC, LMC, and ILMC. 346

6.5 The ESJD and the minimum and mean ESS per second for the thyroid cancer dataset using RMHMC, LMC, and ILMC. 347

6.6 The ESJD, the minimum ESS per second, and the distribution of Kolmogorov-Smirnov statistics for sampling from the multiscale Student-t distribution with $\nu = 5 \times 10^0$ using RMHMC, LMC, and ILMC. 348

6.7 The ESJD, the minimum ESS per second, and the distribution of Kolmogorov-Smirnov statistics for sampling from the multiscale Student-t distribution with $\nu = 5 \times 10^4$ using RMHMC, LMC, and ILMC. 349

6.8 The ESJD and the minimum and mean ESS per second for the stochastic volatility model using HMC, RMHMC, LMC, and ILMC. 350

6.9 The minimum ESS per second and the Kolmogorov-Smirnov ergodicity metric for inferring the posterior distribution of the Fitzhugh-Nagumo posterior distribution. 352
<table>
<thead>
<tr>
<th>Section</th>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>The ergodicity of the geometric Markov chains as a function of the sampling step in the banana-shaped posterior distribution.</td>
<td>406</td>
</tr>
<tr>
<td>7.2</td>
<td>The ESJD, the minimum ESS, and the time-normalized minimum ESS for inference in the banana-shaped posterior distribution.</td>
<td>406</td>
</tr>
<tr>
<td>7.3</td>
<td>The ESJD, the minimum ESS, and the time-normalized minimum ESS for inference in the hierarchical Bayesian logistic regression posterior distribution.</td>
<td>408</td>
</tr>
<tr>
<td>7.4</td>
<td>The ergodicity of the geometric Markov chains as a function of the sampling step in Neal’s funnel distribution.</td>
<td>409</td>
</tr>
<tr>
<td>7.5</td>
<td>The ESJD, the distribution of Kolmogorov-Smirnov statistics, and the time-normalized minimum ESS for inference in Neal’s funnel distribution.</td>
<td>410</td>
</tr>
<tr>
<td>7.6</td>
<td>The ESJD, the minimum ESS, and the time-normalized minimum ESS for inference in the posterior distribution of the stochastic volatility model.</td>
<td>411</td>
</tr>
<tr>
<td>7.7</td>
<td>The ESJD, the minimum ESS, and the time-normalized minimum ESS for inference in the posterior distribution of the log-Gaussian Cox-Poisson model.</td>
<td>412</td>
</tr>
<tr>
<td>7.8</td>
<td>The ergodicity of the geometric Markov chains as a function of the sampling step in the Fitzhugh-Nagumo posterior distribution.</td>
<td>413</td>
</tr>
<tr>
<td>7.9</td>
<td>The ESJD, the minimum ESS, and the time-normalized minimum ESS for inference in the Fitzhugh-Nagumo posterior distribution.</td>
<td>414</td>
</tr>
<tr>
<td>7.10</td>
<td>The ergodicity of the geometric Markov chains as a function of the sampling step in the multiscale Student-(t) distribution.</td>
<td>415</td>
</tr>
<tr>
<td>7.11</td>
<td>The ESJD and the MSJD for inference in the multiscale Student-(t) distribution.</td>
<td>416</td>
</tr>
<tr>
<td>7.12</td>
<td>The distribution of Kolmogorov-Smirnov statistics and the time-normalized minimum ESS for inference in the multiscale Student-(t) distribution.</td>
<td>416</td>
</tr>
</tbody>
</table>
8.1 An illustration of the explicit integration procedure with Lagrange multpliers. ........................................455
8.2 The error in the position variable computed by variations of the explicit and generalized leapfrog integrators in the non-separable Hamiltonian. . . 456
8.3 The deviation in Hamiltonian energy computed by variations of the explicit and generalized leapfrog integrators. ........................................456
8.4 Energy conservation of the explicit integrator, the implicit midpoint, and generalized leapfrog method in the Schwarzschild geodesic example. . . 458
8.5 Trajectories of the explicit integrator, the implicit midpoint, and generalized leapfrog method in the Schwarzschild geodesic. .................................459
8.6 The degree to which volume preservation and reversibility are violated for the explicit integrator with a convergence threshold of $1 \times 10^0$ and $1 \times 10^{-12}$ in the banana-shaped distribution. ........................................459
8.7 Boxplots showing the typical region of Kolmogorov-Smirnov statistics among 10,000 randomly selected one-dimensional subspaces for independent samples and Markov chain samples in the banana-shaped distribution. 460
8.8 The effective sample size of RMHMC procedures sampling from the banana-shaped distribution. ........................................460
8.9 The degree to which volume preservation and reversibility are violated for the explicit integrator with a convergence threshold of $1 \times 10^0$ in Neal’s funnel distribution. .................................463
8.10 Boxplots showing the typical region of Kolmogorov-Smirnov statistics among 10,000 randomly selected one-dimensional subspaces for independent samples and Markov chain samples in Neal’s funnel distribution. . . 463
8.11 The effective sample size of the HMC procedures sampling from Neal’s funnel distribution. ........................................464
### List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Comparison of the implicit midpoint and generalized leapfrog integrators on sampling from the banana-shaped distribution.</td>
</tr>
<tr>
<td>F.1</td>
<td>Results of applying the implicit midpoint integrator to a non-separable Hamiltonian.</td>
</tr>
<tr>
<td>F.2</td>
<td>Visualization of the Hamiltonian energy, the fourth-order shadow Hamiltonian, and the holomorphic semi-shadow for step-sizes $h = 0.01$ and $h = 0.5$.</td>
</tr>
<tr>
<td>F.3</td>
<td>The acceptance probabilities associated to “canonical” HMC and shadow HMC using either the fourth-order truncation of the shadow Hamiltonian or the holomorphic semi-shadow.</td>
</tr>
<tr>
<td>G.1</td>
<td>Illustration of adaptation of the independent Metropolis-Hastings algorithm.</td>
</tr>
<tr>
<td>G.2</td>
<td>Failures of the containment condition in a simple example.</td>
</tr>
<tr>
<td>G.3</td>
<td>Result of the Brownian bridge experiment using adaptive independent Metropolis-Hastings with an affine transformation.</td>
</tr>
<tr>
<td>G.4</td>
<td>Examination of the performance of adaptive MCMC methods on sampling from the multimodal and Neal funnel distributions.</td>
</tr>
<tr>
<td>G.5</td>
<td>Results of the $\phi^4$ field experiment using adaptive independent Metropolis-Hastings with a normalizing flow proposal.</td>
</tr>
<tr>
<td>F.A.1</td>
<td>Examinations of the violations of stationarity that result by maximizing the pseudo-likelihood of the accepted samples as an adaptation mechanism.</td>
</tr>
<tr>
<td>F.A.2</td>
<td>Samples and proposals in the $\phi^4$ field experiment.</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>4.2</td>
<td>Comparison of the implicit midpoint and the naive generalized leapfrog integrators on sampling from Neal's funnel distribution.</td>
</tr>
<tr>
<td>4.3</td>
<td>Comparison of the implicit midpoint generalized leapfrog integrators on the stochastic volatility model.</td>
</tr>
<tr>
<td>4.4</td>
<td>Comparison of the implicit midpoint and generalized leapfrog integrators on sampling from the posterior of the Fitzhugh-Nagumo model.</td>
</tr>
<tr>
<td>4.A.1</td>
<td>Comparison of the implicit midpoint and generalized leapfrog integrators on sampling from the Bayesian logistic regression posterior on the Breast Cancer dataset.</td>
</tr>
<tr>
<td>4.A.2</td>
<td>Comparison of the implicit midpoint and generalized leapfrog integrators on sampling from the Bayesian logistic regression posterior on the Diabetes dataset.</td>
</tr>
<tr>
<td>4.A.3</td>
<td>Comparison of the implicit midpoint and generalized leapfrog integrators on sampling from the Bayesian logistic regression posterior on the heart disease dataset.</td>
</tr>
<tr>
<td>5.A.1</td>
<td>Summary of the posterior distributions we examine in our experimental results regarding the effect of the convergence threshold in RMHMC.</td>
</tr>
<tr>
<td>No.</td>
<td>Description</td>
</tr>
<tr>
<td>-----</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>(Fixed Point Iteration) Procedure for solving the equation ( z = f(z) ) via fixed point iteration to a given tolerance.</td>
</tr>
<tr>
<td>2</td>
<td>(G.L.F.(a)) The procedure for a single step of integrating Hamiltonian dynamics using the generalized leapfrog integrator.</td>
</tr>
<tr>
<td>3</td>
<td>(I.M.(a)) The procedure for a single step of integrating Hamiltonian dynamics using the implicit midpoint integrator.</td>
</tr>
<tr>
<td>4</td>
<td>(G.L.F.(b)) The procedure for a single step of integrating Hamiltonian dynamics using the efficient implementation of the generalized leapfrog integrator.</td>
</tr>
<tr>
<td>5</td>
<td>(I.M.(b)) The procedure for a single step of integrating Hamiltonian dynamics using the implicit midpoint integrator as advocated by Leimkuhler and Reich [2005].</td>
</tr>
<tr>
<td>6</td>
<td>(I.M.(c)) The procedure for a single step of integrating Hamiltonian dynamics using the implicit midpoint integrator based on the discrete Lagrangian [Mackay, 1992; Hairer et al., 2006].</td>
</tr>
<tr>
<td>7</td>
<td>The leapfrog integrator, which is a volume-preserving, reversible, second-order accurate numerical integrator for separable Hamiltonians.</td>
</tr>
<tr>
<td>8</td>
<td>Algorithmic implementation of employing fixed point iterations to solve a fixed point equation.</td>
</tr>
<tr>
<td>Algorithmic implementation of employing Newton’s method to solve a fixed point equation.</td>
<td>271</td>
</tr>
<tr>
<td>Implementation of the generalized leapfrog integrator using either fixed point iteration or Newton’s method to resolve the fixed point equations.</td>
<td>271</td>
</tr>
<tr>
<td>Algorithm implementing the Riemannian manifold Hamiltonian Monte Carlo (RMHMC) Markov chain.</td>
<td>272</td>
</tr>
<tr>
<td>Implementation of Ruppert averaging to identify a threshold that produces a desired number of decimal digits of similarity between two numerical integrators on average.</td>
<td>273</td>
</tr>
<tr>
<td>Pseudo-code implementing the generalized leapfrog integrator for computing approximate solutions to Hamilton’s equations of motion.</td>
<td>356</td>
</tr>
<tr>
<td>Pseudo-code implementing the Lagrangian leapfrog integrator for computing approximate solutions to the Lagrangian equations of motion.</td>
<td>357</td>
</tr>
<tr>
<td>Single step of the Hamiltonian / Lagrangian Monte Carlo transition kernel.</td>
<td>357</td>
</tr>
<tr>
<td>Pseudo-code implementing the inverted Lagrangian leapfrog integrator for computing approximate solutions to the Lagrangian equations of motion.</td>
<td>358</td>
</tr>
<tr>
<td>A second-order explicit numerical integrator of Hamilton’s equations of motion which is not suitable for Hamiltonian Monte Carlo.</td>
<td>448</td>
</tr>
<tr>
<td>Procedure for solving the equation ( f(z) = 0 ) via Newton’s method.</td>
<td>448</td>
</tr>
<tr>
<td>Procedure for solving the equation ( f(z) = 0 ) via Broyden’s method.</td>
<td>448</td>
</tr>
<tr>
<td>A second-order numerical integrator of Hamilton’s equations of motion constrained to the equality manifold that is suitable for Hamiltonian Monte Carlo.</td>
<td>449</td>
</tr>
<tr>
<td>A biased Markov chain technique for generating samples from a target probability distribution using the explicit integrator in algorithm 17.</td>
<td>449</td>
</tr>
</tbody>
</table>
A Markov chain technique for generating samples from a target probability distribution using the explicit integrator with Lagrange multipliers in algorithm 20. As $\delta \to 0$, this procedure faithfully satisfies the detailed balance condition.

Algorithm for sampling from a target distribution by adapting a normalizing flow proposal distribution in the independent Metropolis-Hastings algorithm using the pseudo-likelihood objective.
Chapter 1

Introduction

Central to the problem of Bayesian inference is the task of sampling from the posterior distribution of the model parameters. This may be accomplished via several procedures such as through the use of conjugate priors which we now summarize.

Conjugate Priors  In certain cases, a prior may be identified that is “conjugate” to the likelihood function. In such circumstances, the posterior distribution is in the same family as the prior and analytic sampling may be possible. While theoretically sound and efficient, a limitation of this approach is that conjugate priors may not always be obvious and, when they are, exist typically only for relatively simple Bayesian models.

Markov chain Monte Carlo  In Markov chain Monte Carlo (MCMC) one seeks to establish a Markov chain that converges to the posterior distribution in total variation distance. While it is often possible to construct Markov chains that converge as desired, the speed of convergence may be difficult to characterize in practice.

Variational Bayes  In variational Bayes, one seeks to approximate the posterior via a distribution selected from among some family of approximations. Typically, the approximating family is chosen to have computationally expedient mechanisms for
sampling and computing the density. While this makes sample generation fast, the
fact that variational Bayes is fundamentally an approximation means that certain
attributes of the true posterior may not be reflected in the approximating family.

The focus of this dissertation is on MCMC methods. One of the most popular Markov
chain methods to generate samples from the target posterior is Hamiltonian Monte Carlo
(HMC), a method which has its foundations in classical mechanics. In essence, HMC
solves, approximately, via the use of a numerical integrator, the solution to Hamilton’s
equations of motion. The terminal point of this approximate solution is then taken as the
proposal state of the Markov chain and accepted or rejected according to a Metropolis-
Hastings correction. Advancing the state of the Markov chain in this way has two desir-
able properties in many situations. First, by integrating Hamilton’s equations of motion
over a long trajectory, it is possible to obtain a sequence of states for which the auto-
correlation decreases dramatically. This means that samples produced by the Markov
chain will be “closer to independent” in some sense. Second, the form of the Metropolis-
Hastings correction, viewed in the context of certain analytical properties of the exact
solution to Hamilton’s equations of motion, implies that the acceptance probability should
be high, even for distant proposal points.

Over time, a medley of HMC variations have emerged. One notable among these is
the Riemannian manifold Hamiltonian Monte Carlo (RMHMC), which seeks to incorpo-
rate geometric information about the posterior into the proposal mechanism. By aligning
proposals with the directions of the posterior that exhibit the greatest local variation, one
hopes to traverse the typical region of the posterior more efficiently, thereby, hopefully pro-
ducing less auto-correlated states and faster convergence of the Markov chain. However,
incorporating geometric information in the RMHMC Markov chain comes at a computa-
tional price; in particular, more elaborate forms of geometric numerical integration are
required in order to produce a Markov chain satisfying detailed balance with respect to the
target distribution.

This thesis examines the special role of the numerical integrator in RMHMC and other geometric methods of Markov chain Monte Carlo. Throughout the thesis, we will suggest new methods of numerical integration, evaluate methods of integration that have not been applied in the RMHMC context before, and examine specific computations present in numerical integrators to explore how their implementation on digital computers differs from their mathematical definitions. I further experiment with these algorithms and methodologies on a suite of Bayesian inference benchmarks of varying degrees of sophistication, ranging from sampling multivariate Student-$t$ densities with multiple spatial scales to sampling in a hierarchical log-Gaussian Cox-Poisson process.

Chapter 2 introduces a broad spectrum of preliminaries that are invoked at various points in the dissertation. Included here is background material on numerical errors, Hamiltonian and Lagrangian mechanics, integration on manifolds, stochastic differential equations, information geometry and Markov chains.

In chapter 3 we consider a framework for rigorously thinking about the transition kernel in HMC and its derivative techniques. We leverage change-of-variables to formulate a accept-reject decision for proposals generated via a diffeomorphism that satisfies detailed balance with respect to a target measure.

In chapter 4 we proceed to consider the implicit midpoint integrator for RMHMC. This integrator has unique properties that make it desirable as a method of numerical integration: it has superior stability compared to the leapfrog method as well as the unique property that, for quadratic Hamiltonians, energy is perfectly conserved. On the other hand, the implicit midpoint presents fewer opportunities to cache intermediate computations, thereby producing a less computationally efficient method. Moreover, both the implicit midpoint integrator and the generalized leapfrog integrator involve computing solutions to implicitly defined equations. This chapter evaluates the implicit midpoint integrator in terms of energy conservation and stability, in terms of how the quality of the solution to the
implicit equations impacts detailed balance of the Markov chain, and how computational efficiency affects the time-normalized performance metrics of the Markov chain.

A theme in chapter 4 is that solving implicitly defined equations on a digital computer will produce errors in the satisfaction of detailed balance in the RMHMC Markov chain above the level of machine precision. In chapter 5 we examine this concept in greater detail in order to understand how the precision of the solution affects ergodicity of the Markov chain and the computational efficiency of the method. The precision of the solution is typically measured by a convergence tolerance parameter. In this chapter, we propose mechanisms by which to adapt the convergence tolerance of the fixed point solutions within the context of a given posterior. We also examine a procedure by which to make the RMHMC transition kernel less sensitive to the choice of convergence tolerance.

Chapter 6 changes focus to another Markov chain technique based on Lagrangian mechanics rather than Hamiltonian mechanics called Lagrangian Monte Carlo (LMC). Unlike the numerical integrators employed in HMC and RMHMC, the integrator in LMC is non-volume-preserving, necessitating computation of the Jacobian determinant of the integrator when applying the Metropolis-Hastings accept-reject decision. In practice, this requires computing the determinant of four matrices at each integration step. This chapter makes two contributions when treating the LMC integrator. First, we clarify a misunderstanding in the literature regarding the order of accuracy of this numerical method: it had been previously claimed that the LMC integrator is only first order accurate; in fact, it is second order. Our other contribution is to suggest a simple mechanism by which to reduce the number of matrix determinants in each step from four to two; this can be an important computational acceleration in certain applications.

Chapter 7 examines the geometric ergodicity of geometric methods of Markov chain Monte Carlo, such as RMHMC or LMC. In Euclidean spaces, a single step of HMC and the Metropolis-adjusted Langevin algorithm (MALA) can be constructed to be exactly equivalent to one another. However, it is not true that a single step of RMHMC is equiva-
lent to a single step of the manifold Metropolis-adjusted Langevin algorithm (MMALA). Nevertheless, based on the equivalency in the Euclidean case, we may consider modifying RMHMC (or LMC) by replacing their single-step implementation with MMALA, thereby inheriting geometric ergodicity from MMALA under certain conditions. This chapter also explores conditions under which the original, unmodified versions RMHMC or LMC are geometrically ergodic.

Chapter 8 represents my original research in numerical integration for RMHMC. In this chapter, we consider an integrator that had been proposed for integrating a broad class of Hamiltonians; however, despite attempts in this direction, a naive implementation of RMHMC with this choice of integrator cannot satisfy detailed balance with respect to the target density. Our contribution is to discuss how the numerical integrator in question may be corrected by constraining it to a particular embedded manifold of an expanded phase-space. By enforcing this constraint, one obtains a numerical integrator that is suitable for RMHMC.

This dissertation contains two appendices that contain tangential investigations but which may nevertheless be of broader interest. In chapter F, we investigate the concept of a shadow Hamiltonian for the implicit midpoint integrator which is more precisely conserved by the integrator than the actual Hamiltonian. We derive the shadow Hamiltonian of the implicit midpoint integrator for a broad class of non-canonical Hamiltonian mechanical systems. We also derive a related quantity called the holomorphic semi-shadow which is conserved to the same degree as the shadow Hamiltonian but is particularly simple to compute under the assumption that the Hamiltonian is holomorphic. In chapter G, we abandon Markov chain procedures based on classical mechanics in favor of the independent Metropolis-Hastings sampler. In this work, our motivation is to combine the flexibility of deep generative models with the provable asymptotic ergodicity of MCMC. In this manner, we are able to transform techniques which have been traditionally employed for approximate inference via variational Bayes into correct methods of sampling.
We consider a proposal distribution that is parameterized and state conditions under which the parameters of the proposal distribution may be adapted over the course of sampling without destroying the asymptotic ergodicity of the chain.

We now emphasize the broader contributions and context of this research. In chapters 4, 5 and 8, we examine the inaccuracy of Markov chains in terms of violations of the detailed balance condition, a fundamental criterion that, when present, ensures the stability of Markov chains. Understanding that these inaccuracies can exist (chapters 4 and 8) combined with methods of diagnosing their presence, and alternatives to minimizing their effects (chapter 5) advances the state of the practice of these geometric methods of Bayesian inference. Chapter 6 discusses errors within the context of robustness and examines how certain kinds of numerical integrators offer a degree of protection from violations of detailed balance when other numerical integrators do not. Chapters 4, 6 and 8 explore the numerical integrator used in RMHMC and related geometric methods. In these papers, we seek to understand the degree of freedom one has in the selection of a numerical integrator and the impact this choice has on the Markov chain behavior. In chapter 4, we recall a theory that proves that the implicit midpoint integrator will enjoy a higher acceptance probability than the generalized leapfrog method in Gaussian settings; this is important because the efficiency of a Markov chain is directly affected by the chain’s acceptance rate. Chapter 6 shows how numerical integration of Lagrangian dynamics can be carried out more efficiently by simply exchanging the sequence of integration of the position and velocity variables; we then illustrate how this simple modification can produce more efficient Markov chains in certain situations. Because geometric MCMC procedures such as RMHMC are more computationally expensive than their Euclidean counterparts, understanding how they may be accelerated is an important research direction for advancing their adoption in the Bayesian community of practice. Chapter 7 gives special attention to the question of fast convergence of MCMC methods. Using the theory of inherited geometric ergodicity, we propose a simple modification to RMHMC and LMC
that permits these methods to be imbued with fast convergence rates whenever the manifold Metropolis-adjusted Langevin algorithm is fast. Fast convergence of Markov chains is important not only because one desires samples that closely approximate the target distribution, but also because it implies a central limit theorem for Monte Carlo expectations computed from a single chain. Therefore, it is important from a practitioner’s point of view to have fast convergence. Our work in chapters 4 and 5 also explores ergodicity of the original RMHMC algorithm and gives further emphasis to understanding how errors in detailed balance impact other MCMC performance metrics such as the expected sample size. At the same time, we seek to present a balanced view of our research and we discuss circumstances wherein our proposed methods are at a disadvantage. For instance, in chapter 4 we discuss the relative inefficiency of the implicit midpoint method compared to a well-executed generalized leapfrog implementation; in chapter 6 we explain why inverting the sequence of integration can be ill-advised and result in greater auto-correlation between samples; in chapter 8 we discuss how the theoretical correction to an invalid integrator is not an efficient procedure.
Chapter 2

Preliminaries

2.1 Numerical Errors

Certain components of this work pertain to errors that result when precise mathematical relations are replaced by numerical approximations. In describing these concepts, we draw on Dahlquist and Björck [2003].

2.1.1 Truncation Errors

An important variety of numerical error in this work is truncation error. Truncation errors occur when some infinite procedure, whose limit is the solution to a mathematical problem of interest, is terminated after a finite number of steps. In this case, the premature termination results in non-zero error between the actual solution and the computed value.

Example 1. Consider the definition of the derivative:

\[ f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}. \]  

(2.1)
Consider the following limiting method. For \( h > 0 \), define

\[
f'_h(x) = \frac{f(x + h) - f(x)}{h}. \tag{2.2}
\]

Clearly \( \lim_{h \to 0} f'_h(x) = f'(x) \) from the definition. Consider that we may terminate this infinite sequence at \( h = 10^{-5} \), thereby obtaining the estimate \( f'_{10^{-5}}(x) \) so that the error of the approximation is \( f'(x) - f'_{10^{-5}}(x) \). In fact, this truncation error has an explicit form from Taylor’s remainder theorem:

\[
f'(x) - f'_h(x) = f'(x) - \frac{f(x + h) - f(x)}{h} \tag{2.3}
\]

\[
= f'(x) - f'(x) - hf''(\tilde{x}) \tag{2.4}
\]

\[
= hf''(\tilde{x}), \tag{2.5}
\]

for some \( \tilde{x} \in (x, x + h) \).

**Example 2.** The quality of the finite difference approximation can be improved through the use of a central difference formula. Consider the approximation,

\[
f'_h(x) = \frac{f(x + h/2) - f(x - h/2)}{h}. \tag{2.6}
\]

Then the error is, again using Taylor’s remainder theorem,

\[
f'(x) - \frac{f(x + h/2) - f(x - h/2)}{h} \tag{2.7}
\]

\[
= f'(x) - \frac{1}{h} \left( f(x) + h f'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{6} f'''(\tilde{x}_+) - f(x) + h f'(x) \right) \tag{2.8}
\]

\[
= \frac{h^2}{6} f''(\tilde{x}_+) + \frac{h^2}{6} f''(\tilde{x}_-). \tag{2.9}
\]
for $\bar{x}_+ \in (x, x + h/2)$ and $\bar{x}_- \in (x - h/2, x)$. Hence, instead of having truncation error that is linear in $h$, the central difference formula has error quadratic in $h$.

2.1.2 Finite Precision Number Systems

The preceding examples are true when $f$ can be computed exactly with infinite precision and all elementary algebraic computations (addition, subtraction, multiplication, and division) are carried out exactly. In this case, all error from the finite difference approximation derives from truncation error. However, on modern computing machinery, this is not the only source of error due to the use of finite-precision arithmetic. This raises the important distinction between how numbers are represented mathematically and digitally, and the difference between the implementation of basic mathematical operations on digital systems. The following discussion is drawn from [Trefethen and Bau 1997].

Definition 2.1.1 (Trefethen and Bau [1997]). An idealized finite precision number system $F$ is a subset of the real numbers of the form,

$$F = \left\{ \pm \frac{m}{\beta^e} : \beta^{t-1} \leq m \leq \beta^t - 1, e \in \mathbb{N} \right\}$$

(2.10)

where $t \geq 1$ is the precision, $\beta$ is the base, and $e$ is the exponent.

Example 3. An important case, which is found on many computing devices is $\beta = 2$ and $t = 53$.

[Trefethen and Bau 1997] say that a finite precision number system $F$ has the following property. Let $\text{fl} : \mathbb{R} \to F$ be the function that maps an arbitrary real number to the closest number in the finite precision number system. Then, there is a constant $\epsilon_{\text{mach}} = \frac{1}{2} \beta^{1-t}$ such that for every $x \in \mathbb{R}$, $\text{fl}(x) = (1 + \epsilon)x$ with $|\epsilon| \leq \epsilon_{\text{mach}}$; the quantity $\epsilon$ will depend on $x$, though we will typically suppress this dependence. Therefore, $\epsilon_{\text{mach}}$ is called the machine precision. The quantity $\epsilon_{\text{mach}}$ can be understood as the greatest relative error between an
arbitrary real number and the closest point in the finite precision number system. This can be seen as follows. From the condition \( \beta^{t-1} \leq m \leq \beta^t - 1 \) we see that \( m/\beta^t \in [1/\beta, 1) \).

Therefore, let \( x^* \) be an arbitrary real number in \([1/\beta, 1)\). Then,

\[
\left| x^* \beta^e - \frac{m}{\beta^t} \beta^e \right| \leq \frac{1}{2} \beta^{-t} \beta^e,
\]

(2.11)

which is a bound on the absolute error. The relative error is therefore bounded by,

\[
\frac{\left| x^* \beta^e - \frac{m}{\beta^t} \beta^e \right|}{x^* \beta^e} \leq \frac{1}{2} \beta^{-t} \beta^e
\]

(2.12)

\[
= \frac{1}{2} \frac{\beta^{-t}}{x^*}
\]

(2.13)

\[
\leq \frac{1}{2} \beta^{1-t}
\]

(2.14)

\[
= \epsilon_{\text{mach}}.
\]

(2.15)

Having at this point a notion of a number system, we next require some notion of how mathematical operations are carried out in a finite precision number system. Let \( \star : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \) denote a binary arithmetic operation (such as addition, subtraction, multiplication or division) and let \( \star_{\text{mach}} : \mathbb{F} \times \mathbb{F} \rightarrow \mathbb{F} \) denote its implementation in a finite precision number system. We assume:

**Axiom 1** (Fundamental Axiom of Floating Point Arithmetic [Trefethen and Bau [1997]]).

*Let \( \mathbb{F} \) be an idealized finite precision number system. For every \( x, y \in \mathbb{F} \), and every binary arithmetic operator \( \star \in \{+, -, \times, \div\} \), there exists a number \( \epsilon \), depending on \( x \) and \( y \), such that

\[
x \star_{\text{mach}} y = (1 + \epsilon)(x \star y).
\]

(2.16)

where \(|\epsilon| \leq \epsilon_{\text{mach}}\).*
In other words, the implementations of these elementary arithmetic operations produce outputs that are within $\epsilon_{\text{mach}}$ of the actual value. Let us now see an example of how numerical calculations can be badly behaved.

**Example 4.** Consider computing the function $f(x) = 1 - \cos(x)$ for $x$ close to zero. A Taylor expansion of cosine says that $\cos(x) = 1 - \frac{1}{2}x^2 + O(x^4) \approx 1 - \frac{1}{2}x^2$ for $x$ near zero. Let us assume that the machine implementation of cosine produces a value $\cos_{\text{mach}}(x) = (1 - \frac{1}{2}x^2)(1 + \epsilon_1)$ where $|\epsilon_1| \leq \epsilon_{\text{mach}}$ with $\epsilon_1$ depending on $x$. Therefore the machine implementation of $f$ is

$$1 - \text{mach } \cos_{\text{mach}}(x) = \left(1 - (1 - \frac{1}{2}x^2)(1 + \epsilon_1)\right)(1 + \epsilon_2)$$  \hspace{1cm} (2.17)

where $|\epsilon_2| \leq \epsilon_{\text{mach}}$, with $\epsilon_2$ depending on $\cos_{\text{mach}}(x)$. Multiplying out,

$$\left(1 - (1 - \frac{1}{2}x^2)(1 + \epsilon_1)\right)(1 - \epsilon_2) = \left(1 - 1 + \frac{1}{2}x^2 - \epsilon_1 + \frac{\epsilon_1 x^2}{2}\right)(1 + \epsilon_2)$$  \hspace{1cm} (2.18)

$$= \left(\frac{1}{2}x^2 - \epsilon_1 + \frac{\epsilon_1 x^2}{2}\right)(1 + \epsilon_2)$$  \hspace{1cm} (2.19)

$$= \frac{x^2}{2} - \epsilon_1 + \frac{\epsilon_1 x^2}{2} + \frac{\epsilon_2 x^2}{2} - \epsilon_1 \epsilon_2 + \frac{\epsilon_1 \epsilon_2 x^2}{2}$$  \hspace{1cm} (2.20)

$$\approx \frac{x^2}{2} - \epsilon_1.$$  \hspace{1cm} (2.21)

Approximating $f(x)$ by $x^2/2$ for $x$ near zero, the relative error of this calculation is,

$$\left|\frac{(x^2/2 - \epsilon_1)}{x^2/2} - 1\right| = 2\left|\frac{\epsilon_1}{x^2}\right.$$  \hspace{1cm} (2.22)

Since the error is inversely proportional to $x^2$, we see that, in a relative sense, the error will be large for values of $|x| \leq 10^{-3}$.

**Example 5.** Matters can be even worse for the function $g(x) = \frac{f(x)}{x^2} = \frac{1 - \cos(x)}{x^2}$. For $x$ near zero, several applications of L’Hôpital’s rule shows that $g(x) \approx 1/2$. Let us assume that
the machine implementation of $f$ is computed as before with

$$1 - \text{mach} \cos_{\text{mach}}(x) \approx \frac{x^2}{2} - \epsilon_1. \quad (2.23)$$

Let us now make the optimistic assumption that the remaining operations to compute $g$ (namely the division by $x^2$) can be carried out exactly. In practice this is not true, but it represents a best-case scenario. Thus, the machine implementation of $g$ produces,

$$\frac{1 - \text{mach} \cos_{\text{mach}}(x)}{x^2} = \frac{1}{2} - \frac{\epsilon_1}{x^2}. \quad (2.24)$$

For small $x$, the absolute error of the computed value is

$$\left| \frac{1 - \text{mach} \cos_{\text{mach}}(x)}{x^2} - \frac{1}{2} \right| = \frac{|\epsilon_1|}{x^2}. \quad (2.25)$$

Thus, for $|x| \approx 10^{-8}$ the absolute error of the computation will be approximately 1; the relative error will be two. The results of this computation are poor in both a relative and absolute sense.

**Example 6.** Let us now return to the case of the central finite differences formula in the presence of an idealized finite precision number system. In the best case, the quantity $f(x)$ is represented as $f(x)(1 + \epsilon)$ where $|\epsilon| \leq \epsilon_{\text{mach}}$, with $\epsilon$ depending on $x$. Therefore, in the actual central difference computation, we have

$$\left| \frac{f(x + h/2)(1 + \epsilon_1) - f(x - h/2)(1 + \epsilon_2)}{h} \right| \leq \frac{h^2}{12} \max_{y \in (x-h/2, x+h/2)} |f'''(y)| + \frac{2\epsilon_{\text{mach}}}{h} \max_{y \in (x-h/2, x+h/2)} |f(y)| \quad (2.28)$$

One sees, therefore, that the truncation error has added to it a new term that is proportional
to the machine error and inversely proportional to central difference perturbation size.

While the first term will decrease when $h$ is decreased, the second term will increase.  

How can one choose $h$? One might choose $h$ so that the two error contributions are approximately equal. A very rough calculation might suppose then to set $h^2 \approx \frac{\epsilon_{\text{mach}}}{h}$, or $h^3 \approx \epsilon_{\text{mach}}$. On many computing machines $\epsilon_{\text{mach}} \approx 1.1 \times 10^{-16}$, so that $h \approx 1 \times 10^{-5}$. In many problems, but not all, this is a good choice.

Surprisingly, it is possible to compute numerical derivatives up to machine precision when the function is holomorphic. Eventually, in chapter [F] we will expand our discussion so as to introduce functions of complex-valued arguments. An important concept in complex analysis is that of a holomorphic (equivalently, complex analytic) function.

**Definition 2.1.2.** A function $f : \mathbb{C}^m \to \mathbb{C}$ is holomorphic if, for all $z \in \mathbb{C}^m$ the Taylor series expansion,

$$f(z) + \nabla f(z)^\top (z' - z) + \frac{1}{2} (z' - z)^\top \nabla^2 f(z)(z' - z) + \ldots$$

converges to $f(z')$ where $z' \in \mathbb{C}^m$.

An interpretation of this property is that if $f$ is a holomorphic function then it can be approximated arbitrary well by polynomials within an open ball around each point $z' \in \mathbb{C}^m$. There are many examples of holomorphic functions, such as trigonometric functions, square roots, and reciprocation; however, some functions are not holomorphic, such as the absolute value function. Holomorphic functions have the intriguing property that if we take $z' = z + ihr$ where $f(z) \in \mathbb{R}$, $h \in \mathbb{R}$, and $r \in \mathbb{R}^m$, then we have

$$f(z + ir) = f(z) + ih\nabla f(z)^\top r + \mathcal{O}(h^2) + i\mathcal{O}(h^3).$$

By taking the imaginary part of eq. (2.30) we obtain, $\Im(f(z + ihr)) = h\nabla f(z)^\top r + \mathcal{O}(h^3)$, a third-order (in $h$) accurate approximation of the quantity $h\nabla f(z)^\top r$. In the special
case $f : \mathbb{C} \to \mathbb{C}$, this idea produces a far more accurate approximation formula for the derivative of $f$ without the trade-off we saw in example [6].

**Example 7.** Let $f : \mathbb{C} \to \mathbb{C}$ be a holomorphic function. The quantity $\Im \left( \frac{f(x + ih)}{h} \right)$ is called the **complex step approximation** of the derivative $f'(x)$. In this case, the numerical error satisfies,

$$\Im \left( \frac{f(x + ih)(1 + \epsilon)}{h} \right) = \frac{1}{h} \Im (f(x + ih)(1 + \epsilon))$$

$$= \frac{1}{h} \Im \left( f(x) + ihf'(x) - \frac{h^2}{2} f''(x) + iO(h^3) + O(h^4)(1 + \epsilon) \right)$$

$$= \frac{(1 + \epsilon)}{h} (hf'(x) + O(h^3))$$

$$= (1 + \epsilon)(f'(x) + O(h^2)).$$

Taking $h \approx 10^{-16}$, we have that $\Im \left( \frac{f(x + ih)(1 + \epsilon)}{h} \right) = (1 + \epsilon)f'(x)$, where $|\epsilon| \leq \epsilon_{\text{mach}}$.

### 2.1.3 Human Errors

Another variety of error that will be discussed is human error. As noted by [Dahlquist and Björck, 2003], human errors can emerge from “clerical errors, errors in hand calculation, and misunderstandings. [... W]hen one uses computers, one can expect errors in the program itself, typing errors in entering the data, [and] operator errors.” Therefore, to the extent possible, it is desirable that numerical methods will be robust to errors of this kind.

**Example 8.** As just one example of a human error, we consider the following MATLAB code which divides a power of three successively until the value 3 is attained.

```matlab
x = 3^100;
while x ~= 3
    x = x / 3;
```

disp(x);
pause(0.01);
end

There is no mathematical misunderstanding in our implementation of this algorithm. Surely a power of three is reducible to three by iterated division. However, if one tried to execute this program on a computer, one would find that it would not return, and would, in fact, continue to divide so that the value of $x$ would approach zero. The human misunderstanding instead lies with an incomplete appreciation for the nuances of finite precision arithmetic, which becomes apparent when we show a few lines of output:

81.000000000000043
27.000000000000014
9.000000000000005
3.000000000000002
1.000000000000001
0.333333333333334

We see that the iterations of $x$ are not precisely represented as integers; therefore, the test that $x \approx 3$ will fail even though 3.000000000000002 is close to 3, producing an infinite loop. A more robust implementation might replace the test $x \approx 3$ by $\text{abs}(x - 3) > 2.3e-16$.  

2.2 Vector Fields, Mechanics, and Numerical Methods

A vector field in Euclidean space is a map from $\mathbb{R} \times \mathbb{R}^m$ to $\mathbb{R}^m$, where the first argument may be regarded as time and the second argument as position. In this way, the vector field may not only have spatial dependence but also dependence on a temporal quantity. Vector fields as typically discussed in terms of an initial value problem, are defined as follows.
Definition 2.2.1. Let $\Omega : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ be a smooth function. An initial value problem is a differential equation of the form

$$\frac{d}{dt} z_t = \Omega(t, z_t).$$

(2.35)

where $z_0 \in \mathbb{R}^m$ is given. In an initial value problem, $\Omega$ is called a vector field on $\mathbb{R}^m$, $z_0$ is called the initial value, and the function $z(\cdot) : \mathbb{R} \to \mathbb{R}^m$ is the solution to the initial value problem.

Intuitively, the initial value problem therefore describes how a particle, at fixed initial position $z_0$, moves in time.

The study of initial value problems is wide-spread, ranging in applications. For instance, initial value problems have been applied to the study of $n$-body gravitational dynamics [Dehnen and Read, 2011] and in molecular dynamics [Leimkuhler and Reich, 2005]. In either of these cases, the long-term behavior of the system can be of interest, so that one requires numerical methods which respect the long-term properties of the solution. Initial value problems are also foundational to diffeomorphisms on manifold structures, such as geodesics on Riemannian manifolds and the Darboux theorem on symplectic manifolds [Marsden and Ratiu, 2010]. Perhaps the most famous example of an initial value problem is Newton’s second law of motion, which states that force is equal to the product of mass and acceleration or $F(q_t) = M \ddot{q}_t$; we therefore see that the study of initial value problems can give insights into a broad class of physical systems, since many physical processes can be regarded as solutions to differential equations with prescribed initial conditions. To see how Newton’s second law may be expressed as an initial value problem, write $z_t = (q_t, \dot{q}_t)$ and set $\Omega(z_t) = (\dot{q}_t, M^{-1}F(q_t))$ where $F(q)$ is the force experienced at $q$ and $M$ is the mass of the object. One simply requires initial conditions $q_0$ and $\dot{q}_0$ to regard Newton’s second law as an initial value problem. This view has, over time, given rise to the Lagrangian and Hamiltonian formalisms for understanding classical
mechanics [Abraham and Marsden, 2008], which we shall discuss in greater detail shortly. Indeed, the Hamiltonian formalism has proven to be a rich foundation for the development of Markov chain methods for sampling posterior distributions [Betancourt, 2017; Neal, 2010b]; the requirement of these methods for specialized numerical integrators is a focus of the present work. The Lagrangian formalism has also received interest for sampling [Lan et al., 2015].

2.2.1 Vector Fields and Flow Maps

Definition 2.2.2. The flow map of an initial value problem with vector field \( \Omega \) and initial value \( z_0 \) is a map \( \Phi : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m \) such that \( \Phi(t, z_0) = z_t \), where \( z_t \) depends on \( \Omega \) through definition 2.2.1.

Vector fields \( \Omega : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m \) are called time-inhomogenous because they depend explicitly on the time state of the system. An important special case, which will suffice for many of the circumstances considered in this work, is the notion of a time-homogenous vector field, which is a constant function of time. In this case, we may write that \( \Omega : \mathbb{R}^m \rightarrow \mathbb{R}^m \) is a time-homogenous vector field. Restricting our attention to time-homogenous vector fields has the advantage of allowing us to simplify our notation and definitions. For instance:

Definition 2.2.3. Let \( \Omega : \mathbb{R}^m \rightarrow \mathbb{R}^m \) be a smooth function and let \( z_0 \in \mathbb{R}^m \). A time-homogenous initial value problem is a differential equation of the form,

\[
\frac{d}{dt} z_t = \Omega(z_t) \tag{2.36}
\]

\[
z_t \big|_{t=0} = z_0. \tag{2.37}
\]

In this case, \( \Omega \) is called a time-homogenous vector field. The time-homogenous flow map is the map \( \Phi(t, z_0) = z_t \).
It is always possible to reduce an arbitrary initial value problem into time-homogenous form by appending the time variable $t$ to the state of the system and adding the trivial differential equation $\dot{t} = 1$ to the vector field.

**Proposition 2.2.4** ([Marsden and Ratiu](2010)). Let $\Phi$ be a time-homogenous flow map corresponding to the time-homogenous vector field $\Omega$ as in definition 2.2.3. Fix $t \in \mathbb{R}$ and identify $\Phi_t : \mathbb{R}^m \to \mathbb{R}^m$ with the map $z \mapsto \Phi(t, z)$. The flow map obeys the flow property: $\Phi_{-t}^{-1} = \Phi_t$.

**Proof.** The flow property means $\Phi_{-t} \circ \Phi_t = \text{Id}$. We establish this as follows:

$$\Phi_t(z_0) = z_0 + \int_0^t \Omega(z_s) \, ds \quad (2.38)$$

$$\implies \Phi_{-t}(\Phi_t(z_0)) = \Phi_t(z_0) + \int_0^{-t} \Omega(z_{t+s}) \, ds \quad (2.39)$$

$$\implies \Phi_{-t}(\Phi_t(z_0)) = z_0 + \int_0^t \Omega(z_s) \, ds + \int_t^0 \Omega(z_y) \, dy \quad (2.40)$$

$$\implies \Phi_{-t}(\Phi_t(z_0)) = z_0 + \int_0^t \Omega(z_s) \, ds - \int_0^t \Omega(z_y) \, dy \quad (2.41)$$

$$\implies \Phi_{-t}(\Phi_t(z_0)) = z_0. \quad (2.42)$$

Example 9. Consider the vector field on $\mathbb{R}^2$

$$\Omega(t, (q, p)) = \begin{pmatrix} p \\ -\omega^2 q \end{pmatrix}, \quad (2.43)$$

where $\omega^2 \in \mathbb{R}_+$. Given the initial condition $(q_0, p_0) \in \mathbb{R}^2$, we have the initial value
One checks that the solution of the initial value problem is,

\[
\begin{pmatrix}
q_t \\
p_t
\end{pmatrix} = \begin{pmatrix}
q_0 \cos(\omega t) + \frac{p_0}{\omega} \sin(\omega t) \\
p_0 \cos(\omega t) - q_0 \omega \sin(t \omega)
\end{pmatrix}, \quad (2.45)
\]

and the flow map is

\[
\Phi(t, (q_0, p_0)) = \begin{pmatrix}
\cos(\omega t) & \frac{\sin(t \omega)}{\omega} \\
-\omega \sin(t \omega) & \cos(t \omega)
\end{pmatrix} \begin{pmatrix}
q_0 \\
p_0
\end{pmatrix}. \quad (2.46)
\]

These equations of motion have an important physical interpretation: Imagine a unit-mass weight attached to a spring whose distance from the top of the spring at time \( t \) is \( q_t \) and whose velocity (upward or downward) at time \( t \) is \( p_t \). The quantity \( \omega^2 \) is called the spring constant. The vector field considered in this problem is also called the harmonic oscillator.

Example 9 presents a circumstance wherein the solution and the flow map were available in closed-form for a simple initial value problem. This is a rare circumstance and, in general, analytical solutions to initial value problems will not be available. This has led to the development of numerical methods of integration for producing approximate solutions to initial value problems. Before turning our attention to such methods in the sequel, we first recall how probability distributions evolve under solutions to initial value problems.

**Definition 2.2.5.** Let \( \pi : \mathbb{R}^m \to \mathbb{R}_+ \) be a probability density on \( \mathbb{R}^m \) defined with respect to Lebesgue measure. Let \( \Phi_t : \mathbb{R}^m \to \mathbb{R}^m \) be the flow map of a time-homogenous vector
field $\Omega : \mathbb{R}^m \to \mathbb{R}^m$ as in definition 2.2.3. Let $\pi_t : \mathbb{R}^m \to \mathbb{R}_+$ be the density function of the random variable $\Phi_t(x)$ when $x \sim \pi$. We say that $\Phi_t$ obeys conservation of mass if, for all $A \subseteq \mathbb{R}^m$, we have

$$\int_{\Phi_t(A)} \pi_t(x) \, dx = \int_A \pi(x) \, dx. \quad (2.47)$$

**Corollary 2.2.6.** Differentiating both sides of eq. (2.47) with respect to time immediately yields the following characterization of conservation of mass:

$$\frac{d}{dt} \int_{\Phi_t(A)} \pi_t(x) \, dx = 0. \quad (2.48)$$

The following theorem gives a characterization of how $\pi_t$ must evolve in time. This is called *Reynold’s Transport Theorem*. We will give a proof of this theorem following the strategy of Mack [2021].

**Theorem 2.2.7.** Let $\pi : \mathbb{R}^m \to \mathbb{R}_+$ be a probability density on $\mathbb{R}^m$ defined with respect to Lebesgue measure. Let $\Phi_t : \mathbb{R}^m \to \mathbb{R}^m$ be the vector field flow of a time-homogenous vector field $\Omega : \mathbb{R}^m \to \mathbb{R}^m$ as in definition 2.2.3. Let $\pi_t : \mathbb{R}^m \to \mathbb{R}_+$ be the density function of the random variable $\Phi_t(x)$ when $x \sim \pi$. Then,

$$\frac{d}{dt} \int_{\Phi_t(A)} \pi_t(x) \, dx = \int_{\Phi_t(A)} \frac{\partial \pi_t}{\partial t}(x) + \text{div}(\pi_t(x) \cdot \Omega(x)) \, dx \quad (2.49)$$

We will require the following lemma.

**Lemma 2.2.8.** Let $\Phi_t : \mathbb{R}^m \to \mathbb{R}^m$ be the vector field flow of a time-homogenous vector field $\Omega : \mathbb{R}^m \to \mathbb{R}^m$. Then

$$\frac{d}{dt} \det(\nabla \Phi_t(x)) = \det(\nabla \Phi_t(x)) \cdot \text{div}(\Omega(\Phi_t(x))). \quad (2.50)$$
Proof.

\[
\frac{d}{dt} \det(\nabla \Phi_t(x)) = \det(\nabla \Phi_t(x)) \cdot \text{trace} \left( \left[ \nabla \Phi_t(x) \right]^{-1} \frac{d}{dt} \nabla \Phi_t(x) \right) \tag{2.51}
\]

\[
= \det(\nabla \Phi_t(x)) \cdot \text{trace} \left( \left[ \nabla \Phi_t(x) \right]^{-1} \nabla (\Omega(\Phi_t(x))) \right) \tag{2.52}
\]

\[
= \det(\nabla \Phi_t(x)) \cdot \text{trace} \left( \left[ \nabla \Phi_t(x) \right]^{-1} (\nabla \Omega)(\Phi_t(x)) \cdot \nabla \Phi_t(x) \right) \tag{2.53}
\]

\[
= \det(\nabla \Phi_t(x)) \cdot \text{trace} \left( \nabla \Phi_t(x) [\nabla \Phi_t(x)]^{-1} (\nabla \Omega)(\Phi_t(x)) \right) \tag{2.54}
\]

\[
= \det(\nabla \Phi_t(x)) \cdot \text{trace} ((\nabla \Omega)(\Phi_t(x))) \tag{2.55}
\]

\[
= \det(\nabla \Phi_t(x)) \cdot \text{div}(\Omega(\Phi_t(x))). \tag{2.56}
\]

We now prove Reynold’s Transport Theorem.

Proof of Theorem 2.2.7 By the change-of-variables formula, we have

\[
\int_{\Phi_t(A)} \pi_t(x) \, dx = \int_A \pi_t(\Phi_t(x)) \cdot \det(\nabla \Phi_t(x)) \, dx. \tag{2.57}
\]
Therefore,

$$\frac{d}{dt} \int_{\Phi_t(A)} \pi_t(x) \, dx (2.58)$$

$$= \frac{d}{dt} \int_A \pi_t(\Phi_t(x)) \cdot \det(\nabla \Phi_t(x)) \, dx (2.59)$$

$$= \int_A \frac{d}{dt} \left[ \pi_t(\Phi_t(x)) \cdot \det(\nabla \Phi_t(x)) \right] \, dx (2.60)$$

$$= \int_A \left[ \frac{d}{dt} \pi_t(\Phi_t(x)) \right] \cdot \det(\nabla \Phi_t(x)) + \left[ \pi_t(\Phi_t(x)) \cdot \frac{d}{dt} \det(\nabla \Phi_t(x)) \right] \, dx (2.61)$$

$$= \int_A \det(\nabla \Phi_t(x)) \cdot \left[ \frac{\partial \pi_t}{\partial t}(\Phi_t(x)) + \nabla \pi_t(\Phi_t(x)) \cdot \Omega(\Phi_t(x)) \right. \right. (2.62)$$

$$+ \left. \left. \pi_t(\Phi_t(x)) \cdot \operatorname{div}(\Omega(\Phi_t(x))) \right] \, dx$$

$$= \int_A \det(\nabla \Phi_t(x)) \cdot \left[ \frac{\partial \pi_t}{\partial t}(\Phi_t(x)) + \operatorname{div}(\pi_t(\Phi_t(x)) \cdot \Omega(\Phi_t(x))) \right] \, dx (2.63)$$

$$= \int_{\Phi_t(A)} \frac{\partial \pi_t}{\partial t}(x) + \operatorname{div}(\pi_t(x) \cdot \Omega(x)) \, dx (2.64)$$

where the last equality comes from applying the change-of-variables formula a second time.

\[ \square \]

**Corollary 2.2.9.** When \( \Phi_t \) obeys conservation of mass, the evolution of the density \( \pi_t \) obeys the continuity equation

$$\frac{\partial \pi_t}{\partial t}(x) = -\operatorname{div}(\pi_t(x) \cdot \Omega(x)) (2.65)$$

**Proof.** By setting

$$\frac{d}{dt} \int_{\Phi_t(A)} \pi_t(x) \, dx = 0 (2.66)$$

we have the immediate consequence from theorem [2.2.7] that

$$\int_{\Phi_t(A)} \frac{\partial \pi_t}{\partial t}(x) + \operatorname{div}(\pi_t(x) \cdot \Omega(x)) \, dx = 0. (2.67)$$
Since this is true for every set \( A \subseteq \mathbb{R}^m \), we conclude that the integrand itself must be zero for every \( x \in \mathbb{R}^m \). This yields the result.

### 2.2.2 Numerical Integrators

**Definition 2.2.10.** A single-step integrator \( \hat{\Phi} \) is a map from \( \mathbb{R} \times \mathbb{R} \times \mathbb{R}^m \) to \( \mathbb{R}^m \). The inputs to a single-step integrator are a step-size \( \epsilon \in \mathbb{R} \), a current time \( t \), and a current position \( z \in \mathbb{R}^m \) such that \( \hat{\Phi}(\epsilon, t, z) \in \mathbb{R}^m \). It is notationally convenient to “fix” the step-size \( \epsilon \) and current time \( t \) and to identify \( \hat{\Phi}_{\epsilon,t} : \mathbb{R}^m \to \mathbb{R}^m \) as the map \( z \mapsto \hat{\Phi}(\epsilon, t, z) \).

**Definition 2.2.11.** For \( k \in \mathbb{N} \), a \( k \)-step integrator is the concatenation of \( k \) single-step integrators with incremented current times:

\[
\hat{\Phi}^k(\epsilon, t, z) = \hat{\Phi}_{\epsilon,t+\epsilon} \circ \cdots \circ \hat{\Phi}_{\epsilon,t+\epsilon} \circ \hat{\Phi}_{\epsilon,t}(z). \tag{2.68}
\]

It is perhaps notable that the notion of a numerical integrator has not been defined with respect to any particular initial value problem that one may want to solve. Indeed, as we shall describe next, the suitability of a particular integrator for approximating the solution of a given initial value problem is a concept that should be quantified.

**Definition 2.2.12.** Let \( \Omega \) be a vector field and let \( z_0 \) be an initial condition. Let \( \Phi \) be the flow map of the corresponding initial value problem and let \( \hat{\Phi} \) be a single-step integrator. The integrator \( \hat{\Phi} \) is said to be \( p \)-th order accurate (for \( \Phi \)) if, for every initial condition \( z_0 \) and current time \( t_0 \), we have

\[
\Phi(t_0 + \epsilon, z_{t_0}) - \hat{\Phi}(\epsilon, t_0, z_{t_0}) = \mathcal{O}(\epsilon^{p+1}). \tag{2.69}
\]

The largest \( p \) such that eq. (2.69) holds is called the maximal order of the integrator.

The basic procedure for proving that numerical integrators are at least 1-st order is to expand them, and the flow map, in a Taylor series in \( \epsilon \). Then coefficients of the Taylor
series are compared to identify the maximal order of the method. Let’s examine this via an example using the simplest single-step integrator.

**Definition 2.2.13.** Let \( \Omega \) be a vector field and let \( z_{t_0} \) be an initial condition. **Euler’s method** is a single-step integrator defined as follows:

\[
\hat{\Phi}(\epsilon, t_0, z_{t_0}) = z_{t_0} + \epsilon \cdot \Omega(t_0, z_{t_0}).
\] (2.70)

**Proposition 2.2.14.** Let \( \Omega \) be a vector field and let \( z_{t_0} \) be an initial condition. Let \( \Phi \) be the flow map of the corresponding initial value problem. Euler’s method is at least 1-st order accurate for \( \Phi \).

*Proof.* First we expand the flow map in a Taylor series.

\[
\Phi(t_0 + \epsilon, z_{t_0}) = z_{t_0} + \epsilon \cdot \left( \left. \frac{d}{d\epsilon} \Phi(t_0 + \epsilon, z_{t_0}) \right|_{\epsilon=0} \right) + \mathcal{O}(\epsilon^2)
\] (2.71)

\[
= z_{t_0} + \epsilon \cdot \left( \left. \frac{d}{d\epsilon} z_{t_0+\epsilon} \right|_{\epsilon=0} \right) + \mathcal{O}(\epsilon^2)
\] (2.72)

\[
= z_{t_0} + \epsilon \cdot \Omega(t_0, z_{t_0}) + \mathcal{O}(\epsilon^2)
\] (2.73)

\[
= \hat{\Phi}(\epsilon, t_0, z_{t_0}) + \mathcal{O}(\epsilon^2).
\] (2.74)

Therefore,

\[
\Phi(t_0 + \epsilon, z_{t_0}) - \hat{\Phi}(\epsilon, t_0, z_{t_0}) = \mathcal{O}(\epsilon^2).
\] (2.75)

Hence, Euler’s method is 1-st order accurate. \( \Box \)

Another important integrator with interesting geometric properties is the implicit midpoint method, which is constructed in the following way.

**Definition 2.2.15.** Let \( \Omega \) be a vector field and let \( z_{t_0} \) be an initial condition. The **implicit**
The implicit midpoint differs in an important respect from Euler’s method. Namely, it is an \textit{implicitly-defined} procedure: a single step of the integrator is defined in terms of the result of the single step itself. This means that any implementation of the implicit midpoint method must know how to resolve these implicit relations; this is an important computational theme that appears in this work.

**Proposition 2.2.16.** Let $\Omega$ be a vector field and let $z_{t_0}$ be an initial condition. Let $\Phi$ be the flow map of the corresponding initial value problem. The implicit midpoint method is at least 2-nd order accurate for $\Phi$.

**Proof.** First we expand the flow map in a Taylor series to its second-order term in $\epsilon$. We have,

$$
\frac{d}{d\epsilon} \Phi(t_0 + \epsilon, z_{t_0}) = \Omega(t_0 + \epsilon, z_{t_0+\epsilon}) \tag{2.77}
$$

$$
\frac{d^2}{d\epsilon^2} \Phi(t_0 + \epsilon, z_{t_0}) = \frac{d}{d\epsilon} \Omega(t_0 + \epsilon, z_{t_0+\epsilon}) \tag{2.78}
$$

$$
= \frac{\partial}{\partial t} \Omega(t_0 + \epsilon, z_{t_0+\epsilon}) + \nabla_z \Omega(t_0 + \epsilon, z_{t_0+\epsilon}) \frac{\partial}{\partial \epsilon} z_{t_0+\epsilon} \tag{2.79}
$$

$$
= \frac{\partial}{\partial t} \Omega(t_0 + \epsilon, z_{t_0+\epsilon}) + \nabla_z \Omega(t_0 + \epsilon, z_{t_0+\epsilon}) \Omega(t_0 + \epsilon, z_{t_0+\epsilon}) \tag{2.80}
$$

Therefore,

$$
\Phi(t_0 + \epsilon, z_{t_0}) = z_{t_0} + \epsilon \cdot \Omega(t_0, z_{t_0}) + \frac{\epsilon^2}{2} \left( \frac{\partial}{\partial t} \Omega(t_0, z_{t_0}) + \nabla_z \Omega(t_0, z_{t_0}) \Omega(t_0, z_{t_0}) \right) + \mathcal{O}(\epsilon^3) \tag{2.81}
$$
Subsequently, we expand the implicit midpoint integrator in a Taylor series in $\epsilon$.

\[
\frac{d}{d\epsilon} \Phi(\epsilon, t_0, z_{t_0}) = \Omega \left( t_0 + \frac{\epsilon}{2}, z_{t_0} + \frac{\Phi(\epsilon, t_0, z_{t_0})}{2} \right) + \epsilon \cdot \left( \frac{d}{d\epsilon} \Omega \left( t_0 + \frac{\epsilon}{2}, z_{t_0} + \frac{\Phi(\epsilon, t_0, z_{t_0})}{2} \right) \right)
\]

(2.82)

and

\[
\frac{d^2}{d\epsilon^2} \Phi(\epsilon, t_0, z_{t_0})
= 2 \frac{d}{d\epsilon} \Omega \left( t_0 + \frac{\epsilon}{2}, z_{t_0} + \frac{\Phi(\epsilon, t_0, z_{t_0})}{2} \right) + O(\epsilon)
\]

(2.83)

\[
= \frac{\partial}{\partial t} \Omega \left( t_0 + \frac{\epsilon}{2}, z_{t_0} + \frac{\Phi(\epsilon, t_0, z_{t_0})}{2} \right) + \nabla_z \Omega \left( t_0 + \frac{\epsilon}{2}, z_{t_0} + \frac{\Phi(\epsilon, t_0, z_{t_0})}{2} \right) \frac{d}{d\epsilon} \Phi(\epsilon, t_0, z_{t_0}) + O(\epsilon)
\]

(2.84)

Therefore,

\[
\dot{\Phi}(\epsilon, t_0, z_{t_0}) = z_{t_0} + \epsilon \cdot \Omega(t_0, z_{t_0})
+ \frac{\epsilon^2}{2} \cdot \left( \frac{\partial}{\partial t} \Omega(t_0, z_{t_0}) + \nabla_z \Omega(t_0, z_{t_0}) \Omega(t_0, z_{t_0}) \right) + O(\epsilon^3)
\]

(2.85)

from which we immediately obtain $\Phi(t_0 + \epsilon, z_{t_0}) - \Phi(\epsilon, t_0, z_{t_0}) = O(\epsilon^3)$.

**Definition 2.2.17.** A single-step time-homogenous integrator $\dot{\Phi} : \mathbb{R} \to \mathbb{R}^m$. The inputs to a single-step integrator are a step-size $\epsilon \in \mathbb{R}$ and a current position $z \in \mathbb{R}^m$. As before, we may fix a step-size $\epsilon$ and identify $\dot{\Phi}_\epsilon : \mathbb{R}^m \to \mathbb{R}^m$ as the map $z \mapsto \dot{\Phi}(\epsilon, z)$. A $k$-step time-homogenous integrator is a concatenation of $k$ single-step time-homogenous integrators:

\[
\dot{\Phi}^k(\epsilon, z) = \dot{\Phi}_\epsilon \circ \cdots \circ \dot{\Phi}_\epsilon(z).
\]

(2.86)
Proposition 2.2.18. Let $\Omega : \mathbb{R}^m \rightarrow \mathbb{R}^m$ be a time-homogenous vector field. Let $\Xi : \mathbb{R}^m \rightarrow \mathbb{R}^m$ be a diffeomorphism. Let $z_t$ be the solution to the initial value problem $\dot{z}_t = \Omega(z_t)$ given $z_0$. Let $\Phi_\epsilon : \mathbb{R}^m \rightarrow \mathbb{R}^m$ be the flow map of $z_\epsilon$ and suppose that $\hat{\Phi}_\epsilon$ is a $p$-th order approximation of $\Phi_\epsilon$. Then $\Xi \circ \hat{\Phi}_\epsilon$ is a $p$-th order approximation of $\Xi \circ \Phi_\epsilon$.

Proof. Since $\hat{\Phi}_\epsilon$ is $p$-th order accurate for $\Phi_\epsilon$ we have, by Taylor series expansion, that,

$$\hat{\Phi}_0 = \Phi_0$$

(2.88)

$$\left. \frac{d^k}{d\epsilon^k} \Phi_\epsilon(z_0) \right|_{\epsilon=0} = \left. \frac{d^k}{d\epsilon^k} \Phi_\epsilon(z_0) \right|_{\epsilon=0},$$

(2.89)

for $k = 1, \ldots, p$. Let $f : \mathbb{R} \rightarrow \mathbb{R}^m$, then the Taylor series of expansion of $\Xi \circ f(\epsilon)$ is,

$$\Xi \circ f(\epsilon) = \Xi \circ f(0) + \sum_{k=1}^{p} \epsilon^k g^{[k]} \left( f(0), \frac{d}{d\epsilon} f(0), \ldots, \frac{d^k}{d\epsilon^k} f(0) \right) + \mathcal{O}(\epsilon^{p+1}),$$

(2.90)

where $g^{[k]}$ are functions determining the Taylor series coefficients which depend on the derivatives of $f$. For instance,

$$g^{[1]} \left( f(0), \frac{d}{d\epsilon} f(0) \right) = \nabla \Xi(f(0)) \cdot \frac{d}{d\epsilon} f(0)$$

(2.91)

$$g^{[2]} \left( f(0), \frac{d}{d\epsilon} f(0), \frac{d^2}{d\epsilon^2} f(0) \right) = \nabla^2 \Xi(f(0)) \left( \frac{d}{d\epsilon} f(0), \frac{d}{d\epsilon} f(0) \right) + \nabla \Xi(f(0)) \cdot \frac{d^2}{d\epsilon^2} f(0).$$

(2.92)

By the equality of the Taylor series expansion coefficients in eqs. (2.88) and (2.89) it follows from eq. (2.90) that,

$$\Xi \circ \Phi_\epsilon(z_0) - \Xi \circ \hat{\Phi}_\epsilon(z_0) = \mathcal{O}(\epsilon^{p+1}).$$

(2.93)

This proves that $\Xi \circ \hat{\Phi}_\epsilon$ is also $p$-th order accurate for $\Xi \circ \Phi_\epsilon$. □

An important concept in the design and analysis of numerical integrators is the defi-
tion of an adjoint method.

**Definition 2.2.19.** Let \( \Phi \) be a single-step time-homogenous integrator. The adjoint method is defined by \( \hat{\Phi}_\epsilon^* = \hat{\Phi}_{-\epsilon}^{-1} \). Maps that are their own adjoints (i.e. \( \hat{\Phi}_\epsilon = \hat{\Phi}_{-\epsilon}^{-1} \)) are called symmetric methods.

**Corollary 2.2.20.** Let \( \Phi \) be a time-homogenous flow map. Then \( \hat{\Phi} \) is symmetric.

*Proof.* This follows immediately from the flow property as stated in proposition 2.2.4.

**Lemma 2.2.21.** The concatenation of adjoint methods produces a symmetric method.

*Proof.* Define \( \Psi = \hat{\Phi}_\epsilon^* \circ \hat{\Phi}_\epsilon \). Then

\[
\Psi_\epsilon \circ \Psi_{-\epsilon} = \hat{\Phi}_\epsilon^* \circ \hat{\Phi}_\epsilon \circ \hat{\Phi}_{-\epsilon}^* \circ \hat{\Phi}_{-\epsilon} = \hat{\Phi}_\epsilon^* \circ \hat{\Phi}_{-\epsilon} = \text{Id.}
\]

**Proposition 2.2.22.** Let \( \Omega \) be a time-homogenous vector field. The time-homogenous implicit midpoint integrator defined by

\[
\hat{\Phi}(\epsilon, z) = z + \epsilon \cdot \Omega \left( \frac{z + \hat{\Phi}(\epsilon, z)}{2} \right)
\]

is symmetric.

*Proof.*

\[
\hat{\Phi}_{-\epsilon}(\hat{\Phi}_{\epsilon}(z)) = \hat{\Phi}_{\epsilon}(z) - \epsilon \cdot \Omega \left( \frac{\hat{\Phi}_{\epsilon}(z) + \hat{\Phi}_{-\epsilon}(\hat{\Phi}_{\epsilon}(z))}{2} \right)
\]

\[
= z + \epsilon \cdot \Omega \left( \frac{z + \hat{\Phi}(\epsilon, z)}{2} \right) - \epsilon \cdot \Omega \left( \frac{\hat{\Phi}_{\epsilon}(z) + \hat{\Phi}_{-\epsilon}(\hat{\Phi}_{\epsilon}(z))}{2} \right).
\]
By inspection, one checks that \( \hat{\Phi}_{-\epsilon}(\hat{\Phi}_\epsilon(z)) = z \) solves this implicit equation. This shows that the implicit midpoint method is symmetric.

**Theorem 2.2.23** (Leimkuhler and Reich [2005], Hairer et al. [2006]). Let \( \Phi \) be a time-homogenous flow map and let \( \hat{\Phi} \) be a symmetric single-step time-homogenous integrator. Suppose that \( \hat{\Phi} \) is \( p \)-th order accurate for \( \Phi \) and that this is the maximal order. Then \( p \) is even.

See page 86 in Leimkuhler and Reich [2005] for a proof. The importance of this theorem is that it proves that all symmetric numerical integrators that are at least 1-st order accurate are immediately at least 2-nd order accurate, since the maximal order must be even. As suggested by the proof of proposition 2.2.16 establishing first-order accuracy is often easier than establishing higher-order accuracy. However, if symmetry and first-order accuracy can be deduced, then second-order accuracy follows as an immediate consequence of theorem 2.2.23.

**Definition 2.2.24.** Let \( \Omega \) be a time-homogenous vector field and let \( z_0 \) be an initial condition. The *time-homogenous Euler method* is the single-step time-homogenous integrator defined by,

\[
\hat{\Phi}(\epsilon, z_0) = z_0 + \epsilon \cdot \Omega(z_0) \tag{2.100}
\]

The *time-homogenous implicit Euler method* is defined by,

\[
\hat{\Phi}(\epsilon, z_0) = z_0 + \epsilon \cdot \Omega(\hat{\Phi}(\epsilon, z_0)) \tag{2.101}
\]

**Proposition 2.2.25.** The adjoint of the time-homogenous Euler method is the time-homogenous implicit Euler method.

**Proof.** Fix \( \epsilon \in \mathbb{R} \) and denote the time-homogenous Euler method by \( \hat{\Phi}_\epsilon \) and the time-
homogenous implicit Euler method by $\tilde{\Phi}_\epsilon$. Then we have,

$$
\tilde{\Phi}_{-\epsilon}(\tilde{\Phi}_\epsilon(z_0)) = \tilde{\Phi}_\epsilon(z_0) - \epsilon \cdot \Omega(\tilde{\Phi}_{-\epsilon}(\tilde{\Phi}_\epsilon(z_0)))
$$

(2.102)

$$
= z_0 + \epsilon \cdot \Omega(z_0) - \epsilon \cdot \Omega(\tilde{\Phi}_{-\epsilon}(\tilde{\Phi}_\epsilon(z_0))).
$$

(2.103)

By inspection, one sees that this implicit equation is solved by $\tilde{\Phi}_{-\epsilon}(\tilde{\Phi}_\epsilon(z_0)) = z_0$, which shows the result.

### 2.2.3 Hamiltonian and Lagrangian Mechanics

An important class of vector fields are of the Hamiltonian form.

**Definition 2.2.26.** A matrix $J \in \mathbb{R}^{m \times m}$ is called **skew-symmetric** if $J^\top = -J$.

**Definition 2.2.27.** A **symplectic structure** on $\mathbb{R}^{2m}$ is an invertible, skew-symmetric matrix $J \in \mathbb{R}^{2m \times 2m}$. The negative inverse of $J$ is called the **Poisson structure**, which we denote by $B$; that is,

$$
B = -J^{-1}.
$$

(2.104)

**Lemma 2.2.28.** The Poisson structure is a skew-symmetric matrix.

**Proof.** Let $J$ be a skew-symmetric matrix satisfying. Then

$$
J = -J^\top
$$

(2.105)

$$
\implies \text{Id} = -J^\top \cdot J^{-1}
$$

(2.106)

$$
\implies (J^{-1})^\top = (J^{-1})^\top \cdot -J^\top \cdot J^{-1}
$$

(2.107)

$$
\implies (J^{-1})^\top = (J^\top)^{-1} \cdot -J^\top \cdot J^{-1}
$$

(2.108)

$$
\implies (J^{-1})^\top = -J^{-1}
$$

(2.109)

$$
\implies J^{-1} = -(J^{-1})^\top
$$

(2.110)
Therefore, $-J^{-1} = (J^{-1})^\top$ which is the same as $B = -B^\top$. \hfill \Box

**Definition 2.2.29.** Let $q, p \in \mathbb{R}^m$ and consider $z = (q, p) \in \mathbb{R}^{2m}$. Let $J$ be a symplectic structure on $\mathbb{R}^{2m}$ with corresponding Poisson structure $B$. A **Hamiltonian** is a smooth function $H : \mathbb{R}^{2m} \to \mathbb{R}$ with $z \mapsto H(z)$. When we wish to emphasize the separate roles of $q$ and $p$ we identify $H(z)$ and $H(q, p)$, where in the second case $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$. A **Hamiltonian vector field** is the time-homogenous vector field

$$\Omega(z) = B\nabla_z H(z), \quad (2.111)$$

or, when emphasizing the roles of $q$ and $p$,

$$\Omega((q, p)) = B \begin{pmatrix} \nabla_q H(q, p) \\ \nabla_p H(q, p) \end{pmatrix}. \quad (2.112)$$

This produces the time-homogenous initial value problem describing **Hamilton’s equations of motion**:

$$\frac{d}{dt} z_t = B\nabla_z H(z_t). \quad (2.113)$$

**Example 10.** Let $H$ be a Hamiltonian. The canonical symplectic structure is

$$J = \begin{pmatrix} 0_{m \times m} & \text{Id}_m \\ -\text{Id}_m & 0_{m \times m} \end{pmatrix}. \quad (2.114)$$
In this case $B = J$. Moreover, we recover the (canonical) Hamiltonian vector field,

$$\Omega((q, p)) = \begin{pmatrix} 0_{m \times m} & \text{Id}_m \\ -\text{Id}_m & 0_{m \times m} \end{pmatrix} \begin{pmatrix} \nabla_q H(q, p) \\ \nabla_p H(q, p) \end{pmatrix} \quad (2.115)$$

$$= \begin{pmatrix} \nabla_p H(q, p) \\ -\nabla_q H(q, p) \end{pmatrix}. \quad (2.116)$$

\[\|

**Proposition 2.2.30.** Hamilton’s equations of motion in eq. \(2.113\) possess the following properties: (i) the equations of motion conserve the Hamiltonian $\frac{d}{dt} H(z_t) = 0$ and (ii) the equations of motion are volume preserving in $z$-space since $\text{div}(\dot{z}_t) = 0$.

**Proof.** To show conservation of the Hamiltonian, we have:

$$\frac{d}{dt} H(z_t) = \nabla_z H(z_t)^\top B \nabla_z H(z_t) \quad (2.117)$$

$$= 0. \quad (2.118)$$

Let $\dot{z}_{t,i}$ denote the $i$-th element of the vector field. To show that the vector field is divergence free, we compute:

$$\sum_{i=1}^{2m} \frac{\partial}{\partial z_i} \dot{z}_{t,i} = \sum_{i=1}^{2m} \sum_{j=1}^{2m} B_{ij} \frac{\partial^2}{\partial z_i \partial z_j} H(z) \quad (2.119)$$

$$= \sum_{i=1}^{2m} \sum_{j=i+1}^{2m} B_{ij} \frac{\partial^2}{\partial z_i \partial z_j} H(z) + B_{ji} \frac{\partial^2}{\partial z_j \partial z_i} H(z) \quad (2.120)$$

$$= \sum_{i=1}^{2m} \sum_{j=i+1}^{2m} B_{ij} \frac{\partial^2}{\partial z_i \partial z_j} H(z) - B_{ij} \frac{\partial^2}{\partial z_i \partial z_j} H(z) \quad (2.121)$$

$$= 0, \quad (2.122)$$

since $B$ is a skew-symmetric matrix and there is symmetry of the second derivatives. \(\square\)
**Definition 2.2.31.** Let $\mathcal{J}$ be a symplectic structure on $\mathbb{R}^{2m}$. Let $\Psi : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ be a smooth map. We say that $\Psi$ is a *canonical* map if $\nabla \Psi(z)^\top \mathcal{J} \nabla \Psi(z) = \mathcal{J}$.

**Proposition 2.2.32.** Canonical maps are volume preserving.

*Proof.* By computing the determinant of both sides of the defining relation we obtain,

\[
\det(\nabla \Psi(z)) \cdot \det(\mathcal{J}) \cdot \det(\nabla \Psi(z)) = \det(\mathcal{J}) \quad (2.123)
\]

\[
\implies \det(\nabla \Psi(z))^2 = 1 \quad (2.124)
\]

\[
\implies |\det(\nabla \Psi(z))| = 1. \quad (2.125)
\]

This shows that $\Psi$ is a volume-preserving transformation. \qed

**Proposition 2.2.33.** The concatenation of canonical transformations is canonical.

*Proof.* It suffices to prove the result for the concatenation of two canonical transformations. Let $\Psi_1$ and $\Psi_2$ be two canonical transformations. Their concatenation is $\Psi = \Psi_1 \circ \Psi_2$. Then we have,

\[
(\nabla \Psi_1(\Psi_2(z)) \nabla \Psi_2(z))^\top \mathcal{J} \nabla \Psi_1(\Psi_2(z)) \nabla \Psi_2(z) = \mathcal{J} \quad (2.126)
\]

\[
= \nabla \Psi_2(z)^\top \mathcal{J} \nabla \Psi_1(\Psi_2(z)) = \mathcal{J} \quad (2.127)
\]

\[
= \nabla \Psi_2(z)^\top \mathcal{J} \nabla \Psi_2(z) \quad (2.128)
\]

\[
= \mathcal{J}. \quad (2.129)
\]

\qed

**Theorem 2.2.34** *(Hairer et al. [2006], Poincaré [1899]).* Let $\Phi$ be the flow map of a Hamiltonian vector field. For fixed $t \in \mathbb{R}$, the map $\Phi_t : \mathbb{R}^m \to \mathbb{R}^m$ is canonical.

*Proof.* For $t = 0$, $\Phi_0(z) = z$. Therefore $\nabla \Phi_0(z) = \text{Id}_{2m}$. Hence, $\nabla \Phi_0(z)^\top \mathcal{J} \nabla \Phi_0(z) = \mathcal{J}$. 

34
Now we will establish that

\[
\frac{d}{dt} (\nabla \Phi_t(z)^\top J \nabla \Phi_t(z))
\]

\[= \left( \frac{d}{dt} \nabla \Phi_t(z) \right)^\top J \nabla \Phi_t(z) + \nabla \Phi_t(z)^\top J \left( \frac{d}{dt} \nabla \Phi_t(z) \right) \]  

(2.130)

\[= (B \nabla^2 H(\Phi_t(z)) \nabla \Phi_t(z))^\top J \nabla \Phi_t(z) + \nabla \Phi_t(z)^\top J B \nabla^2 H(\Phi_t(z)) \nabla \Phi_t(z) \]  

(2.131)

\[= \nabla \Phi_t(z)^\top \nabla^2 H(\Phi_t(z)) \nabla \Phi_t(z) - \nabla \Phi_t(z)^\top \nabla^2 H(\Phi_t(z)) \nabla \Phi_t(z) \] + \nabla \Phi_t(z)^\top \nabla^2 H(\Phi_t(z)) \nabla \Phi_t(z) \]  

(2.132)

\[= \nabla \Phi_t(z)^\top \nabla^2 H(\Phi_t(z)) \nabla \Phi_t(z) - \nabla \Phi_t(z)^\top \nabla^2 H(\Phi_t(z)) \nabla \Phi_t(z) \]  

(2.133)

\[= \nabla \Phi_t(z)^\top \nabla^2 H(\Phi_t(z)) \nabla \Phi_t(z) - \nabla \Phi_t(z)^\top \nabla^2 H(\Phi_t(z)) \nabla \Phi_t(z) \]  

(2.134)

\[= 0. \]  

(2.135)

Therefore, for any \( t \in \mathbb{R} \), we obtain \( \nabla \Phi_t(z)^\top J \nabla \Phi_t(z) = J \). \( \square \)

**Definition 2.2.35.** A Hamiltonian \( H \) is called *separable* when it can be written as \( H(q, p) = U(q) + K(p) \) where \( U : \mathbb{R}^m \rightarrow \mathbb{R} \) and \( K : \mathbb{R}^m \rightarrow \mathbb{R} \) are smooth functions. When \( H \) cannot be written in this form, it is called *non-separable*.

**Example 11.** Let \( J \) be the canonical symplectic structure on \( \mathbb{R}^{2m} \). Let \( H \) be a separable Hamiltonian. While exact solutions to Hamilton’s equations of motion for separable Hamiltonians are unavailable in general, it is often possible to integrate “sub-Hamiltonians” exactly. Because \( H \) is separable, it can be written as \( H(q, p) = U(q) + K(p) \); define the functions \( H_{[1]}(q, p) = U(q) \) and \( H_{[2]}(q, p) = K(p) \), which we call sub-Hamiltonians since \( H(q, p) = H_{[1]}(q, p) + H_{[2]}(q, p) \). We can solve the equations of motion for the sub-Hamiltonians exactly. For instance, the equations of motion under \( H_{[1]} \) are,

\[
\begin{pmatrix}
\dot{q}_t \\
\dot{p}_t
\end{pmatrix} = \begin{pmatrix}
\nabla_p H_{[1]}(q_t, p_t) \\
\nabla_q H_{[1]}(q_t, p_t)
\end{pmatrix}
\]

(2.136)

\[
= \begin{pmatrix}
0_m \\
- \nabla_q U(q_t)
\end{pmatrix}.
\]  

(2.137)
Hence, given initial conditions \((q_0, p_0)\) we have \(q_t = q_0 + \int_0^t 0_m \, ds = q_0\) and \(p_t = p_0 - \int_0^t \nabla q U(q_s) \, ds = p_0 - \int_0^t \nabla q U(q_0) \, ds = p_0 - t \nabla q U(q_0)\). Similarly, the equations of motion for the sub-Hamiltonian \(H_2\) are,

\[
\begin{pmatrix}
\dot{q}_t \\
\dot{p}_t
\end{pmatrix} = \begin{pmatrix}
\nabla_p K(p_t) \\
0_m
\end{pmatrix}.
\] (2.138)

The solution to these equations of motion are \(q_t = q_0 + t K(p_0)\) and \(p_t = p_0\).

**Example 12.** Continuing example 11, let \(\Phi_{[1]}\) be the flow map of the time-homogenous Hamiltonian vector field corresponding to \(H_{[1]}\). That is, \(\Phi_{[1]}(\epsilon, (q, p)) = (q, p - \epsilon \nabla_q U(q))\).

Let \(\Phi_{[2]}\) be the analogous flow map corresponding to \(H_{[2]}\). We will show that the integrator \(\hat{\Phi}_\epsilon = \Phi_{[1], \epsilon} \circ \Phi_{[2], \epsilon}\) is first-order accurate for the flow map \(\Phi\) of the Hamiltonian vector field corresponding to \(H = H_{[1]} + H_{[2]}\). Expanding in Taylor series:

\[
\Phi_\epsilon((q, p)) = \begin{pmatrix}
q \\
p
\end{pmatrix} + \epsilon \begin{pmatrix}
\nabla_p K(p) \\
-\nabla_q U(q)
\end{pmatrix} + \mathcal{O}(\epsilon^2).
\] (2.139)

We can write the proposed integrator as,

\[
\hat{\Phi}_\epsilon((q, p)) = \begin{pmatrix}
q + \epsilon \nabla_p K(p) \\
p - \epsilon \nabla_q U(q + \epsilon \nabla_p K(p))
\end{pmatrix}
\] (2.140)

\[
= \begin{pmatrix}
q \\
p
\end{pmatrix} + \epsilon \nabla_p K(p) \left(-\epsilon \left(\nabla q U(q + \epsilon \nabla_p K(p)) + \frac{d}{d\epsilon} \nabla q U(q + \epsilon \nabla_p K(p)) \bigg|_{\epsilon=0}\right) + \mathcal{O}(\epsilon^2)\right)
\] (2.141)

\[
= \begin{pmatrix}
q \\
p
\end{pmatrix} + \epsilon \begin{pmatrix}
\nabla_p K(p) \\
-\nabla_q U(q)
\end{pmatrix} + \mathcal{O}(\epsilon^2).
\] (2.142)

This shows that \(\hat{\Phi}_\epsilon((q, p))\) is indeed first-order accurate.
We now discuss classical mechanics and their relationship to initial value problems and numerical integrators for those problems. We will begin by recalling Lagrange’s formulation of classical mechanics, prove their equivalence to Hamiltonian mechanics, and use the Lagrangian formalism in order to motivate and prove properties of numerical integrators. Our discussion follows the treatment of Hairer et al. [2006].

**Definition 2.2.36.** Let \( L : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be a smooth function and let \( q_t \), regarded as a function \( t \mapsto q_t \), be a smooth map. We define the action integral by

\[
S[q] = \int_{t_0}^{t_1} L(q_t, \dot{q}_t) \, dt
\]

where \( \dot{q}_t = \frac{d}{dt} q_t \).

The key idea of Lagrange is that physical motion extremizes the action integral, a concept that we now define.

**Definition 2.2.37.** Let \( q_0 = q_{t_0} \) and \( q_1 = q_{t_1} \) be given. A map \( q_t \) is said to extremize the action integral if, for every other map \( \delta q : \mathbb{R} \to \mathbb{R}^m \) satisfying \( \delta q_{t_0} = \delta q_{t_1} = 0 \) we have

\[
\left. \frac{d}{d\epsilon} S[q + \epsilon \delta q] \right|_{\epsilon=0} = 0.
\]

**Theorem 2.2.38.** Extremal maps satisfy the Euler-Lagrange equations

\[
\frac{\partial L}{\partial q}(q_t, \dot{q}_t) = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}}(q_t, \dot{q}_t).
\]
Proof. Define the function $y_t(\epsilon) = q_t + \epsilon \delta q_t$. We have,

$$
\frac{d}{d\epsilon} S[y(\epsilon)] = \frac{d}{d\epsilon} \int_{t_0}^{t_1} L(y_t(\epsilon), \dot{y}_t(\epsilon)) \, dt
$$

$$
= \int_{t_0}^{t_1} \frac{d}{d\epsilon} L(y_t(\epsilon), \dot{y}_t(\epsilon)) \, dt
$$

$$
= \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial q}(y_t(\epsilon), \dot{y}_t(\epsilon)) \delta q_t + \frac{\partial L}{\partial \dot{q}}(y_t(\epsilon), \dot{y}_t(\epsilon)) \frac{d}{dt} \delta q_t \right) \, dt
$$

$$
= \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial q}(y_t(\epsilon), \dot{y}_t(\epsilon)) \delta q_t - \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{q}}(y_t(\epsilon), \dot{y}_t(\epsilon)) \right) \delta q_t \right) \, dt
$$

where we have used the fact that $\delta q_{t_0} = \delta q_{t_1} = 0$ in eq. (2.149). Therefore,

$$
\frac{d}{d\epsilon} S[q + \epsilon \delta q] \bigg|_{\epsilon=0} = \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial q}(q_t, \dot{q}_t) \delta q_t - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}}(q_t, \dot{q}_t) \delta q_t \right) \, dt
$$

$$
= \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial q}(q_t, \dot{q}_t) - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}}(q_t, \dot{q}_t) \right) \delta q_t \, dt
$$

$$
= 0,
$$

since $q_t$ is extremal. Because $\delta q_t$ is arbitrary (except for the constraint $\delta q_{t_0} = \delta q_{t_1} = 0$), we conclude that the integrand must itself be zero; that is,

$$
\frac{\partial L}{\partial q}(q_t, \dot{q}_t) = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}}(q_t, \dot{q}_t).
$$

Let us now see that the Euler-Lagrange equations are in correspondence with Hamilton’s equations of motion (with the canonical symplectic structure) for an appropriately defined Hamiltonian.

Definition 2.2.39. Let $L : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be a smooth function. Define the momentum by the formula $p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q})$. In the sequel, we will assume that the map $(q, \dot{q}) \mapsto (q, p)$ is a diffeomorphism.
Theorem 2.2.40. Define the Hamiltonian \( H(q, p) = p^\top \dot{q}(q, p) - L(q, \dot{q}(q, p)) \). The Euler-Lagrange equations are equivalent to Hamilton’s equations of motion for the Hamiltonian \( H(q, p) \).

Proof. The time-derivative of position is,
\[
\frac{\partial H}{\partial p}(q, p) = \dot{q}(q, p) + p^\top \frac{\partial \dot{q}}{\partial p} - \frac{\partial L}{\partial \dot{q}}(q, \dot{q})^\top \frac{\partial \dot{q}}{\partial p} \tag{2.154}
\]
\[
= \dot{q}(q, p) + p^\top \frac{\partial \dot{q}}{\partial p} - p^\top \frac{\partial \dot{q}}{\partial p} \tag{2.155}
\]
\[
= \dot{q}(q, p). \tag{2.156}
\]

The time-derivative of momentum is,
\[
-\frac{\partial H}{\partial q}(q, p) = -p^\top \frac{\partial \dot{q}}{\partial q} + \frac{\partial L}{\partial q}(q, \dot{q}) + \frac{\partial L}{\partial \dot{q}}(q, \dot{q})^\top \frac{\partial \dot{q}}{\partial q} \tag{2.157}
\]
\[
= \frac{\partial L}{\partial q}(q, \dot{q}) \tag{2.158}
\]
\[
= \frac{d}{dt} \frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \tag{2.159}
\]
\[
= \frac{d}{dt} p. \tag{2.160}
\]

2.2.4 Numerical Integrators for Hamiltonian Mechanics

The preceding example demonstrates that numerical integrators can be constructed from the exact flows of sub-Hamiltonians. This procedure can be used to motivate one of the most important numerical integrators: the leapfrog method.

Definition 2.2.41. Let \( H \) be a separable Hamiltonian with \( H(q, p) = U(q) + K(p) \). Define the sub-Hamiltonians \( H_1(q, p) = U(q) \) and \( H_2(q, p) = \frac{1}{2} K(p) \). The Hamiltonian may then be expressed as \( H(q, p) = H_2(q, p) + H_1(q, p) + H_2(q, p) \). Let \( \Phi_1 \) and \( \Phi_2 \) be
the Hamiltonian flow maps of the time-homogenous Hamiltonian vector fields for $H_1$ and $H_2$ under the canonical symplectic structure, respectively. Then the leapfrog method is defined to be the composition $\hat{\Phi}_{\epsilon}((q, p)) = \Phi_{[2], \epsilon} \circ \Phi_{[1], \epsilon} \circ \Phi_{[2], \epsilon}$.

**Lemma 2.2.42.** The leapfrog integrator is at least first-order accurate.

**Proof.** We write the leapfrog method as,

$$\hat{\Phi}_{\epsilon}((q, p)) = \left( \begin{array}{c} q + \epsilon \nabla_p K \left( p - \frac{\epsilon}{2} \nabla_q U(q) \right) \\ p - \frac{\epsilon}{2} \nabla_q U(q) - \frac{\epsilon}{2} \nabla_q U \left( q + \epsilon \nabla_p K \left( p - \frac{\epsilon}{2} \nabla_q U(q) \right) \right) \end{array} \right). \quad (2.161)$$

We have,

$$q + \epsilon \nabla_p K \left( p - \frac{\epsilon}{2} \nabla_q U(q) \right) = q + \epsilon \left( \nabla_p K \left( p - \frac{\epsilon}{2} \nabla_q U(q) \right) \right) + \epsilon \frac{d}{d\epsilon} \nabla_p K \left( p - \frac{\epsilon}{2} \nabla_q U(q) \right) \bigg|_{\epsilon=0} + \mathcal{O}(\epsilon^2) \quad (2.162)$$

$$= q + \epsilon \nabla_p K (p) + \mathcal{O}(\epsilon^2), \quad (2.163)$$

and,

$$\frac{\epsilon}{2} \nabla_q U \left( q + \epsilon \nabla_p K \left( p - \frac{\epsilon}{2} \nabla_q U(q) \right) \right) = \epsilon \left( \frac{1}{2} \nabla_q U \left( q + \epsilon \nabla_p K \left( p - \frac{\epsilon}{2} \nabla_q U(q) \right) \right) \right) + \frac{\epsilon}{2} \frac{d}{d\epsilon} \nabla_q U \left( q + \epsilon \nabla_p K \left( p - \frac{\epsilon}{2} \nabla_q U(q) \right) \right) \bigg|_{\epsilon=0} + \mathcal{O}(\epsilon^2) \quad (2.164)$$

$$= \frac{\epsilon}{2} \nabla_q U (q) + \mathcal{O}(\epsilon^2). \quad (2.165)$$

Therefore,

$$p - \frac{\epsilon}{2} \nabla_q U(q) - \frac{\epsilon}{2} \nabla_q U \left( q + \epsilon \nabla_p K \left( p - \frac{\epsilon}{2} \nabla_q U(q) \right) \right) = p - \epsilon \nabla_q U(q) + \mathcal{O}(\epsilon^2). \quad (2.166)$$
This shows that the leapfrog method is at least first-order accurate.

Lemma 2.2.43. The leapfrog integrator is symmetric.

Proof. Recall that the leapfrog integrator is the composition of time-homogenous Hamiltonian flow maps: \( \hat{\Phi}_\epsilon = \Phi_{[2],\epsilon} \circ \Phi_{[1],\epsilon} \circ \Phi_{[2],\epsilon} \). From corollary 2.2.20, these individual flow maps are symmetric (i.e. \( \Phi^{-1}_{[1],-\epsilon} = \Phi_{[1],\epsilon} \)). Thus,

\[
\hat{\Phi}_{-\epsilon} \circ \hat{\Phi}_\epsilon = \Phi_{[2],-\epsilon} \circ \Phi_{[1],-\epsilon} \circ \Phi_{[2],-\epsilon} \circ \Phi_{[1],\epsilon} \circ \Phi_{[2],\epsilon} = \Phi_{[2],-\epsilon} \circ \Phi_{[1],-\epsilon} \circ \Phi_{[1],\epsilon} \circ \Phi_{[2],\epsilon} = \Phi_{[2],-\epsilon} \circ \Phi_{[2],\epsilon} = \text{Id}. \tag{2.170}
\]

Therefore, the leapfrog integrator is symmetric.

Corollary 2.2.44. The leapfrog integrator is at least second-order accurate.

Proof. Using the fact that the leapfrog integrator is at least first-order accurate and symmetric, second-order accuracy follows immediately from theorem 2.2.23.

Lemma 2.2.45. The leapfrog integrator is a canonical transformation.

Proof. The sub-Hamiltonian flows of the leapfrog integrator are canonical from theorem 2.2.34 and their concatenation, which forms the leapfrog integrator, is therefore canonical from proposition 2.2.33.

We now embark on a construction of a broad class of numerical integrators based on a single unifying principle called the discrete Lagrangian. We will begin by constructing a generating function for symplectic transformations, then continuing to develop Lagrangian mechanics via the action integral, and finally showing how approximations to the action integral will motivate numerical integrators. This will allow us to effortlessly prove that
many numerical integrators are symplectic. To understand this method, we first need to introduce the idea of a generating function.

**Theorem 2.2.46** (Hairer et al. [2006]). Let $S : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be a smooth function. Consider the map $(q, p) \mapsto (Q, P)$ that is implicitly defined by the equations

\begin{align*}
p_j &= -\frac{\partial S}{\partial q_j}(q, Q) \quad \text{for } j = 1, \ldots, m \\
P_j &= \frac{\partial S}{\partial Q_j}(q, Q) \quad \text{for } j = 1, \ldots, m.
\end{align*}

(2.174)

(2.175)

Then the map $(q, p) \mapsto (Q, P)$ is symplectic.

**Proof.** Denote by $Q_q$, $Q_p$, $P_q$, and $P_p$ the Jacobian matrices of $Q$ and $P$ with respect to $q$ and $p$, respectively. The condition for symplecticness is that

\[
\begin{pmatrix}
    Q_q^T & P_q^T \\
    Q_p^T & P_p^T
\end{pmatrix}
\begin{pmatrix}
    0 & \text{Id} \\
    -\text{Id} & 0
\end{pmatrix}
\begin{pmatrix}
    Q_q & Q_p \\
    P_q & P_p
\end{pmatrix}
= 
\begin{pmatrix}
    0 & \text{Id} \\
    -\text{Id} & 0
\end{pmatrix}
\]  

(2.176)

Upon multiplying out terms we find the conditions of symplecticness are:

\[
\begin{pmatrix}
    Q_q^T & P_q^T \\
    Q_p^T & P_p^T
\end{pmatrix}
\begin{pmatrix}
    P_q & P_p \\
    -Q_q & -Q_p
\end{pmatrix}
= 
\begin{pmatrix}
    Q_q^T P_q - P_q^T Q_q & Q_q^T P_p - P_q^T Q_p \\
    Q_p^T P_q - P_p^T Q_p & Q_p^T P_p - P_p^T Q_p
\end{pmatrix}
\]

(2.177)

(2.178)

We will now verify that these relations hold in the case of the given map.

\[
P_i = \frac{\partial S}{\partial Q_i}
\]

(2.179)

\[
\frac{\partial P_i}{\partial p_j} = \sum_{k=1}^{m} \frac{\partial^2 S}{\partial Q_k \partial Q_i} \frac{\partial Q_k}{\partial p_j}
\]

(2.180)
Therefore,

\[
(P_p^\top Q_p)_{ij} = \sum_{k=1}^{m} \frac{\partial P_k}{\partial p_i} \frac{\partial Q_k}{\partial p_j} = \sum_{k=1}^{m} \sum_{l=1}^{m} \frac{\partial^2 S}{\partial Q_l \partial Q_k} \frac{\partial Q_l}{\partial p_i} \frac{\partial Q_k}{\partial p_j} = \sum_{l=1}^{m} \frac{\partial Q_l}{\partial p_i} \sum_{k=1}^{m} \frac{\partial^2 S}{\partial Q_k \partial Q_l} \frac{\partial Q_k}{\partial p_j} = \sum_{l=1}^{m} \frac{\partial Q_l}{\partial p_i} \frac{\partial P_l}{\partial p_j} = (Q_p^\top P_p)_{ij}
\]  

(2.181)

Next we have,

\[
\frac{\partial P_i}{\partial q_j} = \frac{\partial}{\partial q_j} \frac{\partial S}{\partial Q_i} + \sum_{k=1}^{m} \frac{\partial^2 S}{\partial Q_k \partial Q_i} \frac{\partial Q_k}{\partial q_j}
\]  

(2.186)

Moreover, since

\[
p_j = -\frac{\partial S}{\partial q_j}
\]  

(2.187)

\[
\Rightarrow 0 = -\frac{\partial^2 S}{\partial q_i \partial q_j} - \sum_{k=1}^{m} \left( \frac{\partial}{\partial Q_k} \frac{\partial S}{\partial q_j} \right) \frac{\partial Q_k}{\partial q_i}
\]  

(2.188)

\[
\Rightarrow \sum_{k=1}^{m} \left( \frac{\partial}{\partial Q_k} \frac{\partial S}{\partial q_j} \right) \frac{\partial Q_k}{\partial q_i} = -\frac{\partial^2 S}{\partial q_i \partial q_j}
\]  

(2.189)
Therefore,

\[(Q_q^T P_q)_{ij} = \sum_{k=1}^{m} \frac{\partial Q_k}{\partial q_i} \frac{\partial P_k}{\partial q_j}\]  

(2.190)

\[= \sum_{k=1}^{m} \frac{\partial Q_k}{\partial q_i} \left( \frac{\partial}{\partial q_j} \frac{\partial S}{\partial Q_k} + \sum_{l=1}^{m} \frac{\partial^2 S}{\partial Q_l \partial Q_k} \frac{\partial Q_l}{\partial q_j} \right)\]  

(2.191)

\[= -\frac{\partial^2 S}{\partial q_i \partial q_j} + \sum_{l=1}^{m} \sum_{k=1}^{m} \frac{\partial Q_k}{\partial q_i} \frac{\partial^2 S}{\partial Q_l \partial Q_k} \frac{\partial Q_l}{\partial q_j}\]  

(2.192)

\[= \sum_{l=1}^{m} \frac{\partial Q_l}{\partial q_j} \left( \frac{\partial}{\partial q_i} \frac{\partial S}{\partial Q_l} + \sum_{k=1}^{m} \frac{\partial^2 S}{\partial Q_k \partial Q_l} \frac{\partial Q_k}{\partial q_i} \right)\]  

(2.193)

\[= \sum_{l=1}^{m} \frac{\partial Q_l}{\partial q_j} \frac{\partial P_l}{\partial q_i}\]  

(2.194)

\[= (P_q^T Q_q)_{ij}\]  

(2.195)

Finally, we have,

\[\frac{\partial P_i}{\partial p_j} = \sum_{k=1}^{m} \frac{\partial^2 S}{\partial Q_k \partial Q_k} \frac{\partial Q_k}{\partial p_j}\]  

(2.196)

Moreover,

\[p_j = -\frac{\partial S}{\partial q_j}\]  

(2.197)

\[\Longrightarrow \delta_{ij} = -\sum_{k=1}^{m} \left( \frac{\partial}{\partial q_k} \frac{\partial S}{\partial q_j} \right) \frac{\partial Q_k}{\partial p_i}\]  

(2.198)
Now,

\[
(Q_P^T P_q)_{ij} = \sum_{k=1}^{m} \frac{\partial Q_k}{\partial p_i} \frac{\partial P_k}{\partial q_j} = \sum_{k=1}^{m} \frac{\partial Q_k}{\partial p_i} \left( \frac{\partial}{\partial q_j} \frac{\partial S}{\partial Q_k} + \sum_{l=1}^{m} \frac{\partial^2 S}{\partial Q_l \partial Q_k} \frac{\partial Q_l}{\partial q_j} \right) \]

\[
= -\delta_{ij} + \sum_{k=1}^{m} \sum_{l=1}^{m} \frac{\partial^2 S}{\partial Q_l \partial Q_k} \frac{\partial Q_l}{\partial q_j} \frac{\partial Q_k}{\partial p_i} \]

\[
= -\delta_{ij} + \sum_{l=1}^{m} \frac{\partial P_l}{\partial p_i} \frac{\partial Q_l}{\partial q_j} \]

\[
= (P_p^T Q_q - \text{Id})_{ij}
\]

The importance of theorem \(2.2.46\) is that it provides a mechanism by which to “generate” symplectic maps given a smooth function \(S\). Therefore, the function \(S\) is called the \textit{generating function}. In fact, it is known that \textit{all} symplectic maps have a generating function [Hairer et al., 2006]. To understand the important role of generating functions in the development of numerical integrators, we return to the Lagrangian formalism and the action integral. Although the action integral is typically conceptualized as a functional, acting on maps \(t \mapsto q_t\) that connect two endpoints \(q_{t_0}\) and \(q_{t_1}\) at times \(t_0\) and \(t_1\), respectively, we now present a modified definition which we continue to call the action integral.

\textbf{Definition 2.2.47.} Let \(L : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}\) be a smooth function. We define the \textit{modified action integral} by

\[
S(q, Q) = \int_{t_0}^{t_1} L(y_t, \dot{y}_t) \, dt
\]

such that \(y : \mathbb{R} \to \mathbb{R}^m\) is extremal in the sense of definition \(2.2.37\) and \(y_{t_0} = q\) and \(y_{t_1} = Q\).
Unlike the original definition of the action integral, the modified action integral takes as inputs the endpoints and computes the action of an extremal map connecting the two. Consider differentiating $S$ with respect to its first argument.

\[
\frac{\partial S}{\partial q}(q, Q) = \int_{t_0}^{t_1} \frac{\partial L}{\partial y}(y_t, \dot{y}_t) \frac{\partial y_t}{\partial q} + \frac{\partial L}{\partial \dot{y}}(y_t, \dot{y}_t) \frac{\partial \dot{y}_t}{\partial q} \, dt
\]

(2.205)

\[
= \int_{t_0}^{t_1} \frac{\partial L}{\partial y}(y_t, \dot{y}_t) \frac{\partial y_t}{\partial q} + \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial y}(y_t, \dot{y}_t) - \frac{d}{dt} \frac{\partial L}{\partial \dot{y}}(y_t, \dot{y}_t) \right) \frac{\partial \dot{y}_t}{\partial q} \, dt
\]

(2.206)

\[
= \frac{\partial L}{\partial y}(y_{t_1}, \dot{y}_{t_1}) \frac{\partial y_{t_1}}{\partial q} - \frac{\partial L}{\partial y}(y_{t_0}, \dot{y}_{t_0}) \frac{\partial y_{t_0}}{\partial q}
\]

(2.207)

\[
= -\frac{\partial L}{\partial y}(y_{t_0}, \dot{y}_{t_0})
\]

(2.208)

\[
= -p_{t_0},
\]

(2.209)

by the definition of the momentum. Similarly, differentiating the modified action integral with respect to its second argument yields,

\[
\frac{\partial S}{\partial Q}(q, Q) = \int_{t_0}^{t_1} \frac{\partial L}{\partial y}(y_t, \dot{y}_t) \frac{\partial y_t}{\partial Q} + \frac{\partial L}{\partial \dot{y}}(y_t, \dot{y}_t) \frac{\partial \dot{y}_t}{\partial Q} \, dt
\]

(2.211)

\[
= \int_{t_0}^{t_1} \frac{\partial L}{\partial y}(y_t, \dot{y}_t) \frac{\partial y_t}{\partial Q} + \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial y}(y_t, \dot{y}_t) - \frac{d}{dt} \frac{\partial L}{\partial \dot{y}}(y_t, \dot{y}_t) \right) \frac{\partial \dot{y}_t}{\partial Q} \, dt
\]

(2.212)

\[
= \frac{\partial L}{\partial y}(y_{t_1}, \dot{y}_{t_1}) \frac{\partial y_{t_1}}{\partial Q} - \frac{\partial L}{\partial y}(y_{t_0}, \dot{y}_{t_0}) \frac{\partial y_{t_0}}{\partial Q}
\]

(2.213)

\[
= \frac{\partial L}{\partial y}(y_{t_1}, \dot{y}_{t_1}) - \frac{\partial L}{\partial y}(y_{t_0}, \dot{y}_{t_0})
\]

(2.214)

\[
= \frac{\partial L}{\partial \dot{y}}(y_{t_1}, \dot{y}_{t_1})
\]

(2.215)

\[
= p_{t_1}.
\]

(2.216)
Writing \( p = pt_0 \) and \( P = pt_1 \), we see that we have produced the following relations,

\[
p = -\frac{\partial S}{\partial q}(q, Q)
\]
\[\text{(2.217)}\]

\[
P = \frac{\partial S}{\partial Q}(q, Q)
\]
\[\text{(2.218)}\]

where \( S \) is the modified action integral. By theorem 2.2.46, the map \((q, p) \mapsto (Q, P)\) is symplectic.

A broad class of numerical integrators can now be motivated in the following way. As the modified action integral is an integral, it can be approximated by a finite sum. Suppose we have a collection of times \((t_0, t_0 + h, t_0 + 2h, \ldots, t_0 + (n-1)h, t_1)\). We seek to identify a corresponding collection of points \(q_n \in \mathbb{R}^m\) for \(n = 0, \ldots, n-1\) such that,

\[
\int_{t_0}^{t_1} L(q_t, \dot{q}_t) \, dt = \sum_{i=0}^{n-1} L_h(q_i, q_{i+1}),
\]
\[\text{(2.219)}\]

where \( L_h \) is a function approximating the action integral over the time period \((t_0 + ih, t_0 + (i+1)h)\). The function \( L_h \) is called the discrete Lagrangian. We now consider the following idea.

**Definition 2.2.48.** Given a smooth discrete Lagrangian \( L_h : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) which sends \((x, y) \mapsto L_h(x, y)\), we define the numerical integrator associated to the discrete Lagrangian by,

\[
p_n = -\frac{\partial L_h}{\partial x}(q_n, q_{n+1})
\]
\[\text{(2.220)}\]

\[
p_{n+1} = \frac{\partial L_h}{\partial y}(q_n, q_{n+1}).
\]
\[\text{(2.221)}\]

**Corollary 2.2.49.** The numerical integrator associated to the discrete Lagrangian \( L_h \) is a symplectic transformation.

**Proof.** This is an immediate consequence of theorem 2.2.46. \[\square\]
We now investigate some particular choices of discrete Lagrangian and show that it produces familiar integrators.

**Example 13.** Consider the choice of discrete Lagrangian

\[ L_h(q_n, q_{n+1}) = h L \left( q_n, \frac{q_{n+1} - q_n}{h} \right), \quad (2.222) \]

where \( L(q, \dot{q}) = \frac{1}{2} \dot{q}^\top G \dot{q} - U(q) \). The numerical integrator associated to the discrete Lagrangian \( L_h \) is,

\[ p_n = G \left( \frac{q_{n+1} - q_n}{h} \right) + h \nabla U(q_n) \quad (2.223) \]
\[ p_{n+1} = G \left( \frac{q_{n+1} - q_n}{h} \right). \quad (2.224) \]

Upon rearranging terms we find,

\[ q_{n+1} = q_n + h \left( G^{-1} p_n - h G^{-1} \nabla U(q_n) \right) \quad (2.225) \]
\[ = q_n + h G^{-1} p_{n+1} \quad (2.226) \]
\[ p_{n+1} = p_n - h \nabla U(q_n). \quad (2.227) \]

This is automatically a symplectic transformation since it is derived from the principle of the discrete Lagrangian. This is, however, not a symmetric numerical method.

**Example 14.** The following is a generalized version of an example in [Hairer et al. 2006]. Let \( \mathbb{J} \) denote the canonical symplectic structure. Given a Hamiltonian \( H(q, p) = U(q) + \frac{1}{2} p^\top G^{-1}(q)p \), the associated Hamiltonian vector field is,

\[ \dot{z}_t = \mathbb{J} \nabla H(z_t) \quad (2.228) \]

where we have used the identification of \( z = (q, p) \). The implicit midpoint integrator
applied to this Hamiltonian vector field is,

\begin{align*}
  z_{n+1} &= z_n + \epsilon \nabla H(z_{n+1/2}) \\
  z_{n+1/2} &= \frac{z_{n+1} + z_n}{2}.
\end{align*}

(2.229) \quad (2.230)

We may express this integrator in the following way, which was advocated by [Leimkuhler and Reich 2005].

\begin{align*}
  z_{n+1/2} &= z_n + \frac{\epsilon}{2} \nabla H(z_{n+1/2}) \\
  z_{n+1} &= z_{n+1/2} + \frac{\epsilon}{2} \nabla H(z_{n+1/2}).
\end{align*}

(2.231) \quad (2.232)

The Lagrangian associated to this Hamiltonian is \( L(q, \dot{q}) = \frac{1}{2} \dot{q}^\top G(q) \dot{q} - U(q) \). We consider the following discrete Lagrangian,

\begin{align*}
  L_h(q_n, q_{n+1}) &= h L \left( q_{n+1/2}, v_{n+1/2} \right) \\
  q_{n+1/2} &= \frac{q_{n+1} + q_n}{2} \\
  v_{n+1/2} &= \frac{q_{n+1} - q_n}{h}.
\end{align*}

(2.233) \quad (2.234) \quad (2.235)
This produces the following symplectic map:

\[
\begin{align*}
    p_n &= G(q_n + 1/2)v_n + 1/2 - \frac{h}{2}v_n^\top \nabla G(q_n + 1/2)v_n + \frac{h}{2} \nabla U(q_n + 1/2) \\
    &= p_{n+1/2} - \frac{h}{2} p_{n+1/2} G^{-1}(q_n + 1/2) \nabla G(q_n + 1/2) G^{-1}(q_n + 1/2)p_{n+1/2} + \frac{h}{2} \nabla U(q_n + 1/2) \\
    &= p_{n+1/2} = p_n + \frac{h}{2} p_{n+1/2} G^{-1}(q_n + 1/2) \nabla G(q_n + 1/2) G^{-1}(q_n + 1/2)p_{n+1/2} - \frac{h}{2} \nabla U(q_n + 1/2) \\
    &= p_n - \frac{h}{2} \nabla q H(q_n + 1/2; p_{n+1/2})
\end{align*}
\]

where we have employed the transformation \( p_{n+1/2} = G(q_n + 1/2)v_n + 1/2 \). Moreover, we have,

\[
\begin{align*}
    q_n + 1/2 &= q_n + \frac{h}{2} v_n + 1/2 \\
    &= q_n + \frac{h}{2} \left( \frac{q_n + 1/2 - q_n}{h} \right) \\
    &= q_n + \frac{q_n + 1}{2} \\
    &= \frac{q_n + q_{n+1}}{2}.
\end{align*}
\]
which shows \( q_{n+1/2} = q_n + \frac{h}{2} \nabla_p H(q_{n+1/2}, p_{n+1/2}) \). Finally,

\[
q_{n+1} = q_{n+1/2} + \frac{h}{2} v_{n+1/2}
\]

\[
= q_n + hv_{n+1/2}
\]

\[
= q_n + h \left( \frac{q_{n+1} - q_n}{h} \right)
\]

which shows that \( q_{n+1} = q_{n+1/2} + \frac{h}{2} \nabla_p H(q_{n+1/2}, p_{n+1/2}) \). These calculations show that the map \((q_n, p_n) \mapsto (q_{n+1}, p_{n+1})\) can be identified with the characterization of the implicit midpoint integrator given in eqs. (2.231) and (2.232). Therefore, the implicit midpoint is a symplectic map.

**Example 15.** Consider the discrete Lagrangian,

\[
L_h(q_n, q_{n+1}) = \frac{h}{2} L \left( q_n, \frac{q_{n+1} - q_n}{2} \right) + \frac{h}{2} L \left( q_{n+1}, \frac{q_{n+1} - q_n}{h} \right).
\]

Let the Lagrangian be of the form \( L(q, \dot{q}) = \frac{1}{2} \dot{q}^T G \dot{q} - U(q) \). The numerical integrator corresponding to this discrete Lagrangian is,

\[
v_{n+1/2} = \frac{q_{n+1} - q_n}{h}
\]

\[
p_{n+1/2} = G v_{n+1/2}
\]

\[
p_n = \frac{1}{2} \frac{\partial L}{\partial \dot{q}}(q_n, v_{n+1/2}) + \frac{1}{2} \frac{\partial L}{\partial \dot{q}}(q_{n+1}, v_{n+1/2}) - \frac{h}{2} \frac{\partial L}{\partial q}(q_n, v_{n+1/2})
\]

\[
= G v_{n+1/2} + \frac{h}{2} \nabla U(q_n)
\]

\[
p_{n+1/2} = p_n - \frac{h}{2} \nabla U(q_n)
\]

\[
p_{n+1} = \frac{1}{2} \frac{\partial L}{\partial \dot{q}}(q_n, v_{n+1/2}) + \frac{1}{2} \frac{\partial L}{\partial \dot{q}}(q_{n+1}, v_{n+1/2}) + \frac{h}{2} \frac{\partial L}{\partial q}(q_n, v_{n+1/2})
\]

\[
= G v_{n+1/2} - \frac{h}{2} \nabla U(q_n)
\]

Moreover, we have \( q_{n+1} = q_n + hv_{n+1/2} = q_n + hG^{-1} p_{n+1/2} \). Therefore, the sequence of
updates \((q_n, p_n) \mapsto (q_n, p_{n+1/2}) \mapsto (q_{n+1}, p_{n+1/2}) \mapsto (q_{n+1}, p_{n+1})\) can be identified as the leapfrog integrator applied to the Hamiltonian vector field for the Hamiltonian \(H(q, p) = U(q) + \frac{1}{2} p^\top G^{-1} p\). We had already established that the leapfrog integrator was symplectic since it is the composition of exact flows of sub-Hamiltonians, however this provides an alternative demonstration of the claim.

**Example 16.** We will now derive an important integrator for non-separable Hamiltonians using the principle of the discrete Lagrangian. This derivation follows [Hairer et al. 2006].

Define the discrete Lagrangian,

\[
L_h(q_0, q_1) = \frac{h}{2} \left( L(Q_1, \dot{Q}_1) + L(Q_2, \dot{Q}_2) \right)
\]

where \(Q_1 = q_0\) and \(Q_2 = q_0 + \frac{h}{2} \dot{Q}_1 + \frac{h}{2} \dot{Q}_2\) and \((\dot{Q}_1, \dot{Q}_2)\) are implicitly chosen to extremize the discrete Lagrangian under the constraint

\[
q_1 = q_0 + \frac{h}{2} \dot{Q}_1 + \frac{h}{2} \dot{Q}_2.
\]

We introduce the vector of Lagrange multipliers \(\lambda\) and obtain the constraint satisfaction condition by differentiating with respect to \((\dot{Q}_1, \dot{Q}_2)\):

\[
\frac{h}{2} \lambda = \frac{h}{2} \left[ \nabla_q L(Q_1, \dot{Q}_1) + \frac{h}{2} \nabla_q L(Q_2, \dot{Q}_2) \right]
\]

\[
\frac{h}{2} \lambda = \frac{h}{2} \left[ \frac{h}{2} \nabla_q L(Q_2, \dot{Q}_2) + \nabla_q L(Q_2, \dot{Q}_2) \right].
\]
Define the quantities:

\[
\dot{P}_1 = \nabla_q L(Q_1, \dot{Q}_1) \\
\mathbf{P}_1 = \nabla_{\dot{q}} L(Q_1, \dot{Q}_1) \\
\dot{P}_2 = \nabla_q L(Q_2, \dot{Q}_2) \\
\mathbf{P}_2 = \nabla_{\dot{q}} L(Q_2, \dot{Q}_2). 
\] (2.262) (2.263) (2.264) (2.265)

Additionally, for a Hamiltonian constructed in the usual way, \(\dot{P}_1 = -\nabla_q H(Q_1, \mathbf{P}_1)\), \(\dot{P}_2 = -\nabla_q H(Q_2, \mathbf{P}_2)\), \(\dot{Q}_1 = \nabla_p H(Q_1, \mathbf{P}_1)\), and \(\dot{Q}_2 = \nabla_p H(Q_2, \mathbf{P}_2)\). With this notation, the constraint satisfaction condition becomes,

\[
\frac{1}{2} \lambda = \frac{1}{2} P_1 + \frac{h}{4} \dot{P}_2, \\
\frac{1}{2} \lambda = \frac{1}{2} P_2 + \frac{h}{4} \dot{P}_2. 
\] (2.266) (2.267)

The propagation implied by the discrete Lagrangian can then be expressed as,

\[
p_0 = -\nabla_q L_h(q_0, q_1) \\
= -\frac{h}{2} \left[ \nabla_q L(Q_1, \dot{Q}_1) + \nabla_q L(Q_1, \dot{Q}_1) \nabla_{q_0} \dot{Q}_1 \\
+ \nabla_q L(Q_2, \dot{Q}_2)(\text{Id} + \frac{h}{2} \nabla_{q_0} \dot{Q}_1 + \frac{h}{2} \nabla_{q_0} \dot{Q}_2) + \nabla_q L(Q_2, \dot{Q}_2) \nabla_{q_0} \dot{Q}_2 \right] \\
= -\frac{h}{2} \left[ \dot{P}_1 + P_1 \nabla_{q_0} \dot{Q}_1 + \dot{P}_2(\text{Id} + \frac{h}{2} \nabla_{q_0} \dot{Q}_1 + \frac{h}{2} \nabla_{q_0} \dot{Q}_2) + P_2 \nabla_{q_0} \dot{Q}_2 \right] \\
= -\frac{h}{2} \left[ \dot{P}_1 + (\lambda - \frac{h}{2} \dot{P}_2) \nabla_{q_0} \dot{Q}_1 + \dot{P}_2(\text{Id} + \frac{h}{2} \nabla_{q_0} \dot{Q}_1 + \frac{h}{2} \nabla_{q_0} \dot{Q}_2) + (\lambda - \frac{h}{2} \dot{P}_2) \nabla_{q_0} \dot{Q}_2 \right] \\
= -\frac{h}{2} \left[ \dot{P}_1 + \lambda \nabla_{q_0} \dot{Q}_1 + \dot{P}_2 + \lambda \nabla_{q_0} \dot{Q}_2 \right] \\
= -h \left[ \frac{1}{2} \dot{P}_1 + \frac{1}{2} \dot{P}_2 \right] + \lambda
\] (2.268) (2.269) (2.270) (2.271) (2.272) (2.273)
One additionally computes $p_1 = \lambda$. Hence one obtains,

\[ P_1 = p_0 + \frac{h}{2} \dot{P}_1 \]  
(2.274)

\[ = p_0 - \frac{h}{2} \nabla_q H(q_0, P_1) \]  
(2.275)

\[ P_2 = P_1 \]  
(2.276)

\[ p_1 = p_0 + \frac{h}{2} \left[ \dot{P}_1 + \dot{P}_2 \right] \]  
(2.277)

\[ = p_0 - \frac{h}{2} \left[ \nabla_q H(q_0, P_1) + \nabla_q H(q_1, P_2) \right] \]  
(2.278)

\[ q_1 = q_0 + \frac{h}{2} \nabla_p H(q_0, P_1) + \frac{h}{2} \nabla_p H(q_2, P_2) \]  
(2.279)

\[ = q_0 + \frac{h}{2} \nabla_p H(q_0, P_1) + \frac{h}{2} \nabla_p H(q_1, P_1) \]  
(2.280)

These implicit equations describe the generalized leapfrog integrator for the Hamiltonian

\[ H(q, p). \]  

We will now clarify the relationship between how well the discrete Lagrangian approximates the action integral and the resulting order of the numerical method.

**Theorem 2.2.50** (Marsden and West [2001]). Suppose that

\[ L_h(q_0, q_1) = \int_0^h L(q_t, \dot{q}_t) \, dt + h^{r+1} e(q_0, q_1, h), \]  
(2.281)

where \( e(q_0, q_1, h) \) is a bounded function with bounded derivatives. Then the discrete Lagrangian map given in definition 2.2.48 when viewed as a map \( (q_n, p_n) \mapsto (q_{n+1}, p_{n+1}) \), is \( r \)-th order accurate for the Hamiltonian vector field with Hamiltonian \( H(q, p) = p^\top \dot{q}(q, p) - L(q, \dot{q}(q, p)) \).

We now move away from the principle of the discrete Lagrangian and consider quantities that are conserved under Hamilton’s equations of motion. We begin by identifying a distribution \( \pi_t^* \) for which the continuity equation yields \( \frac{\partial \pi_t^*}{\partial t} = 0 \); in this case, we call \( \pi_t^* \) the
stationary distribution. In the case of Hamiltonian vector fields, the stationary distribution is readily identified.

Example 17. Let \( J \) be the canonical symplectic structure and let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be a smooth function. We first observe that the Hamiltonian vector field is divergence free, which is equivalent to the vector field flow being volume preserving. We compute

\[
\text{div}((\dot{q}, \dot{p})) = \sum_{i=1}^{m} \frac{\partial}{\partial q_i} \frac{\partial H}{\partial p_i} (q, p) - \sum_{i=1}^{m} \frac{\partial}{\partial p_i} \frac{\partial H}{\partial q_i} (q, p) 
= \sum_{i=1}^{m} \frac{\partial^2 H}{\partial q_i \partial p_i} (q, p) - \sum_{i=1}^{m} \frac{\partial^2 H}{\partial p_i \partial q_i} (q, p) 
= 0.
\]

Moreover, if we guess that \( \pi^*_t(q, p) \propto \exp(-H(q, p)) \) is stationary (which assumes that the normalizing constant \( \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \exp(-H(q, p)) \, dq \, dp < +\infty \)) then we obtain,

\[
\nabla \pi^*_t(q, p) \cdot (\dot{q}, \dot{p}) \propto \begin{pmatrix}
\nabla_q \exp(-H(q, p)) \\
\nabla_p \exp(-H(q, p))
\end{pmatrix} \cdot \begin{pmatrix}
\nabla_p H(q, p) \\
-\nabla_q H(q, p)
\end{pmatrix}
\]

\[
= -\exp(-H(q, p)) \nabla_q H(q, p) \cdot \nabla_p H(q, p)
+ \exp(-H(q, p)) \nabla_p H(q, p) \cdot \nabla_q H(q, p)
= 0.
\]

Therefore, applying the continuity equation with the form of the stationary distribution we
have guessed leads to,

\[ \frac{\partial \pi^*_t}{\partial t}(x) = -\text{div}(\pi^*_t \cdot (\dot{q}, \dot{p})) \]  
\[ = -\nabla \pi^*_t(q,p) \cdot (\dot{q}, \dot{p}) - \text{div}((\dot{q}, \dot{p})) \]  
\[ = 0, \]  
(2.289)  
(2.290)  
(2.291)

which shows that the prescribed form of \( \pi^*_t \), if it exists, is stationary under the Hamiltonian vector field flow. It is perhaps worth noting that the function \( (q, p) \mapsto \exp(-H(q, p)) \) will be conserved under the Hamiltonian flow and the question of existence of the stationary probability distribution is simply a matter of integrability.

More generally, an argument very similar to the above shows that any function of the form \( (q, p) \mapsto g(H(q, p)) \) will be conserved by the Hamiltonian flow map since the Hamiltonian itself is conserved. Such conserved quantities are given special attention in the study of differential equations and are also related to properties of the numerical integrators themselves. We now make these ideas precise.

**Definition 2.2.51.** Let \( \Omega : \mathbb{R}^m \to \mathbb{R}^m \) be a time-homogenous vector field with vector field flow \( \Phi_t \). A **first integral** of \( \Phi_t \) is a function \( g : \mathbb{R}^m \to \mathbb{R} \) such that \( \frac{d}{dt}g(\Phi_t(x)) = 0 \).

**Example 18.** Let \( H : \mathbb{R}^{2m} \to \mathbb{R} \) be a smooth function and let \( \Omega(z) = J \nabla H(z) \) be the Hamiltonian vector field for \( H \) under the canonical symplectic structure with flow \( \Phi_t \). Then the Hamiltonian is a first integral since,

\[ \frac{d}{dt}H(\Phi_t(z)) = \nabla H(\Phi_t(z)) \cdot J \nabla H(\Phi_t(z)) \]  
\[ = 0. \]  
(2.292)  
(2.293)
A second example of a first integral is the Jacobian determinant of the flow:

\[
\frac{d}{dt} \det(\nabla \Phi_t(z)) = \det(\nabla \Phi_t(z)) \cdot \text{div}(\Omega(\Phi_t(z))) = \det(\nabla \Phi_t(z)) \cdot 0 = 0. \tag{2.294}
\]

Indeed, since \( \det(\nabla \Phi_0(z)) = 1 \), we conclude that the flow has unit Jacobian determinant for all time.

**Definition 2.2.52.** Let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be a smooth function and let \( \Omega : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m \) be the Hamiltonian vector field corresponding to \( H \) under the canonical symplectic structure. The **symplectic Euler-B map** is defined by

\[
p_{n+1} = p_n - \epsilon \nabla_q H(q_n, p_{n+1}) \tag{2.297}
\]
\[
q_{n+1} = q_n + \epsilon \nabla_p H(q_n, p_{n+1}). \tag{2.298}
\]

Similarly, the symplectic Euler-A map is defined by,

\[
p_{n+1} = p_n - \epsilon \nabla_q H(q_{n+1}, p_n) \tag{2.299}
\]
\[
q_{n+1} = q_n + \epsilon \nabla_p H(q_{n+1}, p_n). \tag{2.300}
\]

**Lemma 2.2.53.** The symplectic Euler-A and Euler-B maps are first-order numerical integrators.

**Proof.** For the symplectic Euler-B method, expanding in Taylor series we find,

\[
p_{n+1} = p_n - \epsilon \nabla_q H(q_n, p_n) + O(\epsilon^2) \tag{2.301}
\]
\[
q_{n+1} = q_n + \epsilon \nabla_p H(q_n, p_n) + O(\epsilon^2), \tag{2.302}
\]
which shows that the Euler-B method has first-order accuracy. The proof for the symplectic Euler-A method is nearly identical.

\[ \square \]

**Lemma 2.2.54.** The symplectic Euler-A method is the adjoint method of the symplectic Euler-B method.

**Proof.** Recall that the adjoint method is defined by the relation \( \hat{\Phi}^*_\epsilon = \hat{\Phi}^{-1} \). Given an initial position \((q, p)\), apply the Euler-A method with a negative step-size to obtain,

\[
\begin{align*}
p' &= p + \epsilon \nabla_q H(q', p) \quad (2.303) \\
q' &= q - \epsilon \nabla_p H(q', p). \quad (2.304)
\end{align*}
\]

Now apply the symplectic Euler-B method to \((q', p')\) to obtain,

\[
\begin{align*}
p'' &= p' - \epsilon \nabla_q H(q', p'') \quad (2.305) \\
&= p + \epsilon \nabla_q H(q', p) - \epsilon \nabla_q H(q', p'') \quad (2.306)
\end{align*}
\]

for which \( p'' = p \) is the solution. Moreover,

\[
\begin{align*}
q'' &= q' + \epsilon \nabla_p H(q', p'') \quad (2.307) \\
&= q - \epsilon \nabla_p H(q', p) + \epsilon \nabla_p H(q', p'') \quad (2.308) \\
&= q - \epsilon \nabla_p H(q', p) + \epsilon \nabla_p H(q', p) \quad (2.309) \\
&= q. \quad (2.310)
\end{align*}
\]

\[ \square \]

**Corollary 2.2.55.** The generalized leapfrog integrator (constructed in example 16) is the composition of the Euler-A method with the Euler-B method. Moreover, the generalized leapfrog integrator is second-order.
**Proof.** Denote the update computed by the Euler-B update by,

\[
\bar{p}_{n+1} = p_n - \frac{\epsilon}{2} \nabla_q H(q_n, \bar{p}_{n+1})
\]  
(2.311)

\[
\bar{q}_{n+1} = q_n + \frac{\epsilon}{2} \nabla_p H(q_n, \bar{p}_{n+1}).
\]  
(2.312)

Then the update due to the Euler-A update is,

\[
q_{n+1} = \bar{q}_n + \frac{\epsilon}{2} \nabla_p H(q_n, \bar{p}_{n+1})
\]  
(2.313)

\[
= q_n + \epsilon \nabla_p H(q_{n+1}, \bar{p}_n) + \epsilon \nabla_p H(q_n, \bar{p}_n)
\]  
(2.314)

\[
p_{n+1} = \bar{p}_n - \frac{\epsilon}{2} \nabla_q H(q_{n+1}, \bar{p}_n)
\]  
(2.315)

We therefore recognize the transformation \((q_n, p_n) \mapsto (q_{n+1}, p_{n+1})\) as the generalized leapfrog method. Since the concatenation of adjoint methods produces a symmetric method, second-order of the generalized leapfrog method follows immediately. 

\[\square\]

**2.2.5 Conserved Quantities of Numerical Integrators**

An interesting property is when conserved quantities of a vector field are also conserved by the numerical method. We present two results from [Leimkuhler and Reich](2005) in this direction.

**Proposition 2.2.56** ([Leimkuhler and Reich](2005)). Consider Hamilton’s equations of motion for a smooth Hamiltonian \(H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}\). Let \((q_t, p_t)\) be solutions of Hamilton’s equations of motion (i.e. \(\dot{q}_t = \nabla_p H(q_t, p_t)\) and \(\dot{p}_t = -\nabla_q H(q_t, p_t)\)). Let \(g : \mathbb{R}^m \times \mathbb{R}^m\) be a function of the form

\[
g(q, p) = q^\top A p + b^\top q + c^\top p
\]  
(2.316)

where \(A \in \mathbb{R}^{m \times m}\) is a symmetric matrix. Assume that \(g\) is a first integral of \((q_t, p_t)\)
as defined in definition \ref{2.2.51} (i.e. $\frac{d}{dt}g(q_t, p_t) = 0$). Then the symplectic Euler-A and symplectic Euler-B methods, defined in definition \ref{2.2.52} also conserve eq. (2.316).

**Proof.** If $g$ is a conserved quantity of the dynamics, then $\frac{d}{dt}g(q_t, p_t) = 0$. Applying the chain rule then yields,

\[
\nabla_q g(q_t, p_t) \dot{q}_t + \nabla_p g(q_t, p_t) \dot{p}_t + b^T \dot{q}_t + c^T \dot{p}_t = 0
\]

(2.317)

\[
\Rightarrow p_t^T A \nabla_p H(q_t, p_t) - q_t^T A \nabla_q H(q_t, p_t) + b^T \nabla_p H(q_t, p_t) - c^T \nabla_q H(q_t, p_t) = 0
\]

(2.318)

The point of the derivation in eq. (2.318) is that it implies that for any $(q, p) \in \mathbb{R}^m \times \mathbb{R}^m$ we must have

\[
p^T A \nabla_p H(q, p) - q^T A \nabla_q H(q, p) + b^T \nabla_p H(q, p) - c^T \nabla_q H(q, p) = 0.
\]

(2.319)

Thus, we see that a necessary condition for a function in the form of eq. (2.316) to be a conserved quantity of Hamilton’s equations of motion is that eq. (2.319) holds. Now
consider the Euler-B integrator (definition 2.2.52). We have

\[
g(q_{n+1}, p_{n+1}) = q_{n+1}^\top A p_{n+1} + b^\top q_{n+1} + c^\top p_{n+1}
\]
(2.320)

\[
= q_{n+1}^\top A p_{n+1} + b^\top q_n + e b^\top \nabla_q H + c^\top p_n - c^\top \nabla_p H
\]
(2.321)

\[
= q_n^\top A p_n + e \nabla_p H(q_n, p_{n+1})^\top A p_n - e q_n^\top A \nabla_q H(q_n, p_{n+1})
\]
\[+ e^2 \nabla_p H(q_n, p_{n+1})^\top A \nabla_q H(q_n, p_{n+1}) + b^\top q_n
\]
(2.322)

\[
= q_n^\top A p_n + e \nabla_p H(q_n, p_{n+1})^\top A p_{n+1} - e q_n^\top A \nabla_q H(q_n, p_{n+1})
\]
\[+ b^\top q_n + e b^\top \nabla_q H(q_n, p_{n+1}) + c^\top p_n - c^\top \nabla_p H(q_n, p_{n+1})
\]
(2.323)

\[
= q_n^\top A p_n + b^\top q_n + c^\top p_n
\]
(2.324)

\[
= g(q_n, p_n).
\]
(2.325)

This shows that the Euler-B method conserves first integrals of the prescribed form. The proof for the Euler-A method is similar. \qed

**Corollary 2.2.57.** The generalized leapfrog integrator (constructed in example 16) conserves first integrals of the form

\[
g(q, p) = q^\top A p + b^\top q + c^\top p.
\]

**Proof.** This follows immediately from the fact that the generalized leapfrog integrator is the composition of the Euler-A and Euler-B steps. \qed

**Proposition 2.2.58** (Leimkuhler and Reich [2005]). The implicit midpoint integrator (constructed in example 13) conserves first integrals of the form

\[
g(z) = \frac{1}{2} z^\top A z + b^\top z
\]
(2.327)

where \(A \in \mathbb{R}^{2m \times 2m}\) is a symmetric matrix.
Proof. Let $\Omega : \mathbb{R}^m \to \mathbb{R}^m$ denote the time-homogenous vector field for which the quadratic form in eq. (2.327) is a first integral. Observe that along solutions of the corresponding initial value problem we must have,

$$\frac{d}{dt} g(z_t) = (z_t^T A + b^T)\Omega(z_t).$$  \hfill (2.328)

Given $z_n$, let $z_{n+1} = z_n + \epsilon \nabla H(\bar{z})$ denote the update computed by the implicit midpoint method where $\bar{z} = \frac{z_{n+1} + z_n}{2}$. Combining eq. (2.328) with this equation defining the implicit midpoint update yields,

$$(\bar{z}^T A + b^T)z_{n+1} = (\bar{z}^T A + b^T)z_n + \epsilon (\bar{z}^T A + b^T)\Omega(\bar{z})$$ \hfill (2.329)

$$= (\bar{z}^T A + b^T)z_n.$$ \hfill (2.330)

By expanding eq. (2.330) we obtain,

$$\frac{1}{2}z_{n+1}^T A z_{n+1} + \frac{1}{2}z_n^T A z_n + b^T z_{n+1} = \frac{1}{2}z_{n+1}^T A z_n + \frac{1}{2}z_n^T A z_{n+1} + b^T z_n$$ \hfill (2.331)

$$\implies \frac{1}{2}z_{n+1}^T A z_{n+1} + b^T z_{n+1} = \frac{1}{2}z_{n+1}^T A z_n + b^T z_n.$$ \hfill (2.332)

which verifies the preservation of quadratic first integrals. \hfill \Box

It is interesting to observe that proposition 2.2.58 applies to vector fields in general, not merely those of the Hamiltonian type.

2.2.6 Higher-Order Numerical Methods

Thus far we have only seen examples of numerical integrators that are of first- or second-order accuracy. Integrators with second-order accuracy are standard among the numerical methods employed in Hamiltonian Monte Carlo and related methods, but higher-order methods are available, too. These techniques may be gainfully employed when it is nec-
necessary to obtain a precise simulation of a complex physical system. We now illustrate two methods for fourth order integration.

**Definition 2.2.59.** The three-stage Lobatto IIIA-IIIB integrator is a map \((q_n, p_n) \mapsto (q_{n+1}, p_{n+1})\) constructed as follows.

\[
Q_1 = q_n \\
Q_2 = q_n + \epsilon \left( \frac{5}{24} \nabla q H(Q_1, P_1) + \frac{1}{3} \nabla q H(Q_2, P_2) - \frac{1}{24} \nabla q H(Q_3, P_3) \right) \\
Q_3 = q_n + \epsilon \left( \frac{1}{6} \nabla q H(Q_1, P_1) + \frac{2}{3} \nabla q H(Q_2, P_2) + \frac{1}{6} \nabla q H(Q_3, P_3) \right) \\
P_1 = p_n - \epsilon \left( \frac{1}{6} \nabla q H(Q_1, P_1) - \frac{1}{6} \nabla q H(Q_2, P_2) \right) \\
P_2 = p_n - \epsilon \left( \frac{1}{6} \nabla q H(Q_1, P_1) + \frac{1}{3} \nabla q H(Q_2, P_2) \right) \\
P_3 = p_n - \epsilon \left( \frac{1}{6} \nabla q H(Q_1, P_1) + \frac{5}{6} \nabla q H(Q_2, P_2) \right) \\
q_{n+1} = Q_3 \\
p_{n+1} = p_n - \epsilon \left( \frac{1}{3} \nabla q H(Q_1, P_1) + \frac{2}{3} \nabla q H(Q_2, P_2) + \frac{1}{3} \nabla q H(Q_3, P_3) \right)
\]

The resulting integrator can appear very formidable. However, to gain some intuition for the Lobatto IIIA-IIIB integrator, we can express the generalized leapfrog integrator in
a similar manner. Indeed, the generalized leapfrog integrator can be expressed as

\begin{align}
Q_1 &= q_n \quad (2.341) \\
Q_2 &= q_n + \epsilon \left( \frac{1}{2} \nabla_p H(Q_1, P_1) + \frac{1}{2} \nabla_p H(Q_2, P_2) \right) \quad (2.342) \\
P_1 &= p_n - \epsilon \left( \frac{1}{2} \nabla_q H(Q_1, P_1) \right) \quad (2.343) \\
P_2 &= P_1 \quad (2.344) \\
q_{n+1} &= Q_2 \quad (2.345) \\
p_{n+1} &= p_n - \epsilon \left( \frac{1}{2} \nabla_q H(Q_1, P_1) + \frac{1}{2} \nabla_q H(Q_2, P_2) \right). \quad (2.346)
\end{align}

**Proposition 2.2.60** (Leimkuhler and Reich [2005]). The Lobatto IIIA-IIIB integrator is a fourth-order, symmetric, and symplectic map.

Higher-order numerical integrators may also be constructed from lower order integrators via composition; this is the subject of the following proposition.

**Proposition 2.2.61** (Hairer et al. [2006]). Let \( \hat{\Phi}_\epsilon \) be a single-step numerical method with \( p \)-th order accuracy. Let \( \gamma \in \mathbb{R}^s \) be a vector satisfying the two properties

1. \( \sum_{i=1}^s \gamma_i = 1 \)
2. \( \sum_{i=1}^s \gamma_i^{p+1} = 0. \)

Then the composition method \( \hat{\Psi}_\epsilon = \Phi_{\gamma_s \epsilon} \circ \cdots \circ \Phi_{\gamma_1 \epsilon} \) has \( (p + 1) \)-order.

**Example 19.** This is a modified version of an example in [Hairer et al. 2006] and references therein. Let \( \hat{\Phi}_\epsilon \) denote the generalized leapfrog integrator, which is symmetric and possesses second-order accuracy. Letting \( s = 3 \), and enforcing the constraint \( \gamma_1 = \gamma_3 \) so
that the resulting method is symmetric, it has been proposed to set

\[
\gamma_1 = \frac{1}{2 - \sqrt{2}}
\]

\[
\gamma_2 = \frac{\sqrt{2}}{2 - \sqrt{2}}.
\]

Therefore, by symmetry, the numerical method \( \hat{\Psi}_\epsilon = \hat{\Phi}_{\gamma_1 \epsilon} \circ \hat{\Phi}_{\gamma_2 \epsilon} \circ \hat{\Phi}_{\gamma_3 \epsilon} \) has fourth-order accuracy.

2.2.7 Linear Stability Analysis

Stability analysis can give an indication of how large of a step can be taken using a numerical integrator so that the long-term behavior of the numerical trajectory does not diverge. Stability of numerical integrators is typically characterized in terms of their behavior on the univariate harmonic oscillator, which is a time-homogenous vector field of the form,

\[
\frac{d}{dt} z_t = \begin{pmatrix}
0 & 1 \\
-\omega^2 & 0
\end{pmatrix} z_t,
\]

with initial condition \( z_0 \in \mathbb{R}^2 \). This is the same vector field that was previously described in example 9. Writing \( z = (q, p) \), as in that example, we saw that the exact flow map of the solution could be represented as,

\[
\Phi_t(q, p) = \begin{pmatrix}
\cos(t\omega) & \frac{\sin(t\omega)}{\omega} \\
-\omega \sin(t\omega) & \cos(t\omega)
\end{pmatrix}
\begin{pmatrix}
q \\
p
\end{pmatrix}.
\]

Definition 2.2.62. We say that a numerical integrator \( \hat{\Phi}_\epsilon : \mathbb{R}^2 \to \mathbb{R}^2 \) has \((\epsilon, \omega)\)-linear stability if (i) \( \hat{\Phi} \) is at least first-order accurate for the harmonic oscillator with parameter \( \omega \) and (ii) for any \( z \in \mathbb{R}^2 \) and \( k \in \mathbb{N} \), \( \hat{\Phi}_\epsilon^k(z) \) remains bounded.
Proposition 2.2.63. Suppose that $\hat{\Phi}_\epsilon$ is a linear transformation on $\mathbb{R}^2$ that is diagonalizable over $\mathbb{C}$. Then a sufficient condition for $\hat{\Phi}$ to have $(\epsilon, \omega)$-linear stability is that the eigenvalues of $\hat{\Phi}_\epsilon$ have magnitude inside the unit disk and that $\hat{\Phi}$ is at least first-order accurate for the harmonic oscillator with parameter $\omega$.

Proof. Since $\hat{\Phi}_\epsilon$ is a linear transformation that can be diagonalized over $\mathbb{C}$, it follows that we may express $\hat{\Phi}_\epsilon^k$ as,

$$\hat{\Phi}_\epsilon^k(z) = (Q\Lambda Q^{-1}) \cdots (Q\Lambda Q^{-1}) z\quad (2.351)$$

$$= Q\Lambda^k Q^{-1} z.\quad (2.352)$$

Provided $z \neq 0$, this quantity will remain bounded if and only if $\Lambda^k$ remains bounded. This occurs precisely when the eigenvalues of $\hat{\Phi}_\epsilon$ all lie in the unit disk.

The assumption that $\hat{\Phi}_\epsilon$ is a linear transformation on $\mathbb{R}^2$ and that it is diagonalizable is a reasonable one, since the analytical solution of the harmonic oscillator is of this form. Indeed, we have the following result.

Proposition 2.2.64. For any $\epsilon \in \mathbb{R}$ and $\omega \in \mathbb{R}_+$, the flow map $\Phi_\epsilon$ of the harmonic oscillator vector field has $(\epsilon, \omega)$-linear stability.

Proof. The eigenvalues of the matrix

$$\begin{pmatrix} \cos(\epsilon \omega) & \frac{\sin(\epsilon \omega)}{\omega} \\ -\omega \sin(\epsilon \omega) & \cos(\epsilon \omega) \end{pmatrix} \quad (2.353)$$

are $\lambda_1 = \exp(-i\epsilon \omega)$ and $\lambda_2 = \exp(i\epsilon \omega)$, which both lie on the boundary of the unit disk in $\mathbb{C}$. Hence the analytical flow map is stable for any choice of $\epsilon$ and $\omega$.

Many integrators, however, do not enjoy as strong stability properties as the flow map. Here are two examples from [Leimkuhler and Reich [2005]].
Example 20. The explicit Euler integrator, when applied to the harmonic oscillator, is a linear transformation of the form

\[
\hat{\Phi}_\epsilon = \begin{pmatrix}
1 & \epsilon \\
-\epsilon \omega^2 & 1
\end{pmatrix}.
\] (2.354)

The eigenvalues of this matrix are \( \lambda_1 = 1 + i \omega \epsilon \) and \( \lambda_2 = 1 - i \omega \epsilon \). Since both of these eigenvalues have modulus greater than one, we conclude that Euler’s method does not have \((\epsilon, \omega)\)-linear stability for any \( \epsilon \in \mathbb{R} \setminus \{0\} \) or \( \omega \in \mathbb{R}_+ \).

Example 21. The leapfrog integrator applied to the harmonic oscillator may be written as

\[
\hat{\Phi}_\epsilon = \begin{pmatrix}
1 - \frac{\epsilon^2 \omega^2}{2} & \epsilon \\
-\epsilon \omega^2 \left(1 - \frac{\epsilon^2 \omega^2}{4}\right) & 1 - \frac{\epsilon^2 \omega^2}{2}
\end{pmatrix}.
\] (2.355)

The eigenvalues of this matrix are

\[
\lambda_1 = 1 - \frac{\epsilon^2 \omega^2}{2} + \sqrt{\epsilon^2 \omega^2 \left(\epsilon^2 \omega^2 / 4 - 1\right)}
\] (2.356)

\[
\lambda_1 = 1 - \frac{\epsilon^2 \omega^2}{2} - \sqrt{\epsilon^2 \omega^2 \left(\epsilon^2 \omega^2 / 4 - 1\right)}.
\] (2.357)

These eigenvalues are complex and of unit modulus provided that \( \epsilon^2 \omega^2 < 4 \). Hence we conclude that the leapfrog integrator has \((\epsilon, \omega)\)-linear stability if \( \epsilon^2 \omega^2 < 4 \).

Given the previous two examples, therefore, it is an exceptional fact that the implicit midpoint integrator has the same linear stability properties as the flow map.

**Proposition 2.2.65.** For any \( \epsilon \in \mathbb{R} \) and \( \omega \in \mathbb{R}_+ \), the implicit midpoint integrator applied to the harmonic oscillator vector field has \((\epsilon, \omega)\)-linear stability.
Proof. In this case, the implicit midpoint integrator can be written as

\[
\hat{\Phi}_\epsilon = \left( \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} \right)^{-1} \left( \begin{pmatrix} 1 & \frac{\epsilon}{2} \\ -\omega^2 & 0 \end{pmatrix} \right) \left( \begin{pmatrix} 1 & \frac{\epsilon}{2} \\ -\omega^2 & 0 \end{pmatrix} \right)^{-1}
\] (2.358)

The eigenvalues of this matrix are \( \lambda_1 = -\frac{\epsilon \omega + 2i}{\epsilon - 2i} \) and \( \lambda_2 = -\frac{2 - \epsilon \omega i}{2 + \epsilon \omega i} \). In either case, the eigenvalues are of the form \( \pm(a + bi)/(a - bi) \), which have modulus one. Therefore, the implicit midpoint integrator has \((\epsilon, \omega)\)-linear stability for any choice of \( \omega \) and \( \epsilon \).

\[\Box\]

2.3 Numerics and Integration on Manifolds

The present section is quoted nearly verbatim from [Brofós and Lederman 2020a]; however it has been generalized to include a broader class of Hamiltonian.

2.3.1 Embedded Manifolds

In many cases, a manifold \( M \) can be embedded in a Euclidean space \( \mathbb{R}^m \) as the preimage of a constraint function \( g : \mathbb{R}^m \rightarrow \mathbb{R}^k \) on the level set where \( g \) takes the value zero; that is

\[ M \overset{\text{def.}}{=} \{ q \in \mathbb{R}^m : g(q) = 0 \} \].

We denote the Jacobian of \( g \) at the point \( q \in M \) by \( G(q) \). For any point \( q \in M, G(q) \) is a \( k \times m \) matrix. We assume that \( G(q) \) is full-rank for any \( q \in M \). In the case \( k \leq m \) we say that the dimension of \( M \) is \( k \). Many manifolds of interest may be written in this way such as the sphere, the special orthogonal group, the Stiefel manifold, and tori, among others. We define several important concepts related to embedded manifolds.

Definition 2.3.1. Let \( q \in M \). The tangent space at \( q \), denoted \( T_q M \), is the set of vectors satisfying,

\[ T_q M \overset{\text{def.}}{=} \{ \dot{q} \in \mathbb{R}^m : G(q)\dot{q} = 0 \} \]. (2.359)
where $G$ is the Jacobian of the constraint function $g$.

Motivated by the identity $\dot{q} = \nabla_p H(q, p)$ from classical mechanics on Euclidean spaces, we now apply this concept for manifolds.

**Definition 2.3.2.** Let $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be a smooth function. The *cotangent space* at $q \in M$, denoted $T^*_q M$, is the set of vectors,

$$T^*_q M \overset{\text{def.}}{=} \{ p \in \mathbb{R}^m : G(q) \nabla_p H(q, p) = 0 \}, \quad (2.360)$$

where $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is a smooth function.

The dependence of the cotangent space on the function $H$ is suppressed by convention. The tangent space is a vector space. When $G(q) \nabla_p H(q, p)$ is a linear function of $p$, the cotangent space is also a vector space.

**Definition 2.3.3.** The set of vectors,

$$T^* M \overset{\text{def.}}{=} \{(q, p) \in \mathbb{R}^m \times \mathbb{R}^m : q \in M \text{ and } p \in T^*_q M \} \quad (2.361)$$

is called the *cotangent bundle*.

**Definition 2.3.4.** The embedding of $T^* M$ in $\mathbb{R}^{2m}$ is defined to be the set of vectors

$$\{(q, p) \in \mathbb{R}^{2m} : g(q) = 0 \text{ and } G(q) \nabla_p H(q, p) = 0 \}. \quad (2.362)$$

**Definition 2.3.5 (Linear Maps between Tangent Spaces).** Let $M$ and $N$ be manifolds embedded in $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. Given $u \in T_q M \subset \mathbb{R}^m$, we define the linear map $T_q \Phi : T_q M \to T_{\Phi(q)} N$ as the map $(T_q \Phi)u \overset{\text{def.}}{=} \nabla \Phi(q) \cdot u$.

It is possible to give more abstract definitions for linear maps between tangent spaces to generalize beyond the case of embedded manifolds. However, we eschew this treatment for a simpler definition in definition 2.3.5.
**Definition 2.3.6 (Pullback).** Given a map $\Omega : \underbrace{T_qM \times \cdots \times T_qM}_{k \text{ times}} \rightarrow \mathbb{R}$, its pullback by a smooth function $\Phi : M \rightarrow M$ is the map $\Phi^*\Omega$ defined by $(\Phi^*\Omega)(v_1, \ldots, v_k) \overset{\text{def.}}{=} \Omega((T_q\Phi)v_1, \ldots, (T_q\Phi)v_k)$ where $v_1, \ldots, v_k \in T_qM$.

Formulating Hamiltonian mechanics on embedded manifolds gives a natural method to model physical systems in the presence of constraints. For instance, the motion of a pendulum which is attached by a rigid bar to some center can be viewed as a mechanical system constrained to the sphere. We now present several examples of implicitly defined manifolds on which constrained Hamiltonian dynamics may be constructed.

**Example 22.** Let $g(q) = q^\top q - 1$ for $Q \in \mathbb{R}^m$. This constraint corresponds to the surface of the $(m-1)$-dimensional sphere embedded in $\mathbb{R}^m$.

**Example 23.** Define the function $\text{mat} : \mathbb{R}^{m^2} \rightarrow \mathbb{R}^{m \times m}$ to be the function that assembles a matrix row-wise from a vector. Let $\text{vec}$ be the inverse operation of $\text{mat}$. Let $g : \mathbb{R}^{m^2} \rightarrow \mathbb{R}^{m^2}$ be defined as

$$g(q) = \text{vec}(\text{mat}(q)^\top \text{mat}(q) - \text{Id}). \quad (2.363)$$

The zero-set of this function defines a manifold that is isomorphic to the space of orthonormal matrices. In geometry, this manifold is called the orthogonal group. If we define an additional constraint and a *new* constraint function $\tilde{g} : \mathbb{R}^{m^2} \rightarrow \mathbb{R}^{m^2+1}$ according to

$$\tilde{g}(q) = \begin{pmatrix} g(q) \\ \det(\text{mat}(q)) - 1 \end{pmatrix}, \quad (2.364)$$

then the zero-set of this function defines the *special* orthogonal group, which corresponds to rotation matrices.

**Example 24.** The (compact) Stiefel manifold is an object that generalizes the previous two examples. Let $m, k \in \mathbb{N}$ and define the function $\text{mat} : \mathbb{R}^{nk} \rightarrow \mathbb{R}^{n \times k}$ which manufac-
tures a matrix row-wise from the vector input. Let \( \text{vec} : \mathbb{R}^{k \times k} \to \mathbb{R}^{k^2} \) denote row-wise “unraveling” of the input. Consider the constraint,

\[
g(q) = \text{vec}(\text{mat}(q)^\top \text{mat}(q) - \text{Id}_{k \times k}). \tag{2.365}
\]

This defines the Stiefel manifold: the manifold of orthonormal frames. The case \( n = k \) recovers the orthogonal group while \( k = 1 \) recovers the case of the sphere.

**Example 25.** An example we will employ later is the equality manifold which is implicitly defined as follows. Let \( q \in \mathbb{R}^{2m} \) be partitioned into two components as \( q = (q', q'') \) where \( q', q'' \in \mathbb{R}^m \). The equality manifold is defined by the constraint \( g(q) = q' - q'' \).

### 2.3.2 Differential Forms

Differential forms are an important topic in differential geometry. Nearly any book on differential geometry will contain a detailed discussion of these objects. For instance, Lee [2003], Marsden and Ratiu [2010], Abraham et al. [1988] all contain detailed sections on differential forms.

**Definition 2.3.7.** A matrix \( \mathbb{J} \) is skew-symmetric if \( \mathbb{J}^\top = -\mathbb{J} \). The set of skew-symmetric \( n \times n \) matrices is denoted \( \text{Skew}(n) \).

**Fact.** For a skew-symmetric matrix \( \mathbb{J} \in \text{Skew}(n) \) and a vector \( x \in \mathbb{R}^n \), \( x^\top \mathbb{J} x = 0 \).

**Definition 2.3.8.** The skew-symmetric, bilinear map \( \Omega : T_{(q,p)}T^*M \times T_{(q,p)}T^*M \to \mathbb{R} \) defined by \( \Omega(u, v) = u^\top \mathbb{J} v \) is called a symplectic structure on \( T^*M \) with matrix \( \mathbb{J} \).

**Definition 2.3.9 (Differential k-form (Page 129 of Marsden and Ratiu [2010])).** Let \( M \) be a manifold of dimension \( m \) and let \( q \in M \). A differential \( k \)-form \( \alpha \) on \( M \) \( (k \leq m) \) is a skew-symmetric linear map

\[
\alpha : T_qM \times \cdots \times T_qM \to \mathbb{R}. \tag{2.366}
\]
For a complete appreciation of our theoretical results, an understanding of 1-, 2-, and \( m \)-forms will be required. The most important 1-forms are the coordinate 1-forms.

**Definition 2.3.10.** Let \( M \) be a manifold of dimension \( m \) and let \( q \in M \) with \( q = (q_1, \ldots, q_m) \). The coordinate 1-forms are \( dq_i : T_qM \to \mathbb{R} \) defined by \( dq_i(v) = v_i \) where \( v = (v_1, \ldots, v_m) \in T_qM \).

The wedge product of differential forms is the principle tool by which differential forms are combined to give another differential form.

**Definition 2.3.11.** Let \( S^k \) denote the permutation group on \( k \) elements and let \( \sigma \in S^k \). Two indices \( 1 \leq i < j \leq k \) are said to be *inversions* if \( \sigma(i) > \sigma(j) \).

**Definition 2.3.12.** Let \( S^k \) denote the permutation group on \( k \) elements and let \( \sigma \in S^k \). The quantity \( \text{sign}(\sigma) : S^k \to \{-1, +1\} \) defined by \( \text{sign}(\sigma) = (-1)^r \) where \( r \) is the number of inversions in \( \sigma \).

**Definition 2.3.13.** Let \( \alpha \) be a differential \( k \)-form and \( \beta \) a differential \( l \)-form. The *wedge product* of \( \alpha \) and \( \beta \), denoted \( \alpha \wedge \beta \), is a \((k + l)\)-form defined by

\[
(\alpha \wedge \beta)(v_1, \ldots, v_{k+l}) \overset{\text{def.}}{=} \frac{1}{k!l!} \sum_{\sigma \in S^{k+l}} \text{sign}(\sigma) \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \beta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)})
\]

(2.367)

where \( S^{k+l} \) denotes the permutation group on \( k + l \) elements.

**Lemma 2.3.14.** Let \( dq_i \) and \( dq_j \) be coordinate 1-forms. Let \( u, v \in T_qM \) with \( u = (u_1, \ldots, u_m) \) and \( v = (v_1, \ldots, v_m) \). Then

\[
(dq_i \wedge dq_j)(u, v) = u_iv_j - u_jv_i
\]

(2.368)

**Proof.** Using definition [2.3.10] and the fact that the permutation group on two elements
has only two elements, direct computation yields,

\[(dq_i \wedge dq_j)(u, v) = \frac{1}{i!j!} dq_i(u) dq_j(v) - dq_i(v) dq_j(u) \tag{2.369}\]

\[= u_i v_j - v_i u_j \tag{2.370}\]

\[= u_i v_j - u_j v_i. \tag{2.371}\]

\[
\]

**Fact** (Page 131 in Marsden and Ratiu [2010]). In terms of the coordinate 1-forms, any differential \(k\)-form may be written as

\[
\alpha = \sum_{i_1 < \cdots < i_k} \alpha_{i_1,\ldots,i_k}(q) \ dq_{i_1} \wedge \cdots \wedge dq_{i_k} \tag{2.372}
\]

where \(\alpha_{i_1,\ldots,i_k} : M \to \mathbb{R}\) are smooth functions.

**Definition 2.3.15.** A differential \(k\)-form is called constant when, for all \(i_1 < \cdots < i_k\), the \(\alpha_{i_1,\ldots,i_k}\) in section 2.3.2 are all constant functions.

**Fact** (Wedge Product and Pullback (Page 131 in Marsden and Ratiu [2010])). Let \(\alpha\) be a differential \(k\)-form and \(\beta\) be a differential \(l\)-form on a manifold \(M\). Let \(\Phi : M \to M\) be a smooth function. Then,

\[
\Phi^*(\alpha \wedge \beta) = (\Phi^* \alpha) \wedge (\Phi^* \beta) \tag{2.373}
\]

where \(\Phi^*(\alpha \wedge \beta)\) denotes the pullback (definition 2.3.6) of \(\alpha \wedge \beta\) by \(\Phi\).

**Definition 2.3.16.** A differential \(k\)-form \(\alpha\) is said to be non-vanishing if for every \(q \in M\) there exists \(v_1, \ldots, v_k \in T_q M\) such that \(\alpha(v_1, \ldots, v_k) \neq 0\).

**Definition 2.3.17** (Page 139 in Marsden and Ratiu [2010]). Given a manifold \(M\) of dimension \(m\), a nowhere vanishing differential \(m\)-form on \(M\) is called a volume form.
Fact (Page 399 in Abraham et al. [1988]). The vector space of all constant $m$-forms on $\mathbb{R}^m$ is a vector space of dimension one.

Definition 2.3.18. Let $\Phi : M \to M$ be a smooth map and $V$ a volume form on $M$. Then $\Phi^*V$ is another $m$-form on $M$. The function $\det(\Phi) : M \to \mathbb{R}$ such that

$$\Phi^*V = \det(\Phi)V$$

(2.374)

is called the Jacobian determinant of $\Phi$ where $\Phi^*V$ denotes the pullback (definition 2.3.6) of $V$ by $\Phi$.

Fact (Page 140 in Marsden and Ratiu [2010]). A transformation $\Phi$ is volume preserving for $V$ if and only if $\det(\Phi) = 1$.

It will be convenient to work with vectors of differential 1-forms rather than individual 1-forms. The following definition extends the wedge product of differential 1-forms to vectors of differential 1-forms.

Definition 2.3.19. Let $\mathbf{d}a$ and $\mathbf{d}b$ be $m$-dimensional vectors of differential 1-forms. For instance $\mathbf{d}a = (d_{a1}, \ldots, d_{am})$. The wedge product of such vectors is defined by the relation

$$\mathbf{d}a \wedge \mathbf{d}b \overset{\text{def.}}{=} \sum_{i=1}^{m} d_{ai} \wedge d_{bi}$$

(2.375)

Fact (Page 64 in Leimkuhler and Reich [2005]). Let $\mathbf{d}a$, $\mathbf{d}b$, $\mathbf{d}c$ be $m$-dimensional vectors of differential 1-forms. For instance $\mathbf{d}a = (d_{a1}, \ldots, d_{am})$. The following are properties of the wedge product:

1. Skew-symmetry:

$$\mathbf{d}a \wedge \mathbf{d}b = -\mathbf{d}b \wedge \mathbf{d}a$$

(2.376)
2. Linearity:

\[ da \wedge (r \, db \wedge s \, dc) = r \, da \wedge db + s \, da \wedge dc \] (2.377)

for \( r, s \in \mathbb{R} \).

3. Matrix multiplication: For a matrix \( L \in \mathbb{R}^{m \times m} \),

\[ da \wedge L \, db = L^\top \, da \wedge db. \] (2.378)

4. Annihilation: When \( L \) is a symmetric matrix,

\[ da \wedge L \, da = 0. \] (2.379)

**Lemma 2.3.20.** Let \( q = (q_1, \ldots, q_m) \in M \) and \( p = (p_1, \ldots, p_m) \in T^*_qM \) (where \( T^*_qM \) is the cotangent space defined in definition 2.3.2) and set \( z = (q, p) \in T^*M \). A symplectic structure \( \Omega \) (see definition 2.3.8) with matrix \( J \in \text{Skew}(2m) \) may be written in terms of wedge products as

\[ \Omega = \sum_{i<j} J_{ij} \, dz_i \wedge dz_j \] (2.380)

\[ = \frac{1}{2} \, dz \wedge J \, dz \] (2.381)

**Proof.** Let \( u, v \in T_zT^*M \). The relation eq. (2.380) is standard and may be found in Marsden and Ratiu [2010] on page 147. To prove it, it suffices to use definition 2.3.10 and
Lemma 2.3.14 which yields

\[ \sum_{i<j} \mathbb{J}_{ij} \, dz_i \wedge dz_j(u, v) = \sum_{i<j} \mathbb{J}_{ij} u_i v_j - \mathbb{J}_{ij} u_j v_i \]  \hspace{1cm} (2.382)

\[ = \sum_{i<j} \mathbb{J}_{ij} u_i v_j + \mathbb{J}_{ji} u_j v_i \]  \hspace{1cm} (2.383)

\[ = \sum_{i=1}^{2m} \sum_{j=1}^{2m} \mathbb{J}_{ij} u_i v_j \]  \hspace{1cm} (2.384)

\[ = u^\top \mathbb{J} v \]  \hspace{1cm} (2.385)

\[ = \Omega(u, v) \]  \hspace{1cm} (2.386)

Equation (2.381) follows first from

\[ \sum_{i=1}^{2m} \sum_{j=1}^{2m} \mathbb{J}_{ij} d z_i \wedge d z_j = \sum_{i<j} \mathbb{J}_{ij} d z_i \wedge d z_j + \mathbb{J}_{ji} d z_j \wedge d z_i \]  \hspace{1cm} (2.387)

\[ = \sum_{i<j} \mathbb{J}_{ij} d z_i \wedge d z_j - \mathbb{J}_{ij} d z_j \wedge d z_i \]  \hspace{1cm} (2.388)

\[ = \sum_{i<j} \mathbb{J}_{ij} d z_i \wedge d z_j - d z_j \wedge \mathbb{J}_{ij} d z_i \]  \hspace{1cm} (2.389)

\[ = \sum_{i<j} \mathbb{J}_{ij} d z_i \wedge d z_j + \mathbb{J}_{ij} d z_i \wedge d z_j \]  \hspace{1cm} (2.390)

\[ = \sum_{i<j} 2 \mathbb{J}_{ij} d z_i \wedge d z_j \]  \hspace{1cm} (2.391)
and, using definition 2.3.19 from

\[
\sum_{i<j} \mathbb{J}_{ij} dz_i \wedge dz_j = \frac{1}{2} \sum_{i=1}^{2m} \sum_{j=1}^{2m} \mathbb{J}_{ij} dz_i \wedge dz_j = \frac{1}{2} \sum_{i=1}^{2m} \sum_{j=1}^{2m} dz_i \wedge \mathbb{J}_{ij} dz_j = \frac{1}{2} \sum_{i=1}^{2m} dz_i \wedge \sum_{j=1}^{2m} \mathbb{J}_{ij} dz_j = \frac{1}{2} dz \wedge \mathbb{J} dz
\] (2.392)

The symplectic structures we will consider are constant (see definition 2.3.15) since \( \mathbb{J}_{ij} \) does not depend on \( z \). An important volume form for Hamiltonian mechanics is the Liouville volume form, which is constructed from differential 2-forms.

**Definition 2.3.21** (Page 149 in Marsden and Ratiu [2010]). Let \( M \) be a manifold of dimension \( m \) and let \( \Omega \) be a symplectic 2-form on \( M \). The **Liouville volume form** on \( T^* M \) is the \( 2m \)-form defined by,

\[
\Lambda \overset{\text{def.}}{=} \frac{(-1)^{m(m-1)/2}}{m!} \Omega \wedge \cdots \wedge \Omega \quad \text{(there are } m \text{ copies of } \Omega \text{ in the wedge products)}.
\] (2.396)

When \( \Omega = \Omega_{\text{can}} = dq \wedge dp \), denote the Liouville volume form by \( \Lambda_{\text{can}} \). The Liouville volume form \( \Lambda_{\text{can}}(v_1, \ldots, v_{2m}) \) with \( v_i = (v_i^1, \ldots, v_i^{2m}) \) is proportional to the determinant of the matrix whose \((i, j)\) entry is \( v_i^j \), which, in turn, is the signed volume of parallelepiped spanned by the columns of that matrix.

**Definition 2.3.22.** Let \( \Phi : T^* M \to T^* M \) be a smooth mapping with a smooth inverse. Then \( \Phi \) is called a **diffeomorphism** of \( T^* M \).
Fact (Page 62 in Leimkuhler and Reich [2005]). Let $M$ be a manifold of dimension $m$ with $z \in T^*M$. Let $dz$ be the vector of coordinate 1-forms (see definition 2.3.10). Let $\Phi$ be a smooth function and let $\hat{z} = \Phi(z)$. Then the coordinate 1-forms of $\hat{z}$ are transformations of the coordinate 1-forms of $z$:

\[
d\hat{z}_i = \sum_{j=1}^{2m} \frac{\partial \hat{z}_i}{\partial z_j} dz_j = \sum_{j=1}^{2m} \frac{\partial \Phi(z)_i}{\partial z_j} dz_j
\]

(2.397)

Or, letting $dz = (dz_1, \ldots, dz_{2m})$,

\[
d\hat{z} = \nabla_z \Phi(z)^\top dz
\]

(2.399)

where $\nabla_z \Phi(z)^\top$ is the Jacobian of $\Phi$.

Lemma 2.3.23. Let $\Omega$ be a symplectic structure (definition 2.3.8) with matrix $J$. A map $\Phi : T^*M \to T^*M$ is symplectic with respect to $\Omega$ if and only if

\[
\frac{1}{2} d\hat{z} \wedge J d\hat{z} = \frac{1}{2} dz \wedge J dz
\]

(2.400)

where $d\hat{z} = \nabla_z \Phi(z)^\top dz$.

Proof. A symplectic transformation is one that preserves the symplectic structure under pullback. If $\Phi : T^*M \to T^*M$ then, from definition 2.3.6

\[
(\Phi^* \Omega)(u, v) \overset{\text{def}}{=} \Omega((T_z \Phi)u, (T_z \Phi)v) = \Omega(u, v) \iff \Phi \text{ is symplectic}
\]

(2.401)

for all $u, v \in T_z T^*M$. Letting $u = (u_1, \ldots, u_{2m})$ and $v = (v_1, \ldots, v_{2m})$, in terms of the
matrix $\mathbb{J}$, this is nothing but

\[(\nabla_z \Phi(z)^\top u)^\top \mathbb{J}(\nabla_z \Phi(z)^\top v) = u^\top \mathbb{J} v\]  \hfill (2.402)

or

\[\nabla_z \Phi(z)\mathbb{J} \nabla_z \Phi(z)^\top = \mathbb{J}.\]  \hfill (2.403)

We can now establish that if $\hat{z} = \Phi(z)$ then symplecticness of $\Phi$ is equivalent to conservation of the 2-form. From lemma 2.3.20 $\Omega$ can be written in terms of the wedge product as,

\[\Omega = \frac{1}{2} dz \wedge \mathbb{J} dz\]  \hfill (2.404)

Using section 2.3.2 under the change-of-variables $\hat{z} = \Phi(z)$, the symplectic structure changes to

\[\hat{\Omega} = \frac{1}{2} d\hat{z} \wedge \mathbb{J} dz\]  \hfill (2.405)

\[= \frac{1}{2} \nabla_z \Phi(z)^\top dz \wedge \mathbb{J} \nabla_z \Phi(z)^\top dz.\]  \hfill (2.406)

Using section 2.3.2

\[\hat{\Omega} = \frac{1}{2} dz \wedge \nabla_z \Phi(z)\mathbb{J} \nabla_z \Phi(z)^\top dz\]  \hfill (2.407)

Hence we see that $\hat{\Omega} = \Omega$ when $\nabla_z \Phi(z)\mathbb{J} \nabla_z \Phi(z)^\top = \mathbb{J}$, which conforms to the definition of symplecticness.

\[\square\]

**Lemma 2.3.24.** Let $\Phi(\cdot; t) : T^* M \to T^* M$ be a smooth function. Let $\hat{z}_t = \Phi(z; t)$ be a change-of-variables given $z \in T^* M$ such that $z = \Phi(z; 0)$. Let $\hat{\Omega}_t \overset{\text{def.}}{=} \frac{1}{2} d\hat{z}_t \wedge \mathbb{J} d\hat{z}_t$. Then
\( \Phi(\cdot; t) \) is symplectic with respect to \( \Omega = \frac{1}{2} dz \wedge J \, dz \) if \( \frac{d}{dt} \hat{\Omega}_t = 0 \).

**Proof.** By the fundamental theorem of calculus,

\[
\hat{\Omega}_t - \hat{\Omega}_0 = \int_0^t \frac{d}{ds} \hat{\Omega}_s \, ds. 
\]  
(2.408)

If \( \frac{d}{dt} \hat{\Omega}_t = 0 \) then

\[
\hat{\Omega}_t = \hat{\Omega}_0. 
\]  
(2.409)

Since \( \hat{z}_0 = z, \hat{\Omega}_t = \Omega \). The map \( \Phi(\cdot; t) \) is symplectic by lemma 2.3.23.

**Definition 2.3.25.** A **magnetic symplectic structure** is a symplectic structure assuming the form,

\[
J_{\text{mag}} = \begin{pmatrix} L & \text{Id}_m \\ -\text{Id}_m & 0_m \end{pmatrix} \in \text{Skew}(2m) 
\]  
(2.410)

for some skew-symmetric matrix \( L \in \mathbb{R}^{m \times m} \).

Applying definition 2.3.19 and lemma 2.3.20, the symplectic structure in eq. (2.410) can be expressed as

\[
\sum_{i<j} \mathbb{J}_{ij} dz_i \wedge dz_j = \sum_{i=1}^n dq_i \wedge dp_i + \sum_{i<j} L_{ij} dq_i \wedge dq_j 
\]  
(2.411)

\[
= dq \wedge dp + \frac{1}{2} dq \wedge L \, dq. 
\]  
(2.412)

### 2.3.3 Manifold Hamiltonian Mechanics

Hamiltonian mechanics are classically formulated as differential equations on the cotangent bundle of a smooth manifold \( M \). Formally, given a manifold \( M \), Hamiltonian mechanics give the time evolution of a point \( (q, p) \in T^*M \), often called phase-space in
physics wherein \( p \) is called the momentum. Recall that \( T^*M \) is an embedded manifold from definition 2.3.4.

Our construction of Hamiltonian mechanics requires the specification of an object called the symplectic structure which are reviewed in the case of embedded manifolds in section 2.3.2. One formulation of the symplectic structure uses a matrix associated with it. Let \( J \in \text{Skew}(2m) \) be an invertible, skew-symmetric matrix. Let \( u, v \in T_{(q,p)}T^*M \) (i.e., two vectors, each in the tangent space to the cotangent space \( T^*M \), which is a manifold) with \( u = (u_1, \ldots, u_{2m}) \) and \( v = (v_1, \ldots, v_{2m}) \).

**Definition 2.3.26.** A map \( \Phi : T^*M \to T^*M \) is symplectic if \( \Phi^*\Omega = \Omega \), where \( \Phi^*\Omega \) is the pullback (definition 2.3.6) of \( \Omega \) by \( \Phi \).

Given a symplectic structure on \( T^*M \), we provide a definition of Hamilton’s equations of motion.

**Definition 2.3.27.** Let \( \Omega \) be a symplectic structure on \( T^*M \) and let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be a smooth function; \( H \) is called the Hamiltonian. Let \( (q, p) \in T^*M \) and let \( T_{(q,p)}T^*M \) be the tangent space of \( T^*M \) at \( (q, p) \). The Hamiltonian vector field \( X_H : T^*M \to TT^*M \) satisfies \( \Omega(X_H(q, p), \delta) = (T_{(q,p)}H)\delta \) for all \( \delta \in T_{(q,p)}T^*M \) where \( T_{(q,p)}H \) is the linear map between tangent spaces defined in definition 2.3.5.

**Definition 2.3.28.** The flow of a Hamiltonian vector field \( X_H : T^*M \to TT^*M \) to time \( t \) is the map \( \Phi(\cdot, \cdot; t) : T^*M \to T^*M \) satisfying \( \frac{d}{dt}\Phi(q, p; t) = X_H(\Phi(q, p; t)) \) and \( \Phi(q, p; 0) = (q, p) \) for \( (q, p) \in T^*M \).

**Definition 2.3.29.** Suppose \( (q_t, p_t) = \Phi(q, p; t) \). Since \( \frac{d}{dt}\Phi(q, p; t) = X_H(\Phi(q, p; t)) \), we have derived the equations of motion \( (\dot{q}_t, \dot{p}_t) = X_H((q_t, p_t)) \).

**Fact (Page 209 in Lee [2003]).** Let \( \Phi(\cdot; t) \) be a Hamiltonian vector field flow (definition 2.3.28) to time \( t \). Hamiltonian vector field flows satisfy the **flow property**:

\[
\Phi(\Phi(q, p; t); -t) = (q, p) \quad (2.413)
\]
or, equivalently,

\[ \Phi(\cdot; -t) \circ \Phi(\cdot; t) = \text{Id} \quad (2.414) \]

**Fact** (Page 185 in [Marsden and Ratiu 2010]). Let \( \Phi(\cdot; t) : T^*M \to T^*M \) be the vector field flow (see definition 2.3.28) to time \( t \) of a Hamiltonian vector field \( X_H \) (see definition 2.3.27). Then \( \Phi(\cdot; t) \) is symplectic for every \( t \).

**Fact** (Page 72 in [Marsden and Ratiu 2010]). Let \( \Omega \) be a symplectic 2-form on \( T^*M \). The collection of all maps \( \Phi : T^*M \to T^*M \) such that \( \Phi^*\Omega = \Omega \) forms a group under function composition.

**Definition 2.3.30 (Embedded Cotangent Space).** Let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be a smooth function. Let \( M \) be manifold that can be embedded in \( \mathbb{R}^m \) as the preimage of the zero level set of a constraint function \( g : \mathbb{R}^m \to \mathbb{R}^k \); that is, let \( M = \{ q \in \mathbb{R}^m : g(q) = 0 \} \). To view \( T^*M \) as an embedded sub-manifold of \( \mathbb{R}^{2m} \) means that \( T^*M \) should be identified with the set

\[ \{ (q, p) \in \mathbb{R}^{2m} : g(q) = 0 \text{ and } G(q)\nabla_p H(q, p) = 0 \} \quad (2.415) \]

where \( G(q) \in \mathbb{R}^{k \times m} \) is the Jacobian of the constraint function at \( q \).

**Fact.** View \( T^*M \) as an embedded sub-manifold of \( \mathbb{R}^{2m} \). Given the constraint \( g(q) = 0 \), we may differentiate this constraint with respect to time to obtain a constraint on the velocity. Namely,

\[ \frac{d}{dt} g(q) = G(q)\dot{q} = 0. \quad (2.416) \]

**Corollary 2.3.31.** In Hamiltonian mechanics, \( \dot{q} = \nabla_p H(q, p) \). Hence, \( G(q)\dot{q} = G(q)\nabla_p H(q, p) = 0 \) is the constraint on \( p \).
Fact (Cotangent Space of Embedded Cotangent Bundle (Page 187 in Leimkuhler and Reich [2005])). View $T^*M$ as an embedded sub-manifold of $\mathbb{R}^{2m}$. The embedded cotangent space of $T^*M$, denoted $T^*T^*M$, is a subset of $T^*\mathbb{R}^{2m}$. Let $dq_1, \ldots, dq_m, dp_1, \ldots, dp_m \in T^*\mathbb{R}^{2m}$ be the coordinate 1-forms in the Euclidean space (see definition 2.3.10). The restriction of these differential 1-forms to $T^*T^*M$ implies that they satisfy,

$$G(q) \ dq = 0$$  \hspace{1cm} (2.417)

$$f_q(q, p) \ dq + f_p(q, p) \ dp = 0,$$  \hspace{1cm} (2.418)

where $f(q, p) \overset{\text{def}}{=} G(q)\nabla_p H(q, p)$ is the velocity constraint from section 2.3.3 and corollary 2.3.31 and $f_q(q, p)$ (resp. $f_p(q, p)$) represents its Jacobian with respect to $q$ (resp. $p$).

Lemma 2.3.32. Page 187 in Leimkuhler and Reich [2005] Let $q \in \mathbb{R}^n$. Let $g : \mathbb{R}^n \rightarrow \mathbb{R}^k$ be the constraint function with Jacobian $G(q) \in \mathbb{R}^{k \times n}$. Suppose $g(q) = 0$. Then for any $\mu \in \mathbb{R}^k$,

$$dq \wedge d(G(q)^T \mu) = 0$$  \hspace{1cm} (2.419)

Proof. We have

$$dq \wedge d(G(q)^T \mu) = dq \wedge G(q)^T d\mu + \sum_{i=1}^k dq \wedge \mu_i \Gamma_i dq$$  \hspace{1cm} (2.420)

where $\Gamma_i$ is the Hessian of the $i^{th}$ constraint function. By symmetry of the Hessian and eq. (2.379) from section 2.3.2, the second term is zero. The first term is also zero because $g(q) = 0 \implies G(q)dq = 0$ and since $dq \wedge G(q)^T d\mu = G(q)dq \wedge d\mu = 0$.

Fact. The total force acting on an object is the sum of all individual forces.
Fact. Constraint forces act in the normal direction to the constraint surface. Given a constraint function \( g : \mathbb{R}^m \rightarrow \mathbb{R}^k \), constraint forces are therefore represented by \(-G(q)^\top \lambda\) for \( \lambda \in \mathbb{R}^k \). This is called D’Alembert’s principle.

Fact. The force on a particle \( q \in \mathbb{R}^3 \) under the influence of a magnetic field is given by \( m \times \frac{d}{dt}q \) where \( m \in \mathbb{R}^3 \) represents parameters of the magnetic field and \( \times \) is the vector cross-product; that is,

\[
m \times \frac{d}{dt}q = \begin{pmatrix} 0 & m_1 & -m_2 \\ -m_1 & 0 & m_3 \\ m_2 & -m_3 & 0 \end{pmatrix} \begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{pmatrix}.
\]

This is called the Lorentz force law.

Definition 2.3.33. The magnetic symplectic structure, denoted \( \Omega_{\text{mag}} \), is the symplectic structure with matrix

\[
J_{\text{mag}} = \begin{pmatrix} L & \text{Id}_m \\ -\text{Id}_m & 0_m \end{pmatrix}
\]

where \( L \in \text{Skew}(m) \).

According to Dirac’s theory of constraints [Dirac, 1964], it suffices to embed a manifold-constrained Hamiltonian system in a Euclidean space. Consider the motion on \( T^*M \) determined by,

\[
\dot{q}_t = \nabla_p H(q_t, p_t)
\]

\[
\dot{p}_t = -\nabla_q H(q_t, p_t) - L\nabla_p H(q_t, p_t) - G(q_t)^\top \lambda
\]

\[
g(q_t) = 0
\]

where \( g : \mathbb{R}^m \rightarrow \mathbb{R}^k \) is a constraint function, \( G : \mathbb{R}^m \rightarrow \mathbb{R}^{k \times m} \) is the Jacobian of the
constraint, and $\lambda \in \mathbb{R}^k$ is a vector of Lagrange multipliers. The Lagrange multipliers $\lambda \equiv \lambda(q_t, p_t)$ are uniquely defined by the condition $g(q_t) = 0$ along solutions of eqs. (2.423) to (2.425). These equations of motion correspond to a distinct physical interpretation in $\mathbb{R}^3$ regarding magnetism. To formulate further results, we provide the following definition of a magnetic vector field flow.

**Definition 2.3.34.** Let $\Phi_{\text{mag}}(\cdot, \cdot; t) : T^*M \rightarrow T^*M$ be the vector field flow (see definition 2.3.28) corresponding to the motion given in eqs. (2.423), (2.424) and (2.425).

**Lemma 2.3.35 (Physical Interpretation of Motion).** Let $M = \{q \in \mathbb{R}^3 : g(q) = 0\}$ be a manifold such that $G(q)$ has full-rank. Consider a Hamiltonian of the form $H(q, p) = U(q) + \frac{1}{2b} p^\top p$ with $(q, p) \in T^*M$ and $b \in \mathbb{R}_+$. Then the equations of motion in eqs. (2.423) to (2.425) correspond to the motion of a particle, with mass $b$, simultaneously undergoing potential, magnetic, and manifold constraint forces.

**Proof.** It is common to express potential forces as the negative gradient of some function $U : \mathbb{R}^3 \rightarrow \mathbb{R}$ called the potential function. The Hamiltonian equations of motion for $H(q, p)$ with a magnetic symplectic structure $\Omega_{\text{mag}}$ are:

$$\dot{q}_t = \frac{p_t}{b}$$

$$\dot{p}_t = -\nabla_q U(q_t) - L \frac{p_t}{b} - G(q_t)^\top \lambda$$

$$g(q_t) = 0$$

Now noting that the momentum variables $p_t = b \dot{q}_t$ by substitution we obtain,

$$\dot{q}_t = \dot{q}_t$$

$$b \ddot{q}_t = -\nabla_q U(q_t) + m \times \dot{q}_t + G(q_t)^\top (-\lambda)$$

85
where, since L is a skew-symmetric matrix, we have used section 2.3.3 to identify

\[-L = \begin{pmatrix}
0 & m_1 & -m_2 \\
-m_1 & 0 & m_3 \\
m_2 & -m_3 & 0
\end{pmatrix}.
\]

(2.431)

We have used section 2.3.3 to identify constraint forces and section 2.3.3 to recognize that the sum of these three forces is the total force acting on the particle. Thus we see, by Newton’s second law of motion, that the Hamiltonian equations of motion are equivalent to Newtonian mechanics describing a particle subject to potential, magnetic, and constraint forces.

2.3.4 Numerical Integration

For most Hamiltonian vector fields, even those on Euclidean space as we noted in section 2.2, there do not exist closed-forms for the flows. Therefore, it is necessary to design numerical integrators for Hamiltonian systems. Much of this discussion is in correspondence with the details laid out in sections 2.2.2 and 2.2.4 with suitable generalization to the manifold context.

Definition 2.3.36 (Numerical Integrator). A numerical integrator of a Hamiltonian system with step-size \( \epsilon \in \mathbb{R} \) and number of integration steps \( N \in \mathbb{N} \) is a map \( \hat{\Phi}(\cdot, \cdot; \epsilon, N) : T^*M \rightarrow T^*M \) approximating \( \Phi(\cdot, \cdot; \epsilon \cdot N) \).

While a good approximation is desirable in HMC for high acceptance probabilities, the quality of approximation is of no consequence for the correctness of the sampler. However, it is essential for our formulation of HMC that numerical integrators are symmetric and symplectic, defined as follows.

Definition 2.3.37. A map \( \Phi : T^*M \rightarrow T^*M \) is symmetric if \( \Phi(\Phi(z; -\epsilon); \epsilon) = z \) for all \( z \in T^*M \).
Definition 2.3.38. A numerical integrator \(\hat{\Phi}\) is symmetric if, for fixed \(N\), \(\hat{\Phi}(\cdot; \epsilon, N)\) is a symmetric map for all \(\epsilon\).

Definition 2.3.39. A numerical integrator is *symplectic* if, for fixed \(\epsilon\) and \(N\), the map \(\hat{\Phi}(\cdot; \epsilon, N)\) is symplectic (definition 2.3.26).

Symplectic integrators preserve volume in \(T^*M\) in the following sense.

Definition 2.3.40. A numerical integrator is volume preserving if for any region \(Z \subset T^*M\) with volume \(\text{Vol}(Z)\) the set \(Z' \overset{\text{def.}}{=} \{\hat{\Phi}(q, p; \epsilon, N) : (q, p) \in Z\}\) satisfies \(\text{Vol}(Z) = \text{Vol}(Z')\) for any choice of \(\epsilon\) and \(N\).

Volume on embedded manifolds can be rigorously defined in terms of volume forms (definition 2.3.17). For text covering this construction, we refer the interested reader to [Abraham et al., 1988](#). To enable a simpler treatment, we will not discuss the relationship between volume form and measures on the Borel \(\sigma\)-algebra on \(T^*M\). However, we note that for a given volume form on \(T^*M\) there exists a unique measure on \(\mathcal{B}(T^*M)\) for which the measure of sets agrees with the integral of the set with respect to the volume form.

Definition 2.3.41. Let \(\hat{\Phi}(\cdot; \epsilon, 1) : T^*M \to T^*M\) be a single step numerical integrator (definition 2.3.36) for the Hamiltonian vector field flow \(\Phi(\cdot; \epsilon) : T^*M \to T^*M\) (definition 2.3.28). Then \(\hat{\Phi}\) is said to have *accuracy of order* \(k \in \mathbb{N}\) if for any \((q, p) \in T^*M\) we have

\[
\hat{\Phi}((q, p); \epsilon, 1) - \Phi((q, p); \epsilon) = O(\epsilon^{k+1})
\]

The following fact is the extension of theorem 2.2.23 to the case of Hamiltonian mechanics on the cotangent space.

**Fact** (Page 86 in [Leimkuhler and Reich, 2005](#)). Let \(\hat{\Phi}(\cdot; \epsilon, 1) : T^*M \to T^*M\) be a single step numerical integrator (definition 2.3.36) for the Hamiltonian vector field flow \(\Phi(\cdot; \epsilon) : T^*M \to T^*M\) (definition 2.3.28). Suppose further that \(\hat{\Phi}\) is a symmetric integra-
tor (definition 2.3.38). Then the order of \( \hat{\Phi} \) is even.

### 2.3.5 Uniqueness of the Lagrange Multipliers

**Theorem 2.3.42.** Let \( M = \{ q \in \mathbb{R}^m : g(q) = 0 \} \) be a connected manifold such that \( G(q) \) has full-rank and \( \nabla^2_p H(q, p) \) is invertible. Then the Lagrange multipliers \( \lambda \) in the equations of motion from eqs. (2.423) to (2.425) are uniquely defined.

**Proof.** Write \( g(q) \) in terms of the individual constraint functions by identifying \( g(q) = (g_1(q), \ldots, g_k(q)) \). By definition, \( g(q_t) = 0 \) along a solution of the equations of motion. Therefore,

\[
\frac{d}{dt} g(q_t) = G(q_t) \dot{q}_t = 0.
\]  

(2.433)

Differentiating the constraint twice with respect to time yields,

\[
\frac{d^2}{dt^2} g(q_t) = \frac{d}{dt} G(q_t) \dot{q}_t
\]

(2.434)

\[
= [\nabla G(q_t) \cdot \dot{q}_t] \ddot{q}_t + G(q_t) \dddot{q}_t
\]

(2.435)

\[
= 0
\]

(2.436)

Thus,

\[
\ddot{q}_t = \nabla_q \nabla_p H(q_t, p_t) \cdot \dot{q}_t + \nabla^2_p H(q_t, p_t) \cdot [-\nabla_q H(q_t, p_t) - L \nabla_p H(q_t, p) - G(q_t)^\top \lambda]
\]

(2.437)
and therefore,

$$\begin{align*}
- [\nabla G(q_t) \cdot \dot{q}_t] \dot{q}_t & = G(q_t) \left[ \nabla_q \nabla_p H(q_t, p_t) \cdot \dot{q}_t + \nabla^2_p H(q_t, p_t) \cdot [ - \nabla_q H(q_t, p_t) - L \nabla_p H(q_t, p_t) - G(q_t)^\top \lambda] \right].
\end{align*}$$

(2.438)

Rearranging terms we obtain the expression,

$$\begin{align*}
G(q_t) \nabla^2_p H(q_t, p_t) G(q_t)^\top \lambda & = - [\nabla G(q_t) \cdot \dot{q}_t] \dot{q}_t - G(q_t) \nabla_q \nabla_p H(q_t, p_t) \cdot q_t + G(q_t) \nabla^2_p H(q_t, p_t) [\nabla_q H(q_t, p_t) + L \nabla_p H(q_t, p_t)].
\end{align*}$$

(2.439)

Therefore, $\lambda$ is uniquely defined provided that $G(q_t) \nabla^2_p H(q_t, p_t) G(q_t)^\top$ is invertible. This is achieved when $G(q)$ has full-rank and $\nabla^2_p H(q, p)$ is invertible.

### 2.3.6 Properties of Flows on Embedded Manifolds

In proving these results we will adopt the shorthand notation $H_q$ to denote the partial derivative of $H$ with respect to $q$ regarded as a row vector. The notation $H_{qp}$ denotes the matrix of partial derivatives of $H$ with respect to $q$ and $p$. Other quantities are similarly defined.

**Lemma 2.3.43.** Let $M = \{ q \in \mathbb{R}^m : g(q) = 0 \}$ be a connected manifold such that $G(q)$ has full-rank. View $T^* M$ as an embedded sub-manifold of $\mathbb{R}^{2m}$. Let $\Omega_{mag}$ be the magnetic symplectic structure in the ambient Euclidean space $\mathbb{R}^{2m} \cong \mathbb{R}^m \times \mathbb{R}^m$. Let $H(q, p)$ be a smooth Hamiltonian $H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$. Then $\Phi_{mag}$ (definition 2.3.34) is symmetric: $\Phi_{mag}(\Phi_{mag}(q, p; t); -t) = (q, p)$.

**Proof.** The map $\Phi_{mag} : T^* M \rightarrow T^* M$ is a vector field flow by definition (see definition 2.3.34). By eq. (2.413) in section 2.3.3 it is symmetric. \qed
Lemma 2.3.44. Let $M = \{q \in \mathbb{R}^m : g(q) = 0\}$ be a connected manifold such that $G(q)$ has full-rank. View $T^*M$ as an embedded sub-manifold of $\mathbb{R}^{2m}$. Let $\Omega_{\text{mag}}$ be the magnetic symplectic structure in the ambient Euclidean space $\mathbb{R}^{2m} \cong \mathbb{R}^m \times \mathbb{R}^m$. Let $H(q, p)$ be a smooth Hamiltonian $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$. Then $\Phi_{\text{mag}}(\cdot, \cdot; t)$ is a symplectic transformation (definition 2.3.26) on $T^*M$ for any $t$.

Proof. We want to show that $\Phi_{\text{mag}}$ is a symplectic transformation (definition 2.3.26). By definition 2.3.25, the magnetic symplectic structure can be written in terms of the wedge product as

$$dq \wedge dp + \frac{1}{2} dq \wedge L dq$$

from eq. (2.412). Let $(q_t, p_t) = \Phi_{\text{mag}}(q, p; t)$. Denote $\Omega^t_{\text{mag}} = dq_t \wedge dp_t + \frac{1}{2} dq_t \wedge L dq_t$. From lemma 2.3.24, $\Phi_{\text{mag}}$ is symplectic for the magnetic 2-form $\Omega_{\text{mag}}$ if

$$\frac{d}{dt} \Omega^t_{\text{mag}} = 0.$$  

Hence, our proof strategy will establish $\frac{d}{dt} \Omega^t_{\text{mag}} = 0$ which will imply that $\Phi_{\text{mag}}$ is symplectic.

Lemma 2.3.45.

$$dq_t = H_{qp}(q_t, p_t) dq_t + H_{pp}(q_t, p_t) dp$$

$$dp_t = -H_{qq}(q_t, p_t) dq_t - H_{qp}(q_t, p_t) dp_t - LH_{ql}(q_t, p_t) dq_t - LH_{pp}(q_t, p_t) dp_t - d(G(q_t)^T \lambda)$$

(2.444)
Proof. The equations of motion are

\[
\dot{q}_t = \nabla_p H(q_t, p_t) \quad (2.445)
\]

\[
\dot{p}_t = -\nabla_q H(q_t, p_t) - L \nabla_p H(q_t, p_t) - G(q_t)^T \lambda \quad (2.446)
\]

\[
g(q_t) = 0 \quad (2.447)
\]

Computing the differential yields,

\[
d\dot{q}_t = H_{pq}(q_t, p_t)dq_t + H_{pp}(q_t, p_t)dp_t \quad (2.448)
\]

\[
d\dot{p}_t = -H_{qq}(q_t, p_t)dq_t - H_{qp}(q_t, p_t)dp_t - L H_{pq}(q_t, p_t)dq_t - L H_{pp}(q_t, p_t)dp_t - d(G(q_t)^T \lambda) \quad (2.449)
\]

We use the differentials computed in lemma 2.3.45. The notation $H_{qp}$ denotes the matrix of partial derivatives of $H$ with respect to $q$ and $p$. Symmetry of partial derivatives yields $H_{pq} = H_{qp}^T$. The Hessian matrix with respect to $q$ (resp. $p$) is denoted $H_{qq}$ (resp. $H_{pp}$).

Computing the time derivative of $\Omega_{\text{mag}}^t$, we have that the magnetic symplectic form is
preserved under the solution to the constrained system.

\[
\frac{d}{dt} \Omega_{\text{mag}}^{t} = d\dot{q}_{t} \wedge dp_{t} + dq_{t} \wedge d\dot{p}_{t} + \frac{1}{2} d\dot{q}_{t} \wedge Ldq_{t} + \frac{1}{2} dq_{t} \wedge Ld\dot{q}_{t}
\]  
(2.450)

\[
= d\dot{q}_{t} \wedge dp_{t} + dq_{t} \wedge d\dot{p}_{t} + dq_{t} \wedge Ldq_{t}
\]  
(2.451)

\[
= H_{pq} dq_{t} \wedge dp_{t} + H_{pp} dp_{t} \wedge dp_{t} - dq_{t} \wedge H_{qp} dq_{t} - dq_{t} \wedge L H_{pq} dq_{t}
\]  
(2.452)

\[
- dq_{t} \wedge L H_{pp} dp_{t} - dq_{t} \wedge d(G(q_{t})^{\top} \lambda) + H_{pq} dq_{t} \wedge Ldq_{t} + H_{pp} dp_{t} \wedge Ldq_{t}
\]  

\[
= H_{pq} dq_{t} \wedge dp_{t} +
\]  
(2.453)

\[
- dq_{t} \wedge H_{qp} dp_{t} - dq_{t} \wedge L H_{pq} dq_{t}
\]  
(2.454)

\[
- dq_{t} \wedge L H_{pp} dp_{t} - dq_{t} \wedge d(G(q_{t})^{\top} \lambda) + H_{pq} dq_{t} \wedge Ldq_{t} + H_{pp} dp_{t} \wedge Ldq_{t}
\]  

\[
= 0
\]  
(2.455)

The final equality comes from manipulations of the wedge product using section 2.3.2 and using the fact that \(L\) is a skew-symmetric matrix; in particular, we use eqs. (2.376), (2.378) and (2.379). That \(dq_{t} \wedge d(G(q_{t})^{\top} \lambda) = 0\) from lemma 2.3.32 was also used.

**Lemma 2.3.46.** Let \(M = \{q \in \mathbb{R}^{m} : g(q) = 0\}\) be a connected manifold such that \(G(q)\) has full-rank. View \(T^{\ast}M\) as an embedded sub-manifold of \(\mathbb{R}^{2m}\). Let \(\Omega_{\text{mag}}\) be the magnetic symplectic structure in the ambient Euclidean space \(\mathbb{R}^{2m} \cong \mathbb{R}^{m} \times \mathbb{R}^{m}\). Let \(H(q, p)\) be a smooth Hamiltonian \(H : \mathbb{R}^{m} \times \mathbb{R}^{m} \rightarrow \mathbb{R}\). Then \(H(\Phi_{\text{mag}}(q, p; t)) = H(q, p)\) for any
\((q, p) \in T^*M\) so that the Hamiltonian energy is conserved.

**Proof.** Let \((q_t, p_t) = \Phi_{\text{mag}}(q, p; t)\). To prove that the Hamiltonian is conserved, we verify that the time derivative of \(H(q_t, p_t)\) equals zero.

\[
\frac{d}{dt} H(q_t, p_t) = \nabla_q H(q_t, p_t) \cdot \dot{q}_t + \nabla_p H(q_t, p_t) \cdot \dot{p}_t
\]

\[= \nabla_q H(q_t, p_t) \cdot \nabla_p H(q_t, p_t) - \nabla_p H(q_t, p_t) \cdot \left( L \nabla_p H(q_t, p_t) + G(q_t) \right) \lambda \]

\[= -\nabla_p H(q_t, p_t) \cdot L \nabla_p H(q_t, p_t) - \nabla_p H(q_t, p_t) \cdot G(q_t) \lambda \]

\[= 0 \quad (2.459) \]

by section 2.3.2 using that \(L\) is skew-symmetric and since \(G(q_t) \nabla_p H(q_t, p_t) = G(q_t) \dot{q}_t = 0\) from corollary 2.3.31 and section 2.3.3. Therefore, by the Fundamental Theorem of Calculus:

\[
H(q_t, p_t) - H(q_0, p_0) = \int_0^t \frac{d}{ds} H(q_s, p_s) \, ds = 0. \quad (2.460)
\]

Therefore \(H(q_t, p_t) = H(q_0, p_0)\) and since \((q_0, p_0) = (q, p)\) we have shown that \(H(q_t, p_t) = H(q, p)\). \(\square\)

### 2.3.7 Properties of Numerical Integrators on Embedded Manifolds

We now recall a general method of constructing symmetric, symplectic, and second-order accurate numerical integrators on embedded manifolds for the equations of motion in eqs. (2.423) to (2.425). To accomplish this, we will first require an integrator
\( \Phi_\epsilon : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m \) for the unconstrained equations of motion

\[
\dot{q}_t = \nabla_p H(q_t, p_t) \tag{2.461}
\]

\[
\dot{p}_t = -\nabla_q H(q_t, p_t) - L \nabla_p H(q_t, p_t) \tag{2.462}
\]

**Definition 2.3.47.** Let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be a smooth Hamiltonian. Let \( \Phi_\epsilon : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m \) be a symmetric, symplectic, second-order accurate for the unconstrained Hamiltonian vector field corresponding to \( H \) in eqs. (2.461) and (2.462). At iteration \( n \), let \((q_n, p_n) \in T^* M\). Compute:

\[
\bar{p}_{n+1/2} = p_n - \frac{\epsilon}{2} G(q_n)^\top \mu \tag{2.463}
\]

\[
(q_{n+1}, \bar{p}_{n+1}) = \Phi_\epsilon(q_n, \bar{p}_{n+1/2}) \tag{2.464}
\]

\[
0 = g(q_{n+1}) \tag{2.465}
\]

\[
p_{n+1} = \bar{p}_{n+1} - \frac{\epsilon}{2} G(q_{n+1})^\top \mu' \tag{2.466}
\]

\[
0 = G(q_{n+1})^\top \nabla_p H(q_{n+1}, p_{n+1}) \tag{2.467}
\]

The Lagrange multipliers \( \mu \) and \( \mu' \) are chosen such that eqs. (2.465) and (2.467) are satisfied. Such Lagrange multipliers exist, and are unique, provided \( \epsilon \neq 0 \) is small enough; see [McLachlan et al. 2012].

Note that when eqs. (2.465) and (2.467) are satisfied, \( q_{n+1} \in M \) and \( p_{n+1} \in T^* q_{n+1} M \) but that \( \bar{p}_{n+1/2} \) and \( \bar{p}_{n+1} \) are not guaranteed to respect the manifold constraint.

**Lemma 2.3.48.** Let \( M = \{ q \in \mathbb{R}^m : g(q) = 0 \} \) be a connected manifold such that \( G(q) \) has full-rank. Let \( T^* M \) be an embedded sub-manifold of \( \mathbb{R}^{2m} \) as in definition 2.3.4. Let \( \mu \) and \( \mu' \) be Lagrange multipliers such that eqs. (2.465) and (2.467) are satisfied. Then the map \((q_n, p_n) \mapsto (q_{n+1}, p_{n+1})\) maps \((q_n, p_n) \in T^* M\) to \((q_{n+1}, p_{n+1}) \in T^* M\).

**Proof.** Recall that \( M = g^{-1}(0) \) so that \( q \in M \iff g(q) = 0 \). If \( \mu \) is a Lagrange
multiplier such that eq. (2.465) is satisfied, it is immediate that $q_{n+1} \in M$. From definition 2.3.2 $p \in T^*_qM \iff G(q)\nabla_p H(q, p) = 0$. If $\mu'$ is a Lagrange multiplier satisfying eq. (2.467), then it is immediate that $p_n+1 \in T^*_{q_n+1}M$. Thus, by definition 2.3.4 $(q_{n+1}, p_{n+1}) \in T^*M$.

**Lemma 2.3.49.** Let $g : \mathbb{R}^m \to \mathbb{R}^k$ be a constraint function with full-rank Jacobian $G : \mathbb{R}^m \to \mathbb{R}^{k \times m}$. Let $q_n \in \mathbb{R}^m$ satisfy $g(q_n) = 0$ and let $p_n \in \mathbb{R}^m$. Let $\bar{p}_{n+1/2} = p_n - \frac{\epsilon}{2} G(q_n)^\top \mu$ for $\mu \in \mathbb{R}^k$. Then,

$$dq_n \wedge d\bar{p}_{n+1/2} = dq_n \wedge dp_n$$

(2.468)

**Proof.** By direct calculation using eq. (2.377) from section 2.3.2,

$$dq_n \wedge d\bar{p}_{n+1/2} = dq_n \wedge dp_n - dq_n \wedge d\left(\frac{\epsilon}{2} G(q_n)^\top \mu\right)$$

(2.469)

$$= dq_n \wedge dp_n - \frac{\epsilon}{2} (dq_n \wedge d(G(q_n)^\top \mu))$$

(2.470)

$$= dq_n \wedge dp_n$$

(2.471)

by lemma 2.3.32. Lemma 2.3.32 can be applied since $g(q_n) = 0$ and $\mu \in \mathbb{R}^k$ by assumption.

**Corollary 2.3.50.** For a skew-symmetric matrix $L \in \text{Skew}(m),$

$$dq_n \wedge d\bar{p}_{n+1/2} + \frac{1}{2} dq_n \wedge Ldq_n = dq_n \wedge dp_n + \frac{1}{2} dq_n \wedge Ldq_n$$

(2.472)

**Proof.** Using lemma 2.3.49 add $\frac{1}{2} dq_n \wedge Ldq_n$ to both sides.

**Lemma 2.3.51.** Let $(q_n, \bar{p}_{n+1/2}) \in \mathbb{R}^{2m}$. Let $\Phi : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ be a symplectic transformation with respect to the magnetic symplectic form $\Omega_{mag}$. If $(q_{n+1}, \bar{p}_{n+1}) = \Phi(q_n, \bar{p}_{n+1/2})$
\[ dq_{n+1} \wedge d\bar{p}_{n+1} + \frac{1}{2} dq_{n+1} \wedge Ldq_{n+1} = dq_n \wedge d\bar{p}_{n+1/2} + \frac{1}{2} dq_n \wedge Ldq_n \]  

(2.473)

**Proof.** Since \( \Phi \) is symplectic we have,

\[ \Phi^* \Omega_{\text{mag}} = \Omega_{\text{mag}}. \]  

(2.474)

Using eq. (2.412) from definition 2.3.25 we express the symplecticness of \( \Phi \) using coordinate differential one-forms (definition 2.3.10):

\[ \Phi^* \Omega_{\text{mag}} = \Omega_{\text{mag}} \]  

(2.475)

\[ \implies dq_{n+1} \wedge d\bar{p}_{n+1} + \frac{1}{2} dq_{n+1} \wedge Ldq_{n+1} = dq_n \wedge d\bar{p}_{n+1/2} + \frac{1}{2} dq_n \wedge Ldq_n \]  

(2.476)

**Lemma 2.3.52.** Let \( \mu \) and \( \mu' \) be Lagrange multipliers such that eqs. (2.465) and (2.467) are satisfied. Then the map \((q_n, p_n) \mapsto (q_{n+1}, p_{n+1})\) is symplectic.

**Proof.** At iteration \( n \) of the integrator, assume \((q_n, p_n) \in T^* M\). The map \((q_n, p_n) \mapsto (q_{n+1}, p_{n+1})\) in definition 2.3.47 consists of three steps as follows. Let \((q_n, p_n) \in T^* M\).

1. Set \( \bar{p}_{n+1/2} = p_n - \frac{\epsilon}{2} G(q_n)^\top \mu \).

2. Compute \((q_{n+1}, \bar{p}_{n+1}) = \hat{\Phi}_\epsilon(q_n, \bar{p}_{n+1/2})\).

3. Set \( p_{n+1} = \bar{p}_{n+1} - \frac{\epsilon}{2} G(q_{n+1})^\top \mu' \).

Notice that \( \bar{p}_{n+1/2}, \bar{p}_{n+1} \in \mathbb{R}^m \) but that \((q_{n+1}, p_{n+1}) \in T^* M\) by the choice of Lagrange multipliers \( \mu \) and \( \mu' \). By eq. (2.412) from definition 2.3.25 the magnetic symplectic 2-
form can be written in terms of wedge products as,

\[ dq_n \wedge dp_n + \frac{1}{2} dq_n \wedge Ldq_n \] (2.477)

By lemma 2.3.23 it suffices to show that the integrator conserves the symplectic 2-form on \( T^* M \) under the map \((q_n, p_n) \mapsto (q_{n+1}, p_{n+1})\). Therefore, our proof strategy will be to show that

\[ dq_{n+1} \wedge dp_{n+1} + \frac{1}{2} dq_{n+1} \wedge Ldq_{n+1} = dq_n \wedge dp_n + \frac{1}{2} dq_n \wedge Ldq_n \] (2.478)

Since \( g(q_n) = 0 \) by assumption (since \( q_n \in M \)) and \( \mu \in \mathbb{R}^k \), we may apply corollary 2.3.50 to the first step of the integrator to show that

\[ dq_n \wedge d\bar{p}_{n+1/2} + \frac{1}{2} dq_n \wedge Ldq_n = dq_n \wedge dp_n + \frac{1}{2} dq_n \wedge Ldq_n \] (2.479)

Applying lemma 2.3.51 to \((q_{n+1}, \bar{p}_{n+1})\) in the second step and using the fact that the integrator \( \hat{\Phi} \) is symplectic by assumption shows that,

\[ dq_{n+1} \wedge \bar{p}_{n+1} + \frac{1}{2} dq_{n+1} \wedge Ldq_{n+1} = dq_n \wedge d\bar{p}_{n+1/2} + \frac{1}{2} dq_n \wedge Ldq_n \] (2.480)

\[ = dq_n \wedge dp_n + \frac{1}{2} dq_n \wedge Ldq_n \] (2.481)

Since \( g(q_{n+1}) = 0 \) by construction and since \( \mu' \in \mathbb{R}^k \), applying corollary 2.3.50 a second time to the third step yields

\[ dq_{n+1} \wedge dp_{n+1} + \frac{1}{2} dq_{n+1} \wedge Ldq_{n+1} = dq_{n+1} \wedge \bar{p}_{n+1} + \frac{1}{2} dq_{n+1} \wedge Ldq_{n+1} \] (2.482)

\[ = dq_n \wedge dp_n + \frac{1}{2} dq_n \wedge Ldq_n \] (2.483)

This verifies that the symplectic structure \( \Omega_{\text{can}} \) is preserved. Therefore, the integrator is
Lemma 2.3.53. Let \( \mu \) and \( \mu' \) be Lagrange multipliers such that eqs. (2.465) and (2.467) are satisfied. Let \( \epsilon \) be the integration step-size. Then the map \( (q_n, p_n) \mapsto (q_{n+1}, p_{n+1}) \) in definition 2.3.47 is symmetric under \( \epsilon \mapsto -\epsilon \).

Proof. The integrator in definition 2.3.47 consists of three steps as follows. Let \( (q_n, p_n) \in T^*M \).

1. Set \( \bar{p}_{n+1/2} = p_n - \frac{\epsilon}{2}G(q_n)^\top \mu \).
2. Compute \( (q_{n+1}, \bar{p}_{n+1}) = \hat{\Phi}_\epsilon(q_n, \bar{p}_{n+1/2}) \).
3. Set \( p_{n+1} = \bar{p}_{n+1} - \frac{\epsilon}{2}G(q_{n+1})^\top \mu' \).

Notice that \( \bar{p}_{n+1/2}, \bar{p}_{n+1} \in \mathbb{R}^m \) but that \( (q_{n+1}, p_{n+1}) \in T^*M \) by the choice of Lagrange multipliers \( \mu \) and \( \mu' \). To show that the integration scheme is symmetric, consider beginning from position \( (q_{n+1}, p_{n+1}) \) and applying the three integration steps with a reversed step-size. In the first step, we obtain the update

\[
\bar{p}_{n+1+1/2} = p_{n+1} + \frac{\epsilon}{2}G(q_{n+1})^\top \mu'
\]

where the last equality derives from rearranging the defining relation in the third step. Since the integrator in \( \hat{\Phi} \) is symmetric by assumption, applying the integrator with step-size \( -\epsilon \) maps \( (q_{n+1}, \bar{p}_{n+1}) \) to \( (q_n, \bar{p}_{n+1/2}) \). The third integration step with Lagrange multiplier \( \mu \) yields the update,

\[
p_{n+2} = \bar{p}_{n+1/2} + \frac{\epsilon}{2}G(q_n)^\top \mu
\]

where the last equality derives from rearranging the defining relation in the third step.
By assumption, \((q_n, p_n) \in T^* M\) so that \(g(q_n) = 0\) and \(G(q_n)\nabla_p H(q_n, p_n) = 0\). This completes the reversibility argument.

Recall that the equations of motion for magnetic Hamiltonian dynamics from eqs. (2.423) to (2.425) are

\[
\begin{align*}
\dot{q}_t &= \nabla_p H(q_t, p_t) \\
\dot{p}_t &= -\nabla_q H(q_t, p_t) - L\nabla_p H(q_t, p_t) - G(q_t)^\top\lambda \\
g(q_t) &= 0.
\end{align*}
\]

(2.488) (2.489) (2.490)

The equations of motion may be written in matrix form as,

\[
\begin{pmatrix}
\dot{q}_t \\
\dot{p}_t
\end{pmatrix} =
\begin{pmatrix}
0 & \text{Id} \\
-L & -L
\end{pmatrix}
\begin{pmatrix}
\nabla_q H(q_t, p_t) \\
\nabla_p H(q_t, p_t)
\end{pmatrix} -
\begin{pmatrix}
0 \\
G(q_t)^\top\lambda
\end{pmatrix}.
\]

(2.491)

Notice that eq. (2.491) is the first-order term in the Taylor series expansion of the vector field flow in the time variable:

**Lemma 2.3.54.** Let \(\mu\) and \(\mu'\) be constraint-preserving Lagrange multipliers. The map 
\((q_n, p_n) \mapsto (q_{n+1}, p_{n+1})\) has order at least one.

**Proof of Lemma 2.3.54.** Recall further that the map 
\((q_n, p_n) \mapsto (q_{n+1}, p_{n+1})\) in definition 2.3.47 consists of the following three steps.

1. Set \(\bar{p}_{n+1/2} = p_n - \frac{\epsilon}{2} G(q_n)^\top\mu\).

2. Compute \((q_{n+1}, \bar{p}_{n+1}) = \tilde{\Phi}_\epsilon(q_n, \bar{p}_{n+1/2}).\)

3. Set \(p_{n+1} = \bar{p}_{n+1} - \frac{\epsilon}{2} G(q_{n+1})^\top\mu'.\)
From the fact that $\hat{\Phi}_\epsilon$ is second order by assumption we have that,

\[
\begin{pmatrix}
q_{n+1} \\
\bar{p}_{n+1/2}
\end{pmatrix}
= 
\begin{pmatrix}
q_n \\
p_n - \frac{\epsilon}{2} G(q_n) \top \mu
\end{pmatrix}
+ \epsilon \begin{pmatrix}
0 & \text{Id} \\
-\text{Id} & -L
\end{pmatrix}
\begin{pmatrix}
\nabla_q H(q_n, \bar{p}_{n+1/2}) \\
\nabla_p H(q_n, \bar{p}_{n+1/2})
\end{pmatrix}
+ \mathcal{O}(\epsilon^2)
\]

(2.492)

\[
= 
\begin{pmatrix}
q_n \\
p_n - \frac{\epsilon}{2} G(q_n) \top \mu
\end{pmatrix}
+ \epsilon \begin{pmatrix}
0 & \text{Id} \\
-\text{Id} & -L
\end{pmatrix}
\begin{pmatrix}
\nabla_q H(q_n, p_n) \\
\nabla_p H(q_n, p_n)
\end{pmatrix}
= \bar{p}_{n+1} + \epsilon \begin{pmatrix}
0 \\
G(q_n) \top \mu
\end{pmatrix} + \mathcal{O}(\epsilon^2).
\]

(2.493)

Now expanding $G(q_{n+1})$ as a Taylor series in $\epsilon$ shows $G(q_{n+1}) = G(q_n) + \mathcal{O}(\epsilon)$. Therefore,

\[
p_{n+1} = \bar{p}_{n+1} - \frac{\epsilon}{2} G(q_{n+1}) \top \mu'
\]

(2.495)

\[
= \bar{p}_{n+1} - \frac{\epsilon}{2} G(q_n) \top \mu' + \mathcal{O}(\epsilon^2).
\]

(2.496)

Combining eqs. (2.494) and (2.496) yields,

\[
\begin{pmatrix}
q_{n+1} \\
p_{n+1}
\end{pmatrix}
= 
\begin{pmatrix}
q_n \\
p_n
\end{pmatrix}
+ \epsilon \begin{pmatrix}
0 & \text{Id} \\
-\text{Id} & -L
\end{pmatrix}
\begin{pmatrix}
\nabla_q H(q_n, p_n) \\
\nabla_p H(q_n, p_n)
\end{pmatrix}
= \bar{p}_{n+1} - \frac{\epsilon}{2} G(q_n) \top \left(\frac{\mu + \mu'}{2}\right)
+ \mathcal{O}(\epsilon^2).
\]

(2.497)

Comparing eq. (2.497) and eq. (2.491) with $\lambda = \frac{\mu + \mu'}{2}$ shows that the integrator has order at least one.

It remains to be discussed the uniqueness of the Lagrange multipliers $\mu$ and $\mu'$ appearing in the map $(q_n, p_n) \mapsto (q_{n+1}, p_{n+1})$ in definition 2.3.47. The following result shows that the Lagrange multipliers are uniquely determined when $\epsilon$, the integration step-size, is
sufficiently small. The following result is from Theorem 4.1 in McLachlan et al. [2012].

**Proposition 2.3.55.** Let \( \dot{\Phi}_\epsilon : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m \) be a symmetric, symplectic, second-order integrator and Let \( g : \mathbb{R}^m \to \mathbb{R}^k \) be a constraint function with full-rank Jacobian \( G : \mathbb{R}^m \to \mathbb{R}^{k \times m} \). Let \( (q, p) \in T^* M \). Define,

\[
[(q, p)] = \{ (q, p - G(q)^\top \mu) : \mu \in \mathbb{R}^k \}. 
\] (2.498)

Let \( \text{Proj}_q : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \) be defined by \( \text{Proj}_q(q, p) = q \). Let \( \Phi^q_\epsilon \equiv \text{Proj}_q \circ \dot{\Phi}_\epsilon \) be the projection to the \( q \)-variables of the approximate integrator of the unconstrained dynamics. Then, for \( \epsilon \) sufficiently small, the equation

\[
g(\Phi^q_\epsilon(q^+, p^+)) = 0 \quad \text{such that} \quad (q^+, p^+) \in [(q, p)]
\] (2.499)

has a unique solution in a neighborhood of \( \mu = 0 \).

### 2.3.8 Conservation of Volume

For a closer look at the differential geometry, one might ask, “In what sense does conservation of the symplectic structure imply conservation of volume?”

**Theorem 2.3.56.** Transformations that preserve the symplectic structure under pullback preserve the Liouville volume form from definition [2.3.21]

**Proof.** The Liouville volume form is defined by,

\[
\Lambda \equiv \frac{(-1)^{m(m-1)/2}}{m!} \Omega \wedge \cdots \wedge \Omega
\] (2.500)

If \( \Phi \) is symplectic so that \( \Phi^* \Omega = \Omega \), then using section [2.3.2] immediately implies \( \Phi^* \Lambda = \Lambda \) so that the volume measure is conserved under \( \Phi \).
Theorem 2.3.57. Let $\Phi_{\text{mag}}$ be the magnetic vector field from from definition \[2.3.34\]. Then $\Phi_{\text{mag}}$ preserves the canonical Liouville volume form $\Lambda_{\text{can}}$ from definition \[2.3.21\].

Proof. From theorem \[2.3.56\] $\Phi_{\text{mag}}$ conserves the magnetic Liouville volume form

$$
\Lambda_{\text{mag}} \overset{\text{def}}{=} \left(\frac{-1}{m!}\right)^{m(m-1)/2} \Omega_{\text{mag}} \wedge \cdots \wedge \Omega_{\text{mag}} 
$$

(2.501)

Now recall section \[2.3.2\] which says that the space of volume forms is one-dimensional. Hence any constant (see definition \[2.3.15\]), non-vanishing (see definition \[2.3.16\]) volume form is proportional to any other constant, non-vanishing volume form. Let $\Lambda_{\text{can}} = c \cdot \Lambda_{\text{mag}}$ for some $c \in \mathbb{R}$ with $c \neq 0$. Then,

$$
\Phi^\ast \Lambda_{\text{can}} = \Phi^\ast (c \cdot \Lambda_{\text{mag}}) 
$$

(2.502)

$$
= c \cdot (\Phi^\ast \Lambda_{\text{mag}}) 
$$

(2.503)

$$
= c \cdot \Lambda_{\text{mag}} 
$$

(2.504)

$$
= \Lambda_{\text{can}}. 
$$

(2.505)

By identification, the determinant from definition \[2.3.18\] is $\det(\Phi) = 1$ so that $\Phi$ also conserves volume with respect to $\Lambda_{\text{can}}$ from section \[2.3.2\]. \[ \square \]

2.4 Stochastic Differential Equations

We now present a brief diversion into stochastic differential equations. A thorough treatment of these objects would take us too far afield, however we recall some of the most important results. Our reference for this aside is Øksendal [1992], which is also a self-contained introduction to the theory of stochastic differential equations. Stochastic differential equations, and in particular their discretizations, play an important role in the development of Markov chain Monte Carlo methods based on the local geometry of the target
distribution, which will be discussed later. Stochastic differential equations are an important area related to initial value problems, whose dynamics are governed by a stochastic differential equation, a fundamental example of this being Brownian motion [Øksendal 1992]. Stochastic differential equations have applications in computational finance, such as through the Black-Scholes equation [Shreve 2004] and in Monte Carlo via the Langevin algorithm [Roberts and Tweedie 1996].

2.4.1 Brownian Motion

Definition 2.4.1. Brownian motion \((B_t)_{t \geq 0}\) is a \(\mathbb{R}\)-valued stochastic process possessing the following properties:

1. Brownian motion begins at the origin \(B_0 = 0\).

2. For any finite collection of ordered times \((t_1, \ldots, t_k)\), Brownian motion has a joint Gaussian distribution

\[
\begin{pmatrix}
B_{t_1} \\
\vdots \\
B_{t_k}
\end{pmatrix}
\sim \text{Normal}
\begin{pmatrix}
0 \\
\vdots \\
0
\end{pmatrix},
\begin{pmatrix}
t_1 & t_1 & t_1 & \cdots & t_1 \\
t_1 & t_2 & t_2 & \cdots & t_2 \\
t_1 & t_2 & t_3 & \cdots & t_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
t_1 & t_2 & t_3 & \cdots & t_k
\end{pmatrix}
\tag{2.506}
\]

3. For a given realization of Brownian motion, the sample paths \(t \mapsto B_t\) may be constructed to be continuous.

Lemma 2.4.2. Brownian motion has independent increments; i.e. for \(s < t < v < w\), \(B_t - B_s\) is independent of \(B_w - B_v\).

Proof. From the joint Gaussian distribution, it follows that \(B_t - B_s\) and \(B_w - B_v\) are mean zero Gaussian distributions. We can therefore check if their product is zero in expectation.
in order to verify that the increments are independent.

\[
\mathbb{E}[(B_t - B_s)(B_w - B_v)] = \mathbb{E}[B_t B_w - B_s B_w - B_v B_t + B_s B_v] = t - s - t + s = 0.
\]

\[\text{(2.507)} \quad \text{(2.508)} \quad \text{(2.509)}\]

\[\square\]

**Definition 2.4.3.** Brownian motion in \(\mathbb{R}^m\) is constructed from the concatenation of \(m\) independent \(\mathbb{R}\)-valued Brownian motions in each of the individual dimensions.

**Definition 2.4.4.** A solution to a **stochastic differential equation** is, given initial data \(x_0 \in \mathbb{R}^m\), the solution to the integral equation,

\[
X^i_t = x^i_0 + \int_0^t b^i(s, X_s) \, ds + \int_0^t \sum_{j=1}^m \sigma^i_j(s, X_s) \, dB^j_s
\]

\[\text{(2.510)}\]

where \(b^i : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}\) and \(\sigma^i_j : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}\), \(B_s\) is \(m\)-dimensional Brownian motion, and the stochastic integral should be interpreted in the Itô sense. We call the functions \((b^1, \ldots, b^m)\) the **drift functions** and the functions \((\sigma^1_1, \ldots, \sigma^m_m)\) the **diffusion functions**.

**Remark.** What does it mean to “solve” a stochastic differential equation? Without becoming overly technical, let \(0 = t_0 < t_1, \ldots, t_n = t\) be a partition of the interval \([0, t]\). Then if we inductively define

\[
X^i_k = x^i_0 + \sum_{j=0}^{k-1} b^i(t_j, X_j) \Delta t_j + \sum_{j=0}^{k-1} \sigma^i_j(t_j, X_j) \Delta B_j
\]

\[\text{(2.511)}\]

where \(\Delta B_j \sim \text{Normal}(0, t_j)\) are independent Gaussian increments, can we construct a limiting object for the sequence of random variables \((X_1, \ldots, X_n)\) as \(n \to \infty\) and \(\Delta t_j \to 0\)? If this is indeed possible that we say that the limiting object is the solution of the stochastic differential equation in eq. (2.510).
It is a notational convenience to express this integral equation in “differential form” as

\[ dX^i_t = b^i(t, X_t) \, dt + \sum_{j=1}^{m} \sigma^i_j(t, X_t) \, dB^j_t. \]  
(2.512)

Fact (Theorem 3.2.1 in Øksendal [1992]). Let \( X_t \) be a solution of the stochastic differential equation in eq. (2.510). It is a fact that Itô integrals are martingales; that is, for functions \( f : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R} \) satisfying certain regularity conditions (given in definition 3.1.4 in Øksendal [1992]), we have

\[ \mathbb{E} \left[ \int_0^t \sum_{j=1}^{m} f(t, X_t) \, dB^j_t \right] = 0. \]  
(2.513)

2.4.2 Itô’s Lemma

As in the case of initial value problems, a stochastic differential equation is called time-homogenous when the functions \( b^1, \ldots, b^m \) and \( \sigma_1^1, \ldots, \sigma_m^m \) have no dependency on time. Itô stochastic differential equations are investigated by means of Itô’s lemma.

Lemma 2.4.5 (Itô’s Lemma [Itô 1951, Øksendal 1992]). Let \( X_t \) be the solution to the stochastic differential equation

\[ X^i_t = x^i_0 + \int_0^t b^i(s, X_s) \, ds + \int_0^t \sum_{j=1}^{m} \sigma^i_j(s, X_s) \, dB^j_s. \]  
(2.514)

Let \( g : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^n \) be a smooth function. Then the random variable \( Y_t = g(t, X_t) \) obeys the stochastic differential equation,

\[ dY^k_t = \frac{\partial g^k}{\partial t}(t, X_t) \, dt + \sum_{i=1}^{m} \frac{\partial g^k}{\partial x^i}(t, X_t) \, dX^i_t + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 g^k}{\partial x^i \partial x^j}(t, X_t) \, dX^i_t \, dX^j_t. \]  

(2.515)
2.4.3 Infinitesimal Generators

A fundamental idea in continuous time Markov processes is the infinitesimal generator [Dynkin et al., 2012].

Definition 2.4.6. Let \( f : \mathbb{R}^m \to \mathbb{R}^m \) and let \( X_t \) be a solution to a time-homogenous stochastic differential equation. The infinitesimal generator \( A \) of \( X_t \) is defined by the relation,

\[
Af(x) = \lim_{t \to 0} \frac{\mathbb{E}[f(X_t) | X_0 = x] - f(x)}{t}.
\] (2.516)

Theorem 2.4.7 (Øksendal [1992]). Let \( X_t \) be the solution to the time-homogenous stochastic differential equation

\[
dX^i_t = b^i(X_t) \, dt + \sum_{j=1}^{m} \sigma^i_j(X_t) \, dB^j_t.
\] (2.517)

Define the function \( \Sigma_{ij}(x) = \sum_{k=1}^{m} \sigma_{ik}(x)\sigma_{kj}(x) \). The infinitesimal generator of \( X_t \) is,

\[
Af(x) = \sum_{i=1}^{m} b^i(x) \frac{\partial f}{\partial x_i} (x) + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \Sigma_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} (x).
\] (2.518)

From the infinitesimal generator, we may proceed to deduce the Kolmogorov forward equation, which describes the time evolution of a particle, generated from an initial distribution, under a stochastic differential equation. To establish this result, we require Dynkin’s formula. We begin with a lemma.

Lemma 2.4.8.

\[
\mathbb{E} \left[ \int_0^t A f(X_s) \, ds \bigg| X_0 = x \right] = \int_0^t \mathbb{E} \left[ A f(X_s) \bigg| X_0 = x \right] \, ds.
\] (2.519)
Proof.

\[
E \left[ \int_0^t Af(X_s) \, ds \Big| X_0 = x \right] = E \left[ \lim_{\Delta t \to 0} \sum_k Af(X_{t_k}) \Delta t \Big| X_0 = x \right] \tag{2.520}
\]

\[
= \lim_{\Delta t \to 0} \sum_k E \left[ Af(X_{t_k}) \Big| X_0 = x \right] \Delta t \tag{2.521}
\]

\[
= \int_0^t E \left[ Af(X_s) \Big| X_0 = x \right] \, ds \tag{2.522}
\]

where \(0 = t_1 < t_2 < \cdots < t_{n(\Delta t) - 1} < t_{n(\Delta t)} = t\) are \(\Delta t\)-separated and \(n\) is the fineness of the partition, which depends on the separation.

Proposition 2.4.9 (Dynkin’s Formula). Let \(X_t\) be the solution to the time-homogenous stochastic differential equation

\[
dX^i_t = b^i(X_t) \, dt + \sum_{j=1}^m \sigma^i_j(X_t) \, dB^j_t. \tag{2.523}
\]

Let \(f : \mathbb{R}^m \to \mathbb{R}\) be a smooth function. Then, defining the process \(Z_t = f(X_t)\), we have

\[
E [Z_t | X_0 = x_0] = f(x_0) + \int_0^t E [Af(X_s) | X_0 = x_0] \, ds \tag{2.524}
\]
Proof. From Itô’s lemma (lemma 2.4.5) and the expanding \( dX_t \) via eq. (2.523),

\[
dZ_t = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(X_t) \, dX_t^i + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t) \, dX_t^i \, dX_t^j
\]  
(2.525)

\[
= \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(X_t) \, dt + \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(X_t) \sum_{j=1}^{m} \sigma_j^i(X_t) \, dB_t^j
\]  
(2.526)

\[
+ \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t) \sum_{k=1}^{m} \sigma_k^i(X_t) \, dB_t^k \sum_{l=1}^{m} \sigma_l^j(X_t) \, dB_t^l
\]  
(2.527)

Therefore,

\[
Z_t = Z_0 + \int_0^t \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(X_s) \, ds + \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(X_s) \sum_{j=1}^{m} \sigma_j^i(X_s) \, dB_s^j
\]  
(2.528)

\[
+ \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) \Sigma_{ij}(X_s) \, ds.
\]

Now taking expectations and using the fact that Itô integrals have zero expectation from section 2.4.1,

\[
E[Z_t | X_0 = x] = f(x) + E \left[ \int_0^t A f(X_s) \, ds \bigg| X_0 = x \right]
\]  
(2.529)

\[
+ E \left[ \int_0^t \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(X_t) \sum_{j=1}^{m} \sigma_j^i(X_t) \, dB_t^j \bigg| X_0 = x \right]
\]  
(2.530)

\[
= f(x) + E \left[ \int_0^t A f(X_s) \, ds \bigg| X_0 = x \right]
\]  
(2.531)
Dynkin’s formula in proposition 2.4.9 may also be deduced more generally in the case of stopping times; stopping times are not important for the development here so we eschew discussing these topics further. However, this stronger variation of Dynkin’s formula is carried out in detail in Øksendal [1992].

We now prove the Fokker-Planck equation (also called the Kolmorogov forward equation) [Fokker, 1914, Planck, 1917, Kolmogorov, 1931], which can be considered the stochastic analogue of the continuity equation that was previously derived in the case of deterministic dynamics in corollary 2.2.9. We follow the proof strategy suggested in Øksendal [1992].

Definition 2.4.10. We say that the Itô process $X_t$ has a density $\pi_t : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ with respect to Lebesgue measure if, for every measurable function $f : \mathbb{R}^m \to \mathbb{R}$,

$$
\mathbb{E} [f(X_t) | X_0 = x] = \int_{\mathbb{R}^m} f(y) \pi_t(x, y) \, dy.
$$

In this case, $\pi_t(x, \cdot)$ defines the probability of finding the process in some region at time $t$ given that the initial position of the particle was $x \in \mathbb{R}^m$.

Proposition 2.4.11 (Kolmogorov Forward Equation).

$$
\frac{d}{dt} \pi_t(x, y) = -\sum_{i=1}^{m} \frac{\partial}{\partial y_i} (b_i(y) \pi_t(x, y)) + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2}{\partial y_i \partial y_j} (\Sigma_{ij}(y) \pi_t(x, y)).
$$

Proof. Let $f$ be a compactly supported smooth function on $\mathbb{R}^m$. From the version of Dynkin’s formula in proposition 2.4.9 we have,

$$
\mathbb{E} [f(X_t) | X_0 = x] = f(x) + \int_0^t \mathbb{E} \left[ A f(X_s) \bigg| X_0 = x \right] \, ds.
$$
In the presence of a density, this becomes,

\[
\int_{\mathbb{R}^m} f(y) \pi_t(x, y) \, dy = f(x) + \int_0^t \int_{\mathbb{R}^m} Af(y) \pi_s(x, y) \, dy. \tag{2.535}
\]

By differentiating both sides with respect to time we have,

\[
\int_{\mathbb{R}^m} f(y) \frac{d}{dt} \pi_t(x, y) \, dy = \int_{\mathbb{R}^m} Af(y) \pi_t(x, y) \, dy. \tag{2.536}
\]

By applying integration by parts twice, we have

\[
\int_{\mathbb{R}^m} Af(y) \pi_t(x, y) \, dy = \int_{\mathbb{R}^m} \left( \sum_{i=1}^m b^i(y) \frac{\partial f}{\partial y_i}(y) + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \Sigma_{ij}(y) \frac{\partial^2 f}{\partial y_i \partial y_j}(y) \right) \pi_t(x, y) \, dy \tag{2.537}
\]

\[
= \int_{\mathbb{R}^m} \left( -\sum_{i=1}^m \frac{\partial}{\partial y_i}(x, y)(b^i(y) \pi_t(x, y)) + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \frac{\partial^2}{\partial y_i \partial y_j}(\Sigma_{ij}(y) \pi_t(x, y)) \right) f(y) \, dy. \tag{2.538}
\]

Since \( f \) was arbitrary besides being compactly supported, we conclude,

\[
\frac{d}{dt} \pi_t(x, y) = -\sum_{i=1}^m \frac{\partial}{\partial y_i}(b^i(y) \pi_t(x, y)) + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \frac{\partial^2}{\partial y_i \partial y_j}(\Sigma_{ij}(y) \pi_t(x, y)). \tag{2.539}
\]

\[
\square
\]

### 2.4.4 The Langevin Equation

**Example 26.** An important example of a stochastic differential equation is the *Langevin diffusion*, which, given a smooth probability density \( \pi : \mathbb{R}^m \to \mathbb{R}_+ \) is described by the
equation

\[ dX^i_t = \frac{1}{2} \frac{\partial}{\partial x_i} \log \pi(X_t) \, dt + dB^i_t. \quad (2.541) \]

By inspection, therefore, we identify the drift and diffusion coefficients (defined in definition [2.4.4]) as,

\[ b^i(x) = \frac{1}{2} \frac{\partial}{\partial x_i} \log \pi(X_t) \quad (2.542) \]

\[ \sigma^i_j(x) = \{ i = j \}. \quad (2.543) \]

Inserting these expressions into the Kolmogorov forward equation (proposition [2.4.11]) yields,

\[ \frac{d}{dt} \pi_t(x, y) = -\frac{1}{2} \sum_{i=1}^{m} \frac{\partial}{\partial y_i} \left( \frac{\partial}{\partial y_i} \log \pi(y) \pi_t(x, y) \right) + \frac{1}{2} \sum_{i=1}^{m} \frac{\partial^2}{\partial y_i^2} \pi_t(x, y). \quad (2.544) \]

In the case that \( \pi_t(x, y) = \pi(y) \), then this simplifies as,

\[ \frac{d}{dt} \pi_t(x, y) = -\frac{1}{2} \sum_{i=1}^{m} \frac{\partial}{\partial y_i} \left( \frac{\partial}{\partial y_i} \log \pi(y) \pi(y) \right) + \frac{1}{2} \sum_{i=1}^{m} \frac{\partial^2}{\partial y_i^2} \pi(y) \quad (2.545) \]

\[ = -\frac{1}{2} \sum_{i=1}^{m} \frac{\partial^2}{\partial y_i^2} \pi(y) + \frac{1}{2} \sum_{i=1}^{m} \frac{\partial^2}{\partial y_i^2} \pi(y) \quad (2.546) \]

\[ = 0, \quad (2.547) \]

so that one finds that the distribution \( \pi(y) \) is stationary for the evolution of the Langevin stochastic differential equation.
2.4.5 The Euler-Maruyama Discretization

We next briefly describe methods of numerical integration for stochastic differential equations. The simplest method, and one that is used frequently in Markov chain Monte Carlo methods based on stochastic differential equations, is the Euler-Maruyama integrator [Kloeden and Platen, 2011]. Given the initial data \( x_0 \in \mathbb{R}^m \), let \( X_t \) be the solution of the stochastic differential equation in eq. (2.510) with time-homogenous drift and diffusion coefficients; a single-step of the Euler-Maruyama method computes,

\[
\hat{X}_t^i = \hat{x}_0^i + \sqrt{\epsilon} \sum_{j=1}^{m} \sigma_j^i(x_0) \Delta B_j,
\]

where \( \Delta B_j \sim \text{Normal}(0, 1) \). Iterating this scheme allows one to simulate stochastic differential equations over longer periods.

2.5 Basics of Deterministic and Stochastic Motion on Riemannian Manifolds

Our objective in this section is to give an introduction to differential and Riemannian geometry. From this we hope to indicate the correspondences between deterministic motion and stochastic differential equations within a broader geometric context.

2.5.1 Fundamental Objects in Riemannian Geometry

In section 2.3 we discussed embedded manifolds. In this discussion, we view \( \mathbb{R}^m \) as a manifold but take a more general view by eschewing embedded geometry. In particular, we will formulate more general analogues of concepts we encountered in section 2.3 such as tangent spaces and vector fields, and we will equip manifolds with an additional, specialized structure, called the Riemannian metric that gives rise to additional richness. To
enable a simpler discussion that is relevant to the computational examples given in later chapters, we restrict our attention to the setting wherein a manifold is globally diffeomorphic to $\mathbb{R}^m$.

Consider $\mathbb{R}^m$ endowed with a Riemannian metric $g$; then $M = (\mathbb{R}^m, g)$ is a Riemannian manifold. A diffeomorphism is a smooth, invertible transformation of $M$ to itself with a smooth inverse. As a vector space, an element $x \in \mathbb{R}^m$ can be expressed as $\sum_{i=1}^n x^i e_i$ for coefficients $x^i \in \mathbb{R}$ and linearly independent (standard) basis vectors $e_i \in \mathbb{R}^m$.

**Definition 2.5.1 (Einstein Summation Convention).** In this section, we will adopt the Einstein notation and summarily write $x = x^i e_i$, wherein subscripts are implicitly summed if they should appear as both an upper and lower index.

We will use the notation $\partial_i f \overset{\text{def}}{=} \partial_{x^i} f$ for $f \in C^\infty(\mathbb{R}^m)$ and the notation $(\partial_i)_x \overset{\text{def}}{=} \partial_{x^i}|_x : C^\infty(\mathbb{R}^m) \to \mathbb{R}$ (an operator acting on functions which evaluates the $i$th partial derivative at the point $x \in M$).

**Definition 2.5.2.** Let $M = (\mathbb{R}^m, g)$ be a Riemannian manifold. A tangent space at the
point $x$ is defined to be,

$$T_x M \overset{\text{def.}}{=} \{ a^i (\partial_i)_x : a \in \mathbb{R}^m \}. \quad (2.549)$$

A vector in the tangent space is called a *derivation* as it satisfies linearity and Leibniz’ law (i.e. if $v \in T_x M$ then for functions $f, g \in C^\infty(\mathbb{R}^m)$ we have $v(fg) = fv(g) + gv(f)$).

**Proposition 2.5.3.** The set $((\partial_1)_x, \ldots, (\partial_m)_x)$ are linearly independent.

**Proof.** Consider the operator $O = a^i (\partial_i)_x : C^\infty(\mathbb{R}^m) \to \mathbb{R}$ for $a_1, \ldots, a_m \in \mathbb{R}$. Assume that for every $f \in C^\infty(\mathbb{R}^m)$, we have $O(f) = 0$. We will show that $a_1 = \cdots = a_m = 0$. Consider in particular the function $f_j(x) = x_j$. Hence,

$$O(f_j) = a^i (\partial_i f_j)|_x = a^i \partial_i x_j|_x = a^i \delta_{ij} = a^j. \quad (2.553)$$

But $O(f_j) = a^j = 0$ by assumption. This establishes that $((\partial_1)_x, \ldots, (\partial_m)_x)$ are linearly independent. \hfill \Box

As the $(\partial_i)_x$ are linearly independent, the operators can be viewed as a set of basis vectors and $T_x M$ as a vector space. Moreover:

**Corollary 2.5.4.** $T_x M$ is a vector space with dimension $m$.

**Definition 2.5.5.** The Riemannian metric is a function $g$ which maps points of $M$ to inner products on the tangent spaces; specifically, $g(x) \equiv g_x = \langle \cdot, \cdot \rangle_x : T_x M \times T_x M \to \mathbb{R}$.

By linearity of inner products, it is the action of $g$ on the basis vectors $(\partial_i)_x$ that completely defines the inner product. Indeed:
Proposition 2.5.6. There exist functions $g_{ij} : \mathbb{R}^m \to \mathbb{R}$ such that

$$g_x(u, v) = g_{ij}(x) u^i v^j$$  \hspace{1cm} (2.554)

for $u, v \in T_xM$

Proof. By linearity of the inner product,

$$g_x(u, v) = g_x(\partial^i_x, v^j x (\partial^j_x)) = g_x((\partial^i_x, (\partial^j_x)) u^i v^j. \hspace{1cm} (2.556)$$

Hence, by identification, $g_x((\partial^i_x, (\partial^j_x)) = g_{ij}(x)$.

Definition 2.5.7. The Riemannian metric tensor, a positive definite matrix whose $(i, j)$ entry is specified by $g_{ij}(x) = \langle (\partial^i_x), (\partial^j_x) \rangle_x$.

It is typical to use the notation $g^{ij}$ to denote the $(i, j)$ entry of the inverse matrix of the Riemannian metric tensor. Therefore, using the Einstein summation convention from definition 2.5.1 we have $g^{ki}(x) g_{kl}(x) = \delta_{kl}$.

Definition 2.5.8. Given a vector $v \in T_xM$, the Riemannian length of $v$ is given by $\| v \|_x = \sqrt{\langle v, v \rangle_x} = \sqrt{g_{ij}(x) v^i v^j}$.

Definition 2.5.9. A vector field $Y$ is a smooth map from the manifold to a tangent vector: $x \mapsto Y(x) \in T_xM$. More explicitly, a vector field can be expressed as $Y(x) = Y^i(x)(\partial_i)_x$, where $Y^i \in C^\infty(\mathbb{R}^m)$ are called coordinate functions. Given a smooth function $f : \mathbb{R}^m \to \mathbb{R}$, we denote the action of the vector field $Y$ on $f$ by,

$$Y[f](x) \overset{\text{def}}{=} Y^i(x)(\partial_i f)_x. \hspace{1cm} (2.557)$$

Many objects from differential calculus find their analogues on Riemannian manifolds.
**Definition 2.5.10.** The Riemannian gradient of a function $f \in C^\infty(\mathbb{R}^m)$ is the vector field, denoted $\text{grad} f$ satisfying [Absil et al., 2007],

$$
\partial_i f(x) v^i \overset{\text{def.}}{=} \langle \text{grad} f(x), v \rangle_x
$$

(2.558)

for every $x \in M$ and $v \in T_x M$.

In this discussion we will always denote the Euclidean gradient using $\nabla$ and the Riemannian gradient as $\text{grad}$.

**Proposition 2.5.11.** The Riemannian gradient satisfies,

$$
(\text{grad} f)^i = g^{ij} \partial_j f
$$

(2.559)

**Proof.** From definition 2.5.10 we have,

$$
\partial_x i f(x) v^i = \langle \text{grad} f(x), v \rangle_x
$$

(2.560)

$$
\Rightarrow \partial_i f(x) v^i = (\text{grad} f(x))^i g_{ij} (x) v^j
$$

(2.561)

Now substituting in the claimed expression for $(\text{grad} f(x))^i$ yields,

$$
(\text{grad} f(x))^i g_{ij} (x) v^j = g^{ik}(x) \partial_k f(x) g_{ij} (x) v^j
$$

(2.562)

$$
= g^{ki}(x) \partial_k f(x) g_{ij} (x) v^j
$$

(2.563)

$$
= \delta_{kj} \partial_k f(x) v^j
$$

(2.564)

$$
= \partial_j f(x) v^j
$$

(2.565)

$$
= \partial_i f(x) v^i,
$$

(2.566)

which verifies the claim.

Let $\det(g)$ be a mapping from $x \in M$ to the determinant of Riemannian metric tensor.
Definition 2.5.12. The Riemannian divergence of a vector field is
\[
\text{div}(Y)(x) = \frac{1}{\sqrt{\det(g(x))}} \partial_i \left( \sqrt{\det(g(x))} Y^i(x) \right),
\]
which we see is a mapping \( \text{div}(Y) : M \to \mathbb{R} \).

Definition 2.5.13. The Riemannian generalization of the Laplace operator \( \Delta f \overset{\text{def.}}{=} \sum_{i=1}^m \frac{\partial^2}{\partial x_i^2} f \) is called the Laplace-Beltrami operator and is given by \( \Delta_M f \overset{\text{def.}}{=} \text{div} \ \text{grad} \ f \).

### 2.5.2 Riemannian Measures and Probability Densities

The introduction of a Riemannian metric changes the volume measure associated to \( M \) from the usual definition on \( \mathbb{R}^m \). Let \( \mu \) represent the Riemannian volume measure; in coordinates, the Riemannian volume measure is related to the Lebesgue measure by \( \mu(dx) = \sqrt{\det(g(x))}dx \). The Riemannian volume measure changes probability densities on manifolds by requiring them to integrate to one with respect to the Riemannian volume measure. Denote the set of smooth Riemannian probability densities by \( \mathcal{P}(M) \); for clarity we will denote \( \pi_M \) a density on the manifold and \( \tilde{\pi}_M = \sqrt{\det(g)} \pi_M \) a corresponding density on Euclidean space (this relationship is called the area formula [Federer, 1969]).

Given a density \( \pi_M \in \mathcal{P}(M) \), expectations of functions of random variables become
\[
\mathbb{E}_{x \sim \pi_M} f(x) = \int_M f(x) \pi_M(x) \mu(dx) \overset{(2.568)}{=} \int_{\mathbb{R}^m} f(x) \pi_M(x) \sqrt{\det(g(x))}dx \overset{(2.569)}{=}
\]
An important special case of the Riemannian expectation formula is the choice \( f(x) = \log \pi_M(x) - \log \tilde{\pi}_{M,t}(x) \) for \( \pi_M, \tilde{\pi}_{M,t} \in \mathcal{P}(M) \), which leads to the KL divergence.

Definition 2.5.14 (Riemannian KL Divergence). A divergence measure between densities
on a Riemannian manifold is the Riemannian KL divergence. It is given by,

\[
\mathbb{K}_M(\tilde{\pi}_M \| \pi_M) = \mathbb{E}_{x \sim \tilde{\pi}_M, t} \log \frac{\tilde{\pi}_{M,t}(x)}{\pi_M(x)}
\]  \hspace{1cm} (2.570)

**Theorem 2.5.15** (Riemannian Change-of-Variables [Abraham et al., 1988, Liu and Zhu, 2018]). Let \( \Phi : M \to M \) be a diffeomorphism on \( M \). The density of \( \Phi(x) \) when \( x \sim \pi_M \), which we denote by \( \Phi_# \pi_M \), is given by the Riemannian change-of-variables formula.

\[
\Phi_# \pi_M(x) = \frac{\pi_M(\Phi^{-1}(x)) \cdot \sqrt{\det(g(\Phi^{-1}(x)))}}{\sqrt{\det(g(x))}} \cdot |\det(\nabla_x \Phi^{-1}(x))| \]  \hspace{1cm} (2.571)

**Remark.** one can understand theorem 2.5.15 as saying that if \( x \sim \pi_M \) then the random variable \( \Phi(x) \) has a density function \( \Phi_# \pi_M : M \to \mathbb{R}_+ \); that is, for any Borel set \( A \subseteq \mathbb{R}^m \) the probability that \( \Phi(x) \in A \) is given by,

\[
\Pr(\Phi(x) \in A) = \int_A \Phi_# \pi_M(x) \mu(dx). \]  \hspace{1cm} (2.572)

**Theorem 2.5.16** (Matrix Riemannian Stein’s Identity). Let \( Y \) be a vector field satisfying \( \pi(x) Y^i(x)|_{\partial \mathbb{R}^m} = 0 \) and \( \pi_M(x) \) a density on \( M \). Then we obtain the matrix analogue of Stein’s identity for Riemannian manifolds

\[
\mathbb{E}_{x \sim \pi_M} \left[ Y^i(x) \partial_j \log \pi_M(x) \right] = - \mathbb{E}_{x \sim \pi_M} \left[ \partial_j Y^i(x) - \mathbb{E}_{x \sim \pi_M} \frac{Y^i(x) \partial_j \sqrt{\det(g(x))}}{\sqrt{\det(g(x))}} \right] \]  \hspace{1cm} (2.573)
Proof.

\[
\mathbb{E}_{x \sim \pi_M} \left[ Y^i(x) \partial_j \log \pi_M(x) \right] = \int_M Y^i(x) \partial_j \log(\pi_M(x)) \pi_M(x) \mu(dx) \tag{2.574}
\]

\[
= \int_M Y^i(x) \partial_j \pi_M(x) \frac{\pi_M(x)}{\pi_M(x)} \mu(dx) \tag{2.575}
\]

\[
= \int_M Y^i(x) \partial_j \pi_M(x) \mu(dx) \tag{2.576}
\]

\[
= \int_{\mathbb{R}^m} Y^i(x) \partial_j \pi_M(x) \sqrt{\det(g(x))} dx \tag{2.577}
\]

\[
= - \int_{\mathbb{R}^m} \frac{\sqrt{\det(g(x))}}{\sqrt{\det(g(x))}} \cdot \partial_j \left( Y^i(x) \cdot \sqrt{\det(g(x))} \right) \pi_M(x) dx \tag{2.578}
\]

\[
= - \int_M \frac{\pi_M(x)}{\sqrt{\det(g(x))}} \cdot \partial_j \left( Y^i(x) \cdot \sqrt{\det(g(x))} \right) \mu(dx) \tag{2.579}
\]

\[
= - \int_M \pi_M(x) \cdot \partial_j Y^i(x) + \frac{\pi_M(x)}{\sqrt{\det(g(x))}} \cdot Y^i(x) \cdot \partial_j \sqrt{\det(g(x))} \mu(dx) \tag{2.580}
\]

\[
= - \mathbb{E}_{x \sim \pi_M} \left[ \partial_j Y^i(x) \right] - \mathbb{E}_{x \sim \pi_M} \left[ \frac{Y^i(x) \partial_j \sqrt{\det(g(x))}}{\sqrt{\det(g(x))}} \right] \tag{2.581}
\]

As a matrix, it is possible to consider its matrix trace. Consider moving both terms on the RHS of eq. (2.573) to the LHS so that for all \((i, j)\) the RHS equals zero. By computing the trace of this matrix and identifying terms we arrive at the following corollary.

**Corollary 2.5.17.** Let \(M_{\pi_M}\) be the matrix-valued operator such that

\[
(M_{\pi_M} Y(x))_{ij} = Y^i(x) \partial_j \log \pi_M(x) + \frac{1}{\sqrt{\det(g(x))}} \partial_j \sqrt{\det(g(x))} Y^i(x). \tag{2.582}
\]
Then,

\[
\text{trace} \left( \mathbb{E}_{x \sim \pi_M} M_{\pi_M} Y(x) \right) = \mathbb{E}_{x \sim \pi_M} [Y[\log \pi_M](x) + \text{div}(Y)(x)] \quad (2.584)
\]

\[
= 0. \quad (2.585)
\]

**Proof.** Notice that by the product rule of differentiation we have,

\[
- \mathbb{E}_{x \sim \pi_M} \left[ \partial_j Y^i(x) \right] - \mathbb{E}_{x \sim \pi_M} \left[ \frac{Y^i(x) \partial_j \sqrt{\det(g(x))}}{\sqrt{\det(g(x))}} \right] = - \mathbb{E}_{x \sim \pi_M} \left[ \frac{1}{\sqrt{\det(g(x))}} \partial_j \left( \sqrt{\det(g(x))} Y^i(x) \right) \right] \quad (2.586)
\]

Computing the trace of \((M_{\pi_M} Y)(x)\) has the effect of summing the diagonal entries. Using the Einstein summation convention we have,

\[
\text{trace}(M_{\pi_M} Y(x)) = Y^i(x) \partial_i \log \pi_M(x) + \frac{1}{\sqrt{\det(g(x))}} \partial_i \left( \sqrt{\det(g(x))} Y^i(x) \right). \quad (2.588)
\]

which is the definition of the action of \(Y = Y^i \partial_i\) on \(\log \pi_M\) added to the Riemannian divergence of \(Y\). Taking the expectation and comparing terms with the Riemannian matrix Stein’s identity yields the result. 

\[
\square
\]

### 2.5.3 Riemannian Continuity and Fokker-Planck Equations

The existence of a Riemannian change-of-variables formula enables the study of *continuous time* transformations of densities. When dealing with deterministic dynamics, an important manifold differential equation is the Riemannian continuity equation [Villani 2008, Liu and Zhu 2018].

**Theorem 2.5.18.** Let \(\tilde{\pi}_{M,t}\) be a density of a random variable on \(M\) evolving according to vector field \(Y\). If, moreover, \(A : \mathbb{R} \to M\) is a curve on \(M\) satisfying both \(x = A(0)\) and
\[ \frac{d}{dt} A(t) = Y(A(t)), \] then if \( x \sim \pi_{M,0} \) we may seek the time evolution of the density \( \pi_{M,t} \).

The Riemannian continuity equation states

\[ \frac{d}{dt} \pi_{M,t} = -Y[\pi_{M,t}] - \pi_{M,t} \cdot \text{div}(Y). \] \( (2.589) \)

Notice that the Riemannian continuity equation continues to hold even when the vector field \( Y \) also evolves in time. A stochastic analogue of the continuity equation is the Riemannian Fokker-Planck equation [Qian and Wu, 2013, Risken, 1984]. The Riemannian Fokker-Planck equation describes the time evolution of a density of a random variable described by a stochastic differential equation. In general, a Itô stochastic differential equation (see definition 2.4.4) on the Riemannian manifold \((\mathbb{R}^m, g)\) is specified by

\[ dX_t^i = Y^i(X_t)dt + \sqrt{2} dB_t^i(X_t), \] \( (2.590) \)

where \( Y(x) \in T_xM \) and \( B_t \) is Riemannian manifold Brownian motion, which is expressible in terms of \( n \)-dimensional Euclidean Brownian motion \( B_t \) (see definition 2.4.3) as

\[ dB_t^i(X_t) = \sigma_j^i(X_t) dB_t^j + \frac{1}{2} \left( \frac{1}{\sqrt{\det(g(X_t))}} \partial_j(\sqrt{\det(g(X_t))} \cdot g^{ij}(X_t)) \right) dt \] \( (2.591) \)

The stochastic differential equation for \( \tilde{B}_t \) is constructed such that its infinitesimal generator (see [Oksendal, 1992] and definition 2.4.6) is one-half the Laplace-Beltrami operator (see definition 2.5.13); this is the definition of Brownian motion on Riemannian manifolds [Hsu, 2002]. The notation \( \sigma_j^i \) in eq. (2.591) refers to the \((i, j)\) entry of the square root of the inverse Riemannian metric tensor at the point on the manifold.

**Theorem 2.5.19.** Let \( X_t \) be the solution of the Riemannian Langevin equation in eq. (2.590). Let \( \tilde{\pi}_{M,t} \) denote the density of the randomly evolving particle \( X_t \) at time \( t \); the Riemannian
Fokker-Planck equation states that the time evolution of $\tilde{\pi}_{M,t}$ obeys,

$$\frac{d}{dt} \tilde{\pi}_{M,t} = -Y[\tilde{\pi}_{M,t}] - \tilde{\pi}_{M,t} \cdot \text{div}(Y) + \Delta_M \tilde{\pi}_{M,t}. \quad (2.592)$$

For either the Riemannian continuity equation or the Riemannian Fokker-Planck equation, when the time derivative of the density is zero, the dynamics reach a stationary distribution.

### 2.5.4 Geodesic Motion

**Definition 2.5.20.** Let $g$ be a Riemannian metric and let $\theta, \tilde{\theta} \in \mathbb{R}^m$. Let $\gamma : [a, b] \rightarrow \Theta$ be a differentiable curve connecting $\theta$ and $\tilde{\theta}$ at its endpoints (i.e. $\gamma(a) = \theta$ and $\gamma(b) = \tilde{\theta}$). Then the *Riemannian length* of $\gamma$ is defined by,

$$|\gamma| = \int_a^b \|\gamma'(t)\|_{g(\gamma(t))} \, dt \quad (2.593)$$

where

$$\|v\|_{g(\theta)} = \sqrt{\sum_{i=1}^m \sum_{j=1}^m v_i g_{ij}(\theta) v_j} \quad (2.594)$$

is the Riemannian length from definition 2.5.8.

**Proposition 2.5.21.** The Riemannian length of a curve $\gamma$ does not depend on the parameterization of the curve.

**Proof.** Consider a smooth reparameterization $\omega : [\alpha, \beta] \rightarrow [a, b]$ such that $\omega(\alpha) = a$ and $\omega(\beta) = b$ and for every $t \in [a, b]$ there exists a unique $\tau \in [\alpha, \beta]$ for which $\omega(\tau) = t$. 

122
Then, writing $\kappa(\tau) = \gamma(\omega(\tau))$ we have,

\[
|\kappa| = \int_{\alpha}^{\beta} \|\kappa'(\tau)\| g(\kappa(\tau)) \, d\tau = \int_{\alpha}^{\beta} \|\gamma'(\omega(\tau))\| g(\gamma(\omega(\tau))) \, d\tau = \int_{\alpha}^{\beta} \|\gamma'(\omega(\tau))\| g(\gamma(\omega(\tau))) \omega'(\tau) \, d\tau = \int_{\alpha}^{\beta} \|\gamma'(t)\| g(\gamma(t)) \, dt = |\gamma|.
\]

(2.595)

(2.596)

(2.597)

(2.598)

(2.599)

From our discussion on Lagrangian mechanics, we may interpret eq. (2.593) as an action integral and seek to extremize it.

**Proposition 2.5.22.** Let $g_{ij}$ be a Riemannian metric tensor (definition 2.5.7) and let $g^{ij}$ denote its inverse matrix. The extremal solutions to eq. (2.593) satisfy the equations,

\[
\frac{d^2\gamma_i}{dt^2} = -\Gamma^i_{jk}(\gamma(t)) \frac{d\gamma_j}{dt} \frac{d\gamma_k}{dt}
\]

(2.600)

where

\[
\Gamma^i_{jk}(\theta) = \frac{1}{2} g^{il}(\theta) \left( \frac{\partial g_{lj}}{\partial \theta_k}(\theta) + \frac{\partial g_{lk}}{\partial \theta_j}(\theta) - \frac{\partial g_{jk}}{\partial \theta_l}(\theta) \right)
\]

(2.601)

are the Christoffel symbols.

**Proof.** See [Lee 1997].

Since curves satisfying eq. (2.600) minimize the action integral, which is the length of the curve, they are called geodesic curves. Hamiltonian mechanics can also be used to investigate geodesics.
Theorem 2.5.23. Consider the Hamiltonian 
\[ H(q, p) = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} p_i p_j g^{ij}(q) \]. Then Hamilton’s equations of motion with the canonical symplectic structure correspond to geodesic motion.

Proof. The Hamiltonian equations of motion are,

\[ \dot{q}_a = g^{ai}(q)p_i \] \hspace{1cm} (2.602)

\[ \dot{p}_a = -\frac{1}{2} \sum_{j=1}^{m} p_i p_j \frac{\partial}{\partial q_a} g^{ij}(q) \] \hspace{1cm} (2.603)

\[ = -\frac{1}{2} p_i p_j \frac{\partial}{\partial q_a} g^{ij}(q). \] \hspace{1cm} (2.604)
Therefore,

\[
\ddot{q}_a = g^{aj}(q)\dot{p}_j + \frac{\partial}{\partial q_k} g^{aj}(q) \dot{q}_k \dot{p}_j \quad (2.605)
\]

\[
= g^{aj}(q)\dot{p}_j + \frac{\partial}{\partial q_k} g^{aj}(q) \dot{q}_k g_{jl}(q) \dot{q}_l \quad (2.606)
\]

\[
= -\frac{1}{2} g^{aj}(q) p_i p_k \frac{\partial}{\partial q_j} g^{ik}(q) + \frac{\partial}{\partial q_k} g^{aj}(q) \dot{q}_k g_{jl}(q) \dot{q}_l \quad (2.607)
\]

\[
= -\frac{1}{2} g^{aj}(q) g_{ir}(q) \dot{q}_r g_{ks}(q) \dot{q}_s \frac{\partial}{\partial q_j} g^{ik}(q) + \frac{\partial}{\partial q_k} g^{aj}(q) \dot{q}_k g_{jl}(q) \dot{q}_l \quad (2.608)
\]

\[
= -\frac{1}{2} g^{aj}(q) g_{ir}(q) \dot{q}_r g_{ks}(q) \dot{q}_s \frac{\partial}{\partial q_j} g^{ik}(q) - g^{ax}(q) g^{aj}(q) \frac{\partial}{\partial q_k} g_{xy}(q) \dot{q}_k g_{jl}(q) \dot{q}_l \quad (2.609)
\]

\[
= -\frac{1}{2} g^{aj}(q) g_{ir}(q) \dot{q}_r g_{ks}(q) \dot{q}_s \frac{\partial}{\partial q_j} g^{ik}(q) - g^{ax}(q) \frac{\partial}{\partial q_k} g_{xy}(q) \dot{q}_k \dot{q}_y \quad (2.610)
\]

\[
= -\frac{1}{2} g^{aj}(q) g_{ir}(q) \dot{q}_r g_{ks}(q) \dot{q}_s g^{ix}(q) g^{ky}(q) \frac{\partial}{\partial q_j} g_{xy}(q) - g^{ax}(q) \frac{\partial}{\partial q_k} g_{xy}(q) \dot{q}_k \dot{q}_y \quad (2.611)
\]

\[
= -\frac{1}{2} g^{aj}(q) \frac{\partial}{\partial q_j} g_{xy}(q) \dot{q}_x \dot{q}_y - g^{ax}(q) \frac{\partial}{\partial q_k} g_{xy}(q) \dot{q}_k \dot{q}_y \quad (2.612)
\]

\[
= -\frac{1}{2} g^{aj}(q) \frac{\partial}{\partial q_j} g_{xy}(q) \dot{q}_x \dot{q}_y - g^{aj}(q) \frac{\partial}{\partial q_k} g_{xy}(q) \dot{q}_k \dot{q}_y \quad (2.613)
\]

\[
= -\frac{1}{2} g^{aj}(q) \frac{\partial}{\partial q_j} g_{xy}(q) \dot{q}_x \dot{q}_y - g^{aj}(q) \left( \frac{1}{2} g_{xy}(q) \dot{q}_x \dot{q}_y + \frac{1}{2} \frac{\partial}{\partial q_y} g_{yx}(q) \dot{q}_x \dot{q}_y \right) \quad (2.614)
\]

\[
= -\frac{1}{2} g^{aj}(q) \left( \frac{\partial}{\partial q_j} g_{xy}(q) - \frac{\partial}{\partial q_x} g_{yx}(q) \right) \dot{q}_x \dot{q}_y \quad (2.615)
\]

\[
= -\Gamma^a_{xy}(q) \dot{q}_x \dot{q}_y, \quad (2.616)
\]

which we recognize as the description of a geodesic curve. \(\square\)

Therefore, we see that the equations of motion generated by the Hamiltonian \(H(q, p) = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m p_i p_j g^{ij}(q)\) are geodesics on the manifold. By integrating these equations of motion, we can compute (approximately in the case of numerical integration) manifold geodesics.

**Example 27.** Throughout this work, we will encounter Hamiltonians of the following form. Let \(\pi_E : \mathbb{R}^m \to \mathbb{R}\) be a probability density with respect to Lebesgue measure. Let \(g\) be a Riemannian metric on \(\mathbb{R}^m\); let \(G\) be a matrix whose \((i, j)\)-th entry is \(g_{ij}(q)\). We define the
Riemannian Hamiltonian as a Hamiltonian will the following form,

\[ H(q, p) = -\log \pi_E(q) + \frac{1}{2} \log \text{det}(G(q)) + \frac{1}{2} p^\top G^{-1}(q) p. \] (2.617)

This form of Hamiltonian was first considered for Riemannian manifold Hamiltonian Monte Carlo by Girolami and Calderhead [2011]. The term \( \frac{1}{2} \log \text{det}(G(q)) \) was incorporated such that if we define the density \( \pi_E(q, p) \propto \exp(-H(q, p)) \) (again, with respect to Lebesgue measure), then the conditional distribution of \( p \) given \( q \) is Normal(0, G). However, the discussion in this section allows us to equip this Hamiltonian with a geometric understanding. Namely, the density \( \pi_M(q) = \pi_E(q) / \sqrt{\text{det}(G(q))} \) is the density of the distribution with respect to the Riemannian volume measure. Therefore, we can rewrite \( H(q, p) \) as,

\[ H(q, p) = -\log \pi_M(q) + \frac{1}{2} p^\top G^{-1}(q) p. \] (2.618)

The first term on the right-hand side of eq. (2.618) defines a potential energy that is the negative density of the distribution with respect to the Riemannian volume measure, whereas the second term is the Hamiltonian that generates geodesic motion on the Riemannian manifold \((\mathbb{R}^m, g)\). This interpretation gives geometric insight into the Hamiltonian employed in geometric Markov chain techniques.

### 2.5.5 Correspondence Between Stochastic and Deterministic Dynamics on a Manifold

On a Riemannian manifold, one can identify deterministic initial value problems that produce the same density evolution as a prescribed stochastic differential equation. This is achieved by equating the continuity equation (describing the time evolution for deterministic dynamics) and the Fokker-Planck equation (for stochastic dynamics). Indeed, given
Consider the continuous transformation described by the initial value problem

\[ \dot{x}_t = \text{grad} \log \pi_M(x_t) - \text{grad} \log \tilde{\pi}_{M,t}(x_t). \]  

(2.619)

**Theorem 2.5.24.** Given an initial position \( x_0 \sim \tilde{\pi}_{M,0} \) then the deterministic dynamics eq. (2.619) and the Langevin stochastic dynamics

\[ dX_t = \text{grad} \log \pi_M(X_t) \, dt + \sqrt{2} \, dB_t(X_t) \]  

(2.620)

have precisely the same evolution of the density \( \frac{d}{dt} \tilde{\pi}_{M,t} \). Moreover, the Riemannian dynamics eq. (2.619) are identical to the Euclidean dynamics,

\[ \dot{x}_t = g^{ij}(x_t) \partial_j \log \pi_E(x_t) - g^{ij}(x_t) \partial_j \log \tilde{\pi}_{E,t}(x_t) \]  

(2.621)

where \( g^{ij} \) takes the role of a preconditioner rather than a Riemannian metric tensor. The dynamics eq. (2.621) correspond to the Euclidean stochastic differential equation

\[ dX_t^i = g^{ij}(X_t) \partial_j \log \pi_E(X_t) \, dt + \partial_j g^{ij}(X_t) + \sqrt{2} \sigma^i_j \, dB^j_t(X_t), \]  

(2.622)

whose stationary distribution is \( \pi_E \).

**Proof.** Let \( \sqrt{\det(g(x))} \cdot \tilde{\pi}_{M,t}(x) = \tilde{\pi}_{E,t}(x) \) be the relation between the Riemannian and Euclidean densities. It is immediate that

\[ Y = \text{grad} \log \frac{\pi_M}{\tilde{\pi}_{M,t}} = \text{grad} \log \frac{\pi_E}{\tilde{\pi}_{E,t}} \]  

(2.623)

so that the vector field is “measure invariant” by cancellation of the Riemannian volume.
The Riemannian continuity equation for these dynamics says,

\[
\frac{d}{dt}\tilde{\pi}_{M,t} = -Y[\tilde{\pi}_{M,t}] - \tilde{\pi}_{M,t} \cdot \text{div}_M(Y) \\
= -g^{ij}\partial_j \log \pi_M \partial_i \tilde{\pi}_{M,t} + g^{ij}\partial_j \log \tilde{\pi}_{M,t} \partial_i \tilde{\pi}_{M,t} \\
- \tilde{\pi}_{M,t} \cdot (\Delta_M \log \pi_M - \Delta_M \log \tilde{\pi}_{M,t}).
\]

(2.625)

The Riemannian Fokker-Planck equation for the Riemannian Langevin diffusion is,

\[
\frac{d}{dt}\tilde{\pi}_{M,t} = -\text{div}_M(\tilde{\pi}_{M,t} \cdot \text{grad} \log \pi_M) + \Delta_M \tilde{\pi}_{M,t} \\
= -\text{grad} \log \pi_M[\tilde{\pi}_{M,t}] - \tilde{\pi}_{M,t} \cdot \Delta_M \log \pi_M + \Delta_M \tilde{\pi}_{M,t} \\
= -\text{grad} \log \pi_M[\tilde{\pi}_{M,t}] + \text{grad} \log \tilde{\pi}_{M,t}[\tilde{\pi}_{M,t}] \\
- \tilde{\pi}_{M,t} \cdot \Delta_M \log \pi_M + \tilde{\pi}_{M,t} \cdot \Delta_M \log \tilde{\pi}_{M,t}.
\]

(2.628)

This proves the equivalence of the deterministic dynamics and the Riemannian Langevin diffusion. The equivalence of the Riemannian Langevin diffusion and the Euclidean diffusion in eq. (2.622) was proven in [Xifara et al., 2014]. That this Euclidean diffusion has the same dynamics as \( \dot{x}_t = g^{ij}(x_t)\partial_j \log \pi_E(x_t) \) was established in [Liu et al., 2019]. Finally, it is clear that when \( \tilde{\pi}_{M,t} = \pi_M \) the Riemannian continuity equation for the dynamics is \( \frac{d}{dt}\tilde{\pi}_{M,t} = 0 \); this establishes that \( \pi_M \) is the stationary distribution under the Riemannian volume measure.

Example 28 (Hoffman and Ma [2019]). As a specific example, we consider the special case wherein \( g_{ij}(x) = \delta_{ij} \). In this case, the Langevin diffusion assumes the simple form

\[
dX_t = \nabla \log \pi(X_t) \, dt + \sqrt{2} \, dB_t
\]

(2.629)

where \( \pi : \mathbb{R}^m \rightarrow \mathbb{R}_+ \) is a probability density defined with respect to Lebesgue measure and \( B_t \) is \( m \)-dimensional Brownian motion. According to theorem [2.5.24] the evolution of
$X_t$ is precisely the same density evolution as the deterministic dynamics

$$
\dot{x}_t = \nabla \log \pi(x_t) - \nabla \log \tilde{\pi}_{E,t}(x_t).
$$

(2.630)

The significance of theorem 2.5.24 is that it connects the Riemannian deterministic analogue of a Langevin diffusion to deterministic dynamics on Euclidean space. Then, in Euclidean space, the dynamics correspond to a (non-Langevin) stochastic differential equation. Indeed, eq. (2.622) is the same as that appearing in [Ma et al., 2015] modulo a skew-symmetric matrix. That eq. (2.622) has $\pi_E$ as its stationary distribution is a direct consequence of the reformulation of the Euclidean Fokker-Planck equation in [Shi et al., 2012]. These relationships are summarized in fig. 2.1. A direct relationship (without an intermediate discussion of deterministic dynamics) between eq. (2.620) and eq. (2.622) can be found in [Xifara et al., 2014].

A consequential object of study is the rate at which the KL divergence decreases. It was proven in [Liu and Wang, 2016] that for Euclidean spaces we have the following relationship for densities evolving under vector field dynamics.

**Definition 2.5.25 (Stein’s Operator).** Let $\pi_E(x)$ be a smooth density on $\mathbb{R}^m$ and $Y : \mathbb{R}^m \to \mathbb{R}^m$ a smooth function. The Stein operator is $S_{\pi_E} Y(x) = Y(x) \nabla_x \log \pi_E(x)^\top + \nabla_x Y(x)$, a map from $\mathbb{R}^m$ to $\mathbb{R}^{n \times n}$.

**Theorem 2.5.26 (KL Rate for Euclidean Spaces).** Let $M = (\mathbb{R}^m, \text{Id})$. Then if $Y \in TM$ and $\tilde{\pi}_{E,t}$ evolves under $Y$ we have,

$$
\frac{d}{dt} \mathbb{KL}(\tilde{\pi}_{E,t} \parallel \pi_E) = -\text{trace} \left( \mathbb{E}_{x \sim \tilde{\pi}_{E,t}} [S_{\pi_E} Y(x)] \right)
$$

(2.631)

The corresponding result for vector fields on Riemannian manifolds was proven in [Liu and Zhu, 2018].
Theorem 2.5.27 (KL Rate for Riemannian Manifolds). Let \((M, g)\) be a Riemannian manifold (with or without a global coordinate system) and let \(Y\) be a vector field on \(M\). Then if \(\tilde{\pi}_{M,t}\) evolves under \(Y\) we have

\[
\frac{d}{dt} \KL(\tilde{\pi}_{M,t} \| \pi_M) = - \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} [Y \log \pi](x) + \text{div}(Y)(x) \tag{2.632}
\]

\[
= -\text{trace} \left( \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} M_{\pi_M} Y(x) \right). \tag{2.633}
\]

This theorem allows us to analyze the Riemannian Langevin diffusion eq. (2.619) in terms of the Riemannian KL divergence rate.

Corollary 2.5.28. Let \(Y = \text{grad} \log \pi_M - \text{grad} \log \tilde{\pi}_{M,t}\). Then,

\[
\frac{d}{dt} \KL(\tilde{\pi}_{M,t} \| \pi_M) = - \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} [\text{grad} \log \pi_M \log \pi](x) - \text{grad} \log \tilde{\pi}_{M,t} \log \pi_M](x)
\]

\[
- \text{grad} \log \pi_M \log \tilde{\pi}_{M,t}](x) + \text{grad} \log \tilde{\pi}_{M,t} \log \tilde{\pi}_{M,t}](x)
\]

\[
def = -F_M(\tilde{\pi}_{M,t}, p) \tag{2.635}
\]

Proof. The most important step to verifying this result is the following integration by
\[
\mathbb{E}_{x \sim \tilde{\pi}_M,t} \Delta_M \log \pi_M = \int_{\mathbb{R}^m} \Delta_M \log \pi(x) \sqrt{\det(g(x))} \tilde{\pi}_{M,t}(x) \, dx 
\]
\[
= \int_{\mathbb{R}^m} \partial_i(\sqrt{\det(g(x))} g^{ij}(x) \partial_j \log \pi_M(x)) \tilde{\pi}_{M,t}(x) \, dx \quad (2.636)
\]
\[
= - \int_{\mathbb{R}^m} \sqrt{\det(g(x))} g^{ij}(x) \partial_j \log \pi_M(x) \partial_i \tilde{\pi}_{M,t}(x) \, dx \quad (2.637)
\]
\[
= - \int_{\mathbb{R}^m} \sqrt{\det(g(x))} g^{ij}(x) \partial_j \log \pi_M(x) \partial_i \log \tilde{\pi}_{M,t}(x) \tilde{\pi}_{M,t}(x) \, dx 
\]
\[
= - \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} [g^{ij}(x) \partial_j \log \pi_M(x) \partial_i \log \tilde{\pi}_{M,t}(x)] 
\]
\[
= - \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} \log \pi_M[x] \log \tilde{\pi}_{M,t}[x] 
\]
\[
= - \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} \log \pi_M[\log \tilde{\pi}_{M,t}](x) 
\]

Now, taking \( Y = \log \pi_M - \log \tilde{\pi}_{M,t} \), we obtain from theorem \[ \ref{thm:fisher_divergence} \]
\[
\frac{d}{dt} KL(\tilde{\pi}_{M,t}\|\pi_M) = - \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} \log \pi_M[x] \log \pi_M(x) - \log \tilde{\pi}_{M,t}[\log \pi_M](x) 
\]
\[
+ \Delta_M \log \pi_M(x) - \Delta_M \tilde{\pi}_{M,t}(x) 
\]
\[
= - \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} \log \pi_M[x] \log \pi_M(x) - \log \tilde{\pi}_{M,t}[\log \pi_M](x) 
\]
\[
- \log \pi_M[\log \tilde{\pi}_{M,t}](x) + \log \tilde{\pi}_{M,t}[\log \tilde{\pi}_{M,t}](x) 
\]

\( \square \)

The definition in eq. \[ (2.635) \] is a Riemannian generalization of the Fisher divergence on Euclidean space, despite the rather distinct appearance at first glance. The connection can be clearly seen by examining the expected squared length of \( \dot{x} \) in the tangent spaces of \( M \). Indeed, the rate of decrease of the KL divergence can be additionally characterized by the expected magnitude of the transformation. Intuitively, a closely matched (in terms of KL divergence) approximation \( \tilde{\pi}_{M,t} \) to \( \pi_M \) should require “smaller” transformations. The precise relationship is as follows.
Proposition 2.5.29.

\[
\mathbb{E}_{x \sim \bar{\pi}_{M,t}} \|\dot{x}\|_x^2 = \mathbb{E}_{x \sim \bar{\pi}_{M,t}} [\text{(grad log } \pi_M - \text{grad log } \bar{\pi}_{M,t})] \text{log } \pi_M(x) \\
+ \Delta_M \log \pi_M(x) - \Delta_M \log \bar{\pi}_{M,t}(x) \\
= \mathbb{F}_M(\bar{\pi}_{M,t}, \pi_M)
\]

(2.644)

and hence \[ \frac{d}{dt} \mathcal{KL}(\bar{\pi}_{M,t} || \pi_M) = - \mathbb{E}_{x \sim \bar{\pi}_{M,t}} [\|\dot{x}\|_x^2]. \]

Proof. By definition \( \dot{x} = \text{grad log } \pi_M - \text{grad log } \bar{\pi}_{M,t} \). So by direct calculation,

\[
\mathbb{E}_{x \sim \bar{\pi}_t} \|\dot{x}\|_x^2 = \mathbb{E}_{x \sim \bar{\pi}_t} \langle \dot{x}, \dot{x} \rangle_x \\
= \mathbb{E}_{x \sim \bar{\pi}_t} \langle g^{ij} (\partial_j \log \frac{\pi_M}{\bar{\pi}_t}) \partial_i, g^{ij} (\partial_j \log \frac{\pi_M}{\bar{\pi}_t}) \partial_i \rangle_x \\
(2.646)
\]

Furthermore, let’s examine the expression for the magnitude of the particle velocity further.

\[
\mathbb{E}_{x \sim \bar{\pi}_{M,t}} \langle g^{ij} \partial_j \log \pi_M \partial_i, g^{ij} \partial_j \log \pi_M \partial_i \rangle_x = \mathbb{E}_{x \sim \bar{\pi}_{M,t}} g^{ij}(x) \partial_j \log \pi_M(x) \partial_i \log \pi_M(x) \\
= \mathbb{E}_{x \sim \bar{\pi}_{M,t}} \text{grad log } \pi_M[\text{log } \pi_M](x) \\
(2.648)
\]

(2.649)

(2.650)
and

$$\mathbb{E}_{x \sim \tilde{\pi}_{M,t}} \langle g^{ij} \partial_j \log \pi_M \partial_i, g^{ij} \partial_j \log \tilde{\pi}_{M,t} \partial_i \rangle_x$$  \hspace{1cm} (2.651)

$$= \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} g^{ij}(x) \partial_j \log \pi_M(x) \partial_i \log \tilde{\pi}_{M,t}(x)$$  \hspace{1cm} (2.652)

$$= \int_{\mathbb{R}^m} g^{ij}(x) \partial_j \log \pi_M(x) \partial_i \tilde{\pi}_{M,t}(x) \mu(dx)$$  \hspace{1cm} (2.653)

$$= - \int_{\mathbb{R}^m} \frac{\tilde{\pi}_{M,t}(x)}{\sqrt{\det(g(x))}} \partial_i \left( \sqrt{\det(g(x))} g^{ij}(x) \partial_j \log \pi_M(x) \right) \mu(dx)$$  \hspace{1cm} (2.654)

$$= - \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} \Delta_M \log \pi_M(x)$$  \hspace{1cm} (2.655)

From here it follows:

$$\mathbb{E}_{x \sim \tilde{\pi}_{M,t}} \left\| \dot{x} \right\|^2_x = \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} \langle g^{ij} \partial_j \log \pi_M \partial_i, g^{ij} \partial_j \log \pi_M \partial_i \rangle_x$$

$$- \langle g^{ij} \partial_j \log \pi_M \partial_i, g^{ij} \partial_j \log \tilde{\pi}_{M,t} \partial_i \rangle_x$$

$$- \langle g^{ij} \partial_j \log \tilde{\pi}_{M,t} \partial_i, g^{ij} \partial_j \log \pi_M \partial_i \rangle_x$$

$$+ \langle g^{ij} \partial_j \log \tilde{\pi}_{M,t} \partial_i, g^{ij} \partial_j \log \tilde{\pi}_{M,t} \partial_i \rangle_x$$

$$= \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} [(\text{grad log } \pi_M - \text{grad log } \tilde{\pi}_{M,t})[\log \pi_M](x)$$

$$+ \Delta_M \log \pi_M(x) - \Delta_M \log \tilde{\pi}_{M,t}(x)]$$  \hspace{1cm} (2.656)

$$= \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} [\text{grad log } \pi_M - \text{grad log } \tilde{\pi}_{M,t}][\log \pi_M](x)$$

So we obtain the interesting identity:

$$\frac{d}{dt} \mathbb{KL}(\tilde{\pi}_{M,t} \| \pi_M) = - \mathbb{E}_{x \sim \tilde{\pi}_{M,t}} \left\| \dot{x} \right\|^2_x$$  \hspace{1cm} (2.657)

which connects the physical particle motion to the information-theoretic evolution of the distributions.
2.6 Information Geometry

The field of information geometry was pioneered by Shun-ichi Amari; see *inter alia* Amari and Nagaoka [2000], Amari [2016]. This field treats the problem of endowing the space of probability distributions with a Riemannian manifold structure. Typically, the structure of the manifold has a global coordinate system, such as that considered in section 2.5. The metric assigned to the Riemannian manifold is typically chosen as the Fisher information metric, which we now describe.

Throughout this chapter we will refer to several concepts repeatedly, which we now define.

*Definition* 2.6.1. Let $\Theta$ be an open subset of $\mathbb{R}^m$. A family of probability measures on $\mathbb{R}^k$ is said to be indexed by $\theta$ if there is a bijection between $\Theta$ and the family of probability measures; in this case, we write a member of the family of probability measures as $\Pi_{\theta} : \mathcal{B}(\mathbb{R}^k)$ to emphasize the relation to $\theta \in \Theta$.

*Example* 29. Consider the case of normal distributions which are parameterized by their mean and variance. In this case, take $\Theta = \mathbb{R} \times \mathbb{R}_+$. For $\theta = (\mu, \sigma^2)$, there is a unique probability measure on $\mathbb{R}$ corresponding to a normal distribution with mean $\mu$ and variance $\sigma^2$. This example also illustrates the distinction between the dimension of the parameter space, which is $m = 2$ in this case since $\Theta$ is an open subset of $\mathbb{R}^2$, and the dimension of the space in which the probability measure lives (for normal distributions we have $k = 1$).

*Definition* 2.6.2. A probability measure $\Pi_{\theta} : \mathcal{B}(\mathbb{R}^m) \to [0, 1]$ is said to have a density $\pi_{\theta} : \mathbb{R}^m \to \mathbb{R}_+$ if for every $A \in \mathcal{B}(\mathbb{R}^m)$ we have $\Pi_{\theta}(A) = \int_A \pi_{\theta}(x) \, dx$.

*Remark.* In the sequel, we will assume that if $\{\Pi_{\theta} : \theta \in \Theta\}$ is a family of probability measures indexed by $\Theta$ and that each $\Pi_{\theta}$ has a density $\pi_{\theta}$, then $\pi_{\theta}(x)$ is differentiable with respect to $\theta$ for every $x \in \mathbb{R}^k$. 134
2.6.1 The Fisher Information Metric

Definition 2.6.3. Let $\Pi_\theta$ be a probability measure on $\mathbb{R}^k$ with density $\pi_\theta : \mathbb{R}^k \to \mathbb{R}_+$ with respect to Lebesgue measure. The log-density of the distribution is denoted by $\ell_\theta : x \mapsto \log \pi_\theta(x)$.

Definition 2.6.4. Let $\Pi_\theta$ be a probability measure on $\mathbb{R}^k$ with density $\pi_\theta : \mathbb{R}^k \to \mathbb{R}_+$ where $\theta \in \Theta \subset \mathbb{R}^m$. The Fisher information matrix at $\theta$ is the matrix whose $(i, j)$-th element is,

$$ g_{ij}(\theta) = \mathbb{E}_{x \sim \pi_\theta} \left[ \frac{\partial \ell_\theta(x)}{\partial \theta_i} \frac{\partial \ell_\theta(x)}{\partial \theta_j} \right]. \quad (2.659) $$

Proposition 2.6.5. The Fisher information matrix is positive semi-definite.

Proof. Let $v \in \mathbb{R}^m$ be arbitrary. We have

$$ \sum_{i=1}^m \sum_{j=1}^m g_{ij}(\theta) v_i v_j = \sum_{i=1}^m \sum_{j=1}^m \mathbb{E}_{x \sim \pi_\theta} \left[ \frac{\partial \ell_\theta(x)}{\partial \theta_i} \frac{\partial \ell_\theta(x)}{\partial \theta_j} \right] v_i v_j \quad (2.660) $$

$$ = \mathbb{E}_{x \sim \pi_\theta} \left[ \left( \sum_{i=1}^m \frac{\partial \ell_\theta(x)}{\partial \theta_i} v_i \right) \left( \sum_{j=1}^m \frac{\partial \ell_\theta(x)}{\partial \theta_j} v_j \right) \right] \quad (2.661) $$

$$ = \mathbb{E}_{x \sim \pi_\theta} \left[ \left( \sum_{i=1}^m \frac{\partial \ell_\theta(x)}{\partial \theta_i} v_i \right)^2 \right] \quad (2.662) $$

$$ \geq 0. \quad (2.663) $$

Proposition 2.6.6. Let $\Pi_\theta$ be a probability measure on $\mathbb{R}^k$ with density $\pi_\theta : \mathbb{R}^k \to \mathbb{R}_+$ where $\theta \in \Theta \subset \mathbb{R}^m$. Let $\psi : \mathbb{R}^m \to \mathbb{R}^m$ be a diffeomorphism. Let $\theta = \psi(\xi)$ and consider the reparameterized representation $\pi_\xi$ with density $\pi_\xi \equiv \pi_{\psi(\xi)}$ so that $\ell_\xi(x)$ (in the $\xi$-parameterization) equals $\ell_{\psi(\xi)}(x)$ (in the $\theta$-parameterization). The Fisher information in
the $\xi$-parameterization has the form,

$$g_{ij}(\xi) = \sum_{k=1}^{m} \sum_{l=1}^{m} \frac{\partial \theta_k}{\partial \xi_i} \frac{\partial \theta_l}{\partial \xi_j} g_{kl}(\psi(\xi)).$$  \hspace{1cm} (2.664)

**Proof.**

$$g_{ij}(\xi) = \mathbb{E}_{x \sim \pi_{\xi}} \left[ \frac{\partial \ell_\xi(x)}{\partial \xi_i} \frac{\partial \ell_\xi(x)}{\partial \xi_j} \right]$$

$$= \mathbb{E}_{x \sim \pi_{\xi}} \left[ \left( \sum_{k=1}^{m} \frac{\partial \ell_\theta(x)}{\partial \theta_k} \frac{\partial \theta_k}{\partial \xi_i} \right) \left( \sum_{l=1}^{m} \frac{\partial \ell_\theta(x)}{\partial \theta_l} \frac{\partial \theta_l}{\partial \xi_j} \right) \right]$$

$$= \sum_{k=1}^{m} \sum_{l=1}^{m} \frac{\partial \theta_k}{\partial \xi_i} \frac{\partial \theta_l}{\partial \xi_j} \mathbb{E}_{x \sim \pi_{\xi}} \left[ \frac{\partial \ell_\theta(x)}{\partial \theta_k} \frac{\partial \ell_\theta(x)}{\partial \theta_l} \right]$$

$$= \sum_{k=1}^{m} \sum_{l=1}^{m} \frac{\partial \theta_k}{\partial \xi_i} \frac{\partial \theta_l}{\partial \xi_j} g_{kl}(\psi(\xi)).$$ \hspace{1cm} (2.665, 2.666, 2.667, 2.668)

In eq. (2.666) we have invoked the product rule of differentiation as follows: If we let $
\theta = \psi(\xi)$ then,

$$\frac{\partial \ell_\xi(x)}{\partial \xi_i} = \frac{\partial \ell_\psi(\xi)}{\partial \xi_i}(x)$$

$$= \sum_{k=1}^{m} \frac{\partial \ell_\psi(\xi)}{\partial \theta_k}(x) \frac{\partial \theta_k}{\partial \xi_i}$$

$$= \sum_{k=1}^{m} \frac{\partial \ell_\theta(x)}{\partial \theta_k} \frac{\partial \theta_k}{\partial \xi_i}.$$ \hspace{1cm} (2.669, 2.670, 2.671)

\[\square\]

### 2.6.2 The Geometric Properties of Fisher Information

There are at least two important respects in which the Fisher information reflects the second-order geometry of the distribution. In the first case, we find that the Fisher information is the negative expected value of the Hessian of the distribution. In the second
case, we find that the Taylor expansion of the KL divergence to second-order is a quadratic form in the Fisher information. We now prove both of these statements.

**Proposition 2.6.7.** Let $\Pi_\theta$ be a probability measure on $\mathbb{R}^k$ with density $\pi_\theta: \mathbb{R}^k \to \mathbb{R}^+$ where $\theta \in \Theta \subset \mathbb{R}^m$. Suppose further that $\lim_{\|x\| \to \infty} \pi_\theta(x) = 0$ Then,

$$g_{ij}(\theta) = -\mathbb{E}_{x \sim \pi_\theta} \left[ \frac{\partial^2 \ell_\theta}{\partial \theta_i \partial \theta_j}(x) \right].$$

(2.672)

**Proof.**

$$g_{ij}(\theta) = \mathbb{E}_{x \sim \pi_\theta} \left[ \frac{\partial \ell_\theta}{\partial \theta_i}(x) \frac{\partial \ell_\theta}{\partial \theta_j}(x) \right]$$

(2.673)

$$= \int_{\mathbb{R}^m} \left( \frac{\partial \ell_\theta}{\partial \theta_i}(x) \frac{\partial \ell_\theta}{\partial \theta_j}(x) \right) \pi_\theta(x) \, dx$$

(2.674)

$$= \int_{\mathbb{R}^m} \left( \frac{\partial}{\partial \theta_i} \log \pi_\theta(x) \frac{\partial \ell_\theta}{\partial \theta_j}(x) \right) \pi_\theta(x) \, dx$$

(2.675)

$$= \int_{\mathbb{R}^m} \left( \frac{1}{\pi_\theta(x)} \frac{\partial}{\partial \theta_i} \pi_\theta(x) \frac{\partial \ell_\theta}{\partial \theta_j}(x) \right) \pi_\theta(x) \, dx$$

(2.676)

$$= \int_{\mathbb{R}^m} \frac{\partial \pi_\theta}{\partial \theta_i}(x) \frac{\partial \ell_\theta}{\partial \theta_j}(x) \, dx$$

(2.677)

$$= \pi_\theta(x) \frac{\partial \ell_\theta}{\partial \theta_i}(x) \bigg|_{x \in \partial \mathbb{R}^m} - \int_{\mathbb{R}^m} \pi_\theta(x) \frac{\partial^2 \ell_\theta}{\partial \theta_i \partial \theta_j}(x) \, dx$$

(2.678)

$$= -\mathbb{E}_{x \sim \pi_\theta} \left[ \frac{\partial^2 \ell_\theta}{\partial \theta_i \partial \theta_j}(x) \right].$$

(2.679)

□

**Definition 2.6.8.** Let $\Pi$ and $\tilde{\Pi}$ be two probability measures on $\mathbb{R}^m$ with densities $\pi$ and $\tilde{\pi}$ with respect to Lebesgue measure, respectively. Then the KL divergence, or relative entropy, is defined by,

$$\text{KL}(\pi \| \tilde{\pi}) = \mathbb{E}_{x \sim \pi} \log \frac{\pi(x)}{\tilde{\pi}(x)}.$$  

(2.680)

**Proposition 2.6.9.** Let $\Pi_\theta$ be a probability measure on $\mathbb{R}^k$ with density $\pi_\theta: \mathbb{R}^k \to \mathbb{R}^+$
parameterized by $\theta \in \Theta \subset \mathbb{R}^m$. Let $\Pi_{\tilde{\theta}}$ be another probability measure with density $\pi_{\tilde{\theta}}$ where $\tilde{\theta} \in \Theta$. Then,

$$
\mathbb{K}L(\Pi_\theta \| \Pi_{\tilde{\theta}}) = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} (\theta_i - \tilde{\theta}_i) g_{ij}(\theta)(\theta_j - \tilde{\theta}_j) + \mathcal{O}(\|\theta - \tilde{\theta}\|^3).
$$

(2.681)

**Proof.** The first-order term in the Taylor series is,

$$
\nabla_{\tilde{\theta}} \mathbb{K}L(\Pi_\theta \| \Pi_{\tilde{\theta}}) \bigg|_{\tilde{\theta} = \theta} = -\nabla_{\tilde{\theta}} \int_{\mathbb{R}^m} \pi_{\theta}(x) \log \pi_{\tilde{\theta}}(x) \, dx \bigg|_{\tilde{\theta} = \theta} = -\int_{\mathbb{R}^m} \pi_{\theta}(x) \nabla_{\tilde{\theta}} \log \pi_{\tilde{\theta}}(x) \, dx \bigg|_{\tilde{\theta} = \theta} = -\int_{\mathbb{R}^m} \pi_{\theta}(x) \nabla_{\theta} \log \pi_{\theta}(x) \, dx = 0.
$$

(2.682)

The second-order term in the Taylor series is,

$$
\nabla_{\tilde{\theta}}^2 \mathbb{K}L(\Pi_\theta \| \Pi_{\tilde{\theta}}) \bigg|_{\tilde{\theta} = \theta} = -\nabla_{\tilde{\theta}}^2 \int_{\mathbb{R}^m} \pi_{\theta}(x) \log \pi_{\tilde{\theta}}(x) \, dx \bigg|_{\tilde{\theta} = \theta} = -\int_{\mathbb{R}^m} \pi_{\theta}(x) \nabla_{\tilde{\theta}}^2 \log \pi_{\tilde{\theta}}(x) \, dx \bigg|_{\tilde{\theta} = \theta} = -\int_{\mathbb{R}^m} \pi_{\theta}(x) \nabla_{\theta}^2 \log \pi_{\theta}(x) \, dx = -\mathbb{E}_{x \sim \pi} \nabla_{\theta}^2 \log \pi_{\theta}(x).
$$

(2.683)

Therefore, the Taylor series expansion of the KL divergence is,

$$
\mathbb{K}L(\Pi_\theta \| \Pi_{\tilde{\theta}}) = -\frac{1}{2} (\theta - \tilde{\theta})^\top \left( \mathbb{E}_{x \sim \pi} \nabla_{\theta}^2 \log \pi_{\theta}(x) \right) (\theta - \tilde{\theta}) + \mathcal{O}(\|\theta - \tilde{\theta}\|^3)
$$

(2.689)

$$
= \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} (\theta_i - \tilde{\theta}_i) g_{ij}(\theta)(\theta_j - \tilde{\theta}_j) + \mathcal{O}(\|\theta - \tilde{\theta}\|^3).
$$

(2.690)

□

138
2.6.3 The Natural Gradient Method

The inverse Riemannian metric also functions as a preconditioner in Riemannian spaces, fulfilling a role similar to the inverse Hessian in Newton’s method. When maximizing a function, the natural gradient gives the direction of greatest increase with respect to the Riemannian manifold structure.

**Definition 2.6.10.** Let \( L : \mathbb{R}^m \rightarrow \mathbb{R} \) be a smooth function. Let \( g \) be a Riemannian metric on \( \mathbb{R}^m \) and denote its inverse by \( g^{-1} \). The natural gradient is defined by

\[
\tilde{\partial} L(\theta) = \sum_{j=1}^{m} g^{-1}_{ij}(\theta) \frac{\partial L}{\partial \theta_j}(\theta).
\] (2.692)

**Proposition 2.6.11 (Amari [1998]).** The natural gradient is the direction of greatest increase in \( L \) with respect to the Riemannian metric in the following sense. Denote by \( \nabla_{\delta} L(\theta) = \lim_{\epsilon \to 0} \frac{L(\theta + \epsilon \delta) - L(\theta)}{\epsilon} \) the directional derivative of \( L \) in the direction of \( \delta \). Of all unit vectors (in the Riemannian norm), the one maximizing the directional derivative is proportional to the natural gradient.

**Proof.** Expand \( L(\theta + \epsilon \delta) \) in a Taylor series in \( \epsilon \)

\[
L(\theta + \epsilon \delta) = L(\theta) + \epsilon \nabla L(\theta) \cdot \delta + \mathcal{O}(\epsilon^2).
\] (2.693)

Rearranging terms and passing to the limit, we obtain,

\[
\lim_{\epsilon \to 0} \frac{L(\theta + \epsilon \delta) - L(\theta)}{\epsilon} = \nabla L(\theta) \cdot \delta.
\] (2.694)

Hence \( \nabla_{\delta} L(\theta) = \nabla L(\theta) \cdot \delta \). We wish to maximize \( \nabla_{\delta} L(\theta) \) with respect to \( \delta \) subject to the constraint that \( \sum_{i=1}^{m} \sum_{j=1}^{m} \delta_i g_{ij}(\theta) \delta_j = 1 \). Introducing the Lagrange multiplier \( \lambda \) we
obtain the optimality condition,

\[ \nabla L(\theta) - 2\lambda G(\theta)\delta = 0 \]  

(2.695)

\[ \Rightarrow \delta = \frac{1}{2\lambda} G^{-1}(\theta) \nabla L(\theta), \]  

(2.696)

where we have written \( G(\theta) \) to mean the matrix whose \((i, j)\)-th element is \( g_{ij}(\theta) \). By inspection, one sees that the \( i \)-th element of the optimal \( \delta \) is proportional to the natural gradient, with \( \lambda \) being equal to one-half of \( \| G^{-1}(\theta) \nabla L(\theta) \| \) so that \( \delta \) is a unit vector. Introducing this value of delta into the directional derivative yields,

\[ \nabla_{\delta} L(\theta) = \frac{1}{2\lambda} \nabla L(\theta) \cdot G^{-1}(\theta) \nabla L(\theta) \]  

(2.697)

\[ \geq 0, \]  

(2.698)

which confirms that the natural gradient is indeed the direction of greatest increase. \( \square \)

This fact motivates the use of the natural gradient in optimization schemes, such as Fisher scoring. Indeed, a sequence of estimators defined recursively through the formula,

\[ \theta_{n+1} = \theta_n + G^{-1}(\theta_n) \nabla \ell(\theta_n) \]  

(2.699)

has been used to maximize the log-likelihood of numerous statistical models, such as being the default procedure for finding the maximum likelihood estimate in generalized linear models in \( \mathbb{R} \).

The choice of a Riemannian metric endows the space \( \mathbb{R}^m \) with a distance function, called the Riemannian distance. This distance was investigated in the case of the Fisher information metric in [Rao 1992].
2.6.4 The SoftAbs Metric

It may come to pass that the Fisher information metric is not feasible to use in a computational setting for a number of reasons. For instance, it may be that its definition given in definition 2.6.4 involves an integral over $\mathbb{R}^m$ that cannot be exactly computed. Alternatively, one may find that in attempting to invoke proposition 2.6.7 that the expected Hessian of the log-density may not be convex, in which case it is inadmissible as a metric. For these reasons, it has been of computational interest to seek a general procedure by which one may produce a metric capturing second-order geometry of a target distribution. These considerations led Betancourt [2012] to develop the SoftAbs metric, which is constructed as follows.

Let $\pi : \mathbb{R}^m \to \mathbb{R}_+$ be a probability density. Let $H(\theta) = \nabla^2 \log \pi(\theta)$ be the Hessian of the log-density. Because the Hessian is a real, symmetric matrix, it has an eigen-decomposition,

$$H(\theta) = Q(\theta)\Lambda(\theta)Q(\theta)^T,$$  \hspace{1cm} (2.700)

where $Q(\theta)$ is a matrix of eigenvectors (depending on $\theta$) and $\Lambda(\theta) = \text{diag}(\lambda_1, \ldots, \lambda_m)$ is a diagonal matrix of eigenvalues (also depending on $\theta$). Note that the negative Hessian at $\theta$ need not be positive definite, so that the Hessian cannot itself be used as a Riemannian metric on $\mathbb{R}^m$.

**Remark.** Given a fixed $x \in \mathbb{R}^k$, we could consider the negative Hessian of $\log \pi_{\theta}(x)$ with respect to $\theta$. Unlike the Fisher information, which is the expected value of negative Hessian averaging over the randomness in $x|\theta$ as in proposition 2.6.7, it is easily seen from definition 2.6.4 that the Fisher information is defined to be the expected outer product of score vectors; this can be used to show that the Fisher information is a positive definite matrix as in proposition 2.6.5. However, simply computing the negative Hessian at a fixed
value of $x$ may not produce a positive definite matrix.

However, Betancourt [2012] consider the following transformation of the eigenvalues:

$$\text{softabs}(\lambda) = \lambda \cdot \frac{\exp(\alpha \lambda) + \exp(-\alpha \lambda)}{\exp(\alpha \lambda) - \exp(-\alpha \lambda)}. \quad (2.701)$$

By plotting this function with respect to $\lambda$ for fixed $\alpha > 0$, one finds that it is a smooth approximation to the absolute value function, wherein the sharpness of the approximation near zero is controlled by the parameter $\alpha$. One can show that $\text{softabs}(\lambda) \geq 1/\alpha$ for every $\lambda \in \mathbb{R}$. Hence, one may construct a positive definite matrix as follows:

$$\tilde{H}(\theta) = Q(\theta) \tilde{A}(\theta) Q(\theta)^\top \quad (2.702)$$

$$\tilde{A}(\theta) = \text{diag}(\text{softabs}(\lambda_1(\theta)), \ldots, \text{softabs}(\lambda_m(\theta))). \quad (2.703)$$

This positive definite matrix may then be regarded as a Riemannian metric on $\mathbb{R}^m$ and gives a general-purpose mechanism by which to produce a metric that captures second-order behavior of the target density $\pi$. The SoftAbs metric was developed with applications to Riemannian manifold Hamiltonian Monte Carlo foremost in mind; Riemannian manifold HMC will be discussed in detail in chapters 4 to 6. For this purpose, it is necessary to obtain the following derivatives (with dependencies on $\theta$ suppressed for clarity):

$$\frac{\partial}{\partial \theta_i} \log \det(\tilde{H}) = \text{trace} \left( Q(\tilde{A}^{-1} \circ J) Q^\top \frac{\partial}{\partial \theta_i} H \right) \quad (2.704)$$

$$J_{ij} = \begin{cases} \frac{\partial}{\partial \lambda_j} \lambda_i \coth(\alpha \lambda_i) & \text{if } \lambda_i = \lambda_j \\ \frac{\lambda_i \coth(\alpha \lambda_i) - \lambda_j \coth(\alpha \lambda_j)}{\lambda_i - \lambda_j} & \text{if } \lambda_i \neq \lambda_j \end{cases} \quad (2.705)$$

$$\frac{\partial}{\partial \theta_i}(p^\top \tilde{H} p) = -\text{trace} \left( Q DJDQ^\top \frac{\partial}{\partial \theta_i} H \right) \quad (2.706)$$

$$D = \tilde{A}^{-1} Q^\top p, \quad (2.707)$$
and where $A \circ B$ denotes the element-wise product of $A$ and $B$.

## 2.7 Markov Chain Monte Carlo


### 2.7.1 Markov Chain Transition Kernels

Let $X$ be a set with Borel $\sigma$-algebra $\mathcal{B}(X)$. The fundamental object of a Markov chain is the transition kernel, which is defined as follows.

**Definition 2.7.1** ([Robert and Casella (2005)](https://www.springer.com/gp/book/9780387212328)). A transition kernel is a map $K : X \times \mathcal{B}(X) \to [0, 1]$ satisfying the following two properties.

1. For every fixed $x \in X$, $K(x, \cdot) : \mathcal{B}(X) \to \mathbb{R}$ is a probability measure.
2. For every fixed $A \in \mathcal{B}(X)$, $K(\cdot, A)$ is $\mathcal{B}(X)$-measurable.

**Definition 2.7.2.** The Markov chain with transition kernel $K$ is defined to be the stochastic process $(X_n)_{n \in \mathbb{N}}$ with joint probability distribution,

\[
Pr(X_1 \in A_1, \ldots, X_n \in A_n | X_0 = x) = \int_{A_1} \cdots \int_{A_{n-1}} K(y_{n-1}, A_n) K(x, dy_1) \cdots K(y_2, dy_1). \tag{2.708}
\]

We adopt the notation $K^n(x, A) = Pr(X_n \in A | X_0 = x)$.

**Lemma 2.7.3.**

\[
K^n(x, A) = \int_{X} K^{n-1}(y, A) K(x, dy) \tag{2.709}
\]
Proof.

\[ K^n(x, A_n) = \Pr(X_1 \in \mathcal{X}, \ldots, X_{n-1} \in \mathcal{X}, X_n \in A_n | X_0 = x) \quad (2.710) \]

\[ = \int_{\mathcal{X}} \cdots \int_{\mathcal{X}} K(y_{n-1}, A_n)K(x, dy_1) \cdots K(y_{n-2}, dy_{n-1}) \quad (2.711) \]

\[ = \int_{\mathcal{X}} K(x, dy_1) \int_{\mathcal{X}} \cdots \int_{\mathcal{X}} K(y_{n-1}, A_n)K(y_1, dy_2) \cdots K(y_{n-2}, dy_{n-1}) \quad (2.712) \]

\[ = \int_{\mathcal{X}} K(x, dy_1) \Pr(X_{n-1} \in A_n | X_0 = y_1) \quad (2.713) \]

\[ = \int_{\mathcal{X}} K(x, dy_1)K^{n-1}(y_1, A) \quad (2.714) \]

This result allows us to produce the Chapman-Kolmogorov equation.

**Proposition 2.7.4.**

\[ K^{n+m}(x, A) = \int_{\mathcal{X}} K^n(y, A)K^m(x, dy) \quad (2.715) \]

Proof.

\[ K^{n+m}(x, A) = \int_{\mathcal{X}} K(x, dy_1)K^{n+m-1}(y_1, A) \quad (2.716) \]

\[ = \int_{\mathcal{X}} K(x, dy_1) \int_{\mathcal{X}} K(y_1, dy_2)K^{n+m-2}(y_2, A) \quad (2.717) \]

\[ = \int_{\mathcal{X}} \cdots \int_{\mathcal{X}} K(x, dy_1) \cdots K(y_{m-1}, dy_m)K^n(y_m, A) \quad (2.718) \]

\[ = \int_{\mathcal{X}} K^m(x, dy_m)K^n(y_m, A) \quad (2.719) \]

\[ \square \]
2.7.2 Stationarity and Detailed Balance

Definition 2.7.5. Consider a Markov chain with transition kernel $K$. A probability distribution $\Pi$ is said to be stationary for $K$ if, for any $A \in \mathcal{B}(\mathcal{X})$, we have

$$\Pi(A) = \mathbb{E}_{x \sim \Pi} K(x, A). \tag{2.720}$$

Lemma 2.7.6. If $K$ is stationary for $\Pi$ then so is $K^n$.

Proof. We first establish the base case $n = 1$. Since $K$ is stationary for $\Pi$ and $K = K^1$, we immediately establish that $K^1$ is stationary for $\Pi$. Now suppose that $K^n$ is stationary for $\Pi$. Then, by induction and applying the Chapman-Kolmogorov equation,

$$\mathbb{E}_{x \sim \Pi} K^{n+1}(x, A) = \mathbb{E}_{x \sim \Pi} \int_{\mathcal{X}} K^n(x, dy) K(y, A) \tag{2.721}$$

$$= \int_{\mathcal{X}} \Pi(dy) K(y, A) \tag{2.722}$$

$$= \Pi(A). \tag{2.723}$$

Within the context of Markov chain Monte Carlo, stationarity is often established via the property of detailed balance, which we now define.

Definition 2.7.7. A Markov chain with transition kernel $K$ is said to satisfy detailed balance with respect to the probability measure $\Pi$ if

$$\int_A K(x, B) \Pi(dx) = \int_B K(x, A) \Pi(dx). \tag{2.724}$$

for any $A, B \in \mathcal{B}(\mathcal{X})$. 
Lemma 2.7.8. When $K$ satisfies detailed balance with respect to $\Pi$, it follows that $\Pi$ is the stationary distribution of $K$.

Proof. Let $B = \mathcal{X}$. Then,

\[
\Pi(A) = \int_A K(x, \mathcal{X})\Pi(dx)
\]

(2.725)

\[
= \int_\mathcal{X} K(x, A)\Pi(dx)
\]

(2.726)

\[
= \mathbb{E}_{x \sim \Pi}[K(x, A)].
\]

(2.727)

\[\square\]

2.7.3 The Total Variation Distance

Definition 2.7.9. Let $\mu$ and $\nu$ be probability measures. The total variation norm is defined by

\[
\|\mu - \nu\|_{TV} = 2 \sup_{A \in \mathcal{B}(\mathcal{X})} |\mu(A) - \nu(A)|.
\]

(2.728)

We now give an alternative characterization of the total variation norm; we follow the proof strategy in [Roberts and Rosenthal, 2004].

Lemma 2.7.10.

\[
\|\mu - \nu\|_{TV} = \sup_{|h| \leq 1} \left| \int_{\mathcal{X}} h(x) \mu(dx) - \int_{\mathcal{X}} h(x) \nu(dx) \right|
\]

(2.729)
Proof. For any $A \in \mathcal{B}(\mathcal{X})$, we have that

$$2 |\mu(A) - \nu(A)| = 2 \left| \int_A \mu(dx) - \int_A \nu(dx) \right|$$

(2.730)

$$= 2 \left| \int_{\mathcal{X}} 1 \{x \in A\} \mu(dx) - \int_{\mathcal{X}} 1 \{x \in A\} \nu(dx) \right|$$

(2.731)

$$\leq 2 \sup_{0 \leq h \leq 1} \left| \int_{\mathcal{X}} h(x) \mu(dx) - \int_{\mathcal{X}} h(x) \nu(dx) \right|$$

(2.732)

$$= \sup_{0 \leq h \leq 2} \left| \int_{\mathcal{X}} (h(x) - 1) \mu(dx) - \int_{\mathcal{X}} (h(x) - 1) \nu(dx) \right|$$

(2.733)

$$= \sup_{-1 \leq h \leq 1} \left| \int_{\mathcal{X}} h(x) \mu(dx) - \int_{\mathcal{X}} h(x) \nu(dx) \right|$$

(2.734)

Taking the supremum over $A$ then reveals that

$$\|\mu - \nu\|_{TV} \leq \sup_{|h| \leq 1} \left| \int_{\mathcal{X}} h(x) \mu(dx) - \int_{\mathcal{X}} h(x) \nu(dx) \right|.$$  

(2.736)

On the other hand, let $f : \mathcal{X} \rightarrow [-1, 1]$. Define the measure $\rho = \mu + \nu$ and set $g = \frac{d\mu}{d\rho}$ and $h = \frac{d\nu}{d\rho}$. Then,

$$\left| \int_{\mathcal{X}} f(x) \mu(dx) - \int_{\mathcal{X}} f(x) \nu(dx) \right|$$

(2.737)

$$= \left| \int_{\mathcal{X}} f(x)g(x) \rho(dx) - \int_{\mathcal{X}} f(x)h(x) \rho(dx) \right|$$

(2.738)

$$= \left| \int_{\mathcal{X}} f(x)(g(x) - h(x)) \rho(dx) \right|$$

(2.739)

$$\leq \int_{g>h} (g - h) \rho(dx) + \int_{h<g} (h - g) \rho(dx)$$

(2.740)

$$= |\mu(A) - \nu(A) + (1 - \nu(A)) - (1 - \mu(A))|$$

(2.741)

$$= 2 |\mu(A) - \nu(A)|$$

(2.742)

where $A = \{x \in \mathcal{X} : g(x) > h(x)\}$. Therefore, taking supremum over functions $f$ and
sets $A$, we obtain,

$$
\sup_{|f| \leq 1} \left| \int_X f(x) \mu(dx) - \int_X f(x) \nu(dx) \right| \leq \|\mu - \nu\|_{TV}.
$$

(2.743)

This proves the equality. □

### 2.7.4 Irreducibility and Small Sets

**Definition 2.7.11.** Let $\psi : \mathfrak{B}(\mathcal{X})$ be a measure (not necessarily a probability measure). A Markov chain with transition kernel $K$ is called $\psi$-irreducible if for every $A \in \mathfrak{B}(\mathcal{X})$ with $\psi(A) > 0$ there exists $n \in \mathbb{N}$ such that $K^n(x, A) > 0$ for every $x \in \mathcal{X}$.

In other words, $\psi$-irreducibility means that, for every Borel subset to which $\psi$ assigns non-zero measure, it is possible to reach this subset from an arbitrary initial position in a finite number of steps. Let us now define the time of first entrance for a Markov chain:

$$
\tau_A = \inf\{n \geq 1 : X_n \in A\}.
$$

(2.744)

**Lemma 2.7.12.** There exists $n \in \mathbb{N}$ such that $K^n(x, A) > 0$ if and only if $\Pr(\tau_A < \infty | X_0 = x) > 0$.

**Proof.**

$\implies$ Then there exists $n \in \mathbb{N}$ such that $\Pr(X_n \in A | X_0 = x) > 0$. Hence $\Pr(\tau_A < \infty) \geq \Pr(\tau_A \leq n) > 0$.

$\iff$ Then there exists a non-zero probability of entering $A$ in a finite number of steps.

Thus, there exists $n \in \mathbb{N}$ such that $K^n(x, A) > 0$.

□

**Corollary 2.7.13.** A Markov chain with transition kernel $K$ is $\psi$-irreducible if for every $A \in \mathfrak{B}(\mathcal{X})$ with $\psi(A) > 0$ we have $\Pr(\tau_A < \infty | X_0 = x) > 0$. 

148
A related concept to the time of first entry is the expected number of passages of a Markov chain through a state \( A \in \mathcal{B}(\mathcal{X}) \):

\[
\eta_A = \sum_{n=0}^{\infty} 1\{X_n \in A\}.
\]

(2.745)

**Lemma 2.7.14** (Theorem 6.15 [Robert and Casella 2005]). A Markov chain with transition kernel \( K \) is \( \psi \)-irreducible if and only if \( \mathbb{E}[\eta_A | X_0 = x] > 0 \) for every \( A \in \mathcal{B}(\mathcal{X}) \) for which \( \psi(A) > 0 \).

**Proof.**

\( \implies \) Since \( K \) is \( \psi \)-irreducible, for every \( A \) with \( \psi(A) > 0 \) we have that there exists \( n \in \mathbb{N} \) such that \( K^n(x, A) > 0 \). By the monotone convergence theorem,

\[
\mathbb{E}[\eta_A | X_0 = x] = \sum_{k=0}^{\infty} \mathbb{E}[1\{X_k \in A\} | X_0 = x]
\]

\[
= \sum_{k=0}^{\infty} \mathbb{P}(X_k \in A | X_0 = x)
\]

\[
= \sum_{k=0}^{\infty} K^k(x, A)
\]

\[
\geq K^n(x, A)
\]

\[
> 0.
\]

(2.750)

\( \leftarrow \) Then, we have

\[
\mathbb{E}[\eta_A | X_0 = x] = \sum_{n=0}^{\infty} K^n(x, A).
\]

(2.751)

Now since \( \mathbb{E}[\eta_A | X_0 = x] > 0 \) by assumption, we conclude that at least one term in the sum must be non-zero. Hence, there exists \( n \in \mathbb{N} \) such that \( K^n(x, A) > 0 \).

\( \square \)
Definition 2.7.15. Let $K$ be a Markov chain transition kernel. A set $A \in \mathcal{B}(\mathcal{X})$ is called small if there exists an integer $m$, a probability measure $\nu_m$ and constant $\epsilon_m > 0$ such that

$$K^m(x, B) \geq \epsilon_m \nu_m(B) \quad (2.752)$$

for every $B \in \mathcal{B}(\mathcal{X})$ and every $x \in A$.

The usefulness of a notion of a small set derives, in part, from the coupling inequality; in describing this inequality and its applications to Markov chains, we follow Jiang et al. [2021].

Proposition 2.7.16. Let $X$ and $Y$ be two random variables on the space $\mathcal{X}$. Then for any $A \in \mathcal{B}(\mathcal{X})$, we have,

$$|\Pr(X \in A) - \Pr(Y \in A)| \leq \Pr(X \neq Y). \quad (2.753)$$

Proof.

$$|\Pr(X \in A) - \Pr(Y \in A)| = |\Pr(X \in A, X = Y) + \Pr(X \in A, X \neq Y) - \Pr(Y \in A, X = Y) - \Pr(Y \in A, X \neq Y)| \quad (2.754)$$

$$= |\Pr(X \in A, X \neq Y) - \Pr(Y \in A, X \neq Y)| \quad (2.755)$$

$$\leq \Pr(X \neq Y). \quad (2.756)$$

\[ \square \]

Corollary 2.7.17. Let $\Pi_X$ and $\Pi_Y$ denote the probability measures associated to $\mathcal{X}$-valued random variables $X$ and $Y$, respectively. Then,

$$\|\Pi_X - \Pi_Y\|_{TV} \leq 2 \Pr(X \neq Y). \quad (2.757)$$
Proof. Taking the supremum over all sets $A \in \mathcal{B}(\mathcal{X})$ yields,

$$\sup_{A \in \mathcal{B}(\mathcal{X})} |\Pi_X(A) - \Pi_Y(A)| \leq \Pr(X \neq Y). \quad (2.758)$$

Multiplying both sides by two (to be consistent with the definition of total variation norm) yields the result. \square

In section 2.7.5 we will discuss how small sets relate to the ergodicity of Markov chains. To preface this discussion, we require the following decomposition from Athreya and Ney [1978].

**Proposition 2.7.18.** When $A$ is a small set (definition 2.7.15), the transition kernel, for $x \in A$, may be expressed as,

$$K^m(x, B) = \epsilon \nu(B) + (1 - \epsilon)K'(x, B) \quad (2.759)$$

where

$$K'(x, B) = \frac{K^m(x, B) - \epsilon \nu(B)}{1 - \epsilon}. \quad (2.760)$$

where $K'(x, \cdot)$ is a probability measure.

Proof.

$$\epsilon \nu(B) + (1 - \epsilon)K'(x, B) = \epsilon \nu(B) + (1 - \epsilon)\frac{K^m(x, B) - \epsilon \nu(B)}{1 - \epsilon} \quad (2.761)$$

$$= \epsilon \nu(B) + K^m(x, B) - \epsilon \nu(B) \quad (2.762)$$

$$= K^m(x, B). \quad (2.763)$$
To show that $K'(x, \cdot)$ is a probability measure, observe that

$$K'(x, \emptyset) = \frac{K^m(x, \emptyset) - \epsilon \nu(\emptyset)}{1 - \epsilon} = 0$$  \hspace{1cm} \text{(2.765)}

$$K'(x, \mathcal{X}) = \frac{K^m(x, \mathcal{X}) - \epsilon \nu(\mathcal{X})}{1 - \epsilon} = \frac{1 - \epsilon}{1 - \epsilon} = 1.$$  \hspace{1cm} \text{(2.767)}

Let $(B_1, B_2, \ldots)$ be pairwise disjoint. Countable additivity follows from,

$$K'(x, \bigcup B_i) = \frac{K^m(x, \bigcup B_i) - \epsilon \nu(\bigcup B_i)}{1 - \epsilon} = \sum_i K^m(x, B_i) - \epsilon \sum_i \nu(B_i)$$  \hspace{1cm} \text{(2.770)}

$$= \sum_i \frac{K^m(x, B_i) - \epsilon \nu(B_i)}{1 - \epsilon} = \sum_i K'(x, B_i).$$  \hspace{1cm} \text{(2.772)}

The fact that $K'(x, \cdot)$ is non-negative on $A$ follows immediately from the smallness condition.

The significance of this decomposition is that when the chain enters the small set, the next state can be generated from the probability distribution $\nu$ with probability $\epsilon$. When this occurs, the chain “forgets its past” so that the subsequent state is independent of all previous states, a result stronger than the Markov property. In the Markov chain literature, this is called regeneration.

**Definition 2.7.19** (Robert and Casella [2005], Meyn and Tweedie [1993], Doob and Doob [1953]). A Markov chain is said to satisfy the **Doeblin condition** if $\mathcal{X}$ is small.

Another important condition in Markov chain theory is the notion of aperiodicity.
Definition 2.7.20 (Meyn and Tweedie [1993]). A Markov chain with transition kernel \( K \) is called aperiodic if for every small set \( A \) with \( \psi(A) > 0 \), the greatest common divisor of the set of \( m \) for which eq. (2.752) holds is one.

Lemma 2.7.21. If the greatest common divisor of the set of \( m \) for which eq. (2.752) holds is one for some small set, then the greatest common divisor is one for all small sets.

Proof. Let \( A \) be a small set for which the greatest common divisor is one. Let \( B \) be another small set which satisfies eq. (2.752) for \( \tilde{m} \). By the Chapman-Kolmogorov equations, for any \( m \) for which eq. (2.752) holds for \( A \),

\[
K^{\tilde{m} + m}(x, C) = \int_X K^m(y, C)K^{\tilde{m}}(x, dy) \\
\geq \int_A K^m(y, C)K^{\tilde{m}}(x, dy) \\
\geq \int_A \epsilon_m \nu_m(C)K^{\tilde{m}}(x, dy) \\
\geq \epsilon_m \nu_m(C)K^{\tilde{m}}(x, A) \\
\geq \epsilon_m \delta_{\tilde{m}} \mu_{\tilde{m}}(A)\nu_m(C).
\]

Hence \( B \) is small. Since the greatest common divisor of the set of \( m \) for which eq. (2.752) holds for \( A \) is one, it follows that the same is true of the small set \( B \).

\[\square\]

2.7.5 Ergodicity

Definition 2.7.22. A Markov chain with transition kernel \( K \) is said to be ergodic for the distribution \( \Pi \) from initial position \( x \) if

\[
\lim_{n \to \infty} \|K^n(x, \cdot) - \Pi(\cdot)\|_{TV} = 0.
\]

We may now state an important result in the theory of Markov chains, which justifies their use as a mechanism by which to generate samples from a target distribution.
Theorem 2.7.23 (Robert and Casella [2005], Meyn and Tweedie [1993]). Let \( \Pi \) be a probability measure on a space \( \mathcal{X} \). Let \( K : \mathcal{X} \times \mathcal{B}(\mathcal{X}) \to \mathbb{R}_+ \) be a Markov chain transition kernel which is aperiodic, \( \Pi \)-irreducible, and has \( \Pi \) as a stationary distribution. Then, for \( \Pi \)-almost all initial conditions \( x \in \mathcal{X} \), \( K \) is ergodic for the distribution \( \Pi \) from initial position \( x \).

Theorem 2.7.24. Let \( K \) be a Markov chain transition kernel with stationary distribution \( \Pi \) that satisfies the Doeblin condition (definition 2.7.19) for \( m \in \mathbb{N} \). Then, for every \( x \in \mathcal{X} \) and \( d \in \mathbb{N} \),

\[
\|K^{dm}(x, \cdot) - \Pi(\cdot)\|_{TV} \leq 2(1 - \epsilon)^d. \tag{2.779}
\]

Proof. This follows from a coupling argument. Let \( Y_0 \sim \Pi \). Because \( \Pi \) is the stationary distribution of the Markov chain, it follows that if we generate \( Y_{md} \sim K^m(Y_{m(d-1)}, \cdot) \) then, in fact, \( Y_{md} \sim \Pi \) for every \( d \in \mathbb{N} \). Similarly, let \( X_0 = x \) and generate \( X_{md} \sim K^m(X_{m(d-1)}, \cdot) \). The two sequences \( (X_{md}) \) and \( (Y_{md}) \) can be coupled in the following manner. If \( X_{md} = Y_{md} \), then sample \( Z \sim K^m(X_{md}, \cdot) \) and set \( X_{m(d+1)} = Y_{m(d+1)} = Z \).

If \( X_{md} \neq Y_{md} \), then we employ proposition 2.7.18. In particular, with probability \( \epsilon \), generate \( Z \sim \nu \) and set \( X_{m(d+1)} = Y_{m(d+1)} = Z \), at which point the two chains are coupled. Otherwise, with probability \( 1 - \epsilon \), generate \( X_{m(d+1)} \sim K'(X_{md}, \cdot) \) and \( Y_{m(d+1)} \sim K'(Y_{md}, \cdot) \), independently. For each \( d \in \mathbb{N} \), the probability that the chains have not coupled is at most \( (1 - \epsilon)^d \). Therefore, from the coupling inequality,

\[
\|K^{dm}(x, \cdot) - \Pi(\cdot)\|_{TV} \leq 2(1 - \epsilon)^d. \tag{2.780}
\]

This result shows that at multiples of \( m \), the chain is approaching the target distribution in total variation. But how do we know that total variation does not increase arbitrarily for
The following proposition provides the answer.

**Proposition 2.7.25** (Meyn and Tweedie [1993]). Let $K$ be a Markov transition kernel with stationary distribution $\Pi$. Then, for any $x \in X$,

$$\|K^n(x, \cdot) - \Pi(\cdot)\|_{TV} \leq \rho^n$$

is non-increasing in $n$.

**Proof.** We use the alternative representation of the total variation distance given in lemma 2.7.10. We have,

$$\|K^{n+1}(x, \cdot) - \Pi(\cdot)\|_{TV} \leq \rho^n$$

**Definition 2.7.26.** A Markov chain with transition kernel $K$ is said to be *geometrically ergodic* for the distribution $\Pi$ from initial position $x$ if

$$\|K^n(x, \cdot) - \Pi(\cdot)\|_{TV} \leq V(x)\rho^n$$

where $V : X \to [1, \infty)$ and $\rho \in [0, 1)$. 

155
From theorem 2.7.24 and proposition 2.7.25 we see that Markov chains that satisfy the Doeblin condition are geometrically ergodic.

2.7.6 Metropolis-Hastings Transition Kernels

Let us now see some important examples of Markov chain transition kernels which are based on correcting samples from a “proposal density” using a specific accept-reject test.

**Definition 2.7.27.** Let $x \in \mathcal{X}$ and let $\pi$ and $\tilde{\pi}(\cdot|x)$ be probability densities on $\mathcal{X}$ with respect to a dominating measure $\mu$. A Markov chain transition kernel $K$ is said to be of Metropolis-Hastings type if it has the form,

$$K(x, A) = \int_A \alpha(x, y) \tilde{\pi}(y|x) \, \mu(dy) + (1 - \rho(x)) 1\{x \in A\}, \quad (2.789)$$

where

$$\alpha(x, y) = \min \left\{1, \frac{\pi(y) \tilde{\pi}(x|y)}{\pi(x) \tilde{\pi}(y|x)} \right\} \quad (2.790)$$

$$\rho(x) = \int_{\mathcal{X}} \alpha(x, y) \frac{\tilde{\pi}(y|x)}{\pi(x)} \, \mu(dy), \quad (2.791)$$

where, in this case, $\tilde{\pi}$ is called the proposal density, $\alpha$ the acceptance probability, and $\rho$ the rejection rate.

Sampling from a Markov chain of Metropolis-Hastings type can therefore be achieved by sampling $\tilde{X}_{n+1} \sim \tilde{\pi}(\cdot|X_n)$, computing the acceptance probability $\alpha(X_n, \tilde{X}_{n+1})$ and setting $X_{n+1} = \tilde{X}_{n+1}$ if $U < \alpha(X_n, \tilde{X}_{n+1})$ and $X_{n+1} = X_n$ otherwise, where $U \sim \text{Uniform}(0, 1)$.

**Proposition 2.7.28** ([Mengersen and Tweedie 1996]). Let $K$ be the transition kernel of a Metropolis-Hastings Markov chain with proposal density $\tilde{\pi}$. Suppose that $\text{Supp}(\pi) = \mathcal{X}$ and that $\text{Supp}(\tilde{\pi}(\cdot|x)) = \mathcal{X}$ for every $x \in \mathcal{X}$. Suppose further that $\pi$ is continuous on $\mathcal{X}$.
and \((x, y) \mapsto \tilde{\pi}(y|x)\) is continuous on \(X \times X\). Then every compact set is small.

**Proof.** Let \(C\) be a compact set. Then, by positivity, \(a = \sup_{x \in C} \pi(x) < \infty\) and \(b = \inf_{x, y \in C} \tilde{\pi}(y|x) > 0\) and both limits are achieved. We will create a measure that is concentrated on \(C\) for which eq. (2.752) is satisfied. Let \(A \in \mathcal{B}(C)\) be arbitrary. Define the guaranteed acceptance region within \(C\) from \(x\) by,

\[
A_C(x) = \left\{ y \in C : \frac{\pi(y)\tilde{\pi}(x|y)}{\pi(x)\tilde{\pi}(y|x)} \geq 1 \right\} \tag{2.792}
\]

and the possible rejection region within \(A\) from \(x\) by \(R_C(x) = A \setminus A_C(x)\). Thus,

\[
K(x, A) \geq \int_A \min \left\{ 1, \frac{\pi(y)\tilde{\pi}(x|y)}{\pi(x)\tilde{\pi}(y|x)} \right\} \pi(y|x) \mu(dy) \tag{2.793}
\]

\[
= \int_{A_C(x)} \tilde{\pi}(y|x) \mu(dy) + \int_{R_C(x)} \frac{\pi(y)\tilde{\pi}(x|y)}{\pi(x)} \mu(dy) \tag{2.794}
\]

\[
= \int_{A_C(x)} \frac{\pi(y)}{\pi(y)} \tilde{\pi}(y|x) \mu(dy) + \int_{R_C(x)} \frac{\pi(y)\tilde{\pi}(x|y)}{\pi(y)} \mu(dy) \tag{2.795}
\]

\[
\geq \frac{b}{a} \int_{A_C(x)} \pi(y) \mu(dy) + \frac{b}{a} \int_{R_C(x)} \pi(y) \mu(dy) \geq \frac{b}{a} \int_{A} \pi(y) \mu(dy) \tag{2.796}
\]

\[
= \frac{b}{a} \Pi(A) \tag{2.797}
\]

\[
= \frac{b}{a} \Pi(C) \left( \frac{\Pi(A)}{\Pi(C)} \right) \tag{2.798}
\]

This shows that the set \(C\) is small for the probability measure \(A \mapsto \Pi(A)/\Pi(C)\) concentrated on the set \(C\).

**Corollary 2.7.29.** Under the conditions of proposition 2.7.28, since eq. (2.752) holds for \(m = 1\), it follows immediately that the Metropolis-Hastings Markov chain is aperiodic.

**Lemma 2.7.30.** A Markov chain transition kernel of Metropolis-Hastings type satisfies detailed balance with respect to the distribution \(\Pi\) whose density is \(\pi\).
Proof. First, observe that
\[
\int_B (1 - \rho(x)) 1 \{x \in A\} \pi(x) \mu(dx) = \int_A (1 - \rho(x)) 1 \{x \in B\} \pi(x) \mu(dx). \tag{2.800}
\]

Moreover,
\[
\int_B \left( \int_A \alpha(x, y) \tilde{\pi}(y|x) \mu(dy) \right) \pi(x) \mu(dx) = \int_B \int_A \min \left\{ \tilde{\pi}(y|x), \pi(y) \tilde{\pi}(x|y) \right\} \mu(dy) \mu(dx) \tag{2.801}
\]
\[
= \int_B \int_A \min \left\{ 1, \frac{\pi(x) \tilde{\pi}(y|x)}{\pi(y) \tilde{\pi}(x|y)} \right\} \pi(y) \tilde{\pi}(x|y) \mu(dx) \mu(dy) \tag{2.802}
\]
\[
= \int_A \left( \int_B \alpha(y, x) \tilde{\pi}(x|y) \mu(dy) \right) \pi(y) \mu(dy). \tag{2.803}
\]

From these two equations, we have,
\[
\int_B K(x, A) \pi(x) \mu(dx) = \int_A K(x, B) \pi(x) \mu(dx). \tag{2.805}
\]
\[
\square
\]

Corollary 2.7.31. Under the conditions of proposition 2.7.28, the Metropolis-Hastings Markov chain is ergodic from theorem 2.7.23.

Definition 2.7.32. An independent Metropolis-Hastings sampler is a Markov chain of Metropolis-Hastings type with proposal density \( \tilde{\pi}(y|x) \equiv \tilde{\pi}(y) \) (note the independence of the proposal density of \( x \)).

Proposition 2.7.33. The independent Metropolis-Hastings sampler is geometrically ergodic if there exists a constant \( M < \infty \) such that \( \pi(x) \leq M \tilde{\pi}(x) \) for every \( x \in X \).
Proof. For any \( x \in \mathcal{X} \) and \( A \in \mathcal{B}(\mathcal{X}) \),

\[
K(x, A) \geq \int_A \alpha(x, y)\bar{\pi}(y)\mu(dy)
= \int_A \min \left\{ \bar{\pi}(y), \frac{\pi(y)\bar{\pi}(x)}{\pi(x)} \right\} \mu(dy)
\geq \int_A \min \left\{ \bar{\pi}(y), \frac{\pi(y)}{M} \right\} \mu(dy)
= \frac{1}{M} \Pi(A).
\]

Hence the Doeblin condition is satisfied and geometric ergodicity follows. \( \square \)

Definition 2.7.34. A Markov chain transition kernel of Metropolis-Hastings type is called symmetric when, for every \( x, y \in \mathcal{X} \), we have that \( \bar{\pi}(x|y) = \bar{\pi}(y|x) \) and in this case the acceptance probability is,

\[
\alpha(x, y) = \min \left\{ 1, \frac{\pi(y)}{\pi(x)} \right\}.
\]

Definition 2.7.35. A random walk Metropolis-Hastings sampler is a symmetric Markov chain of Metropolis-Hastings type with proposal density \( \bar{\pi}(y|x) = \text{Normal}(y|x, \sigma^2) \), where \( \sigma^2 \) is a parameter of the method.

The random walk Metropolis-Hastings sampler can then be understood in the following way. Proposals are generated by running \( m \)-dimensional Brownian motion for \( \sigma \) units of time from the initial position \( x \); Brownian motion is the exact solution of the stochastic differential equation \( dX^i_t = dB^i_t \). The terminal position of the Brownian motion is then accepted or rejected according to the Metropolis-Hastings criterion. Under certain conditions, the random walk Metropolis-Hastings sampler will be geometrically ergodic; the following result and terminology is from Mengersen and Tweedie [1996], Robert and Casella [2005].

Definition 2.7.36. A continuous density \( \pi : \mathbb{R} \to \mathbb{R}_+ \) is said to have log-concave tails if
there exists $\alpha > 0$ and $x^* > 0$ such that

$$\log \pi(x) - \log \pi(y) \geq \alpha |x - y|$$

whenever $y > x > x^*$ or $y < x < -x^*$.

**Theorem 2.7.37** (Mengersen and Tweedie [1996]). Let $\pi$ have log-concave tails and suppose $\pi(x) = \pi(-x)$. Then the random walk Metropolis-Hastings sampler is geometrically ergodic.

These results only apply in the one-dimensional case, perhaps limiting the usefulness of the random walk Metropolis sampler in more general scenarios.

A disadvantage of the random walk Metropolis-Hastings sampler is that, because it is merely a corrected version of Brownian motion, it does not incorporate local information about the posterior into its proposals. This can be alleviated by considering discretizations of the Langevin stochastic differential equation as an alternative proposal mechanism.

**Definition 2.7.38.** Let $\pi$ be a smooth target distribution. The *Metropolis-adjusted Langevin Algorithm* (MALA) is a Markov chain of Metropolis-Hastings type with proposal density

$$\tilde{\pi}(y|x) = \text{Normal} \left( y|x + \frac{\epsilon^2}{2} \nabla \log \pi(x), \epsilon^2 \text{Id} \right).$$

This proposal can be recognized as a Euler-Maruyama discretization to time $\epsilon^2$ of the Langevin stochastic differential equation with

$$b^i(x) = \frac{1}{2} \frac{\partial}{\partial x_i} \log \pi(x)$$

and

$$\sigma^i_j = \delta_{i=j}.$$

The MALA algorithm may also be equipped with a geometric ergodicity theory. Before stating the result, we require two pieces of terminology, both from Roberts and Tweedie.
Definition 2.7.39. The guaranteed acceptance region at $x$ of a Markov chain of Metropolis-Hastings type is the set,

$$A(x) = \{ y \in \mathcal{X} : \pi(x) \tilde{\pi}(y|x) \leq \pi(y) \tilde{\pi}(x|y) \}.$$  \hfill (2.815)

Definition 2.7.40. Let $I(x) = \{ y \in \mathcal{X} : \|y\| \leq \|x\| \}$. The guaranteed acceptance region at $x$ is said to converge inward if

$$\lim_{\|x\| \to \infty} \int_{(A(x) \cup I(x)) \cap (A(x) \cap I(x))} \tilde{\pi}(y|x) \mu(dy) = 0.$$  \hfill (2.816)

Theorem 2.7.41 \textbf{(Roberts and Tweedie \cite{1996})}. Let $K$ be the Markov chain transition kernel of the Metropolis-adjusted Langevin algorithm. If the guaranteed acceptance region at $x$ converges inward, and if,

$$\liminf_{\|x\| \to \infty} \left( \|x\| - \|x + \frac{\epsilon^2}{2} \nabla \log \pi(x)\| \right) > 0,$$  \hfill (2.817)

then $K$ has geometric convergence toward the distribution with density $\pi$. On the other hand, if

$$\lim_{\|x\| \to \infty} \|\nabla \log \pi(x)\| = 0,$$  \hfill (2.818)

then $K$ is not geometrically ergodic.

We conclude this section by motivating a broad class of Metropolis-Hastings Markov chains based on the discretization of Hamiltonian mechanics.

Example 30. In non-canonical Hamiltonian dynamics the equations of motion are gov-
\[
\begin{pmatrix}
\dot{q}_t \\
\dot{p}_t
\end{pmatrix} =
\begin{pmatrix}
\mathbf{E}(q_t) & \mathbf{F}(q_t) \\
-\mathbf{F}^\top(q_t) & \mathbf{R}(q_t)
\end{pmatrix}
\begin{pmatrix}
\nabla_q H(q_t, p_t) \\
\nabla_p H(q_t, p_t)
\end{pmatrix},
\] (2.819)

where \(E(q)\) and \(R(q)\) are skew-symmetric matrices. When the Hamiltonian is of the form
\[
H(q, p) = U(q) + \frac{1}{2}p^\top p,
\]
this further simplifies to
\[
\begin{pmatrix}
\dot{q}_t \\
\dot{p}_t
\end{pmatrix} =
\begin{pmatrix}
\mathbf{E}(q_t) & \mathbf{F}(q_t) \\
-\mathbf{F}^\top(q_t) & \mathbf{R}(q_t)
\end{pmatrix}
\begin{pmatrix}
\nabla_q U(q_t) \\
\nabla_p U(q_t)
\end{pmatrix}.
\] (2.820)

While obtaining an exact solution to these dynamics seems impossible, we can settle for a first-order approximation. Indeed, by the Fundamental Theorem of Calculus,
\[
q_T = q_0 + \int_0^T \left[ \mathbf{E}(q_t) \nabla_q U(q_t) + \mathbf{F}(q_t)p_t \right] dt
\] (2.821)

and
\[
p_T = p_0 + \int_0^T \left[ -\mathbf{F}^\top(q_t) \nabla_q U(q_t) + \mathbf{R}(q_t)p_t \right] dt.
\] (2.822)

Thus,
\[
q_T = q_0 + \int_0^T \left[ \mathbf{E}(q_t) \nabla_q U(q_t) + \mathbf{F}(q_t) \left( p_0 + \int_0^t \left[ -\mathbf{F}^\top(q_s) \nabla_q U(q_s) + \mathbf{R}(q_s)p_s \right] ds \right) \right] dt
\] (2.823)
Now take $\sqrt{\delta} \ll 1$. A first-order approximation yields,

$$
q_{\sqrt{\delta}} \approx q_0 + \sqrt{\delta} \mathbf{E}(q_0) \nabla_q U(q_0) - \frac{\delta}{2} \mathbf{F}(q_0) \mathbf{F}^\top(q_0) \nabla_q U(q_0) + \frac{\delta}{2} \mathbf{F}(q_0) \mathbf{R}(q_0) p_0 + \sqrt{\delta} \mathbf{F} p_0
$$

(2.824)

$$
= q_0 + \left( \sqrt{\delta} \mathbf{E}(q_0) - \frac{\delta}{2} \mathbf{F}(q_0) \mathbf{F}^\top(q_0) \right) \nabla_q U(q_0) + \left( \frac{\delta}{2} \mathbf{F}(q_0) \mathbf{R}(q_0) + \sqrt{\delta} \mathbf{F}(q_0) \right) p_0
$$

(2.825)

$$
def. \hat{q}_{\sqrt{\delta}}.
$$

(2.826)

When $p_0 \sim \text{Normal}(0, \mathbf{I})$ then,

$$
\hat{\mu} \equiv q_0 + \left( \sqrt{\delta} \mathbf{E}(q_0) - \frac{\delta}{2} \mathbf{F}(q_0) \mathbf{F}^\top(q_0) \right) \nabla_q U(q_0)
$$

(2.827)

$$
\hat{\Sigma} \equiv \left( \frac{\delta}{2} \mathbf{F}(q_0) \mathbf{R}(q_0) + \sqrt{\delta} \mathbf{F}(q_0) \right) \left( \frac{\delta}{2} \mathbf{F}(q_0) \mathbf{R}(q_0) + \sqrt{\delta} \mathbf{F}(q_0) \right)^\top
$$

(2.828)

$$
\hat{q}_{\sqrt{\delta}} \sim \text{Normal} \left( \hat{\mu}, \hat{\Sigma} \right).
$$

(2.829)

This gives a proposal density suitable for MCMC. There are some interesting special cases of eq. (2.829).

1. Let $\mathbf{E} = \mathbf{R} = 0$ and let $\mathbf{F} = \mathbf{I}$. Then

$$
\hat{q}_{\sqrt{\delta}} = q_0 - \frac{\delta}{2} \nabla_q U(q_0) + \sqrt{\delta} p_0.
$$

(2.830)

Since $p_0 \sim \text{Normal}(0, \mathbf{I})$, this is just the Euler-Maruyama discretization of the Langevin diffusion $d q_t = -\nabla_q U(q_t) \, dt + dB_t$.

2. If $\mathbf{E} = \mathbf{R} = 0$ and $\mathbf{F} \equiv \mathbf{F}(q)$ then

$$
\hat{q}_{\sqrt{\delta}} = q_0 - \frac{\delta}{2} \mathbf{F}(q_0) \mathbf{F}(q_0)^\top \nabla_q U(q_0) + \sqrt{\delta} \mathbf{F}(q_0) p_0.
$$

(2.831)
In this case, eq. (2.831) has two possible interpretations. First, if $F$ is constant as a function of $q$, then $G = (FF^\top)^{-1}$ is also constant and eq. (2.831) corresponds to a discretization of the Riemannian Langevin diffusion

$$dq_t = -G^{-1}\nabla_q U(q_t) \, dt + \sqrt{G^{-1}} \, dB_t \quad (2.832)$$

with $G^{-1}\nabla_q U(q)$ the Riemannian natural gradient and $\sqrt{G^{-1}} \, dB_t$ the Riemannian Brownian motion. On the other hand, even when $G(q)$ is not flat, eq. (2.831) instead corresponds to the proposal distribution of the simplified manifold Langevin MCMC procedure described in [Girolami and Calderhead 2011].

Thus, we see that eq. (2.829) generalizes some well-known methods for MCMC based on Langevin diffusion.
Chapter 3

Diffeomorphism Monte Carlo

In this chapter, our purpose is to construct and consider a family of MCMC techniques based on diffeomorphisms (smooth functions with smooth inverses). We call this method 

\textit{diffeomorphism Monte Carlo}. We present material as a natural means by which to formally discuss Hamiltonian Monte Carlo and its variations as Markov chains. It turns out that the transition kernels of diffeomorphism Monte Carlo contain the transition kernels of Metropolis-Hastings type, though this will not be evident at first glance. We begin by stating the fundamental assumptions of this chapter.

3.1 Construction of the Diffeomorphism Monte Carlo Transition Kernel

\textit{Assumption} 3.1.1. Assume that $\mathcal{X}$ and $\mathcal{Y}$ are manifold structures with non-negative measures $\mu : \mathcal{B}(\mathcal{X}) \rightarrow \mathbb{R}_+$ and $\nu : \mathcal{B}(\mathcal{Y}) \rightarrow \mathbb{R}_+$. Assume moreover that given a diffeomorphism $\Phi : \mathcal{X} \rightarrow \mathcal{Y}$ there exists a function, denoted $|\det(\nabla \Phi(\cdot))| : \mathcal{X} \rightarrow \mathbb{R}$, called the Jacobian determinant, such that

$$
\int_{\mathcal{Y}} f(y) \, \nu(dy) = \int_{\mathcal{X}} f(\Phi(x)) |\det(\nabla \Phi(x))| \, \mu(dx).
$$

(3.1)
For a thorough treatment of integration on manifolds, we refer the interested reader to Abraham et al. [1988].

The key definition appearing in this chapter, which is referenced elsewhere, is definition 3.1.4. This definition provides a formal understanding of the Markov chain transition kernel appearing in Hamiltonian Monte Carlo and other methods based on smooth, invertible transformations. Critically, the resulting transition kernel satisfies detailed balance, which is proved in proposition 3.1.5. We then go further to consider a special case that is nonetheless prevalent in applications of diffeomorphism Monte Carlo, wherein the state space is partitioned into two components, \( q \) and \( p \), and wherein the density of the stationary distribution \( \pi(q, p) = \pi(q)\pi(p|q) \) is such that \( \pi(p|q) \) is easy to sample. In this regime, we discuss marginal transition kernels of diffeomorphism Monte Carlo in proposition 3.1.9 and how they, too, satisfy detailed balance with respect to the marginal distribution whose density is \( \pi(q) \). We begin with a lemma.

**Lemma 3.1.2.** Let \( \mathcal{B}(\mathcal{X}) \) denote the Borel \( \sigma \)-algebra on \( \mathcal{X} \). Let \( A, B \in \mathcal{B}(\mathcal{X}) \) and let \( \Phi \) be a diffeomorphism from \( \mathcal{X} \) to itself. Let \( \pi : \mathcal{X} \to \mathbb{R}_+ \) be a probability density on \( \mathcal{X} \) with respect to measure \( \mu \). Then, for \( \beta \in (0, 1) \),

\[
\int_{B \cap \Phi^{-1}(A)} \min \left\{ 1, \frac{1 - \beta \pi(\Phi(x))}{\beta \pi(x)} \left| \det(\nabla \Phi(x)) \right| \right\} \pi(x) \mu(dx) = \int_{A \cap \Phi(B)} \min \left\{ 1, \frac{1 - \beta \pi^{-1}(x)}{\beta \pi(x)} \left| \det(\nabla \Phi^{-1}(x)) \right| \right\} \pi(x) \mu(dx)
\]

(3.2)
Proof.

\[
\int_{B \cap \Phi^{-1}(A)} \min \left\{ \frac{1}{\beta} \frac{\pi(x)}{\pi(\Phi(x))}, \beta \frac{\pi(\Phi(x))}{\pi(x)} \right\} \pi(x) \mu(dx) \tag{3.4}
\]

\[
= \int_{B \cap \Phi^{-1}(A)} \min \left\{ \frac{\pi(x)}{\det(\nabla \Phi(x))}, \beta \frac{1 - \beta}{\pi(\Phi(x))} \right\} \det(\nabla \Phi(x)) \mu(dx) \tag{3.5}
\]

\[
= \int_{B \cap \Phi^{-1}(A)} \min \left\{ \frac{\pi(x)}{\det(\nabla \Phi(x))}, \beta \frac{1 - \beta}{\pi(\Phi(x))} \right\} \det(\nabla \Phi(x)) \mu(dx) \tag{3.6}
\]

\[
= \int_{\Phi(B) \cap A} \frac{\pi(\Phi^{-1}(x))}{\beta} \left\{ \det(\nabla \Phi^{-1}(x)) \right\} \pi(x) \mu(dx) \tag{3.7}
\]

\[
= \int_{A \cap \Phi(B)} \frac{1}{\beta} \left( \frac{\pi(\Phi^{-1}(x))}{\pi(x)} \right) \left\{ \det(\nabla \Phi^{-1}(x)) \right\} \pi(x) \mu(dx). \tag{3.8}
\]

By swapping the role of \( \Phi \) and \( \Phi^{-1} \) and of \( \beta \) and \( 1 - \beta \) we immediately prove:

\[\square\]

**Corollary 3.1.3.** Under the same conditions as lemma 3.1.2 we have,

\[
\int_{B \cap \Phi^{-1}(A)} \min \left\{ \frac{\beta}{1 - \beta} \frac{\pi(\Phi^{-1}(x))}{\pi(x)} \right\} \pi(x) \mu(dx) \tag{3.9}
\]

\[
= \int_{A \cap \Phi^{-1}(B)} \min \left\{ \frac{\beta}{1 - \beta} \frac{\pi(\Phi(x))}{\pi(x)} \right\} \pi(x) \mu(dx) \tag{3.10}
\]

**Definition 3.1.4.** Let \( \Phi : \mathbb{R}^m \to \mathbb{R}^m \) be a diffeomorphism and let \( \beta \in (0, 1) \). The diffeo-
morphism Monte Carlo transition kernel has the form,

\[
K(x, A) = \beta \left( \min \left\{ 1, \frac{1 - \beta \pi(\Phi(x))}{\beta \pi(x)} |\det(\nabla \Phi(x))| \right\} \mathbf{1}\{\Phi(x) \in A\} \\
+ \left( 1 - \min \left\{ 1, \frac{1 - \beta \pi(\Phi(x))}{\beta \pi(x)} |\det(\nabla \Phi(x))| \right\} \right) \mathbf{1}\{x \in A\} \right) \\
+ (1 - \beta) \left( \min \left\{ 1, \frac{\beta \pi(\Phi^{-1}(x))}{1 - \beta \pi(x)} |\det(\nabla \Phi^{-1}(x))| \right\} \mathbf{1}\{\Phi^{-1}(x) \in A\} \\
+ \left( 1 - \min \left\{ 1, \frac{\beta \pi(\Phi^{-1}(x))}{1 - \beta \pi(x)} |\det(\nabla \Phi^{-1}(x))| \right\} \right) \mathbf{1}\{x \in A\} \right) \\
\right)
\tag{3.11}
\]

The diffeomorphism Monte Carlo transition kernel can therefore be understood to attempt to make a transition using \( \Phi \) with probability \( \beta \) and otherwise to attempt a transition using \( \Phi^{-1} \). In the case when \( \beta = 1/2 \), we call the diffeomorphism Monte Carlo symmetric.

**Proposition 3.1.5.** The diffeomorphism Monte Carlo transition kernel satisfies detailed balance (definition 2.7.7) with respect to the density \( \pi \).

**Proof.** Let \( A, B \in \mathcal{B}(\mathcal{X}) \). From lemma 3.1.2 we have,

\[
\beta \int_B \min \left\{ 1, \frac{1 - \beta \pi(\Phi(x))}{\beta \pi(x)} |\det(\nabla \Phi(x))| \right\} \mathbf{1}\{\Phi(x) \in A\} \pi(x) \mu(dx) \\
= \beta \int_A \min \left\{ 1, \frac{\beta \pi(\Phi^{-1}(x))}{1 - \beta \pi(x)} |\det(\nabla \Phi^{-1}(x))| \right\} \mathbf{1}\{\Phi^{-1}(x) \in B\} \pi(x) \mu(dx) \\
= (1 - \beta) \int_A \min \left\{ 1, \frac{\beta \pi(\Phi^{-1}(x))}{1 - \beta \pi(x)} |\det(\nabla \Phi^{-1}(x))| \right\} \mathbf{1}\{\Phi^{-1}(x) \in B\} \pi(x) \mu(dx)
\tag{3.12}
\]

\[
\int_B \min \left\{ 1, \frac{\beta \pi(\Phi^{-1}(x))}{1 - \beta \pi(x)} |\det(\nabla \Phi^{-1}(x))| \right\} \mathbf{1}\{\Phi^{-1}(x) \in B\} \pi(x) \mu(dx)
\tag{3.13}
\]

\[
\int_A \min \left\{ 1, \frac{\beta \pi(\Phi^{-1}(x))}{1 - \beta \pi(x)} |\det(\nabla \Phi^{-1}(x))| \right\} \mathbf{1}\{\Phi^{-1}(x) \in B\} \pi(x) \mu(dx)
\tag{3.14}
\]
Similarly from corollary 3.1.3,

\[(1 - \beta) \int_B \min \left\{ 1, \frac{\beta}{1 - \beta} \frac{\pi(\Phi^{-1}(x))}{\pi(x)} |\det(\nabla \Phi^{-1}(x))| \right\} 1 \{ \Phi^{-1}(x) \in A \} \pi(x) \mu(dx) \]

\[(3.15)\]

\[(1 - \beta) \int_A \min \left\{ \frac{\beta}{1 - \beta} \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} 1 \{ \Phi(x) \in B \} \pi(x) \mu(dx) \]

\[(3.16)\]

\[= \beta \int_A \min \left\{ 1, \frac{1 - \beta}{\beta} \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} 1 \{ \Phi(x) \in B \} \pi(x) \mu(dx) \]

\[(3.17)\]

From these two equations it immediately follows that the diffeomorphism Monte Carlo transition kernel satisfies detailed balance.

In the case that the diffeomorphism \( \Phi \) is also an involution (i.e. \( \Phi = \Phi^{-1} \)), then the resulting method is called involutive Monte Carlo [Neklyudov et al., 2020].

**Proposition 3.1.6.** Let \( \Phi \) be a a diffeomorphism that is also an involution. Let \( K \) be the Markov chain transition kernel for symmetric (i.e. \( \beta = 1/2 \)) diffeomorphism Monte Carlo. Then the transition kernel has the form,

\[ K(x, A) = \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} 1 \{ \Phi(x) \in A \} \]

\[+ \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} \right) 1 \{ x \in A \} \]

\[(3.18)\]
Proof. In this case, some elementary simplifications yield,

\[
K(x, A) = \frac{1}{2} \left( \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} 1 \{\Phi(x) \in A\} 
\right.
\]

\[
+ \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} \right) 1 \{x \in A\} 
\]

\[
+ \frac{1}{2} \left( \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} 1 \{\Phi(x) \in A\} 
\right.
\]

\[
+ \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} \right) 1 \{x \in A\} 
\]

\[
= \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} 1 \{\Phi(x) \in A\} 
\]

\[
+ \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} \right) 1 \{x \in A\} 
\]

(3.19)

(3.20)

Making the choice \( \beta = 1/2 \) may seem arbitrary in involutive Monte Carlo, but it is actually the optimal choice in a certain sense. We first require a lemma.

Lemma 3.1.7. Let \( x \in \mathbb{R}_+ \). A function of the form

\[
L(\beta) = \beta \min \left\{ 1, \frac{1 - \beta}{\beta} x \right\} + (1 - \beta) \min \left\{ 1, \frac{\beta}{1 - \beta} x \right\} 
\]

(3.21)

is maximized when \( \beta = 1/2 \). When \( x = 1 \), \( \beta = 1/2 \) is the unique maximizer.

Proof. Write \( L(\beta) \) as

\[
L(\beta) = \min \{\beta, (1 - \beta)x\} + \min \{1 - \beta, \beta x\} 
\]

(3.22)

We have several cases to consider. First suppose that \( \beta \leq (1 - \beta)x \) and \( 1 - \beta \leq \beta x \). In
In this case,

\[ L(\beta) = \beta + 1 - \beta \]  \hspace{1cm} (3.23)

\[ = 1 \]  \hspace{1cm} (3.24)

so any choice of \( \beta \) (satisfying the above conditions) will do. In this case we have \( x \geq \max \{ \beta/(1 - \beta), (1 - \beta)/\beta \} \). The quantity on the right-hand side is not less than one, which is achieved when \( \beta = 1/2 \). Similarly, consider the case \( \beta \geq (1 - \beta)x \) and \( \beta x \leq 1 - \beta \). In this case,

\[ L(\beta) = (1 - \beta)x + \beta x \]  \hspace{1cm} (3.25)

\[ = 1. \]  \hspace{1cm} (3.26)

In this case, we have \( x \leq \min \{ \beta/(1 - \beta), (1 - \beta)/\beta \} \). The right-hand side is not greater than one, which is achieved when \( \beta = 1/2 \). Suppose \( \beta \leq (1 - \beta)x \) and \( \beta x \leq 1 - \beta \). In this case,

\[ L(\beta) = \beta + \beta x. \]  \hspace{1cm} (3.27)

The above conditions on \( \beta \) imply \( \beta/(1 - \beta) \leq x \leq (1 - \beta)/\beta \), so that \( \beta \in (0, 1/2] \) or else these bounds do not hold. Since \( x \) is positive, the function \( L \) is maximized by choosing \( \beta \) as large as possible, in which case \( \beta = 1/2 \) is optimal. Similarly suppose \( \beta \geq (1 - \beta)x \) and \( 1 - \beta \leq \beta x \). In this case

\[ L(\beta) = (1 - \beta)x + 1 - \beta. \]  \hspace{1cm} (3.28)

The conditions on \( \beta \) imply \( (1 - \beta)/\beta \leq x \leq \beta/(1 - \beta) \), for which we require \( \beta \in [1/2, 1) \). The function \( L \) is therefore maximized by setting \( \beta \) as small as possible so that \( \beta = 1/2 \)
is optimal.

To see that \( \beta = 1/2 \) is in fact the only optimal choice, suppose that \( x = 1 \). In this case the function \( L \) has the form,

\[
L(\beta) = \begin{cases} 
2\beta & \text{for } \beta \in (0, 1/2] \\
2(1 - \beta) & \text{for } \beta \in [1/2, 0).
\end{cases}
\] (3.29)

In the first case, the function is increasing for \( \beta \in (0, 1/2] \). In the second case, the function is decreasing for \( \beta \in [1/2, 0) \). Therefore, \( \beta = 1/2 \) is the optimal choice.

**Proposition 3.1.8.** In involutive Monte Carlo, the choice \( \beta = 1/2 \) is the only choice that consistently maximizes the acceptance probability.

**Proof.** Let \( \Phi \) be a smooth involution but suppose \( \beta \in (0, 1) \) is a free parameter. By examining the Markov chain transition kernel of diffeomorphism Monte Carlo in definition 3.1.4 with \( \Phi = \Phi^{-1} \), one observes that the probability of transitioning to the state \( \Phi(x) \) is

\[
A(\beta) = \beta \min \left\{ 1, \frac{1 - \beta}{\beta} \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} + (1 - \beta) \min \left\{ 1, \frac{\beta}{1 - \beta} \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\}.
\] (3.30)

We therefore regard eq. (3.30) as the acceptance probability for the proposal state \( \Phi(x) \) given the current state \( x \). We recognize the acceptance probability as a function of the form in lemma 3.1.7 for which the choice \( \beta = 1/2 \) produces the maximal acceptance probability.

An important limitation of diffeomorphism Monte Carlo is that, in general, it can suffer from complete failures of ergodicity as the subsequent example demonstrates.

**Example 31.** Consider the smooth involution \( \Phi(x) = -x \), for which \( |\det(\nabla \Phi(x))| = 1 \).
Given an initial position $x$, the sequence of samples computed by the involutive Monte Carlo transition kernel will be a sequence of $x$ and $-x$. This shows the important distinction between detailed balance and stationarity, and ergodicity of the Markov chain.

In practical settings, it is typical to consider diffeomorphism Monte Carlo on augmented state spaces and with an augmented target probability measure. The next proposition shows that such constructions respect detailed balance on the original, unaugmented space, too.

**Proposition 3.1.9.** Let $\pi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$ be a probability density with marginal density $\mathcal{X} \ni q \mapsto \pi(q) \in \mathbb{R}_+$ with respect to the measure $\nu$ on $\mathcal{B}(\mathcal{X})$ and conditional density $\mathcal{Y} \ni p \mapsto \pi(p|q) \in \mathbb{R}_+$ with respect to the measure $\mu$ on $\mathcal{B}(\mathcal{Y})$ such that the joint density satisfies $\pi(q, p) = \pi(p|q)\pi(q)$. Let $K : (\mathcal{X} \times \mathcal{Y}) \times \mathcal{B}(\mathcal{X} \times \mathcal{Y}) \to [0, 1]$ be a Markov chain transition kernel for which $\pi(q, p)$ satisfies detailed balance. Then the marginal transition kernel on $\mathcal{X}$,

$$\tilde{K}(q, A) = \int_{\mathcal{Y}} K((q, p), (A, \mathcal{Y}))\pi(p|q) \mu(dp)$$

(3.31)

satisfies detailed balance with respect to $\pi(q)$.

**Proof.** Let $B \in \mathcal{B}(\mathcal{X})$. Then,

$$\int_B \tilde{K}(q, A)\pi(q) \nu(dq) = \int_B \left( \int_{\mathcal{Y}} K((q, p), (A, \mathcal{Y}))\pi(p|q) \mu(dp) \right) \pi(q) \nu(dq)$$

(3.32)

$$= \int_B \int_{\mathcal{Y}} K((q, p), (A, \mathcal{Y}))\pi(p|q) \mu(dp) \nu(dq)$$

(3.33)

$$= \int_A \int_{\mathcal{Y}} K((q, p), (B, \mathcal{Y}))\pi(p|q) \mu(dp) \nu(dq)$$

(3.34)

$$= \int_A \left( \int_{\mathcal{Y}} K((q, p), (B, \mathcal{Y}))\pi(p|q) \mu(dp) \right) \pi(q) \nu(dq)$$

(3.35)

$$= \int_A \tilde{K}(q, B)\pi(q) \nu(dq),$$

(3.36)
which shows detailed balance.

**Corollary 3.1.10.** Given a diffeomorphism on $\mathcal{X} \times \mathcal{Y}$ and a joint density $\pi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+$, diffeomorphism Monte Carlo (definition 3.1.4) satisfies detailed balance with respect to $\pi(q, p)$ from proposition 3.1.5. Hence, the marginal transition kernel of diffeomorphism Monte Carlo in eq. (3.31) satisfies detailed balance with respect to the density $q \mapsto \int_{\mathcal{Y}} \pi(q, p) \mu(dp)$.

Involutive Monte Carlo actually generalizes a broad class of Markov transition kernels. For instance, all transition kernels of Metropolis-Hastings type can be characterized as instances of the marginal chain of involutive Monte Carlo.

**Example 32 ([Neklyudov et al., 2020]).** Let $\pi : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$ be a probability density with marginal density $\mathcal{X} \ni q \mapsto \pi(q) \in \mathbb{R}_+$ and conditional density $\mathcal{X} \ni p \mapsto \tilde{\pi}(p|q) \in \mathbb{R}_+$ with respect to the measure $\mu$ on $\mathcal{B}(\mathcal{X})$ such that $\pi(q, p) = \tilde{\pi}(p|q) \pi(q)$. Consider the involution $\Phi(q, p) = (p, q)$, which is volume preserving and therefore has unit Jacobian determinant. For $A, B \in \mathcal{B}(\mathcal{X})$, the transition kernel of involutive Monte Carlo becomes,

$$K((q, p), (A, B)) = \min \left\{ 1, \frac{\pi(p)\tilde{\pi}(q|p)}{\pi(q)\tilde{\pi}(p|q)} \right\} \mathbf{1}\{(p, q) \in A \times B\} + \left( 1 - \min \left\{ 1, \frac{\pi(p)\tilde{\pi}(q|p)}{\pi(q)\tilde{\pi}(p|q)} \right\} \right) \mathbf{1}\{(q, p) \in A \times B\}.$$ \hspace{1cm} (3.37)

Therefore, the marginal transition kernel from proposition 3.1.9 is

$$\tilde{K}(q, A) = \int_{\mathcal{X}} K((q, p), (A, \mathcal{X})) \tilde{\pi}(p|q) \mu(dp)$$

$$= \int_A \min \left\{ 1, \frac{\pi(p)\tilde{\pi}(q|p)}{\pi(q)\tilde{\pi}(p|q)} \right\} \tilde{\pi}(p|q) \mu(dp)$$

$$+ \left( 1 - \int_{\mathcal{X}} \min \left\{ 1, \frac{\pi(p)\tilde{\pi}(q|p)}{\pi(q)\tilde{\pi}(p|q)} \right\} \tilde{\pi}(p|q) \mu(dp) \right) \mathbf{1}\{q \in A\}.$$ \hspace{1cm} (3.39)

This is the transition kernel of a Markov chain of Metropolis-Hastings type.

**Definition 3.1.11.** We define the momentum flip operator as the map $F(q, p) \overset{\text{def}}{=} (q, -p)$. 

174
Definition 3.1.12. Let $\pi(q, p) = \pi(p|q)\pi(q)$ be a probability density on $\mathbb{R}^m \times \mathbb{R}^m \cong \mathbb{R}^{2m}$.

Let $\hat{\Phi}_\epsilon : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ be a symmetric and symplectic numerical method with step-size $\epsilon$.

Let $k \in \mathbb{N}$ be the number of integration steps. Let $F$ be the momentum flip operator in definition 3.1.11. The Hamiltonian Monte Carlo transition kernel is the transition kernel of involutive Monte Carlo with smooth involution $F \circ \hat{\Phi}_\epsilon^k$; that is, when $A \in \mathcal{B}(\mathbb{R}^{2m})$,

$$K((q, p), A) = \alpha(q, p) \mathbf{1} \left\{ F \circ \hat{\Phi}_\epsilon^k(q, p) \in A \right\}$$

$$+ \quad (1 - \alpha(q, p)) \mathbf{1} \left\{ x \in A \right\}$$

where the quantity

$$\alpha(q, p) = \min \left\{ 1, \frac{\pi(F \circ \hat{\Phi}_\epsilon^k(q, p))}{\pi(q, p)} \right\}$$

is called the acceptance rate. Since, in this case $\left| \det(\nabla \hat{\Phi}_\epsilon^k(z)) \right| = 1$ for every $z \in \mathbb{R}^{2m}$ by symplecticness of $\hat{\Phi}_\epsilon$.

Corollary 3.1.13. Let $K$ be the transition kernel of Hamiltonian Monte Carlo (definition 3.1.12). Then the marginal transition kernel $\tilde{K}$ on $\mathbb{R}^m$ defined by,

$$\tilde{K}(q, A) = \int_{\mathbb{R}^m} K((q, p), (A, \mathbb{R}^m)) \pi(p|q) \, dp$$

satisfies detailed balance with respect to $\pi(q)$.

Proof. This follows immediately from proposition 3.1.9 since $K$ satisfies detailed balance with respect to $\pi(q, p)$.

Notice that, until this point, we have not actually specified which vector field is being integrated in Hamiltonian Monte Carlo (see definition 3.1.12); we have merely assumed that $\hat{\Phi}_\epsilon$ is symmetric and symplectic. One natural choice is the Hamiltonian vector field
corresponding to the Hamiltonian $H(q, p) = \log \pi(q, p) = \log \pi(q) + \log \pi(p|q)$ (or this Hamiltonian plus some additive constant) because, in this case, we may expect that the integrator $\hat{\Phi}_\epsilon$ will *approximately* conserve the Hamiltonian. This follows because the actual Hamiltonian flow map conserves the Hamiltonian and $\hat{\Phi}_\epsilon$ will approximate the actual flow map (see e.g. definition 2.2.2 proposition 2.2.16 and corollary 2.2.55) with second-order error because it is symmetric from theorem 2.2.23. The composition of the integrator with the momentum flip (definition 3.1.11) will not degrade conservation of the Hamiltonian provided $\pi(p|q) = \pi(-p|q)$ (e.g. $\pi(p|q) = \text{Normal}(p; 0, G(q))$ where $G(q)$ is a positive definite matrix). As seen from the form of the acceptance rate (eq. (3.42)) in Hamiltonian Monte Carlo, conservation of the Hamiltonian will imply a large acceptance probability, leading to more frequent Markov chain transitions. However, this is not the only choice of Hamiltonian vector field and one may decide that $\hat{\Phi}_\epsilon$ should approximate the solution of some other Hamiltonian vector field; this is the distinction between the *acceptance Hamiltonian*, which has the form $H(q, p) = \log \pi(q, p) = \log \pi(q) + \log \pi(p|q)$, and the *guidance Hamiltonian*, which may be arbitrarily chosen, as discussed in Duane et al. [1987]. For instance, in Brubaker et al. [2012], guidance Hamiltonians corresponding to geodesic motion are considered to motivate a Markov chain whose random walk behavior incorporates geometric information. However, in most cases, choosing the guidance Hamiltonian and acceptance Hamiltonian to be equal is a compelling selection.

More generally, we can consider reversible numerical integrators that are not necessarily symplectic (and therefore not volume preserving). Provided that such an integrator is a smooth diffeomorphism, it is possible to construct a Markov chain transition kernel satisfying detailed balance using definition 3.1.4 or proposition 3.1.6 if the diffeomorphism can be easily made into an involution. This is the subject of Lagrangian Monte Carlo [Lan et al. 2015].

In implementations of Hamiltonian Monte Carlo, it is common to construct the transition kernels as a mixture over the number of integration steps [Neal 2010b Betancourt 2017].
In particular, in definition 3.1.12 we have specified an involution by \( F \circ \hat{\Phi}_\epsilon^k \), where \( k \) is the number of integration steps. Let us introduce the notation \( K(\cdot, \cdot; \epsilon, k) \) to mean the transition kernel of Hamiltonian Monte Carlo with integration step-size \( \epsilon \) and \( k \) integration steps. Given a sequence \( (\alpha_1, \alpha_2, \ldots) \) satisfying \( \alpha_k \geq 0 \) for \( k = 1, 2, \ldots \) and \( \sum_{k=1}^{\infty} \alpha_k = 1 \), we may construct the mixture transition kernel

\[
K(\cdot, \cdot; \epsilon) = \sum_{k=1}^{\infty} \alpha_k K(\cdot, \cdot; \epsilon, k),
\]

(3.44)

which can be interpreted as the transition kernel of Hamiltonian Monte Carlo with a randomized number of integration steps, wherein the probability mass assigned to each number of integration steps \( k \) is specified by \( \alpha_k \). It is also possible to consider step-size randomization, which is a component of the analytical apparatus in Bou-Rabee and Sanz-Serna [2017]. In fact, the randomization of the number of integration steps plays an important role in the available geometric ergodicity theory of Hamiltonian Monte Carlo.

Theorem 3.1.14 (Livingstone et al. [2016]). Consider a Hamiltonian of the form \( H(q, p) = -\log \pi(q) + \frac{1}{2} p^\top p \), where \( q, p \in \mathbb{R}^m \), and consider Hamiltonian Monte Carlo with involution \( F \circ \hat{\Phi}_\epsilon^k : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m \), where \( \hat{\Phi}_\epsilon \) is the leapfrog integrator applied to the Hamiltonian \( H(q, p) \). Suppose that the number of integration steps \( k \) is a random variable for which \( \alpha_1 \overset{\text{def}}{=} \Pr(k = 1) > 0 \), and consider the Markov chain with transition kernel eq. (3.44). Suppose further that for any \( (q, p) \in \mathbb{R}^{2m} \) we have that there exists \( s < \infty \) such that

\[
\mathbb{E}_k [\exp(s \| q \|)] < \infty,
\]

(3.45)

where \( (\tilde{q}, \tilde{p}) = F \circ \hat{\Phi}_\epsilon^k(q, p) \). Let \( \mu_{\epsilon,k}(q, p) = \tilde{q} - k \epsilon p \). Then Hamiltonian Monte Carlo is
geometrically ergodic if for each $k$ and for some $\delta \in (1/2, 1]$

$$\limsup_{\|q\| \to \infty, \|p\| \leq \|q\|^\delta} (\|\mu_{e,k}(q, p)\| - \|q\|) < -\sqrt{2k\epsilon} \frac{\Gamma((m + 1)/2)}{\Gamma(m/2)}$$  \hspace{1cm} (3.46)

and if for each $k$

$$\lim_{\|q\| \to \infty} \int_{R(q) \cap I(q)} P_{e,k}(q, d\tilde{q}) = 0$$  \hspace{1cm} (3.47)

where,

$$P_{e,k}(q, A) = \int_{\mathbb{R}^m} 1 \left\{ \Phi_e^k(q, p) \in A \times \mathbb{R}^m \right\} \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} p^\top p \right) dp \hspace{1cm} (3.48)$$

$$R(q) = \left\{ \tilde{q} \in \mathbb{R}^m : \min \left\{ 1, \frac{\exp(-H(\tilde{q}, \tilde{p}))}{\exp(-H(q, p))} \right\} < 1 \right\}$$  \hspace{1cm} (3.49)

$$I(q) = \{ \tilde{q} \in \mathbb{R}^m : \|\tilde{q}\| \leq \|q\| \}.$$  \hspace{1cm} (3.50)

### 3.2 Variations of Diffeomorphism Monte Carlo

We now give some examples of Markov chain transition kernels that have appeared in the literature as instances of diffeomorphism Monte Carlo.

**Example 33.** So far, we have seen examples of Hamiltonian Monte Carlo that employ integrators that can be easily turned into involutions by negating the momentum at the end of the trajectory. This is not always the case, however. We may, for instance, consider the Hamiltonian vector field,

$$\dot{z}_t = B \nabla H(z_t),$$  \hspace{1cm} (3.51)

where $H : \mathbb{R}^{2m} \to \mathbb{R}$ is an arbitrary smooth Hamiltonian and $B \in \mathbb{R}^{2m \times 2m}$ is an arbitrary, skew-symmetric, and invertible matrix. These equations of motion are called non-
canonical Hamiltonian dynamics [Marsden and Ratiu, 2010]. Non-canonical Hamiltonian
dynamics have been proposed in the Markov chain literature for sampling; see Brofos
and Lederman [2020b], Tripuraneni et al. [2017]. The transition kernel is constructed as
follows. Let \( \hat{\Phi}_\epsilon \) be a symmetric and symplectic numerical method that is a second-order
approximation to eq. (3.51); for instance the implicit midpoint integrator is a suitable
choice in general, though other integrators are possible, too, such as Scovel’s method in
the case of magnetic dynamics [Leimkuhler and Reich, 2005]. However, unlike canonical
HMC, integrators of non-canonical Hamiltonian dynamics cannot be augmented with a
momentum flip in order to produce an involution. Indeed, Tripuraneni et al. [2017] go to
great lengths to augment the state space such that an involution may be computed. Nev-
evertheless, a more direct approach, which was pursued implicitly by Brofos and Lederman
[2020b], is to observe that multistep methods \( \hat{\Phi}_\epsilon^k \) are diffeomorphisms, with smooth in-
verse \( \hat{\Phi}_\epsilon^{-k} \) (since these integrators are symmetric methods), thereby facilitating the use of
definition 3.1.4 in order to construct a Markov chain transition kernel satisfying detailed
balance. Writing \( z = (q, p) \), suppose in particular that \( H(q, p) = -\log \pi(q) + \frac{1}{2} p^\top p \)
and let \( \pi(q, p) \propto \exp(-H(q, p)) \). The Markov chain transition kernel of non-canonical
Hamiltonian Monte Carlo is,

\[
K((q, p), A) = \frac{1}{2} \left( \min \left\{ 1, \frac{\pi(\hat{\Phi}_\epsilon^k(q, p))}{\pi(x)} \right\} \mathbf{1} \left\{ \hat{\Phi}_\epsilon^k(q, p) \in A \right\} 
+ \left( 1 - \min \left\{ 1, \frac{\pi(\hat{\Phi}_\epsilon(q, p))}{\pi(x)} \right\} \right) \mathbf{1} \left\{ (q, p) \in A \right\} 
+ \frac{1}{2} \left( \min \left\{ 1, \frac{\pi(\hat{\Phi}_\epsilon^{-k}(q, p))}{\pi(q, p)} \right\} \mathbf{1} \left\{ \hat{\Phi}_\epsilon^{-k}(q, p) \in A \right\} 
+ \left( 1 - \min \left\{ 1, \frac{\pi(\hat{\Phi}_\epsilon^{-k}(q, p))}{\pi(q, p)} \right\} \right) \mathbf{1} \left\{ (q, p) \in A \right\} \right) 
\]

(3.52)

This provides an example of a Markov chain transition kernel that is assembled from a
diffeomorphism but not from an involution.
Example 34. The framework of diffeomorphism Monte Carlo also presents the opportunity to implement variations of Hamiltonian Monte Carlo that have not, to the best of our knowledge, been considered previously in the literature. Let $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be a smooth Hamiltonian. The vast majority of implementations of Hamiltonian Monte Carlo utilize the leapfrog integrator in the case wherein $H$ is a separable Hamiltonian and the generalized leapfrog integrator (definition 2.2.41) in the case where $H$ is non-separable (derived from a variational principle in example 16). Combined with the negation of the momentum variable, this produces an involution. As an alternative, one may consider constructing a diffeomorphism from the Euler-A integrator, whose inverse is the Euler-B integrator with a negated step-size (see definition 2.2.52 and lemma 2.2.54). In comparison to the leapfrog method, which is second-order due to its symmetry, the Euler-A and Euler-B methods are only first-order numerical integrators from lemma 2.2.53. However, the Euler-A and Euler-B methods enjoy a computational advantage in that a single step only requires the solution to a single fixed point equation, whereas the generalized leapfrog integrator must solve two fixed point equations in a single step. In some cases, it may be desirable to trade-off greater computational accuracy for a faster implementation.

Example 35. Dutta and Bhattacharya [2014] consider a variation of diffeomorphism Monte Carlo wherein the diffeomorphism assumes the special form $\Phi(q, p) = (\Psi(q, p), p)$. In this case, $\Psi : \mathcal{X} \times \mathcal{Y} \to \mathcal{X}$ is a smooth map which, for fixed $p$, is invertible; i.e. there exists $\Psi^{-1}(\cdot, p)$ such that $\Psi^{-1}(\Psi(q, p), p) = q$. This method was given the name “Markov Chain Monte Carlo Based on Deterministic Transformations.”
Chapter 4

Evaluating the Implicit Midpoint Integrator for Riemannian Manifold Hamiltonian Monte Carlo

This chapter is adapted from [Brofos and Lederman [2021a]] with additional exposition and experimentation. This work is a collaboration with Roy Lederman. The vast majority of the exposition is due to me, as is all of the experimentation.

Abstract. Riemannian manifold Hamiltonian Monte Carlo is traditionally carried out using the generalized leapfrog integrator. However, this integrator is not the only choice and other integrators yielding valid Markov chain transition operators may be considered. In this work, we examine the implicit midpoint integrator as an alternative to the generalized leapfrog integrator. We discuss advantages and disadvantages of the implicit midpoint integrator for Hamiltonian Monte Carlo, its theoretical properties, and an empirical assessment of the critical attributes of such an integrator for Hamiltonian Monte Carlo: energy conservation, volume preservation, and reversibility. Empirically, we find that while leapfrog iterations are faster, the implicit midpoint integrator has better energy conservation, leading to higher acceptance rates, as well as better conservation of volume and better reversibility, arguably yielding a more accurate sampling procedure.
Figure 4.1: Trajectories computed by the leapfrog integrator and the implicit midpoint integrator for a quadratic, separable Hamiltonian $H(q, p) = \frac{1}{2} q^T q + \frac{1}{2} p^T p$. Both integrators are stable but the leapfrog integrator deviates from the level sets of the Hamiltonian (faint blue circles) whereas every iterate of the implicit midpoint integrator lies on the same energy level set as the previous one.

4.1 Introduction

Riemannian manifold Hamiltonian Monte Carlo (RMHMC) is a powerful algorithm for sampling from Bayesian posterior distributions [Girolami and Calderhead, 2011]. Given a log-posterior function $\mathcal{L} : \mathbb{R}^m \rightarrow \mathbb{R}$ and a Riemannian metric $G : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times m}$ (a mapping satisfying the condition that $G(q)$ is positive definite for each $q \in \mathbb{R}^m$), RMHMC considers the Hamiltonian dynamics corresponding to the Hamiltonian,

$$H(q, p) = -\mathcal{L}(q) + \frac{p^T G^{-1}(q)p}{2} + \frac{\log \det(G(q))}{2}.$$  \hspace{1cm} (4.1)

See section 2.5 for a review of Riemannian manifolds with global coordinate systems. Riemannian metrics are incorporated into HMC in order to precondition dynamics and more efficiently explore the distribution. Irrespective of the choice of metric, the form of the Hamiltonian in eq. (4.1) corresponds to a Gibbs distribution proportional to $\exp(-H(q, p))$ and tractable conditional distribution $p|q \sim \text{Normal}(0, G(q))$. However, the form of this Hamiltonian is such that it cannot be written as the sum of two functions, each a func-
tion of \( q \) or \( p \) alone; such a Hamiltonian is called “non-separable.” The presence of a non-separable Hamiltonian presents unique challenges for numerical integration.

The leapfrog integrator (see e.g. definition 2.2.41) and its variants are a ubiquitous choice for the numerical integration of Hamiltonian mechanics for HMC; for instance see Brubaker et al. [2012], Tripuraneni et al. [2017], Neal [2010a], Byrne and Girolami [2013], Betancourt [2012], Girolami and Calderhead [2011], among many others. It may, therefore, not be apparent that numerical integrators other than the leapfrog method are applicable to HMC, provided that they exhibit two properties:

1. The integrator has a unit Jacobian determinant so that it preserves volume in \((q, p)\)-space.

2. The integrator is symmetric under negation of the momentum variable.

These properties are sufficient to prove that HMC satisfies detailed balance, which in turn establishes that the stationary distribution has density proportional to \( \exp(-H(q, p)) \); see Neal [2010a], Bishop [2006]. Detailed balance also follows immediately from the discussion of diffeomorphism Monte Carlo detailed in chapter 3. Detailed balance implies stationarity of the target distribution (lemma 2.7.8).

Why does the choice of numerical integrator matter? There are at least three reasons.

1. Numerical integrators differ with respect to energy conservation and stability. (For a review of linear stability analysis, we refer the reader to section 2.2.7) The acceptance probability of HMC depends on the energy conservation and the ability of the HMC proposal to use large integration steps depends on stability.

2. Numerical integrators may only satisfy properties (i) and (ii) above approximately, particularly if the integrators are defined as solutions to implicitly-defined equations. These approximate solvers of implicit equations will be discussed in more details below. Typically, the error of these methods will depend on a convergence tolerance.
δ used to find fixed-points of the integration step (algorithm [1]). For a non-zero convergence tolerance, the degree to which properties (i) and (ii) are violated will depend on the integrator and the tolerance.

3. Numerical integrators will differ in their efficiency in the sense that there may be structural properties of the Hamiltonian system that the integrator exploits. More efficient integrators will exhibit higher effective sample sizes *per second* when used in HMC.

The contribution of this work is to compare and contrast the generalized leapfrog integrator with the implicit midpoint method in application to RMHMC. We consider RMHMC because the non-separable Hamiltonian necessitates elaborate integration schemes which require solving implicitly-defined equations; this is in contrast to Euclidean HMC with constant $G$ which produces a separable Hamiltonian that can be integrated explicitly. First, we compare the two integrators on the energy conservation, volume preservation, and reversibility as discussed in reason (a). Second, we study the breakdown of *exact* satisfaction of properties (i) and (ii) in implicitly-defined integrators as described in reason (b). Third, we consider multiple variants of the generalized leapfrog and implicit midpoint integrators that exhibit different efficiencies, relevant to reason (c). We conclude that the implicit midpoint integrator exhibits superior energy conservation, conservation of volume, and symmetry compared to the generalized leapfrog integrator. We explore inference in sophisticated Bayesian inference tasks wherein the implicit midpoint integrator is competitive with, or exceeds, the time-normalized performance of the generalized leapfrog method. We therefore argue that the implicit midpoint integrator is a procedure worth consideration in RMHMC.
4.2 Background

For Bayesian inference tasks, $\mathcal{L}$ is the sum of the log-likelihood and the log-prior; in this circumstance, the typical form of the Riemannian metric $G$ is the sum of the Fisher information of the log-likelihood and negative Hessian of the log-prior; this choice of Riemannian metric is motivated by information geometry; see [Amari, 2016] and section 2.6.

The Hamiltonian in eq. (4.1) leads to the equations of motion,

\[
\dot{q}_i = \sum_{j=1}^{m} G^{-1}_{ij}(q)p_j \quad (4.2)
\]

\[
\dot{p}_i = -\frac{\partial}{\partial q_i} \mathcal{L}(q) - \frac{1}{2} \text{trace}\left( G^{-1}(q) \frac{\partial}{\partial q_i} G(q) \right) + \frac{1}{2} p^T G^{-1}(q) \frac{\partial^2}{\partial q_i^2} G(q) G^{-1}(q)p \quad (4.3)
\]

As stated in section 4.1, the standard integrator for RMHMC is the (generalized) leapfrog integrator. A naive implementation of a single step of the generalized leapfrog integrator with step-size $\epsilon$ and initial position $(q, p)$ is presented in algorithm 2. Notice that eqs. (4.7) and (4.8) are implicitly defined in the sense that the quantities appearing on the left-hand side also appear on the right-hand side; these equations are typically solved to a given tolerance $\delta \geq 0$ (in the sense defined in the fixed point iteration algorithm algorithm 1). When $\delta = 0$, the generalized leapfrog integrator satisfies properties (i) and (ii), however, in practice, the tolerance is often chosen to be larger than machine precision in order to reduce the number of fixed point iterations; therefore properties (i) and (ii) are no longer satisfied accurately.

The implicit midpoint method, an alternative to the generalize leapfrog integrator, is presented in algorithm 3; the implicit midpoint integrator also involves the solution to an implicitly-defined eq. (4.10). When $\delta = 0$, it is well-known that the implicit midpoint integrator satisfies property (i); see [Leimkuhler and Reich, 2005]. It also satisfies property
Algorithm 1 (Fixed Point Iteration) Procedure for solving the equation \( z = f(z) \) via fixed point iteration to a given tolerance.

1: **Input**: Function \( f : \mathbb{R}^m \rightarrow \mathbb{R}^m \), initial guess \( z \in \mathbb{R}^m \), fixed point convergence tolerance \( \delta \geq 0 \).

2: Set \( \Delta z = \infty \) and \( z' = z \).

3: While: \( \Delta z > \delta \) compute

\[
    z'' = f(z') \\
    \Delta z = \max_{i \in \{1, ..., m\}} |z''_i - z'_i| \\
    z' = z''
\]  

4: **Return**: \( z' \in \mathbb{R}^m \).

Algorithm 2 (G.L.F.(a)) The procedure for a single step of integrating Hamiltonian dynamics using the generalized leapfrog integrator.

1: **Input**: Hamiltonian \( H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \), initial position and momentum variables \((q, p) \in \mathbb{R}^m \times \mathbb{R}^m\), integration step-size size \( \epsilon \in \mathbb{R} \), fixed-point convergence tolerance \( \delta \geq 0 \).

2: Use algorithm 1 with tolerance \( \delta \) and initial guess \( p \) to solve for \( \bar{p} \),

\[
    \bar{p} \overset{\text{def.}}{=} p - \frac{\epsilon}{2} \nabla_q H(q, \bar{p}) \quad \text{subject to} \quad f(\bar{p}) 
\]  

3: Use algorithm 1 with tolerance \( \delta \) and initial guess \( q \) to solve for \( q' \),

\[
    q' \overset{\text{def.}}{=} q + \frac{\epsilon}{2} \left( \nabla_p H(q, \bar{p}) + \nabla_p H(q', \bar{p}) \right) \quad \text{subject to} \quad f(q') 
\]  

4: Compute the explicit update

\[
    p' \overset{\text{def.}}{=} \bar{p} - \frac{\epsilon}{2} \nabla_q H(q', \bar{p}) 
\]

5: **Return**: \((q', p') \in \mathbb{R}^m \times \mathbb{R}^m\).

(ii) for Hamiltonians of the form eq. (4.1); see section 4.A.1.

We turn now to discussing a theoretical property of numerical integrators related to conserved quantities.

**Definition 4.2.1.** Let \( z = (q, p) \) where the time evolution of \( q_i \) is given by eq. (4.2) and
Algorithm 3 (I.M.(a)) The procedure for a single step of integrating Hamiltonian dynamics using the implicit midpoint integrator.

1: **Input:** Hamiltonian $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, initial position and momentum variables $(q, p) \in \mathbb{R}^m \times \mathbb{R}^m$, integration step-size size $\epsilon \in \mathbb{R}$, fixed-point convergence tolerance $\delta \geq 0$.

2: Use algorithm [1] with tolerance $\delta$ and initial guess $(q, p)$ to solve for $(q', p')$

\[
\begin{pmatrix} q' \\ p' \end{pmatrix} \equiv \begin{pmatrix} q \\ p \end{pmatrix} + \epsilon \begin{pmatrix} \nabla_p H(\bar{q}, \bar{p}) \\ -\nabla_q H(\bar{q}, \bar{p}) \end{pmatrix}
\]

\[f(q', p')\]

where $\bar{q} \equiv (q' + q)/2$ and $\bar{p} \equiv (p' + p)/2$.

3: **Return:** $(q', p') \in \mathbb{R}^m \times \mathbb{R}^m$.

of $p_i$ by eq. (4.3) for $i = 1, \ldots, m$. A conserved quantity of $z$ is a real-valued function $z \mapsto \mathcal{G}(z)$ for which $\frac{d}{dt} \mathcal{G}(z) = 0$.

Conserved quantities are reviewed in detail for several numerical integrators in section 2.2.5. It is important to notice that definition 4.2.1 is a statement about the underlying dynamics and has nothing to do with the integrator used to approximate these dynamics. The properties of the integrators will be discussed in the next paragraph. For Hamiltonian systems, the canonical example of a conserved quantity is the Hamiltonian energy itself; see [Marsden and Ratiu, 2010]. Hamiltonian flows are also symplectic [Hairer et al., 2006] which implies conservation of volume (in the same sense as that of property (i)).

A numerical integrator cannot preserve all of the conserved quantities as the underlying ODE, but it may be able to conserve some simple ones. The following two results may be found in [Leimkuhler and Reich, 2005] and proofs are given in proposition 2.2.58 and corollary 2.2.57.

**Theorem 4.2.2.** Let $z = (q, p)$. The generalized leapfrog integrator (algorithm 2) with $\delta = 0$ preserves any conserved quantity of the form $\mathcal{G}(z) = q^\top A p + b^\top z$ where $A \in \mathbb{R}^{m \times m}$ is a symmetric matrix and $b \in \mathbb{R}^{2m}$.

**Theorem 4.2.3.** Let $z = (q, p)$. The implicit midpoint integrator (algorithm 3) with $\delta = 0$
preserves any conserved quantity of the form $\mathcal{G}(z) = z^\top A z + b^\top z$ where $A \in \mathbb{R}^{2m \times 2m}$ is a symmetric matrix and $b \in \mathbb{R}^{2m}$.

Notice that theorem 4.2.3 contains a strictly more general class of conserved quantity than theorem 4.2.2. We come now to a hypothesis that would justify the consideration of the implicit midpoint integrator within the context of HMC. Before stating the hypothesis, we provide some initial motivation for how the implicit midpoint integrator performs in the presence of a quadratic Hamiltonian.

**Proposition 4.2.4.** Let $H(q, p) \equiv H(z) = z^\top A z$ be a quadratic Hamiltonian. Then, for any step-size, the proposals generated by Hamiltonian Monte Carlo using the implicit midpoint integrator with $\delta = 0$ will be accepted.

**Proof.** The implicit midpoint integrator conserves quadratic first integrals (see theorem 4.2.3) and, in this case, the Hamiltonian energy is itself a quadratic function. □

Note, however, that perfect conservation of the Hamiltonian energy does not imply that the implicit midpoint integrator is the exact solution of the Hamilton’s equations of motion. Nevertheless, proposition 4.2.4 suggests an important difference between the generalized leapfrog integrator and the implicit midpoint method in terms of their conservation properties. Although Bayesian posterior distributions are unlikely to be Gaussian, it is widely accepted that Gaussian approximations are useful. Such notions materialize, for example, in the central limit theorem and the Laplace approximation. We therefore speculate that the Gaussian case may be useful for providing intuition for the more general case we examine in our experimental results. Provided the posterior is approximately Gaussian, therefore, this leads us to the following hypothesis.

**Hypothesis.** The implicit midpoint algorithm will exhibit higher acceptance probabilities than the generalized leapfrog integrator for the same step-size.

If true, and if the fixed point iterations required by the implicit midpoint procedure are not too burdensome relative to the generalized leapfrog integrator, then the higher
acceptance rate may produce more favorable effective sample sizes for the Markov chain whose transitions are computed using the implicit midpoint algorithm. In this scenario, the implicit midpoint integrator may be worth consideration as an alternative to the generalized leapfrog integrator.

We wish to emphasize that the behavior of the implicit midpoint integrator in the presence of a quadratic Hamiltonian is not a definitive explanation of all differences in sampling behaviors that may arise when using it as a transition operator in RMHMC. However, we believe that the setting of quadratic Hamiltonians, corresponding to Gaussian densities, can provide helpful intuition. In section 4.4 we will turn to the empirical evaluation of the implicit midpoint method to examine the extent to which this alternative integrator offers an advantage over the generalized leapfrog method in the non-Gaussian regime.

4.3 Related Work

Most relevant to our discussion is Pourzanjani and Petzold [2019]. In this work, the authors examine the relationship between the (non-generalized) leapfrog integrator and the implicit midpoint integrator; the authors make the argument that the implicit midpoint integrator is more stable in the presence of posteriors whose dimensions exhibit large differences in their variability (“multi-scale”). The presence of multi-scale posterior dimensions necessitates a small step-size for the leapfrog integrator, which is found to be unnecessary for the implicit midpoint algorithm. As the authors note, however, “RMHMC uses local Hessian evaluations of the potential energy surface to adaptively change this step-size based on the local curvature;” therefore, their experiments instead focus on the circumstance where a constant mass matrix is utilized, corresponding to Euclidean HMC with no local adaptation of the step-size. Indeed, as observed in Martens [2020], the Fisher information captures the second-order geometry of the posterior and actually exhibits properties that make it preferable to the Hessian of the posterior in optimization. Further details on
how the Fisher information metric captures second-order geometry of the distribution are detailed in section 2.6. Therefore, the present work differs from Pourzanjani and Petzold [2019] in its focus on Riemannian geometry wherein the metric compensates (at least locally) for multi-scale dimensions; moreover, our empirical analysis of reversibility and volume preservation is, to the best of our knowledge, novel. Before proceeding to the experimental results, we note that stability alone cannot account for the high acceptance rate enjoyed by the implicit midpoint integrator: even in the regime wherein the generalized leapfrog integrator is stable, it is not able to perfectly conserve the Hamiltonian energy as the implicit midpoint integrator does. This phenomenon is visualized in fig. 4.1.

### 4.4 Experimental Results

We turn now to evaluating the implicit midpoint integrator in several Bayesian inference tasks. We consider inference in a banana-shaped posterior, sampling from Neal’s funnel distribution, a stochastic volatility model, and Bayesian inference in the Fitzhugh-Nagumo differential equation model. We have additional experimental results in our appendices. In section 4.A.2, we seek to verify theorem 4.2.3 in the presence of a truly quadratic Hamiltonian. In section 4.A.3, we examine Bayesian inference in a logistic regression posterior. To define a stopping condition for the fixed point iterations used by the implicit midpoint and generalized leapfrog methods, we demand that the absolute change in each coordinate be less than a threshold (the use of relative change in each coordinate could produce distinct results); we let $\delta \in \{1 \times 10^{-9}, 1 \times 10^{-6}, 1 \times 10^{-3}\}$ when considering errors in reversibility and volume preservation which are measured according to the procedure detailed in section 4.A.4. When reporting performance metrics such as effective sample size, we report results corresponding to a threshold of $\delta = 1 \times 10^{-6}$. We implemented all methods in 64-bit precision using NumPy and SciPy [Harris et al. 2020, Virtanen et al. 2020]. We compute effective sample sizes (ESS) using Kumar et al. [2019]. Code for our ex-
experiments can be found at https://github.com/JamesBrofos/Evaluating-the-Implicit-Midpoint-Integrator.

4.4.1 Summary of Integrators

We consider two variants of the generalized leapfrog method and two variants of the implicit midpoint integrator, which we summarily describe as follows.

**G.L.F.(a)** An implementation of the generalized leapfrog integrator as presented in algorithm 2.

**G.L.F.(b)** An implementation of the generalized leapfrog integrator that caches repeated calculations and which is specific to Hamiltonians in the form of eq. (4.1). See algorithm 4 in section 4.A.5. G.L.F.(b) is mathematically identical to G.L.F.(a), but this implementation avoids some redundant computation. Differences between the outputs of G.L.F.(a) and G.L.F.(b) are due to random seeds and machine error in computation.

**I.M.(a)** An implementation of the implicit midpoint integrator as presented in algorithm 3.

**I.M.(b)** An implementation of the implicit midpoint integrator that implicitly computes the midpoint followed by an explicit Euler step, as advocated by Leimkuhler and Reich [2005]. See algorithm 5 in section 4.A.5.

**I.M.(c)** An implementation of the implicit midpoint integrator that is derived from the discrete Lagrangian approach detailed in Hairer et al. [2006] for the following Lagrangian

\[ L(q, \dot{q}) = \mathcal{L}(q) - \frac{\log \det(G(q))}{2} + \frac{\dot{q}^\top G(q) \dot{q}}{2}. \]  

(4.11)

See also Mackay [1992] and algorithm 6 in section 4.A.5 This integrator is derived from the principle of the discrete Lagrangian in example 13.
The banana-shaped distribution was proposed in a discussion to Girolami and Calderhead [2011] as a representative example of the ridge-like posterior structure that can manifest in non-identifiable models. The banana-shaped distribution is defined by the following generative model.

\[
y_i|\theta_1, \theta_2 \sim \text{Normal}(\theta_1 + \theta_2^2, \sigma_y^2) \quad \text{for } i = 1, \ldots, n
\]  
\[
\theta_i \sim \text{Normal}(0, \sigma_\theta^2) \quad \text{for } i \in \{1, 2\}.
\]  

Table 4.1: Banana-shaped distribution metrics. Comparison of the implicit midpoint and generalized leapfrog integrators on sampling from the banana-shaped distribution. To assess performance of the sampler, we measure the effective sample size (ESS) and present per-second timing comparisons for the mean and minimum ESS. The hypothesis that the implicit midpoint integrator should exhibit better energy conservation is captured in the acceptance probability of the Markov chain. Results are averaged over ten trials.

In all of our implementations, we use fixed point iterations in order to find solutions to implicitly-defined relations. This is the approach advocated by Hairer et al. [2006]. Additional details are presented in section 4.A.5.

4.4.2 Banana-Shaped Distribution

The banana-shaped distribution was proposed in a discussion to Girolami and Calderhead [2011] as a representative example of the ridge-like posterior structure that can manifest in non-identifiable models. The banana-shaped distribution is defined by the following generative model.

\[
y_i|\theta_1, \theta_2 \sim \text{Normal}(\theta_1 + \theta_2^2, \sigma_y^2) \quad \text{for } i = 1, \ldots, n
\]  
\[
\theta_i \sim \text{Normal}(0, \sigma_\theta^2) \quad \text{for } i \in \{1, 2\}.
\]  

<table>
<thead>
<tr>
<th>Step Size</th>
<th>Num. Steps</th>
<th>Method</th>
<th>Acc. Prob.</th>
<th>Time (Sec.)</th>
<th>Mean ESS</th>
<th>Min. ESS</th>
<th>Mean ESS / Sec.</th>
<th>Min. ESS / Sec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>5</td>
<td>G.L.F.(a)</td>
<td>0.62 ± 0.01</td>
<td>400.33 ± 6.83</td>
<td>486.32 ± 17.89</td>
<td>286.28 ± 16.09</td>
<td>1.21 ± 0.04</td>
<td>0.71 ± 0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G.L.F.(b)</td>
<td>0.61 ± 0.01</td>
<td>145.99 ± 2.20</td>
<td>491.07 ± 22.03</td>
<td>301.90 ± 16.21</td>
<td>3.36 ± 0.13</td>
<td>2.07 ± 0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(a)</td>
<td>0.98 ± 0.00</td>
<td>102.37 ± 1.13</td>
<td>884.09 ± 27.39</td>
<td>620.26 ± 30.08</td>
<td>8.65 ± 0.31</td>
<td>6.07 ± 0.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(b)</td>
<td>0.98 ± 0.00</td>
<td>95.28 ± 2.28</td>
<td>857.61 ± 24.99</td>
<td>619.91 ± 27.09</td>
<td>9.03 ± 0.28</td>
<td>6.53 ± 0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(c)</td>
<td>0.98 ± 0.00</td>
<td>83.28 ± 0.95</td>
<td>884.03 ± 29.03</td>
<td>620.28 ± 38.29</td>
<td>10.63 ± 0.37</td>
<td>7.58 ± 0.48</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>G.L.F.(a)</td>
<td>0.50 ± 0.01</td>
<td>615.14 ± 6.51</td>
<td>1038.47 ± 25.68</td>
<td>778.80 ± 30.92</td>
<td>1.69 ± 0.05</td>
<td>1.27 ± 0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G.L.F.(b)</td>
<td>0.49 ± 0.01</td>
<td>231.55 ± 2.48</td>
<td>1027.78 ± 25.19</td>
<td>782.85 ± 14.87</td>
<td>4.44 ± 0.12</td>
<td>3.38 ± 0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(a)</td>
<td>0.98 ± 0.00</td>
<td>194.82 ± 2.39</td>
<td>3018.50 ± 70.23</td>
<td>2518.65 ± 83.19</td>
<td>15.51 ± 0.39</td>
<td>12.94 ± 0.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(b)</td>
<td>0.98 ± 0.00</td>
<td>172.12 ± 1.45</td>
<td>3025.14 ± 55.19</td>
<td>2540.88 ± 99.95</td>
<td>17.58 ± 0.34</td>
<td>14.76 ± 0.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(c)</td>
<td>0.98 ± 0.00</td>
<td>159.35 ± 2.12</td>
<td>3080.80 ± 63.15</td>
<td>2641.56 ± 102.88</td>
<td>19.34 ± 0.33</td>
<td>16.55 ± 0.51</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>G.L.F.(a)</td>
<td>0.13 ± 0.00</td>
<td>2133.63 ± 64.84</td>
<td>1928.82 ± 24.27</td>
<td>79.12 ± 14.55</td>
<td>0.09 ± 0.01</td>
<td>0.04 ± 0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G.L.F.(b)</td>
<td>0.14 ± 0.00</td>
<td>786.40 ± 18.50</td>
<td>247.58 ± 28.90</td>
<td>119.23 ± 18.68</td>
<td>0.31 ± 0.03</td>
<td>0.15 ± 0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(a)</td>
<td>0.95 ± 0.00</td>
<td>938.79 ± 15.17</td>
<td>4173.70 ± 199.89</td>
<td>3207.59 ± 113.94</td>
<td>4.47 ± 0.25</td>
<td>3.43 ± 0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(b)</td>
<td>0.95 ± 0.00</td>
<td>834.63 ± 13.42</td>
<td>3928.27 ± 159.13</td>
<td>3158.40 ± 93.76</td>
<td>4.73 ± 0.24</td>
<td>3.80 ± 0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(c)</td>
<td>0.95 ± 0.00</td>
<td>758.45 ± 8.59</td>
<td>4359.09 ± 507.74</td>
<td>3093.59 ± 140.71</td>
<td>5.77 ± 0.68</td>
<td>4.09 ± 0.20</td>
</tr>
</tbody>
</table>

192
Figure 4.2: Banana-shaped distribution detailed balance errors. Comparison between the error in symmetry and error in volume preservation properties of the implicit midpoint integrator and the generalized leapfrog integrator on the banana-shaped distribution. We observe that the implicit midpoint integrator tends to produce transitions whose median reversibility and volume preservation can be an order of magnitude, or more, better than the generalized leapfrog integrator. Each point is the median of one-hundred measurements of symmetry and reversibility shown for each of the ten trials. For $\epsilon = 1/10$ and fifty integration steps, the generalized leapfrog integrator exhibits severely divergent behavior. The implicit midpoint is represented by the symbol ($\cdot$) and the generalized leapfrog by the symbol ($+$).

For the banana-shaped distribution, the Riemannian metric is

$$G(\theta_1, \theta_2) = \begin{pmatrix}
n \sigma_y^2 + \frac{1}{\sigma_\theta^2} & \frac{2n\theta_2}{\sigma_\theta^2} \\
\frac{2n\theta_2}{\sigma_\theta^2} & \frac{4n\theta_2^2}{\sigma_\theta^4} + \frac{1}{\sigma_\theta^2}
\end{pmatrix}. \quad (4.14)$$

In our experiments, we take $n = 100$. We generate observations $\{y_1, \ldots, y_{100}\}$ from the banana-shaped distribution by setting $\theta_1 = 1/2$, $\theta_2 = 1/\sqrt{2}$, and $\sigma_y = \sigma_\theta = 2$. We then attempt to sample the posterior distribution of $(\theta_1, \theta_2)$ using RMHMC when integration is performed using the implicit midpoint algorithm or the generalized leapfrog method. We consider two step-sizes $\{0.01, 0.1\}$ and a number of integration steps in $\{5, 10, 50\}$. We attempt to draw 10,000 samples from the posterior. Each of these configurations is replicated ten times.

Results are shown in table 4.1 demonstrating that the I.M.(a) and (b) integrators are able to maintain high energy conservation at step-sizes for which the G.L.F.(a) and (b)
variants cannot. As a consequence, Markov chains using I.M.(a) or I.M.(b) are able to achieve very high effective sample sizes (ESS); moreover, because the cost of evaluating the gradients of the banana-shaped posterior is not too large, these Markov chains also exhibit superior performance on the timing comparisons. We find that I.M.(a) and I.M.(b) perform similarly. In addition to energy conservation, an essential component of HMC are volume preservation and reversibility (recall properties (i) and (ii) from section 4.1).

Using the samples drawn by the Markov chains with either integrator, we may compute numerical estimates of the degree to which these properties are satisfied. We give a detailed description of the volume preservation and reversibility metrics in section 4.A.4. We use one-hundred randomly selected samples generated from the Markov chains in order to compute these statistics. Results showing the median reversibility versus the median difference from unit Jacobian are shown in fig. 4.2. These results show that the median symmetry and volume preservation of the implicit midpoint integrator is approximately an order of magnitude more faithfully preserved than is the case for the generalized leapfrog method.

Table 4.2: Neal’s funnel distribution metrics. Comparison of the implicit midpoint and the naive generalized leapfrog integrators on sampling from Neal’s funnel distribution. We see that the implicit midpoint integrator is able to take large steps and produce an effective sample size that outperforms the generalized leapfrog integrator even in the time-normalized performance.
Figure 4.3: Neal’s funnel distribution detailed balance. Comparison of the degree to which the implicit midpoint and generalized leapfrog integrators violate reversibility and volume preservation on Neal’s funnel distribution. We observe that the implicit midpoint integrator exhibits better symmetry and volume preservation. The implicit midpoint is represented by the symbol (·) and the generalized leapfrog by the symbol (+).

4.4.3 Hierarchical Neal’s Funnel Distribution

As an example of a hierarchical Bayesian posterior, we consider Neal’s funnel distribution defined by,

\[ x_i \sim \text{Normal}(0, \exp(-v)) \quad \text{for } i = 1, \ldots, 10 \]

\[ v \sim \text{Normal}(0, 9). \]

Due to the hierarchical structure of the distribution, the Hessian of the distribution is not convex and therefore cannot be used to construct a Riemannian metric on its own. Instead, we follow the approach proposed in Betancourt [2012] and adopt the SoftAbs transformation of the Hessian in order to construct a positive definite Riemannian metric. This allows us to sample all variables of the hierarchical distribution jointly. For RMHMC, we consider an integration step-size in \{0.1, 0.2, 0.5\} and we attempt to draw 10,000 samples of \((x_1, \ldots, x_{10}, v)\) from Neal’s funnel distribution.

Results are presented in table 4.2. For the largest step-size, the I.M.(a) and (b) integrators are able to maintain high acceptance probabilities. Markov chains using the I.M.(a)
Table 4.3: Stochastic volatility posterior metrics. Comparison of the implicit midpoint generalized leapfrog integrators on the stochastic volatility model. We see that the implicit midpoint integrator is competitive on the effective sample size and timing comparisons. We also evaluate the median reversibility and volume preservation of the implicit midpoint and the generalized leapfrog integrators on the stochastic volatility model with $\delta = 1 \times 10^{-6}$. Here we see that the implicit midpoint integrator enjoys better conservation of volume and reversibility.

<table>
<thead>
<tr>
<th>Method</th>
<th>Acc. Prob.</th>
<th>Time (Sec.)</th>
<th>Mean ESS</th>
<th>Min. ESS</th>
<th>Mean ESS (Sec.)</th>
<th>Min. ESS (Sec.)</th>
<th>Volume Preservation</th>
<th>Symmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Median 90th.-Per.</td>
<td>Median 90th.-Per.</td>
</tr>
<tr>
<td>G.L.F.(a)</td>
<td>0.78 ± 0.0</td>
<td>2616.70 ± 16.5</td>
<td>294.51 ± 10.7</td>
<td>121.79 ± 7.7</td>
<td>0.11 ± 0.0</td>
<td>0.05 ± 0.0</td>
<td>7.2e−07</td>
<td>3.1e−06</td>
</tr>
<tr>
<td>G.L.F.(b)</td>
<td>0.78 ± 0.0</td>
<td>2339.78 ± 9.5</td>
<td>297.82 ± 9.2</td>
<td>132.79 ± 6.6</td>
<td>0.13 ± 0.0</td>
<td>0.06 ± 0.0</td>
<td>6.7e−07</td>
<td>2.8e−06</td>
</tr>
<tr>
<td>I.M.(a)</td>
<td>0.80 ± 0.0</td>
<td>2761.61 ± 8.3</td>
<td>309.10 ± 10.5</td>
<td>133.32 ± 7.8</td>
<td>0.11 ± 0.0</td>
<td>0.05 ± 0.0</td>
<td>8.7e−08</td>
<td>2.2e−07</td>
</tr>
<tr>
<td>I.M.(b)</td>
<td>0.80 ± 0.0</td>
<td>2710.47 ± 8.4</td>
<td>296.72 ± 9.2</td>
<td>127.35 ± 7.3</td>
<td>0.11 ± 0.0</td>
<td>0.05 ± 0.0</td>
<td>1.7e−07</td>
<td>4.5e−07</td>
</tr>
<tr>
<td>I.M.(c)</td>
<td>0.80 ± 0.0</td>
<td>2867.27 ± 17.2</td>
<td>312.34 ± 10.7</td>
<td>129.35 ± 7.9</td>
<td>0.11 ± 0.0</td>
<td>0.05 ± 0.0</td>
<td>1.6e−07</td>
<td>4.1e−07</td>
</tr>
</tbody>
</table>

4.4.4 Stochastic Volatility Model

While Neal’s funnel is a hierarchical distribution, it is not sampled in a hierarchical manner, instead sampling all variables jointly using the SoftAbs Riemannian metric [Betancourt 2012]. Here, we consider a stochastic volatility model whose posterior includes the stochastic volatilities as well as latent hyperparameters of the model; we will sample these variables using an alternating Gibbs procedure. Following [Girolami and Calderhead 2011], the stochastic volatility model is defined, for $t = 1, \ldots, T$, by

\[ y_t | \beta, x_t \sim \text{Normal}(0, \beta^2 e^{x_t}) \]  

\[ x_t | x_{t-1}, \phi, \sigma^2 \sim \text{Normal}(\phi x_{t-1}, \sigma^2) \]  

196
where \( x_1 \sim \text{Normal}(0, \sigma^2/(1 - \phi^2)) \), \( \pi(\beta) \propto 1/\beta \), \( \sigma^2 \sim \text{Inv-}\chi^2(10, 0.05) \), and \((\phi + 1)/2 \sim \text{Beta}(20, 1.5)\). The sampler proceeds by alternating between sampling the conditional posteriors of \((x_1, \ldots, x_T|(y_1, \ldots, y_T), \phi, \beta, \sigma^2\) and \((\phi, \beta, \sigma^2|(x_1, \ldots, x_T), (y_1, \ldots, y_T)\).

The Riemannian metric of this first posterior is constant with respect to \((x_1, \ldots, x_T)\); therefore, sampling is carried out using the standard leapfrog integrator. The second distribution has a position-dependent Riemannian metric, necessitating the use of implicitly-defined integrators; here, we compare the implicit midpoint and generalized leapfrog integrators. For details of the Riemannian structures of the conditional posteriors, see [Girolami and Calderhead 2011].

In our experiments, we set \( T = 1,000 \) and use fifty integration steps with a step-size of 0.1 to sample \((x_1, \ldots, x_T|(y_1, \ldots, y_T), \phi, \beta, \sigma^2\) and six integration steps with a step-size of 0.5 to sample \((\phi, \beta, \sigma^2|(x_1, \ldots, x_T), (y_1, \ldots, y_T)\). We seek to sample 20,000 times from the posterior and use a burn-in period of 10,000 iterations. We repeat this experiment one-hundred times for each integrator. Effective sample size metrics and measures of the volume preservation and symmetry are are presented in table 4.3 for the parameters \( \phi, \beta, \) and \( \sigma^2 \). We find that the I.M.(a) and (b) integrators are comparable to the G.L.F.(a) and (b) methods in terms of their time-normalized performance. However, volume preservation and symmetry are better for the I.M.(a) and (b) integrators.

### 4.4.5 Fitzhugh-Nagumo ODE Model

The Fitzhugh-Nagumo ordinary differential equation is a model of neural spiking activity. It is described by two time-varying measurements whose dynamics obey,

\[
\dot{v} = \theta_3 \left( v - \frac{v^3}{3} + r \right) \tag{4.19}
\]

\[
\dot{r} = -\left( \frac{v - \theta_1 + \theta_2 r}{\theta_3} \right) \tag{4.20}
\]
If we equip the parameters dynamics of eqs. (4.19) and (4.20) and the assumed noise distribution in order to sample $\sigma$ have been corrupted by i.i.d. Gaussian noise with a known standard deviation of $\sigma = 1/2$. Consider the setting wherein one has 200 observations of the Fitzhugh-Nagumo dynamics with the generalized leapfrog integrator’s computational advantages.

Figure 4.4: *Fitzhugh-Nagumo detailed balance errors.* Comparison of the degree to which the implicit midpoint and generalized leapfrog integrators violate reversibility and volume preservation on the Fitzhugh-Nagumo posterior. We observe that for the chosen convergence tolerances, the implicit midpoint integrator tends to exhibit better reversibility and volume preservation except for the smallest tolerance where the generalized leapfrog has slightly better volume preservation but worse reversibility. The implicit midpoint is represented by the symbol (·) and the generalized leapfrog by the symbol (+).

<table>
<thead>
<tr>
<th>Step Size</th>
<th>Num. Steps Method</th>
<th>Acc. Prob.</th>
<th>Time (Sec.)</th>
<th>Mean ESS</th>
<th>Min. ESS</th>
<th>Mean ESS / Sec.</th>
<th>Min. ESS / Sec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1 G.L.F.(a)</td>
<td>0.74 ± 0.01</td>
<td>4441.16 ± 117.73</td>
<td>276.39 ± 13.48</td>
<td>235.39 ± 10.55</td>
<td>0.06 ± 0.00</td>
<td>0.05 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>G.L.F.(b)</td>
<td>0.73 ± 0.01</td>
<td>1343.63 ± 76.12</td>
<td>272.72 ± 7.96</td>
<td>235.52 ± 8.99</td>
<td>0.21 ± 0.01</td>
<td>0.18 ± 0.01</td>
</tr>
<tr>
<td></td>
<td>I.M. (a)</td>
<td>0.95 ± 0.00</td>
<td>4523.49 ± 160.00</td>
<td>227.01 ± 13.47</td>
<td>199.19 ± 13.79</td>
<td>0.05 ± 0.00</td>
<td>0.04 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>I.M. (b)</td>
<td>0.95 ± 0.00</td>
<td>4486.29 ± 147.33</td>
<td>248.11 ± 8.74</td>
<td>225.67 ± 9.11</td>
<td>0.06 ± 0.00</td>
<td>0.05 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>I.M. (c)</td>
<td>0.94 ± 0.00</td>
<td>4268.94 ± 103.11</td>
<td>220.45 ± 5.25</td>
<td>193.48 ± 5.89</td>
<td>0.05 ± 0.00</td>
<td>0.05 ± 0.00</td>
</tr>
<tr>
<td>2</td>
<td>2 G.L.F.(a)</td>
<td>0.74 ± 0.01</td>
<td>8100.90 ± 423.17</td>
<td>1089.36 ± 55.17</td>
<td>967.30 ± 75.02</td>
<td>0.14 ± 0.01</td>
<td>0.11 ± 0.01</td>
</tr>
<tr>
<td></td>
<td>G.L.F.(b)</td>
<td>0.75 ± 0.01</td>
<td>1912.73 ± 48.99</td>
<td>1109.78 ± 71.50</td>
<td>917.52 ± 98.66</td>
<td>0.58 ± 0.04</td>
<td>0.48 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>I.M. (a)</td>
<td>0.94 ± 0.00</td>
<td>8704.37 ± 250.22</td>
<td>1387.96 ± 40.53</td>
<td>1212.91 ± 42.68</td>
<td>0.16 ± 0.01</td>
<td>0.14 ± 0.01</td>
</tr>
<tr>
<td></td>
<td>I.M. (b)</td>
<td>0.94 ± 0.00</td>
<td>8421.40 ± 139.60</td>
<td>1366.07 ± 56.59</td>
<td>1213.01 ± 55.18</td>
<td>0.16 ± 0.00</td>
<td>0.14 ± 0.03</td>
</tr>
<tr>
<td></td>
<td>I.M. (c)</td>
<td>0.94 ± 0.00</td>
<td>8084.16 ± 274.83</td>
<td>1343.91 ± 69.42</td>
<td>1171.60 ± 71.75</td>
<td>0.17 ± 0.01</td>
<td>0.15 ± 0.01</td>
</tr>
<tr>
<td>5</td>
<td>5 G.L.F.(a)</td>
<td>0.70 ± 0.01</td>
<td>20096.64 ± 384.16</td>
<td>188.59 ± 27.78</td>
<td>90.00 ± 19.15</td>
<td>0.01 ± 0.00</td>
<td>0.00 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>G.L.F.(b)</td>
<td>0.71 ± 0.01</td>
<td>5138.31 ± 300.83</td>
<td>150.03 ± 23.06</td>
<td>102.32 ± 22.10</td>
<td>0.03 ± 0.00</td>
<td>0.02 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>I.M. (a)</td>
<td>0.94 ± 0.00</td>
<td>22069.48 ± 650.03</td>
<td>954.94 ± 93.17</td>
<td>818.16 ± 91.14</td>
<td>0.04 ± 0.01</td>
<td>0.04 ± 0.01</td>
</tr>
<tr>
<td></td>
<td>I.M. (b)</td>
<td>0.94 ± 0.00</td>
<td>20575.78 ± 566.21</td>
<td>1095.88 ± 130.14</td>
<td>821.77 ± 80.15</td>
<td>0.05 ± 0.01</td>
<td>0.04 ± 0.00</td>
</tr>
<tr>
<td></td>
<td>I.M. (c)</td>
<td>0.94 ± 0.00</td>
<td>22971.96 ± 1548.11</td>
<td>854.30 ± 61.71</td>
<td>712.87 ± 56.85</td>
<td>0.04 ± 0.00</td>
<td>0.03 ± 0.00</td>
</tr>
</tbody>
</table>

Table 4.4: *Fitzhugh-Nagumo posterior metrics.* Comparison of the implicit midpoint and generalized leapfrog integrators on sampling from the posterior of the Fitzhugh-Nagumo model. In this example, the higher acceptance probability of the implicit midpoint integrator did not produce a performance increase for a single integration step. The implicit midpoint integrator becomes super-efficient in the two-step regime, but cannot compete with the generalized leapfrog integrator’s computational advantages.

Consider the setting wherein one has 200 observations of the Fitzhugh-Nagumo dynamics at equally-spaced times between zero and ten. Assume moreover that these observations have been corrupted by i.i.d. Gaussian noise with a known standard deviation of $\sigma = 1/2$. If we equip the parameters $\theta_1$, $\theta_2$, and $\theta_3$ with standard normal priors, we may use the dynamics of eqs. (4.19) and (4.20) and the assumed noise distribution in order to sample...
the posterior of \((\theta_1, \theta_2, \theta_3)\). Let \((v_n, r_n)\) be the solution of the Fitzhugh-Nagumo ODE at the \(n^{th}\) time period. For the Fitzhugh-Nagumo differential equation model, the \((i, j)\)-entry of the metric is,

\[
G_{ij}(\theta_1, \theta_2, \theta_3) = \frac{1}{\sigma^2} \left( \sum_{n=1}^{200} \frac{\partial v_n}{\partial \theta_i} \frac{\partial v_n}{\partial \theta_j} + \frac{\partial r_n}{\partial \theta_i} \frac{\partial r_n}{\partial \theta_j} \right) + \delta_{ij}.
\]

(4.21)

In generating data from the Fitzhugh-Nagumo model, we set \(\theta_1 = \theta_2 = 0.2\) and \(\theta_3 = 3\). The dynamics are integrated using SciPy’s `odeint` function and gradients are approximated by forward sensitivity analysis as in [Girolami and Calderhead, 2011]. We consider an integration step-size of \(\epsilon = 1\) and a number of integration steps in \(\{1, 2, 5\}\); each configuration is replicated ten times. We sample 1,000 times from the posterior.

We expect the Fitzhugh-Nagumo ODE model to favor the generalized leapfrog integrator because of the complexity of evaluating the log-posterior, the gradient of the log-posterior, the Riemannian metric, and the gradient of the Riemannian metric, each of which involves solving a system of differential equations. Therefore, the caching behavior associated to the G.L.F.(b) integrator gives it an advantage here. Table 4.4 shows the results of inferences in the Fitzhugh-Nagumo posterior. We observe that for a single-step, the I.M.(a) and (b) integrators appears to perform somewhat worse than the G.L.F.(a) and (b) variants, even on the measures of ESS that ignore timing; this occurs despite the larger acceptance rate enjoyed by the implicit midpoint integrator. For two integration steps, the inferences produced by I.M.(a) and (b) become super-efficient; however, G.L.F.(a) and (b) are also efficient and the computational advantage of the (b) variant cause it to have superior performance in the timing metrics. For the largest number of steps, the performance of G.L.F.(a) and (b) deteriorates so that the I.M.(a) and (b) integrators outperform them even on the timing comparison.

We also evaluate the degree to which the numerical integrator possesses the properties of symmetry and volume preservation. The results are shown in fig. 4.4. We see that the
implicit midpoint integrator offers a clear advantage in numerical symmetry, and performs better on volume preservation as well.

4.5 Conclusion

This work has considered the implicit midpoint integrator as a substitute for the generalized leapfrog integrator for use in RMHMC. Inspired by the theory of the conserved quantities of numerical integrators, we hypothesized that the implicit midpoint integrator would have better energy conservation in posterior distributions that are approximately Gaussian. Hamiltonian Monte Carlo requires that its integrators are volume preserving and reversible; we give numerical assessments of the extent to which these properties are present in implementations of these integrators, which differ from their theoretical representation when a convergence tolerance is used to halt a fixed point iteration. We find that the implicit midpoint integrator has superior energy conservation, conservation of volume, and reversibility across several Bayesian inference tasks. In three of the four example applications, the implicit midpoint integrator met or exceeded the time-normalized performance of the generalized leapfrog integrator. This, combined with its better volume preservation and reversibility, leads us to conclude that it is a method worth consideration when implementing RMHMC.
Appendix 4.A

Appendices to Chapter 4

4.A.1 Momentum Negation Symmetry of Implicit Midpoint

Lemma 4.A.1.1. Given a Hamiltonian $H$ in the form of eq. (4.1), step-size $\epsilon$, and initial position $(q, p)$, compute $(q', p')$ according to algorithm 3 with $\delta = 0$. If one then computes $(q'', p'')$ from initial position $(q', -p')$ using algorithm 3 a second time (with the same Hamiltonian, step-size, and $\delta = 0$), then $q'' = q$ and $-p'' = p$.

Lemma 4.A.1.1 establishes that the implicit midpoint integrator is suitable for HMC in that it satisfies properties (i) and (ii).

Proof. Consider the initial condition $(q, p)$ and a fixed step-size of $\epsilon$. For the Riemannian manifold Hamiltonian Monte Carlo, the implicit midpoint integrator computes the
following updates:

\begin{align}
q'_i &= q_i + \epsilon \left( \sum_{j=1}^{m} G^{-1}_{ij} \left( \frac{q' + q}{2} \right) \left( \frac{p'_j + p_j}{2} \right) \right) \\
p'_i &= p_i + \epsilon \left( \frac{1}{2} \text{trace} \left( G^{-1} \left( \frac{q' + q}{2} \right) \frac{\partial}{\partial q_i} G \left( \frac{q' + q}{2} \right) \right) + \frac{1}{2} \left( \frac{p' + p}{2} \right)^\top G^{-1} \left( \frac{q' + q}{2} \right) \frac{\partial}{\partial q_i} G \left( \frac{q' + q}{2} \right) G^{-1} \left( \frac{q' + q}{2} \right) \left( \frac{p' + p}{2} \right) \right) \right.
\end{align}

What we want to show is that if we compute \((q', p')\), negate the momentum \((q', p') \rightarrow (q', -p')\), and apply the implicit midpoint integrator a second time, then we arrive at \((q, -p)\). Thus, we need to establish that \((q, -p)\) is a fixed point of the relations,

\begin{align}
q''_i &= q'_i + \epsilon \left( \sum_{j=1}^{m} G^{-1}_{ij} \left( \frac{q'' + q'}{2} \right) \left( \frac{p''_j + (-p'_j)}{2} \right) \right) \\
p''_i &= -p'_i + \epsilon \left( \frac{1}{2} \text{trace} \left( G^{-1} \left( \frac{q'' + q'}{2} \right) \frac{\partial}{\partial q_i} G \left( \frac{q'' + q'}{2} \right) \right) + \frac{1}{2} \left( \frac{p'' + (-p')}{2} \right)^\top G^{-1} \left( \frac{q'' + q'}{2} \right) \frac{\partial}{\partial q_i} G \left( \frac{q'' + q'}{2} \right) G^{-1} \left( \frac{q'' + q'}{2} \right) \left( \frac{p'' + (-p')}{2} \right) \right) .
\end{align}

Plugging in we obtain,

\begin{align}
q' + \epsilon \left( \sum_{j=1}^{m} G^{-1}_{ij} \left( \frac{q + q'}{2} \right) \left( \frac{(-p_j) + (-p'_j)}{2} \right) \right) &= q' - \epsilon \left( \sum_{j=1}^{m} G^{-1}_{ij} \left( \frac{q + q'}{2} \right) \left( \frac{p_j + p'_j}{2} \right) \right) \\
&= q_i
\end{align}
by rearranging eq. (4.A.1). For notational simplicity let us define

\[
U(q) \overset{\text{def.}}{=} -\frac{\partial}{\partial q_i} \mathcal{L}(q) - \frac{1}{2} \text{trace} \left( G^{-1}(q) \frac{\partial}{\partial q_i} G(q) \right) \tag{4.A.7}
\]

\[
R(q) \overset{\text{def.}}{=} \frac{1}{2} G^{-1}(q) \frac{\partial}{\partial q_i} G(q) G^{-1}(q) \tag{4.A.8}
\]

so that

\[
p''_i = -p'_i + \epsilon \left( U \left( \frac{q'' + q}{2} \right) + \left( \frac{p'' + (-p')}{2} \right) \right)^\top R \left( \frac{q'' + q}{2} \right) \left( \frac{p'' + (-p')}{2} \right) \tag{4.A.9}
\]

Plugging in, we obtain,

\[
- p'_i + \epsilon \left( U \left( \frac{q + q}{2} \right) + \left( \frac{-p + (-p')}{2} \right) \right)^\top R \left( \frac{q + q}{2} \right) \left( \frac{-p + (-p')}{2} \right) \tag{4.A.10}
\]

\[
= - p'_i + \epsilon \left( U \left( \frac{q + q}{2} \right) + \left( \frac{p + p'}{2} \right)^\top R \left( \frac{q + q}{2} \right) \left( \frac{p + p'}{2} \right) \right) \tag{4.A.11}
\]

\[
= - p_i \tag{4.A.12}
\]

which follows from negating eq. (4.A.2) and rearranging.

This shows that the implicit midpoint integrator can be easily combined with a momentum flip in order to produce an involution, thereby producing a Markov chain transition kernel satisfying detailed balance according to involutive Monte Carlo as detailed in proposition 3.1.6.
4.A.2 Experimentation with a Quadratic Hamiltonian

We consider using HMC to draw samples from a Gaussian distribution in two dimensions. In particular, we aim to sample from the joint distribution of position and momentum defined by,

\[
q \sim \text{Normal} \begin{pmatrix} 1/2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 & 1/2 \\ 1/2 & 2 \end{pmatrix} \tag{4.A.13}
\]

\[
p \sim \text{Normal} \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 1/2 \\ 0 & 2 \end{pmatrix}^{-1} \tag{4.A.14}
\]

These distributions correspond to the quadratic Hamiltonian

\[
H(q, p) = \frac{1}{2}q^T \Sigma^{-1} q + \frac{1}{2}p^T \Sigma p.
\]

This Hamiltonian can be interpreted in the Riemannian manifold setting as sampling from the posterior \( q \sim \text{Normal}(\mu, \Sigma) \) and the constant metric \( G(q) = \Sigma^{-1} \). The corresponding Hamiltonian is quadratic and therefore theorem 4.2.3 applies. We expect perfect conservation of the Hamiltonian regardless of step-size. To evaluate the conservation of the

![Histograms showing energy conservation for different step-sizes](image)

Figure 4.A.1: Comparison between the energy conservation of the implicit midpoint integrator and the generalized leapfrog integrator on a quadratic Hamiltonian. We observe that for every step-size, the energy conservation of the implicit midpoint method is in the neighborhood of \( 1 \times 10^{-10} \) whereas the generalized leapfrog has energy conservation that degrades with larger steps.
Hamiltonian energy, we consider drawing \((q, p)\) from their joint distribution and integrating Hamilton’s equations of motion for ten integration steps. We consider integration step-sizes in \(\{0.01, 0.1, 1.0\}\). We then compare the initial Hamiltonian energy to the Hamiltonian energy at the terminal point of the integrator. We repeat this procedure 10,000 times and show the results in fig. 4.A.1 where the absolute difference in Hamiltonian energy is shown as a histogram. This experiment clearly shows that the implicit midpoint integrator has excellent conservation of the quadratic Hamiltonian energy and is orders of magnitude better than the generalized leapfrog integrator. Note that for a separable Hamiltonian, as is the case here, the steps of the generalized leapfrog integrator reduce to the standard leapfrog method.

### 4.A.3 Bayesian Logistic Regression

Binary classification is a ubiquitous task in the data sciences and logistic regression is the most popular algorithm for obtaining probabilistic estimates of class membership. Bayesian logistic regression simply equips each of the linear coefficients in the logistic regression model with a prior distribution. We consider Bayesian logistic regression as defined by the following generative model:

\[
y_i | x_i, \beta \sim \text{Bernoulli}(\sigma(x_i^T \beta)) \quad \text{for } i = 1, \ldots, n \tag{4.A.15}
\]

\[
\beta_i \sim \text{Normal}(0, 1) \quad \text{for } i = 1, \ldots, k, \tag{4.A.16}
\]

where \(x_i \in \mathbb{R}^k\) is vector of explanatory variables and \(\sigma : \mathbb{R} \rightarrow (0, 1)\) is the sigmoid function. For the logistic regression model, let \(x \in \mathbb{R}^{n \times m}\) represent the matrix of features. The Riemannian metric formed by the sum of the Fisher information and the negative Hessian of the log-prior is \(G(\beta) = x^T \Lambda x + \text{Id}\) where \(\Lambda\) is a diagonal matrix whose \(i^{th}\) diagonal entry is \(\sigma(x_i^T \beta)(1 - \sigma(x_i^T \beta))\).
We consider sampling from the posterior distribution of the linear coefficients for a breast cancer, heart disease, and diabetes dataset. We consider integration step-sizes in \{1/10, 1\} and a number of integration steps in \{5, 10, 50\}; each configuration of step-size and number of steps is replicated ten times and we attempt to draw 10,000 samples from the posterior. Results are presented in tables 4.A.1 to 4.A.3. For the smaller step-size, both integrators enjoy very high acceptance rates and similar performance when not adjusted for timing; when adjusted for timing, the generalized leapfrog is often the better choice in the presence of a small step-size. For the larger step-size, only the implicit midpoint integrator is able to maintain a high acceptance rate; occasionally, the implicit midpoint is able to produce the optimal mean ESS and minimum ESS per second.

4.A.4 Volume Preservation and Symmetry Metrics

Here we describe how we compute metrics related to volume preservation and symmetry. Let \((q_1, \ldots, q_n)\) be samples generated by Hamiltonian Monte Carlo with numerical integrator \(\Phi : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m \times \mathbb{R}^m\); each \(q_i\) is an element of \(\mathbb{R}^m\).

Reversibility

For each sample \(q_i\), generate \(p_i|q_i \sim \text{Normal}(0, G(q_i))\) and compute \((q_i', p_i') = \Phi(q_i, p_i)\). Now compute \((q_i'', -p_i'') = \Phi(q_i', -p_i').\) The violation of reversibility is defined by

\[
\sqrt{\| (q_i, p_i') - (q_i'', p_i'') \|^2_2}. \tag{4.A.17}
\]

If the numerical integrator is reversible, this norm will be zero. In our metrics, we report the median violation of reversibility.
Table 4.A.1: Logistic regression on breast cancer dataset metrics. Comparison of the implicit midpoint and generalized leapfrog integrators on sampling from the Bayesian logistic regression posterior on the Breast Cancer dataset. For the larger step-size, the acceptance rate of the generalized leapfrog integrator completely deteriorates whereas the implicit midpoint integrator is more robust. The implicit midpoint integrator is able to achieve super-efficient sampling for a step-size of $\epsilon = 1$ and ten integration steps.

### Volume Preservation

Let $z = (q, p)$ and identify $\Phi(z) \equiv \Phi(q, p)$. Define $f_j(z) = \frac{\Phi(z_1, \ldots, z_j + \eta/2, \ldots, z_{2m}) - \Phi(z_1, \ldots, z_j - \eta/2, \ldots, z_{2m})}{\eta}$ for $\eta = 1 \times 10^{-5}$ (except for the Fitzhugh-Nagumo model where we set $\eta = 1 \times 10^{-3}$ for numerical reasons), which is the central difference formula that approximates $\frac{\partial}{\partial z_j} \Phi(z)$. We compute the approximation to the Jacobian of $\Phi$ by constructing,

$$
\nabla \Phi(z) \approx F(z) \overset{\text{def.}}{=} \left( f_1(z) \quad f_2(z) \quad \cdots \quad f_{2m}(z) \right) \in \mathbb{R}^{2m \times 2m}. \quad (4.A.18)
$$
If the numerical integrator is volume preserving, this difference will be zero. In our metrics, we report the median violation of volume preservation.

Table 4.A.2: Logistic regression on diabetes dataset metrics. Comparison of the implicit midpoint and generalized leapfrog integrators on sampling from the Bayesian logistic regression posterior on the Diabetes dataset.

For each sample \( q_i \), generate \( p_i|q_i \sim \text{Normal}(0, G(q_i)) \) and set \( z_i = (q_i, p_i) \). The violation of volume preservation is defined by

\[
||\det(F(z_i))|| - 1. 
\]

(4.A.19)
<table>
<thead>
<tr>
<th>Step Size</th>
<th>Num. Steps</th>
<th>Method</th>
<th>Acc. Prob.</th>
<th>Mean ESS</th>
<th>Min. ESS</th>
<th>Mean ESS / Sec.</th>
<th>Min. ESS / Sec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>5</td>
<td>G.L.F.(a)</td>
<td>1.00 ± 0.00</td>
<td>65.34 ± 9.93</td>
<td>54.18 ± 16.32</td>
<td>3.86 ± 0.05</td>
<td>3.19 ± 0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G.L.F.(b)</td>
<td>1.00 ± 0.00</td>
<td>65.56 ± 6.98</td>
<td>53.31 ± 11.07</td>
<td>9.45 ± 0.15</td>
<td>2.75 ± 0.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(a)</td>
<td>1.00 ± 0.00</td>
<td>66.14 ± 4.07</td>
<td>56.04 ± 8.36</td>
<td>4.23 ± 0.04</td>
<td>2.07 ± 0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(b)</td>
<td>1.00 ± 0.00</td>
<td>66.05 ± 6.83</td>
<td>51.75 ± 15.55</td>
<td>4.78 ± 0.08</td>
<td>3.73 ± 0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(c)</td>
<td>1.00 ± 0.00</td>
<td>65.59 ± 6.25</td>
<td>54.81 ± 13.39</td>
<td>4.91 ± 0.05</td>
<td>4.10 ± 0.10</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>G.L.F.(a)</td>
<td>0.99 ± 0.00</td>
<td>301.18 ± 16.31</td>
<td>265.45 ± 40.27</td>
<td>9.13 ± 0.17</td>
<td>8.05 ± 0.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G.L.F.(b)</td>
<td>0.99 ± 0.00</td>
<td>302.32 ± 18.34</td>
<td>268.55 ± 45.81</td>
<td>24.54 ± 0.28</td>
<td>21.83 ± 0.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(a)</td>
<td>1.00 ± 0.00</td>
<td>301.84 ± 16.66</td>
<td>272.01 ± 27.65</td>
<td>9.82 ± 0.07</td>
<td>8.86 ± 0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(b)</td>
<td>1.00 ± 0.00</td>
<td>302.96 ± 13.91</td>
<td>268.39 ± 43.94</td>
<td>11.42 ± 0.10</td>
<td>10.12 ± 0.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(c)</td>
<td>1.00 ± 0.00</td>
<td>302.41 ± 18.40</td>
<td>272.08 ± 31.33</td>
<td>11.87 ± 0.12</td>
<td>10.68 ± 0.11</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>G.L.F.(a)</td>
<td>0.99 ± 0.00</td>
<td>493.18 ± 14.56</td>
<td>450.03 ± 42.68</td>
<td>3.14 ± 0.02</td>
<td>2.86 ± 0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G.L.F.(b)</td>
<td>0.99 ± 0.00</td>
<td>495.56 ± 14.79</td>
<td>443.26 ± 54.96</td>
<td>8.84 ± 0.07</td>
<td>7.91 ± 0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(a)</td>
<td>1.00 ± 0.00</td>
<td>510.34 ± 20.29</td>
<td>462.57 ± 72.19</td>
<td>3.44 ± 0.01</td>
<td>3.12 ± 0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(b)</td>
<td>1.00 ± 0.00</td>
<td>508.77 ± 21.30</td>
<td>464.54 ± 54.89</td>
<td>4.02 ± 0.02</td>
<td>3.67 ± 0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(c)</td>
<td>1.00 ± 0.00</td>
<td>507.61 ± 15.17</td>
<td>460.57 ± 49.33</td>
<td>4.18 ± 0.03</td>
<td>3.79 ± 0.04</td>
</tr>
<tr>
<td>1.0</td>
<td>5</td>
<td>G.L.F.(a)</td>
<td>0.05 ± 0.00</td>
<td>25.97 ± 3.33</td>
<td>7.21 ± 1.36</td>
<td>0.08 ± 0.01</td>
<td>0.02 ± 0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G.L.F.(b)</td>
<td>0.05 ± 0.00</td>
<td>29.48 ± 2.90</td>
<td>9.96 ± 2.03</td>
<td>0.37 ± 0.04</td>
<td>0.13 ± 0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(a)</td>
<td>0.70 ± 0.00</td>
<td>4335.06 ± 63.32</td>
<td>3250.56 ± 138.84</td>
<td>6.53 ± 0.12</td>
<td>4.90 ± 0.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(b)</td>
<td>0.70 ± 0.00</td>
<td>4352.91 ± 83.82</td>
<td>3300.84 ± 186.04</td>
<td>5.31 ± 0.19</td>
<td>4.95 ± 0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(c)</td>
<td>0.70 ± 0.00</td>
<td>4455.35 ± 51.17</td>
<td>3372.85 ± 76.75</td>
<td>7.20 ± 0.14</td>
<td>5.45 ± 0.15</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>G.L.F.(a)</td>
<td>0.03 ± 0.00</td>
<td>44.35 ± 4.24</td>
<td>12.20 ± 1.90</td>
<td>0.12 ± 0.01</td>
<td>0.03 ± 0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G.L.F.(b)</td>
<td>0.02 ± 0.00</td>
<td>44.07 ± 8.26</td>
<td>10.09 ± 1.91</td>
<td>0.51 ± 0.11</td>
<td>0.11 ± 0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(a)</td>
<td>0.66 ± 0.00</td>
<td>9217.63 ± 305.79</td>
<td>5320.85 ± 661.55</td>
<td>5.45 ± 0.21</td>
<td>3.16 ± 0.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(b)</td>
<td>0.66 ± 0.00</td>
<td>9752.37 ± 315.58</td>
<td>5172.40 ± 382.12</td>
<td>5.72 ± 0.20</td>
<td>3.05 ± 0.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(c)</td>
<td>0.66 ± 0.00</td>
<td>9421.09 ± 664.29</td>
<td>5257.13 ± 448.88</td>
<td>5.90 ± 0.44</td>
<td>3.29 ± 0.29</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>G.L.F.(a)</td>
<td>0.01 ± 0.00</td>
<td>19.19 ± 3.58</td>
<td>4.62 ± 0.55</td>
<td>0.05 ± 0.02</td>
<td>0.01 ± 0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G.L.F.(b)</td>
<td>0.01 ± 0.00</td>
<td>24.34 ± 7.90</td>
<td>5.81 ± 1.41</td>
<td>0.21 ± 0.07</td>
<td>0.05 ± 0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(a)</td>
<td>0.58 ± 0.00</td>
<td>2818.14 ± 244.09</td>
<td>1287.74 ± 212.68</td>
<td>0.22 ± 0.02</td>
<td>0.10 ± 0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(b)</td>
<td>0.58 ± 0.01</td>
<td>3028.35 ± 370.09</td>
<td>1445.54 ± 329.95</td>
<td>0.25 ± 0.03</td>
<td>0.12 ± 0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I.M.(c)</td>
<td>0.56 ± 0.01</td>
<td>2672.53 ± 429.87</td>
<td>1308.80 ± 386.88</td>
<td>0.22 ± 0.04</td>
<td>0.11 ± 0.03</td>
</tr>
</tbody>
</table>

Table 4.A.3: Logistic regression on heart disease dataset metrics. Comparison of the implicit midpoint and generalized leapfrog integrators on sampling from the Bayesian logistic regression posterior on the heart disease dataset.

### 4.A.5 Implementation of Integrators

#### 4.A.5.1 Implementations of the Generalized Leapfrog Integrator

As the purpose of this research is to compare two integrators for Hamiltonian Monte Carlo, we wish to be precise about how these numerical methods have been implemented.

The first implementation of the generalized leapfrog integrator is presented in algorithm 2. The system eqs. (4.7) to (4.9) was described as “naive” because it appears to ignore important structural properties of the equations of motion in eqs. (4.2) and (4.3) that would accelerate a step of the generalized leapfrog integrator. For instance, in eq. (4.7),
the metric $G(q)$ may be precomputed because it is an invariant of the fixed-point relation. Another example is that $G^{-1}(q)\bar{p}$ is an invariant quantity of eq. (4.8) and need not be recomputed in each fixed-point iteration. Thus, we see that the generalized leapfrog integrator, when efficiently implemented, has an important computational advantage in that certain invariant quantities can be “cached” when finding fixed points. A more complicated implementation of the generalized leapfrog integrator with caching is presented in given in algorithm 4. We stress that algorithms 2 and 4 perform the same calculation.

### 4.A.5.2 Implementations of the Implicit Midpoint Integrator

**Algorithm 4 (G.L.F.(b))** The procedure for a single step of integrating Hamiltonian dynamics using the efficient implementation of the generalized leapfrog integrator.

1. **Input:** Log-posterior $L : \mathbb{R}^m \to \mathbb{R}$, Riemannian metric $G : \mathbb{R}^m \to \mathbb{R}^{m \times m}$, initial position and momentum variables $(q, p) \in \mathbb{R}^m \times \mathbb{R}^m$, integration step-size size $\epsilon \in \mathbb{R}$.
2. Precompute $G^{-1}(q), \frac{\partial}{\partial q_i}G(q)$, and define
   \[
   A_i \overset{\text{def}}{=} \frac{\partial}{\partial q_i}L(q) + \frac{1}{2}\text{trace}\left(G^{-1}(q)\frac{\partial}{\partial q_i}G(q)\right) \quad (4.A.20)
   \]
   for $i = 1, \ldots, m$.
3. Use algorithm 1 with tolerance $\delta$ and initial guess $p$ to solve for $\bar{p}$,
   \[
   \bar{p} = \underbrace{p - \frac{\epsilon}{2} \left( A_1 - \frac{1}{2}(G^{-1}(q)\bar{p})^\top \frac{\partial}{\partial q_1}G(q)(G^{-1}(q)\bar{p}), \ldots, A_m - \frac{1}{2}(G^{-1}(q)\bar{p})^\top \frac{\partial}{\partial q_m}G(q)(G^{-1}(q)\bar{p}) \right)^\top}_{f(\bar{p})}. \quad (4.A.21)
   \]
4. Precompute $G(q)\bar{p}$.
5. Use algorithm 1 with tolerance $\delta$ and initial guess $q$ to solve for $q'$,
   \[
   q' = q + \frac{\epsilon}{2} \left(G^{-1}(q)\bar{p} + G^{-1}(q')\bar{p}\right). \quad (4.A.22)
   \]
6. Compute $p'$ using eq. (4.9), which is an explicit update.
7. **Return:** $(q', p') \in \mathbb{R}^m \times \mathbb{R}^m$. 

210
Algorithm 5 (I.M.(b)) The procedure for a single step of integrating Hamiltonian dynamics using the implicit midpoint integrator as advocated by [Leimkuhler and Reich, 2005].

1: **Input:** Hamiltonian $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, initial position and momentum variables $(q, p) \in \mathbb{R}^m \times \mathbb{R}^m$, integration step-size size $\epsilon \in \mathbb{R}$, fixed-point convergence tolerance $\delta \geq 0$.

2: Use algorithm 1 with tolerance $\delta$ and initial guess $(q, p)$ to solve for $(\bar{q}, \bar{p})$,

$$\begin{pmatrix} \bar{q} \\ \bar{p} \end{pmatrix} \triangleq \begin{pmatrix} q \\ p \end{pmatrix} + \frac{\epsilon}{2} \left( \begin{pmatrix} \nabla_p H(\bar{q}, \bar{p}) \\ -\nabla_q H(\bar{q}, \bar{p}) \end{pmatrix} \right).$$  

(4.A.23)

3: Compute the explicit update

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \begin{pmatrix} q \\ p \end{pmatrix} + \frac{\epsilon}{2} \left( \begin{pmatrix} \nabla_p H(\bar{q}, \bar{p}) \\ -\nabla_q H(\bar{q}, \bar{p}) \end{pmatrix} \right).$$  

(4.A.24)

4: **Return:** $(q', p') \in \mathbb{R}^m \times \mathbb{R}^m$.

Algorithm 6 (I.M.(c)) The procedure for a single step of integrating Hamiltonian dynamics using the implicit midpoint integrator based on the discrete Lagrangian [Mackay, 1992, Hairer et al., 2006].

1: **Input:** Log-posterior $L : \mathbb{R}^m \to \mathbb{R}$, metric $G : \mathbb{R}^m \to \mathbb{R}^{m \times m}$, derivatives of the metric $\nabla G : \mathbb{R}^m \to \mathbb{R}^{m \times m \times m}$, initial position and momentum variables $(q, p) \in \mathbb{R}^m \times \mathbb{R}^m$, integration step-size size $\epsilon \in \mathbb{R}$, fixed-point convergence tolerance $\delta \geq 0$.

2: Use algorithm 1 with tolerance $\delta$ and initial guess $(q, G^{-1}(q)p)$ to solve for $(\bar{q}, \bar{v})$,

$$\begin{pmatrix} \bar{q} \\ \bar{v} \end{pmatrix} \triangleq \begin{pmatrix} q \\ G^{-1}(q)p \end{pmatrix} + \frac{\epsilon}{2} \left( G^{-1}(\bar{q}) \left[ \nabla L(\bar{q}) + \bar{v}^\top \nabla G(\bar{q}) \bar{v} + \frac{1}{2} \text{trace} (G^{-1}(\bar{q}) \nabla G(\bar{q})) \right] \right).$$

(4.A.25)

3: Compute the explicit update

$$\begin{pmatrix} \bar{q}' \\ \bar{v}' \end{pmatrix} \triangleq \begin{pmatrix} \bar{q} \\ \bar{v} \end{pmatrix} + \frac{\epsilon}{2} \left( G^{-1}(\bar{q}) \left[ \nabla L(\bar{q}) + \bar{v}^\top \nabla G(\bar{q}) \bar{v} + \frac{1}{2} \text{trace} (G^{-1}(\bar{q}) \nabla G(\bar{q})) \right] \right).$$  

(4.A.26)

4: Set $p' = G(\bar{q})\bar{v}'$.

5: **Return:** $(q', p') \in \mathbb{R}^m \times \mathbb{R}^m$.

In addition to the implementation of the implicit midpoint method described in algorithm 3, we also consider a variant advocated by [Leimkuhler and Reich, 2005] and present
the implementation in algorithm 5. The essential difference between algorithms 3 and 5 is whether or not the implicitly-defined update computes the terminal point of the step (algorithm 3) or the midpoint of the step (algorithm 5). We stress that algorithms 3 and 5 perform the same calculation when $\delta = 0$, which is easily verified by plugging eq. (4.A.23) into eq. (4.A.24) and comparing to eq. (4.10). Unlike the generalized leapfrog integrator, the implicit midpoint integrator does not enjoy the ability to cache intermediate computations. We show the implementation of the implicit midpoint integrator based on the discrete Lagrangian formulation in algorithm 6.

4.A.6 Proposal Distance of the Implicit Midpoint

**Proposition 4.A.6.1.** Consider the harmonic oscillator differential equation defined by,

$$\frac{d}{dt} z = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} z.$$  \hspace{2cm} (4.A.27)

Identifying $z = (q, p)$, let $R_{\text{lf}}^\epsilon : \mathbb{R}^2 \to \mathbb{R}$ be the transformation of the $q$-variable computed by the leapfrog integrator with step-size $\epsilon$ and let $R_{\text{im}}^\epsilon : \mathbb{R}^2 \to \mathbb{R}$ be the transformation of the $q$-variable computed by the implicit midpoint integrator with step-size $\epsilon$. Then

$$|R_{\text{lf}}^\epsilon (q, p) - q| = \left( 1 + \frac{\epsilon^2 \omega^2}{4} \right) |R_{\text{im}}^\epsilon (q, p) - q|.$$  \hspace{2cm} (4.A.28)

**Proof.** The propagator matrix of the leapfrog integrator on the harmonic oscillator is given by,

$$\begin{pmatrix} q'_{\text{lf}} \\ p'_{\text{lf}} \end{pmatrix} = \begin{pmatrix} 1 - \frac{\epsilon^2 \omega^2}{2} & \epsilon \\ -\epsilon \omega^2 & 1 - \frac{\epsilon^2 \omega^2}{2} \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}. \hspace{2cm} (4.A.29)$$

212
The propagator matrix of the implicit midpoint integrator is likewise given by,

\[
\begin{pmatrix}
q_{\text{im}}' \\
p_{\text{im}}'
\end{pmatrix}
= \frac{1}{1 + \epsilon^2 \omega^2 / 4}
\begin{pmatrix}
1 - \frac{\epsilon^2 \omega^2}{4} & \epsilon \\
-\epsilon \omega^2 & 1 - \frac{\epsilon^2 \omega^2}{4}
\end{pmatrix}
\begin{pmatrix}
q \\
p
\end{pmatrix}.
\] (4.A.30)

Therefore,

\[
\mathcal{R}_{\text{lf}}^\epsilon(q, p) = q - q\frac{\epsilon^2 \omega^2}{2} + \epsilon p
\] (4.A.31)

\[
\mathcal{R}_{\text{im}}^\epsilon(q, p) = \frac{1}{1 + \epsilon^2 \omega^2 / 4}
\left(q - q\frac{\epsilon^2 \omega^2}{4} + \epsilon p\right)
\] (4.A.32)

Moreover,

\[
\mathcal{R}_{\text{lf}}^\epsilon(q, p) - q = -q\frac{\epsilon^2 \omega^2}{2} + \epsilon p
\] (4.A.33)

\[
\mathcal{R}_{\text{im}}^\epsilon(q, p) - q = \frac{1}{1 + \epsilon^2 \omega^2 / 4}
\left(q - q\frac{\epsilon^2 \omega^2}{4} + \epsilon p\right) - \left(\frac{1 + \epsilon^2 \omega^2}{4}\right) q
\] (4.A.34)

\[
= \frac{1}{1 + \epsilon^2 \omega^2 / 4}
\left(-q\frac{\epsilon^2 \omega^2}{2} + \epsilon p\right)
\] (4.A.35)

\[
= \frac{1}{1 + \epsilon^2 \omega^2 / 4}
(\mathcal{R}_{\text{lf}}^\epsilon(q, p) - q)
\] (4.A.36)

Rearranging terms and taking the absolute value gives the conclusion.

This result is important because it gives a plausible explanation for the phenomenon that the higher acceptance probability enjoyed by the implicit midpoint integrator does not always guarantee a larger effective sample size. The reason for this is that the updates computed by the leapfrog method give larger transitions in the \(q\)-variable than the transitions generated by the implicit midpoint method. How much larger? On the harmonic oscillator, the distance is modulated by a factor of \(1 + \frac{\epsilon^2 \omega^2}{4}\). The leapfrog integrator is only stable when \(\epsilon^2 \omega^2 < 4\); if we take four as an upper bound on this quantity for both integrators, then the proposal distance of the implicit midpoint method can be a little as
half that computed by the leapfrog integrator.

**Proposition 4.A.6.2.** Suppose we wish to draw samples from the Gaussian distribution Normal($0, \Sigma$). A sensible corresponding Hamiltonian is

$$H(q, p) = \frac{1}{2} q^\top \Sigma^{-1} q + \frac{1}{2} p^\top \Sigma p. \quad (4.A.37)$$

Or, writing $z = (q, p)$ as the concatenation of the position and momentum variable,

$$H(z) = \frac{1}{2} z^\top \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & \Sigma \end{pmatrix} z \quad (4.A.38)$$

As $|\epsilon| \to \infty$, Hamiltonian Monte Carlo that uses the implicit midpoint integrator is not even ergodic for the Gaussian distribution and is, in fact, periodic.

**Proof.** Let $\mathbb{J} \in \mathbb{R}^{2m \times 2m}$ be an invertible, skew-symmetric matrix. One step of the implicit midpoint integrator computes

$$z' = \left(\text{Id}_{2m} - \frac{\epsilon}{2} \mathbb{J} \mathbb{A}\right)^{-1} \left(\text{Id}_{2m} + \frac{\epsilon}{2} \mathbb{J} \mathbb{A}\right) z \quad (4.A.39)$$

$$= \left(\frac{\epsilon}{2}\right)^{-1} \left(\frac{2}{\epsilon} \text{Id}_{2m} - \mathbb{J} \mathbb{A}\right)^{-1} \left(\frac{\epsilon}{2}\right) \left(\frac{2}{\epsilon} \text{Id}_{2m} + \mathbb{J} \mathbb{A}\right) z \quad (4.A.40)$$

$$= \left(\frac{2}{\epsilon} \text{Id}_{2m} - \mathbb{J} \mathbb{A}\right)^{-1} \left(\frac{2}{\epsilon} \text{Id}_{2m} + \mathbb{J} \mathbb{A}\right) z. \quad (4.A.41)$$

Thus,

$$\lim_{|\epsilon| \to \infty} \left(\frac{2}{\epsilon} \text{Id}_{2m} - \mathbb{J} \mathbb{A}\right)^{-1} \left(\frac{2}{\epsilon} \text{Id}_{2m} + \mathbb{J} \mathbb{A}\right) = - (\mathbb{J} \mathbb{A})^{-1} \mathbb{J} \mathbb{A} \quad (4.A.42)$$

$$= - \text{Id}_{2m} \quad (4.A.43)$$

Thus, in the limit of large step-sizes, $z' = -z$. Writing $z' = (q', p')$, this formula says that
the proposals computed by the implicit midpoint method bounce back and forth between $q$ and $-q$ on repeated applications of the implicit midpoint integrator. Such a transition is obviously not ergodic, nor does the transition even depend on the momentum variable.
Chapter 5

On Numerical Considerations in Riemannian Manifold Hamiltonian Monte Carlo

This chapter is adapted from [Brofos and Lederman 2021b]. This work is a collaboration with Roy Lederman. The vast majority of the exposition is due to me, as is all of the experimentation.

Abstract. Riemannian manifold Hamiltonian Monte Carlo (RMHMC) is a sampling algorithm that seeks to adapt proposals to the local geometry of the posterior distribution. The specific form of the Hamiltonian used in RMHMC necessitates implicitly-defined numerical integrators in order to sustain reversibility and volume-preservation, two properties that are necessary to establish detailed balance of RMHMC. In practice, these implicit equations are solved to a non-zero convergence tolerance via fixed-point iteration. However, the effect of these convergence thresholds on the ergodicity and computational efficiency properties of RMHMC are not well understood. The purpose of this research is to elucidate these relationships through numerous case studies. Our analysis reveals circumstances wherein the RMHMC algorithm is sensitive, and insensitive, to these convergence tolerances. Our empirical analysis examines several aspects of the computation: (i) we examine the ergodicity of the RMHMC Markov chain by employing statistical methods
for comparing probability measures based on collections of samples; (ii) we investigate the degree to which detailed balance is violated by measuring errors in reversibility and volume-preservation; (iii) we assess the efficiency of the RMHMC Markov chain in terms of time-normalized ESS. In each of these cases, we investigate the sensitivity of these metrics to the convergence threshold and further contextualize our results in terms of comparison against Euclidean HMC. We propose a method by which one may select the convergence tolerance within a Bayesian inference application using techniques of stochastic approximation. In addition, we examine Newton’s method as an alternative to fixed point iterations, and find that it can eliminate much of the sensitivity of RMHMC to the convergence threshold.

5.1 Introduction

Bayesian inference provides a theoretical basis for reasoning under uncertainty in statistical applications. In particular, let \( q \in \mathbb{R}^m \) be a parameter vector of interest and let \( \mathcal{L}(q) \) denote the log-density of \( q \); in many applications of interest \( \mathcal{L}(q) \) will be, up to an additive constant, the sum of the log-likelihood of data given \( q \) and a log-density representing a prior belief over \( q \). A central task in Bayesian inference is the estimation of expectations of functions of the random variable \( q \). However, the integrals required for the computation of the expectation may be intractable, for example in the case of high dimension \( m \). Instead, in order to compute expectations over the distribution over \( q \), we seek to generate samples from the distribution of \( q \) and form a Monte Carlo expectation; this idea has led to the development of Markov chain Monte Carlo (MCMC), which requires only that \( \mathcal{L}(q) \) be known up to an additive constant. MCMC sampling methods are the subject of many textbooks; an overview can be found in, inter alia, [Brooks et al.] [2011]. In MCMC, one establishes a Markov chain whose stationary distribution has density proportional to \( \exp(\mathcal{L}(q)) \); by running the Markov chain for a “large” number of steps, one obtains a
sequence of auto-correlated samples whose marginal distributions have converged – one hopes – to the distribution of interest. We refer the interested reader to [Meyn and Tweedie 1993] for a discussion of the ergodicity properties of Markov chains.

Hamiltonian Monte Carlo (HMC) [Duane et al., 1987, Betancourt, 2017, Neal, 2010b] is a MCMC sampling procedure that seeks to efficiently explore the typical set of a smooth probability density by integrating Hamilton’s equations of motion. A variant of HMC that is the focus of this work is Riemannian manifold Hamiltonian Monte Carlo (RMHMC), which seeks to exploit the local geometry of the posterior distribution in order to align proposals along directions of the posterior that exhibit the greatest local variation. The de facto standard numerical integrator for RMHMC is the generalized leapfrog method [Betancourt, 2012, Girolami and Calderhead, 2011, Cobb et al., 2019, Brofos and Lederman, 2021a], which involves implicitly defined steps. Only when these implicit steps are solved “perfectly” is the generalized leapfrog method reversible (assumption 5.2.15) and volume preserving (assumption 5.2.16) in the sense required to establish stationarity of the Markov chain (theorem 5.2.18 and corollary 5.2.19). These implicit steps are typically resolved to a finite precision via fixed point iteration, thereby necessitating the practitioner to choose a suitable convergence tolerance for their posterior. This begs the questions:

1. To what extent is detailed balance violated in practice with variable convergence thresholds?

2. How much do non-zero convergence thresholds effect the ergodicity of the RMHMC Markov chains?

3. How does the choice of convergence threshold impact the computational efficiency and time-normalized performance of the sampler?

In the case of violations of detailed balance, we emphasize that the resulting errors may be significantly above the machine precision, indicating that the errors we study are not due
to the limitations of finite precision arithmetic but instead caused by the errors introduced by truncating the fixed point iteration after a specified level of convergence is achieved. The role of these convergence tolerances in the generalized leapfrog method, and its effect on the ergodicity of the Markov chain, does not appear to have been thoroughly evaluated in the literature. Therefore, the purpose of this work is to elucidate these relationships, paying special attention to the manner in which these convergence tolerances impact the detailed balance condition in RMHMC, for which specialized integrators such as the generalized leapfrog is the principle motivation. In our experimental evaluation, we find that there tends to be a convergence threshold after which point no further improvement in the ergodicity of the Markov chain can be obtained, even though detailed balance may be more precisely enforced by decreasing the convergence threshold. Since sample ergodicity, rather than detailed balance, is what matters in Bayesian analysis, computational benefits – such as faster sampling – emerge by careful selection of this convergence parameter. We propose one mechanism, based on Ruppert averaging, by which one may choose a convergence tolerance for a given posterior.

An additional contribution of this work lies in its examination of Newton’s method as an alternative to fixed point iterations with special attention to the reversibility, volume-preservation, and ergodicity of the resulting method. Newton’s method has an important computational property that, within its radius of convergence, it identifies the solution faster than fixed point iterations. This faster order of convergence can eliminate much of the sensitivity of RMHMC to the specification of the convergence tolerance; however, Newton’s method is more computationally burdensome than fixed point iterations. Our analysis illuminates circumstances wherein RMHMC implemented with Newton’s method possesses desirable computational properties, and metrics under which fixed point iterations are preferred.

The structure of the chapter is as follows. section 5.2 is a review of relevant concepts from classical mechanics, Markov chains, Riemannian manifold Hamiltonian Monte
Carlo, and numerical methods. We begin in section 5.2.1 with a summary of existing work on the generalized leapfrog integrator and fixed point iterations, which are referenced directly in this study. The remainder of section 5.2 provides a broader context on classical mechanics, Markov chains, fixed point iterations and Newton’s method, Riemannian manifold Hamiltonian Monte Carlo, and stochastic approximation. Section 5.2.3 elaborates on the importance of reversibility (assumption 5.2.15) and volume preservation (assumption 5.2.16) in establishing detailed balance of Markov chains based on numerical integration. In section 5.3 we review related work on RMHMC, with special attention toward those works that have investigated the computational efficiency of the technique. Section 5.4 states the research question about the importance of these thresholds in the RMHMC algorithm and discusses the metrics we use in our experiments to study these questions; we additionally propose a method based on stochastic approximation to select a convergence threshold. Section 5.4 continues to introduce the use of Newton’s method in RMHMC as a mechanism to reduce the sensitivity of the RMHMC algorithm to the convergence threshold. More detailed information about the algorithms can be found in section 5.A.5. Section 5.5 focuses on experimentation, which, together with section 5.A.8 consider non-identifiable models, hierarchical Bayesian logistic regression, hierarchical models with all posterior quantities sampled jointly, a hierarchical stochastic volatility and log-Gaussian Cox-Poisson model, Bayesian inference in stochastic differential equations, and a multiscale Student-\(t\) distribution.

### 5.2 Preliminaries

The purpose of this section is to summarize existing results about RMHMC. For brevity, section 5.2.1 summarizes the essential results used directly in this chapter. The remainder of the section restates additional results that presents the nuances of RMHMC discussed in this chapter in a broader context.
5.2.1 Essentials for Riemannian Manifold Hamiltonian Monte Carlo

Definition 5.2.1. Let \( \Pi \) be a target probability distribution with log-density denoted \( \mathcal{L} : \mathbb{R}^m \rightarrow \mathbb{R} \). Let \( G \) be a map from \( \mathbb{R}^m \) to the set of positive definite matrices (i.e. \( G(q) \in \mathbb{R}^{m \times m} \) is positive definite for every \( q \in \mathbb{R}^m \)). The Riemannian Hamiltonian \( H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \) is defined by

\[
H(q, p) = -\mathcal{L}(q) + \frac{1}{2} \log \det(G(q)) + \frac{1}{2} p^\top G^{-1}(q)p.
\] (5.1)

Hamiltonians such as the one described in definition 5.2.1 give rise to equations of motion, which are described in detail in section 5.2.2. In most cases, there will not be a closed-form solution to these equations of motion. Instead, methods of numerical integration are employed to compute approximate solutions. We now describe the generalized leapfrog integrator, arguably the most popular integrator, which is the focus of our analysis.

Definition 5.2.2. Let \( H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \) be a smooth function and fix \( \epsilon \in \mathbb{R} \). The generalized leapfrog integrator is the map \((\tilde{q}, \tilde{p}) = \hat{\Phi}_\epsilon(q, p) \) defined by

\[
\dot{p} = p - \frac{\epsilon}{2} \nabla_q H(q, \tilde{p}) \tag{5.2}
\]

\[
\dot{q} = q + \frac{\epsilon}{2} (\nabla_p H(q, \tilde{p}) + \nabla_p H(\tilde{q}, \tilde{p})) \tag{5.3}
\]

\[
\dot{\tilde{p}} = \tilde{p} - \frac{\epsilon}{2} \nabla_q H(\tilde{q}, \tilde{p}). \tag{5.4}
\]

We note that eq. (5.2) and eq. (5.3) are implicit equations, where \( \dot{p} \) or \( \dot{q} \) appear on both sides. Therefore, an implementation of the algorithm in definition 5.2.1 requires a numerical solver for these equations; the issues arising from this solver are discussed in this chapter. eq. (5.2) is called the implicit update to the momentum and eq. (5.3) is called the implicit update to the position.
Definition 5.2.3. The Riemannian manifold Hamiltonian Monte Carlo Markov chain is defined inductively as follows. Let \( q_n \in \mathbb{R}^m \) be the state of the Markov chain at the \( n \)-th step. Sample \( p_n | q_n \sim \text{Normal}(0, G(q)) \). Let \( H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \) be as in definition 5.2.1.

Then apply the generalized leapfrog integrator (definition 5.2.2) with step-size \( \epsilon \in \mathbb{R} \) and \( k \in \mathbb{N} \) integration steps to \( H \) and obtain \( (q'_{n+1}, p'_{n+1}) = \hat{\Phi}_k^\epsilon(q_n, p_n) \), where \( \hat{\Phi}_k^\epsilon \) denotes \( k \) applications of the integrator. With probability \( \min\{1, \exp(H(q_n, p_n) - H(q'_{n+1}, p'_{n+1}))\} \), set the new state to be \( q_{n+1} = q'_{n+1} \); otherwise, remain at the current state and set \( q_{n+1} = q_n \).

It has been noted here and elsewhere [Neal, 2010b, Bishop, 2006, Girolami and Calderhead, 2011] that the correctness of Hamiltonian Monte Carlo is predicated on the fact that the generalized leapfrog integrator is reversible and volume-preserving. Reversibility and volume preservation are defined rigorously in assumption 5.2.15 and assumption 5.2.16 respectively, and the fact that the generalized leapfrog is reversible and volume preserving is established in proposition 5.2.10. The precise reason why reversibility and volume-preservation are important to the correctness of Hamiltonian Monte Carlo is treated in section 5.2.3, specifically theorem 5.2.18. As we noted, \( \tilde{p} \) in eq. (5.2) and \( \tilde{q} \) in eq. (5.3) are both solutions of fixed point equations. This brings us to an important distinction between how the generalized leapfrog method is discussed mathematically and how it would typically be implemented numerically on a digital computer. In particular, an implementation of the generalized leapfrog method must incorporate a procedure for solving the equations eqs. (5.2) and (5.3). Critically, numerical methods of solving fixed point equations rely on a convergence parameter \( \delta \) that controls the quality of the computed solution. To emphasize the role of the convergence tolerance, we adopt the notation \( \hat{\Phi}_k^\epsilon,\delta \) to mean \( k \) steps of the numerical integrator with convergence tolerance \( \delta \). We may now describe how the generalized leapfrog integrator would actually be implemented.

Definition 5.2.4. Let \( H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \) be a smooth function, fix \( \epsilon \in \mathbb{R} \) and \( \delta \in \mathbb{R}_+ \). An implementation of the generalized leapfrog integrator with step-size \( \epsilon \) and convergence
tolerance $\delta$ is a map $\hat{\Phi}^{\epsilon,\delta} : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m$ constructed as follows. For the implicit update to the momentum (eq. (5.2)), define the equation

$$g_p(\tilde{p}) = p - \frac{\epsilon}{2} \nabla_q H(q, \tilde{p}). \tag{5.5}$$

Solve the fixed point equation $\tilde{p} = g_p(\tilde{p})$ to tolerance $\delta$, with convergence measured in $\| \cdot \|_{\infty}$, obtaining the approximate solution $\tilde{p}^*$. Methods of finding solutions of fixed point equations are treated thoroughly in section 5.2.4, traditionally, RMHMC is implemented using fixed point iterations (definition 5.2.22).

Similarly, for the implicit update to the position, define the function,

$$g_q(\tilde{q}) = q + \frac{\epsilon}{2}(\nabla_p H(q, \tilde{p}^*) + \nabla_p H(\tilde{q}, \tilde{p}^*). \tag{5.6}$$

Solve the fixed point equation $\tilde{q} = g_q(\tilde{q})$ to tolerance $\delta$, with convergence measured in $\| \cdot \|_{\infty}$, obtaining the approximate solution $\tilde{q}^*$. Finally, compute

$$\tilde{p}^* = \tilde{p}^* - \frac{\epsilon}{2} \nabla_q H(\tilde{q}^*, \tilde{p}^*). \tag{5.7}$$

We then define the map $\hat{\Phi}^{\epsilon,\delta}(q, p) = (\tilde{q}^*, \tilde{p}^*)$.

We come now to the observation that motivates the entire study.

**Observation 5.2.5.** For $\delta > 0$, when $\hat{\Phi}^k_\epsilon$ (definition 5.2.2) is replaced by $\hat{\Phi}^{k,\epsilon,\delta}_\epsilon$ (definition 5.2.4) in Riemannian manifold Hamiltonian Monte Carlo (definition 5.2.3), $\hat{\Phi}^{k,\epsilon,\delta}_\epsilon$ no longer satisfies exact reversibility or volume preservation. The degree to which these properties are violated is controllable by the convergence parameter $\delta$.

The preceding observation motivates us to investigate the numerical issues that emerge from the use of non-zero convergence thresholds in RMHMC, with particular attention given to how this parameter affects ergodicity, Markov chain performance metrics; we
also examine methods by which to select the convergence threshold and mechanisms by which to reduce the sensitivity of RMHMC to the convergence parameter.

The remainder of section 5.2 provides the broader context and the details omitted from this section for brevity.

### 5.2.2 Hamiltonian Mechanics

Here we will review the foundations of Hamiltonian mechanics; for a broader perspective on the importance of Hamiltonian mechanics in physics and for a thorough treatment grounded in differential geometry, we refer the interested reader to Marsden and Ratiu [2010]. Formally, given a smooth function $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, Hamilton’s equations of motion are described by the initial value problem,

$$
\dot{q}_t = \nabla_p H(q_t, p_t) \\
\dot{p}_t = -\nabla_q H(q_t, p_t),
$$

with initial condition $(q_0, p_0) \in \mathbb{R}^m \times \mathbb{R}^m$. Physically, the solutions $q_t$ and $p_t$ are referred to as the position and momentum variables, respectively, and together $(q_t, p_t)$ represents a point in phase space. Hamilton’s equations of motion possess several special properties, which we now recall.

**Assumption 5.2.6.** Let $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be a smooth function. Assume that $\nabla_p H(q, -p) = \nabla_p H(q, p)$ and $\nabla_q H(q, -p) = \nabla_q H(q, p)$.

**Proposition 5.2.7.** Let $H$ be a function satisfying assumption 5.2.6. Then the following properties hold: (i) energy is conserved over the course of the trajectory so that $\frac{d}{dt} H(q_t, p_t) = 0$; (ii) volume in $\mathbb{R}^m \times \mathbb{R}^m$ is conserved under the evolution prescribed by Hamilton’s equations of motion; i.e. $\text{div}((\dot{q}_t, \dot{p}_t)) = 0$; and (iii) the equations of motion are reversible under negation of the momentum in the sense that if $(q_t, p_t)$ are solutions...
to Hamilton’s equations of motion given in eqs. (5.8) and (5.9) then, for fixed \( \tau \in \mathbb{R} \), the trajectories \( \tilde{q}_s = q_{\tau-s} \) and \( \tilde{p}_s = -p_{\tau-s} \) are also solutions to Hamilton’s equations.

Proofs of these three properties are given in section 5.A.1. Since reversibility by momentum flip will play a critical role in our analysis, we give this operation its own notation.

**Definition 5.2.8.** Let \( q \) and \( p \) be vectors in \( \mathbb{R}^m \). The **momentum flip operator** is the map \( F(q, p) = (q, -p) \) regarded as a map from \( \mathbb{R}^m \times \mathbb{R}^m \) to itself.

**Definition 5.2.9.** Let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be a smooth function and fix \( \epsilon \in \mathbb{R} \). The implicit midpoint integrator is the map \( (\tilde{q}, \tilde{p}) = \hat{\Phi}_\epsilon(q, p) \) defined by,

\[
\begin{pmatrix}
\tilde{q} \\
\tilde{p}
\end{pmatrix} = \begin{pmatrix}
q \\
p
\end{pmatrix} + \epsilon \begin{pmatrix}
\nabla_p H(\tilde{q}, \tilde{p}) \\
\nabla_q H(\tilde{q}, \tilde{p})
\end{pmatrix}
\tag{5.10}
\]

where

\[
\tilde{q} = \frac{\tilde{q} + q}{2} \tag{5.11}
\]
\[
\tilde{p} = \frac{\tilde{p} + p}{2}. \tag{5.12}
\]

Given a numerical integrator \( \hat{\Phi}_\epsilon \), we denote \( k \) successive applications of the integrator by \( \hat{\Phi}_\epsilon^k = \hat{\Phi}_\epsilon \circ \cdots \circ \hat{\Phi}_\epsilon \) \((k \text{ times})\). Both the generalized leapfrog (definition 5.2.2) and the implicit midpoint definition 5.2.9 integrators approximate Hamiltonian dynamics with third-order local error in the sense that \( \| \hat{\Phi}_\epsilon(q, p) - (q_\epsilon, p_\epsilon) \| = \mathcal{O}(\epsilon^3) \).

**Proposition 5.2.10.** Suppose that assumption 5.2.6 holds for a Hamiltonian \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \). Then the generalized leapfrog integrator applied to \( H \) with step-size \( \epsilon \) conserves volume in phase-space and is reversible under negation of the momentum in the sense that \( F \circ \hat{\Phi}_\epsilon^k \circ F = \text{Id} \).

A proof is given in section 5.A.2. An analogous result for the implicit midpoint integrator was established in [Brofos and Lederman 2021a].
5.2.3 Markov Chains

We now discuss Markov chains on \( \mathbb{R}^m \). We denote by \( \mathcal{B}(\mathbb{R}^m) \) the Borel \( \sigma \)-algebra on \( \mathbb{R}^m \). Markov chains may be characterized completely by their transition kernel, which we now define.

Definition 5.2.11. A Markov chain transition kernel is a map \( K : \mathbb{R}^m \times \mathcal{B}(\mathbb{R}^m) \rightarrow [0, 1] \) with the properties [Robert and Casella 2005],

1. For any \( A \in \mathcal{B}(\mathbb{R}^m) \), \( K(\cdot, A) \) is measurable.

2. For any \( x \in \mathbb{R}^m \), \( K(x, \cdot) \) is a probability measure.

A Markov chain is then constructed from the Markov chain transition kernel by the relation \( X_{n+1} | X_n = x_n \sim K(x_n, \cdot) \). We denote by \( K^n(x_0, \cdot) \) the \( n \)-step transition law; i.e. the distribution of \( X_n | X_0 = x_0 \). In the context of MCMC, a Markov chain is used to generate approximate samples from a target density. We measure distances between probability measures via the total variation distance and call a Markov chain ergodic if the total variation distance is zero asymptotically.

Definition 5.2.12. Let \( \Pi \) and \( \Xi \) be two probability measures on \( \mathbb{R}^m \). The total variation distance between \( \Pi \) and \( \Xi \) is defined by

\[
\|\Pi(\cdot) - \Xi(\cdot)\|_{TV} = \sup_{A \in \mathcal{B}(\mathbb{R}^m)} |\Pi(A) - \Xi(A)|. \tag{5.13}
\]

Definition 5.2.13. Let \( K : \mathbb{R}^m \times \mathcal{B}(\mathbb{R}^m) \rightarrow \mathbb{R}_+ \) be a Markov chain transition kernel and let \( \Pi \) be a probability distribution on \( \mathbb{R}^m \). We say that \( K \) is ergodic for \( \Pi \) if,

\[
\lim_{n \rightarrow \infty} \|K^n(x, \cdot) - \Pi(\cdot)\|_{TV} = 0. \tag{5.14}
\]
Definition 5.2.14. Let $K : \mathbb{R}^m \times \mathcal{B}(\mathbb{R}^m) \to [0, 1]$ be a Markov chain transition kernel. Then $K$ is said to satisfy detailed balance with respect to the probability distribution $\Pi : \mathcal{B}(\mathbb{R}^m) \to \mathbb{R}_+$ if $\Pr_{z \sim \Pi} [K(z, A)] = \Pi(A)$ for every $z \in \mathbb{R}^m$ and $A \in \mathcal{B}(\mathbb{R}^m)$.

We now describe the transition kernel that will be of particular interest to us. The Markov chain is constructed from a smooth transformation $\Phi : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ which must satisfy the following assumptions.

Assumption 5.2.15. Let $\Phi : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ be a diffeomorphism. Assume that $\Phi$ is reversible; i.e. $\Phi \circ \Phi = \text{Id}$.

Assumption 5.2.16. Let $\Phi : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ be a diffeomorphism. Assume that $\Phi$ is volume preserving; i.e. for every $z \in \mathbb{R}^{2m}$ we have $|\text{det}(\nabla \Phi(z))| = 1$.

Definition 5.2.17. Let $\Pi : \mathcal{B}(\mathbb{R}^{2m}) \to \mathbb{R}_+$ be a probability distribution with density $\pi : \mathbb{R}^{2m} \to \mathbb{R}_+$. Fix $k \in \mathbb{N}$ and $\epsilon \in \mathbb{R}$. Let $\hat{\Phi}_\epsilon$ be a symplectic numerical integrator (such as described in definitions 5.2.2 and 5.2.9) and let $\Psi^k = F \circ \hat{\Phi}^k$ where $F$ is the momentum flip operator. The transition kernel of Hamiltonian Monte Carlo is

$$K((q, p), A \times B) = \min \left\{1, \frac{\pi(\Psi^k(q, p))}{\pi(q, p)} \right\} \mathbf{1} \{\Psi^k(q, p) \in A \times B\}
+ \left(1 - \min \left\{1, \frac{\pi(\Psi^k(q, p))}{\pi(q, p)} \right\} \right) \mathbf{1} \{(q, p) \in A \times B\},$$

(5.15)

where $A, B \in \mathcal{B}(\mathbb{R}^m)$.

Theorem 5.2.18. Let $K$ be the Markov chain transition kernel of Hamiltonian Monte Carlo with map $\Psi^k : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$. Then $K$ satisfies detailed balance with respect to the distribution $\Pi$ if assumptions 5.2.15 and assumption 5.2.16 hold for $\Psi^k$.

A proof of this result is given in section 5.A.3. The following result follows immediately from detailed balance.

227
Corollary 5.2.19. Let $K$ be the Markov chain transition kernel of Hamiltonian Monte Carlo with map $\Psi^k : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m}$. Under the conditions of assumption 5.2.15 and assumption 5.2.16, $\Pi$ is the stationary distribution of $K$.

Remark. Reversibility and volume-preservation are immediately seen to be very important properties; if they do not hold, there is no guarantee that the target distribution $\Pi$ is invariant under the Markov chain transition kernel. Failure of invariance may mean that the total variation distance $\|K^n(x, \cdot) - \Pi(\cdot)\|_{TV}$ can actually increase as a function of $n$, which is impossible when $\Pi$ is stationary for $K$. Practically, this means that one’s sample quality could conceivably degrade as the Markov chain is run.

Remark. We have not yet described how to construct either the probability measure $\Pi$ or the numerical integrator $\hat{\Phi}_\epsilon$ appearing in definition 5.2.17. In practice, the density $\pi$ of $\Pi$ is set to have the form $\pi(q, p) \propto \exp(-H(q, p))$ where $L(q) = \log \int_{\mathbb{R}^m} \pi(q, p) \, dp$ for an appropriate choice of Hamiltonian $H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$; this guarantees that the marginal distribution of $q$ generated by the Markov chain at stationarity is the target density. In this case, $\hat{\Phi}_\epsilon$ may be chosen to integrate the Hamiltonian vector field corresponding to the Hamiltonian $H$.

Indeed, since samples from $\pi(q) = \exp(L(q))$ are of principle interest from a Bayesian inference perspective, it is worthwhile to consider the marginal transition kernel of a Markov chain defined on $\mathbb{R}^m \times \mathbb{R}^m$.

Definition 5.2.20. Let $\pi(q, p)$ be the joint distribution of $(q, p)$ and let $\pi(p|q)$ denote the conditional distribution of $p$ given $q$. Consider the transition kernel $K$ given in definition 5.2.17, the marginal transition kernel is defined by,

$$\tilde{K}(q, A) = \int_{\mathbb{R}^m} K((q, p), A \times \mathbb{R}^m) \pi(p|q) \, dp$$

(5.16)

where $A \in \mathcal{B}(\mathbb{R}^m)$. 

228
In other words, the marginal transition kernel results from sampling \( p \) from its conditional distribution given a current state \( q \), sampling \((\tilde{q}, \tilde{p}) \sim K((q, p), \cdot)\), discarding \( \tilde{p} \) and taking \( \tilde{q} \) as the subsequent state of the Markov chain. From the condition that \( K \) satisfies detailed balance with respect to \( \pi(q, p) \), it is an easy verification that \( \tilde{K} \) satisfies detailed balance with respect to \( \pi(q) \).

### 5.2.4 Resolution of Fixed Point Equations

**Definition 5.2.21.** Let \( g : \mathbb{R}^m \to \mathbb{R}^m \). A fixed point \( z^* \in \mathbb{R}^m \) of \( g \) is a solution to the equation \( z = g(z) \).

In the present work, we will consider two mechanisms by which to resolve the solution of the fixed point equation.

**Definition 5.2.22.** Fix a convergence tolerance \( \delta > 0 \). The method of *fixed point iteration* approximately solves the fixed point iteration by generating a sequence \( z_{n+1} = g(z_n) \) given an initial candidate solution \( z_0 \in \mathbb{R}^m \). The sequence is terminated at the first \( n \) for which \( \|z_{n+1} - z_n\|_\infty < \delta \), for which we take \( z_* = z_{n+1} \) as an approximate solution of the fixed point equation.

**Definition 5.2.23.** Let \( g : \mathbb{R}^m \to \mathbb{R}^m \) be smooth. Fix a convergence tolerance \( \delta > 0 \). *Newton’s method* approximately solves the fixed point iteration by generating a sequence \( z_{n+1} = z_n - (\text{Id} - \nabla g(z_n))^{-1} g(z_n) \). The sequence is terminated at the first \( n \) for which \( \|z_{n+1} - z_n\|_\infty < \delta \), for which we take \( z_* = z_{n+1} \) as an approximate solution of the fixed point equation.

**Remark.** The use of the norm \( \cdot \|_\infty \) to measure convergence is not the only choice, since in \( \mathbb{R}^m \), \( \cdot \|_\infty, \cdot \|_2, \) and \( \cdot \|_1 \) are all equivalent in the sense that \( \|z\|_\infty \leq \|z\|_2 \leq \|z\|_1 \leq m\|z\|_\infty \). Therefore, convergence to tolerance \( \delta \) in \( \cdot \|_\infty \) implies convergence to tolerance \( m\delta \) in \( \cdot \|_1 \) and \( \cdot \|_2 \).
Definition 5.2.24. Let \((z_0, z_1, z_2, \ldots)\) be an \(\mathbb{R}^m\)-valued sequence converging to some \(z^* \in \mathbb{R}^m\). The sequence is said to converge with order \(k\) to the value \(z^*\) if,

\[
\lim_{n \to \infty} \frac{\|z_{n+1} - z^*\|_\infty}{\|z_n - z^*\|_\infty^k} < L
\]

for some \(L \in \mathbb{R}\).

Under conditions on the function \(g\) and initial condition \(z_0\), fixed point iterations converge with order one and Newton’s method converges with order two. This is made precise in the following two theorems from [Fletcher, 1987].

Theorem 5.2.25. Let \(g : \mathbb{R}^m \to \mathbb{R}^m\) be a smooth function and suppose that \(z^*\) is a solution of the fixed point equation \(z^* = g(z^*)\). Suppose that \(\|\nabla g(z^*)\|_{\text{op}} < 1\). Then there is a neighborhood \(U\) of \(z^*\) in which fixed point iterations (definition 5.2.22) converge with order at least one provided that \(z_k \in U\) for some \(k \in \mathbb{N}\).

Theorem 5.2.26. Let \(g : \mathbb{R}^m \to \mathbb{R}^m\) be smooth. Define the function \(r(z) = z - g(z)\) and note that if \(r(z^*) = 0\) then \(z^*\) is a solution of the fixed point equation \(z = g(z)\). Writing \(r(z) = (r_1(z), \ldots, r_m(z))\), suppose further that \(\frac{\partial r_i}{\partial z_j}(z)\) is a Lipschitz function for \(i, j \in \{1, \ldots, m\}\). Then there is a neighborhood \(U\) of \(z^*\) in which Newton’s method (definition 5.2.23) converges with order at least two provided that \(z_k \in U\) for some \(k \in \mathbb{N}\).

5.2.5 Riemannian Manifold Hamiltonian Monte Carlo

One difficulty that arises in HMC is that if the density has multiple spatial scales, employing the generalized leapfrog integrator with a Hamiltonian in the form of \(H(q, p) = -L(q) + \frac{1}{2}p^\top p\) will produce severe oscillations in the directions of smallest variation. This can be alleviated by employing a preconditioner that continuously adapts the momentum to the local variation of the posterior [Girolami and Calderhead, 2011]. Riemannian manifold Hamiltonian Monte Carlo (RMHMC) is a MCMC sampling procedure that seeks
to adapt proposals to the directions of greatest variation in the posterior locally. This is accomplished by introducing second-order geometric information into the Hamiltonian Monte Carlo transition operator; second-order information is typically represented by the sum of the Fisher information matrix of the log-likelihood and the negative Hessian of the log-prior, or by the Hessian of the log-density $\mathcal{L}(\theta)$. As in HMC, the proposal generated in RMHMC is obtained by numerically integrating Hamilton’s equations of motion corresponding to the Hamiltonian in definition 5.2.1. It is easily verified that the Hamiltonian in eq. (5.1) satisfies assumption 5.2.6; it follows, therefore, from proposition 5.2.10 that the generalized leapfrog integrator can be combined with the momentum flip operator to produce an involution suitable for use in the HMC transition kernel given in definition 5.2.17. This form of Hamiltonian is motivated by geometric principles; the Hamiltonian $H(q, p) = \frac{1}{2} p^T G(q)^{-1} p$ corresponds to co-geodesic motion on $(\mathbb{R}^m, G)$ [Calin and Chang, 2004]. Thus, physically, $\mathcal{L}(q) + \frac{1}{2} \log \det(G(q))$ represents a potential energy function that causes motion to deviate from the co-geodesics of the manifold. As conceived by Girolami and Calderhead [2011], when the density $\pi(q, p)$ is proportional to $\exp(-H(q, p))$, the term $\frac{1}{2} \log \det(G(q))$ is included in the Hamiltonian so that we obtain the conditional distribution $p|q \sim \text{Normal}(0_m, G(q))$. The marginal distribution of $q$ has density $\pi(q) \propto \exp(\mathcal{L}(q))$. The quadratic form in $p$ causes the form of the Hamiltonian to be “non-separable;” this means that the Hamiltonian is not expressible as the sum of two functions, one a function of the position alone and the other a function of the momentum alone. As a result, this Hamiltonian cannot be integrated by the (non-generalized) leapfrog method in a way that preserves reversibility and volume preservation. The form
of eq. (5.1) produces the following equations of motion:

\[ \dot{q}_t = \nabla_p H(q_t, p_t) = G(q_t)^{-1} p_t \]  
\[ \dot{p}_t = -\nabla_q H(q_t, p_t) = \nabla_q \mathcal{L}(q_t) - \frac{1}{2} \text{trace} \left( G(q_t)^{-1} \nabla G(q_t) \right) + \frac{1}{2} p_t^\top G(q_t)^{-1} \nabla G(q_t) G(q_t)^{-1} p_t, \]  

(5.18)  
(5.19)

where \( \text{trace}(G(q) \nabla G(q)) \) is an \( m \)-dimensional vector whose \( k \)-th element is \( \text{trace}(G(q)^{-1} \frac{\partial}{\partial q_k} G(q)) \).

Similarly, the \( k \)-th element of \( p^\top G(q)^{-1} \nabla G(q) G(q)^{-1} p \) is \( p^\top G(q)^{-1} \frac{\partial}{\partial q_k} G(q) G(q)^{-1} p \).

\textbf{Definition 5.2.27.} The transition kernel for Riemannian manifold Hamiltonian Monte Carlo (RMHMC) is defined as follows. Let \( \hat{\Phi}^k \epsilon \) be the \( k \)-step generalized leapfrog integrator with step-size \( \epsilon \in \mathbb{R} \) applied to the Hamiltonian \( H(q, p) \) in eq. (5.1). Denote by \( \pi(q, p) \) the density proportional to \( \exp(-H(q, p)) \). The RMHMC transition kernel is an instance of definition 5.2.17 with \( \hat{\Phi}^k \epsilon \) and \( \pi(q, p) \) as defined herein.

\textbf{Remark.} Unlike an implementation of the leapfrog integrator, whose reversibility and volume-preservation are assured up to machine precision, the errors in these properties seen in an implementation of the generalized leapfrog method will, even for “small” values of \( \delta \) (e.g. \( \delta = 1 \times 10^{-6} \), which measures absolute deviation between fixed point iterations), be far above the IEEE standard for floating point relative error. In the sequel, we will denote by \( \tau_{\text{mach}} \) the machine error, which is approximately \( 2.2 \times 10^{-16} \) under the IEEE standard; we note that \( \delta \) measures error in fixed point iterations in an absolute sense, whereas \( \tau_{\text{mach}} \) is the greatest relative error between an arbitrary real number and its closest approximation within a finite precision number system (see section 2.1). Therefore, \( \delta \) and \( \tau_{\text{mach}} \) should not be directly compared. Nevertheless, the degree to which an implementation of a numerical integrator violates reversibility and volume-preservation may be well in excess of the machine error to which one may have grown accustomed.

\textbf{Remark.} Implementations of RMHMC have differed greatly in the convergence tolerance, indicating a lack of consensus on how this quantity should be treated. For in-
stance [Cobb et al. 2019] uses a default convergence tolerance of $1 \times 10^{-20}$ with a convergence norm of $\max_{i=1,\ldots,m} |z_i|^2$. Brubaker et al. [2012] measures convergence in the norm $\max_{i=1,\ldots,m} |z_i|$ and uses a default convergence tolerance of $1 \times 10^{-6}$. Graham [2020] uses a convergence threshold of $1 \times 10^{-9}$ with a convergence norm defined by $\max_{i=1,\ldots,m} |z_i|$. Brofos and Lederman [2021a] search over convergence tolerances in the set $\{1 \times 10^{-3}, 1 \times 10^{-6}, 1 \times 10^{-3}\}$ with convergence measured in the norm $\max_{i=1,\ldots,m} |z_i|$ but without regard to which of these convergence tolerances should be preferred from a computational standpoint. Girolami and Calderhead [2011] eschewed convergence tolerances altogether, instead recommending between four and six fixed point iterations per step. We will use the method of measuring convergence tolerance since it clearly defines the accuracy of the solution that we seek to obtain. Moreover, unlike a prescribed number of iterations, a convergence tolerance can avoid unnecessary additional iterations once the threshold has been achieved.

5.2.6 Stochastic Approximation

Stochastic approximation is a mechanism for finding the solution of a non-linear equation when the equation itself is measured with error. Formally, we give the following problem statement.

**Problem.** Let $B : \mathbb{R} \to \mathbb{R}$ be an unknown function and let $\kappa \in \mathbb{R}$. Suppose there exists a random variable $L_\xi \in \mathbb{R}$ for which $\mathbb{E}[L_\xi] = B(\xi)$. Using observations of $L_\xi$, identify a value of $\xi$ for which $B(\xi) = \kappa$.

In the case where $B$ is directly computable, solving the equation $B(\xi) = \kappa$ could be accomplished via a line search or, under suitable smoothness assumptions, via (quasi-)Newton methods. The fact that $B(\xi)$ is only measured stochastically via $L_\xi$ complicates matters. Robbins and Monro [1951] proposed a sequential method by which to identify $\xi$ such that $B(\xi) = \kappa$ which depends only on a sequence $L(\xi_n)$.  

233
Theorem 5.2.28. Consider the sequence of random variables defined recursively by

\[ \xi_{n+1} = \xi_n - \gamma_n(L(\xi_n) - \kappa) \]  

(5.20)

where \( \{\gamma_n\}_{n=1}^\infty \) is a sequence of \( \mathbb{R} \)-valued step-sizes. Suppose that the following conditions hold:

1. \( \sum_{i=1}^\infty \gamma_i = \infty \) and \( \sum_{i=1}^\infty \gamma_i^2 < \infty \).
2. The function \( B \) is monotonically increasing.
3. There exists \( \xi \in \mathbb{R} \) such that \( B(\xi) = \kappa \).
4. The function \( B \) is differentiable in a neighborhood of \( \xi \) and \( \frac{d}{d\xi} B(\xi) > 0 \).

Then \( \mathbb{E}(\xi_n - \xi)^2 \to 0 \).

The following method, called Ruppert averaging, takes the method of Robbins and Monro [1951] further and establishes an asymptotic optimality property.

Theorem 5.2.29 (Ruppert [1988]). Consider a step-size sequence of the form \( \gamma_n = Dn^{-\omega} \) for \( \omega \in (1/2, 1) \) and \( D \in \mathbb{R}_+ \). Under the assumptions of theorem 5.2.28 and with the sequence \( \{\xi_n\}_{n=1}^\infty \) constructed as in eq. (5.20), construct the average sequence,

\[ \bar{\xi}_{n+1} = \frac{n}{n+1} \bar{\xi}_n + \frac{1}{n+1} \xi_{n+1}. \]

(5.21)

The sequence \( \{\bar{\xi}_n\}_{n=1}^\infty \) is asymptotically efficient as an estimator of \( \xi \) in the sense that \( \bar{\xi}_n \) has the smallest asymptotic variance among all consistent estimators of \( \xi \).

Methods similar to this averaging procedure have been considered previously in application to MCMC; perhaps most notable among these is the dual averaging method of Hoffman and Gelman [2014], which seeks to tune the step-size of HMC in order to achieve a desired acceptance probability.
5.3 Related Work

The method of RMHMC was originally conceived in [Girolami and Calderhead, 2011], though the fundamental idea of using adaptive geometry to inform the proposals can be found earlier in [Zlochin and Baram, 2000]. This latter work used the (non-generalized) leapfrog integrator to compute proposals, which [Girolami and Calderhead, 2011] criticized as lacking reversibility and volume-preservation, and therefore failing to maintain detailed balance with respect to the target distribution. Several attempts at resolving the implicit updates used in applying the generalized leapfrog integrator. [Lan et al., 2015] made progress in this direction by transforming from Hamiltonian to Lagrangian dynamics, which eliminates fixed point iterations but at a cost of requiring the determinant of an $m \times m$ matrix at every step of the integrator; moreover, the resulting integrator purports to have a larger global error rate than integration methods based on Hamiltonian mechanics. It is also not obvious how to incorporate general-purpose metrics such as SoftAbs [Betancourt, 2012] into the Lagrangian framework. More recently, [Cobb et al., 2019] proposed an explicit integrator for RMHMC by “coupling” two copies of the numerical trajectory; however, this method also suffers from the lack of a theory of reversibility and volume-preservation. In previous work by the authors [Brofos and Lederman, 2021a], we examined the use of the implicit midpoint integrator as an alternative to the generalized leapfrog method, though the implicit midpoint method lacks the computational efficiency properties of the generalized leapfrog method. In the present work, we focus our attention on the generalized leapfrog integrator since it is the de facto standard integrator used in RMHMC and to enable a simplified analysis of the threshold effects under consideration.

Yet another approach to improving the performance of RMHMC was explored by [Zhang and Sutton, 2014], which proposes to utilize a Riemannian metric with a specific “alternating block-wise” structure that facilitates the use of the non-generalized leapfrog method.
The importance of volume preservation and symmetry of the proposal operator is often stressed in the proofs of detailed balance for HMC (see, inter alia, [Neal 2010b], [Lelièvre et al. 2019]) proposed a check for violations of reversibility in the case of embedded manifolds, wherein the numerical integrator involves solving for Lagrange multipliers.

5.4 Analytical Apparatus and Algorithms

Observation 5.2.5 states that when the generalized leapfrog integrator $\hat{\Phi}_\epsilon$ (definition 5.2.2) is implemented with a non-zero convergence tolerance, yielding the map $\hat{\Phi}_{\epsilon,\delta}$ (definition 5.2.4), the properties of reversibility in assumption 5.2.15 and volume preservation in assumption 5.2.16 are no longer exactly conserved. Therefore, the conclusion of detailed balance of the RMHMC Markov chain (definition 5.2.3) deduced in theorem 5.2.18 no longer holds. This observation introduces a wealth of questions that we now seek to examine.

**How should one determine the convergence threshold?** Each fixed point iteration in RMHMC incurs a computational cost of $O(m^3)$; see section 5.A.4 for a discussion of the computational complexity of the implicit updates. Therefore, it seems reasonable to alleviate the computational burden by selecting a non-zero threshold that does not degrade the ergodicity of the RMHMC algorithm. We investigate methods by which one may adapt the convergence threshold, similar to how step-sizes are adapted in the No-U-Turn sampler [Hoffman and Gelman 2014].

**To what extent are the ergodicity and the computational efficiency of the RMHMC sampler affected by these convergence tolerances?** The question of ergodicity is important because the objective of a sampler is to produce a Markov chain that converges to the target density; if the choice of threshold produces detectable and substantial differences between the large-sample distribution of the chain and the target distribution, then ergodicity has been meaningfully violated. The computational efficiency is also important: if a
larger threshold produces minute differences in the large-sample distribution of the chain, then one would like to characterize the computational savings associated to this larger threshold relative to a more stringent one. This will allow us to quantify and compare, among other properties, the effective sample size per unit of computation in the RMHMC algorithm with varying thresholds.

**Can one devise implementations of RMHMC that are less sensitive to the convergence threshold?** Standard implementations of the generalized leapfrog integrator use fixed point iterations (definition 5.2.22) to find the solutions to the implicit update to the momentum (eq. (5.2)) or the implicit update to position (eq. (5.3)); this is the approach advocated by Hairer et al. [2006], Leimkuhler and Reich [2005] and employed in Betancourt [2012], Girolami and Calderhead [2011], Lan et al. [2015], Cobb et al. [2019] and many others. However, by utilizing a solver of fixed point equations such as Newton’s method (definition 5.2.23), which enjoys second-order convergence, the sensitivity of the generalized leapfrog integrator to the choice of threshold may be significantly reduced relative to the standard implementation using fixed point iterations. We propose and evaluate generalized leapfrog integrator employing Newton’s method in application to RMHMC.

We next elaborate on each of these themes before proceeding to experimental evaluation in section 5.5.

### 5.4.1 Applying Newton’s Method in Riemannian Manifold Hamiltonian Monte Carlo

As noted above, standard implementations of the generalized leapfrog integrator (definition 5.2.4) use fixed point iterations (definition 5.2.22) in order to find solutions to the fixed point equations eqs. (5.2) and (5.3) defining the generalized leapfrog integrator. As noted in theorem 5.2.25 fixed point iterations have convergence of order one. Convergence of order one means that one can expect that, with each fixed point iteration, the error of the
computed solution is proportional to the error of the previous solution. An algorithmic description of fixed point iteration is given in algorithm 8 in section 5.A.5. By contrast, Newton’s method (definition 5.2.23) converges with order two as stated in theorem 5.2.26; one expects that, with each Newton iteration, the error of the computed solution is proportional to the squared error of the previous solution. As a result, the actual accuracy of the solution obtained by Newton’s method may be significantly better than the convergence threshold would suggest. An algorithm implementing Newton’s method is provided in algorithm 9 in section 5.A.5.

The stronger order of convergence enjoyed by Newton’s method leads one to expect that an implementation of RMHMC wherein the generalized leapfrog integrator is implemented with Newton, rather than fixed point, iterations will exhibit decreased sensitivity to the convergence threshold. Algorithm 10 in section 5.A.5 gives an implementation of the generalized leapfrog algorithm making clear that the fixed point equations in eqs. (5.2) and (5.3) may be solved to tolerance $\delta$ using either fixed point iteration or Newton’s method. Using either of these fixed point equation solvers, the generalized leapfrog integrator may then be employed in the RMHMC Markov chain, an algorithmic implementation of which is given in algorithm 11 in section 5.A.5. In section 5.5.3 we turn to the empirical evaluation of this modified implementation of RMHMC.

5.4.2 Identification of Threshold via Stochastic Approximation

We propose a mechanism to adapt the convergence threshold during sampling as a component of a burn-in phase. Specifically, we propose to adapt the threshold to achieve a prescribed average number of decimal digits of similarity with a numerical integrator using a strict convergence tolerance (such as $1 \times 10^{-10}$). In order to apply the techniques in section 5.2.6, we must construct the random variable $L_\xi$ and function $B(\xi)$ described in section 5.2.6.
**Definition 5.4.1.** Let $\tau_{\text{mach}}$ denote the machine precision (typically $\tau_{\text{mach}} \approx 2.2 \times 10^{-16}$). Let $\delta$ and $\delta'$ be two convergence thresholds and let $\hat{\Phi}_{\epsilon,\delta}$ denote the generalized leapfrog integrator with step-size $\epsilon$ and convergence tolerance $\delta$ as described in definition 5.2.4. Define the quantity

\[
G_{\epsilon,k}(q,p,\delta,\delta') = \begin{cases} 
\log_{10} \max \left( \|\hat{\Phi}_{\epsilon,\delta}^k(q,p) - \hat{\Phi}_{\epsilon,\delta'}^k(q,p)\|_2, \tau_{\text{mach}} \right) & \text{if } \delta > \delta' \\
\log_{10} \tau_{\text{mach}} & \text{otherwise.}
\end{cases}
\]  

(5.22)

This is the negative number of decimal digits of similarity between the numerical integrators with step-size $\epsilon$ and number of integration steps $k$, but with differing convergence tolerances $\delta$ and $\delta'$. This measure treats thresholds that are less than or equal to the baseline as equivalent; for thresholds greater than the baseline, we measure the number of decimal digits of similarity between their respective integrators. We may wish to choose $\delta$ to produce a given number of decimal digits of similarity between these two integrators.

**Definition 5.4.2.** Fix $\delta' \in \mathbb{R}_+$. Given a target probability measure $\Pi$ with density $\pi(q) = \exp(\mathcal{L}(q))$, let $q \sim \Pi$ and $p|q \sim \text{Normal}(0_m, G(q))$. Define the random variable

\[
L_\xi \equiv L_\xi(q,p; 10^\xi, \delta') = G_{\epsilon,k}(q,p,10^\xi,\delta')
\]

(5.23)

where $G_{\epsilon,k}(q,p,\delta,\delta')$ is defined as in definition 5.4.1. We then define the function

\[
B(\xi) = \mathbb{E}_{q \sim \Pi} \mathbb{E}_{p \sim \text{Normal}(0_m,G(q))} L_\xi,
\]

(5.24)

The function $B(\xi)$ defined in definition 5.4.2 is the negative average number of decimal digits of similarity between the numerical integrators $\hat{\Phi}_{\epsilon,\delta}^k$ and $\hat{\Phi}_{\epsilon,\delta'}^k$ where $\delta = 10^\xi$. We may wish to choose $\xi$ to produce a given number of decimal digits of similarity between these two integrators; we denote the desired negative number of decimal digits of similarity by $\kappa$. Hence, we seek $B(\xi) = \kappa$. The role of $\delta'$ is to specify a convergence tolerance that
is assumed to be strict enough for the target posterior.

We have now formulated the problem of threshold selection in the form of section 5.2.6. We therefore employ Ruppert averaging as described in theorem 5.2.29 in order to produce sequences \((\xi_n)\) defined in eq. (5.20) and \(\bar{\xi}_n\) defined in eq. (5.21). In our experiments, we set \(\delta' = 1 \times 10^{-10}\), which is assumed to represent a convergence tolerance strict enough for all target distributions considered here, and \(\omega = 3/4\) and \(D = 1\), where \(\omega\) and \(D\) produce the step-size \(\gamma_n = Dn^{-\omega}\) in accordance with theorem 5.2.29; these parameters were found to produce reasonable behavior. Convergence may be faster for variations of these hyperparameters. We consider a maximum value of 1,000 for \(n\), the number of steps.

Ideally, we would want to choose a value of the threshold that produces a desired error in reversibility (assumption 5.2.15) and volume preservation (assumption 5.2.16) relative to the convergence threshold \(\delta'\). The number of decimal digits of similarity between transition integrators is only a proxy for the violation of reversibility and volume preservation. However, this is a measure that is simple to compute and deploy in practice; in contrast, a measure to directly match a prescribed average violation of volume preservation would be significantly more expensive due to the use of finite differences to approximate the Jacobian. Moreover, if one can hypothesize a minimum scale of the posterior distribution, then this scale can be used to guide how many decimal digits of similarity one should require on average from the RMHMC numerical integrator. Therefore, we suggest the number of decimal digits of similarity as an appropriate proxy metric for threshold selection.

In our experimental evaluation, we make an effort to check that the assumptions of the averaging procedure, enumerated in theorem 5.2.28 are satisfied. To accomplish this, we compute a Monte Carlo approximation of \(B(\xi)\) for one-hundred logarithmically-spaced values of \(\delta = 1 \times 10^6\) between \(1 \times 10^{-10}\) and \(1 \times 10^{-1}\). In practice, these Monte Carlo approximations of \(B(\xi)\) appear to be monotonically increasing, to have a value of \(\xi\) satisfying \(B(\xi) = \kappa\), and look smooth.

In our experiments, we compute a Monte Carlo approximation of \(B(\xi)\) in which we use
i.i.d. samples from the target distribution $\Pi$ consistent with the definition of $B(\xi)$ given in eq. (5.24). In practice, however, one will not have access to i.i.d. samples from $\Pi$; nevertheless, one may employ a Markov chain in order to generate approximate samples from $\Pi$. Indeed, in our experiments, when reporting sequences $(\bar{L}_n)$ and $(\bar{\xi}_n)$, in order to mirror the actual practice of the technique, we instead draw samples using an RMHMC Markov chain with convergence threshold $\delta'$. Therefore, the value of $\bar{\xi}_n$ after 1,000 iterations may not precisely match the apparent solution of the equation $B(\xi) = \kappa$, though typically the two values are close. As a specific example, in figs. 5.4a and 5.4b we employ samples of $q$ generated by a Markov chain, whereas in fig. 5.4c we use i.i.d. samples from the target distribution. In algorithm 12 in section 5.A.5 we provide an algorithm implementing this approach for identifying a threshold that produces a desired number of decimal digits of similarity between two integrators.

5.4.3 Measures and Metrics for Markov Chains

In the experiments in section 5.5 we use several metrics to measure the reversibility and volume preservation (assumption 5.2.15 and assumption 5.2.16), as well as the quality of the samples in the sense of their ergodicity and computational efficiency. The detailed procedures are described in section 5.A.7; in summary, when measuring reversibility, we simply apply the integrator forward, negate the momentum, and integrate again and measure distance from the initial position. Volume preservation is measured via finite-differences. Ergodicity is computed by comparing against i.i.d. samples, when available, using sliced Wasserstein and Kolmogorov-Smirnov distances, maximum mean discrepancy, and methods based on comparing the convergence of multiple chains along specific dimensions of the posterior. In addition to measuring ergodicity in this manner, we compute the time-normalized ESS in order to assess the relative efficiency of the sampling procedure. Computational effort is measured by the number of fixed point iterations re-
quired to resolve the implicit updates to position and momentum. We seek to measure the similarity of RMHMC transition kernels by comparing the expected distance (in $q$-space) between transitions computed using variable thresholds.

5.5 Experiments

In this section, we present our empirical analysis of the role of thresholds in RMHMC. Throughout our experiments, we carefully control the seed of the pseudo-random number generator used in sampling the random momentum and in applying the Metropolis-Hastings accept-reject criterion. As a result of this experimental design, we may assess the causal effect of adjusting the convergence threshold. We summarize some critical aspects of our experimentation in table 5.A.1. Code implementing these experiments may be found at https://github.com/JamesBrofos/Thresholds-in-Hamiltonian-Monte-Carlo. We present a significantly expanded experimental evaluation in section 5.A.8 wherein we include additional experiments on a stochastic volatility model, a Cox-Poisson model, a multi-scale Student-$t$ distribution, a banana-shaped distribution, and hierarchical Bayesian logistic regression.

Throughout our experimental results, one will observe “diminishing returns” as the convergence threshold is decreased toward zero. We expect this behavior as solutions of fixed point equations converge as the threshold decreases to zero.

5.5.1 Neal’s Funnel Distribution

Neal’s funnel distribution [Neal, 2003] is a hierarchical distribution constructed as follows $v \sim \text{Normal}(0, 9)$ and $x_i | v \sim \text{Normal}(0, e^{-v})$ for $i = 1, \ldots, 10$. One sees by inspection that this distribution is trivial to sample analytically. However, the purpose of Neal’s funnel distribution is to provide an example of a distribution which HMC struggles to sample. Indeed, for large values of $v$, the conditional distribution $x_i | v$ becomes increasingly
concentrated near zero, producing the eponymous funnel shape. Without preconditioning, HMC is unable to penetrate this narrow funnel. Neal’s funnel distribution is also challenging because it represents a distribution in which no global preconditioning is apparent. Therefore, in applying RMHMC to this task, we follow Betancourt [2012] and employ the SoftAbs Riemannian metric.

In our experiments we consider RMHMC with varying thresholds and with an integration step-size set to $\epsilon = 0.2$ and a maximum number of integration steps equal to twenty-five. We also compare RMHMC against HMC with eight integration steps and an integration step-size $\epsilon \in \{0.001, 0.01, 0.1, 0.2\}$; the parameters of RMHMC and HMC were chosen based off the discussion in Betancourt [2012]. When assessing ergodicity in Neal’s funnel distribution, results are reported in fig. 5.2b; one observes that the weakest threshold produces a chain whose similarity to the target distribution is approximately the same as HMC with step-size 0.1 or 0.2, yielding around 1.5 digits of similarity in the Kolmogorov-Smirnov statistics along a randomly chosen subspace. When the threshold is decreased to $1 \times 10^{-2}$, around 2.5 digits of similarity are obtained for a randomly chosen one-dimensional subspace. All of the thresholds smaller than $1 \times 10^{-2}$ produce nearly indistinguishable measures of ergodicity as measured by the Kolmogorov-Smirnov statistic.
Figure 5.2: Neal’s funnel distribution sampling metrics and ergodicity. Visualization of the sample quality of the variables in Neal’s funnel distribution as measured by the ESS per second (fig. 5.2a) and the distribution of Kolmogorov-Smirnov (KS) statistics over random one-dimensional subspaces (fig. 5.2b). In measuring KS, we also compare against i.i.d. samples, for which the average KS statistic (over random one-dimensional subspaces) is also shown. Distributions over ESS are computed by splitting a Markov chain of length 1,000,000 into twenty contiguous sequences of length 50,000.

along a random subspace. Although employing RMHMC with a threshold of $1 \times 10^{-3}$ typically exhibits only two digits of reversibility and volume preservation, the transition kernels exhibit a similarity of around 2.5 digits. This level of performance can be achieved with three or four fixed point iterations on average for each implicit update, compared with the eleven or twelve required by a stronger convergence tolerance of $1 \times 10^{-9}$, which
Figure 5.3: *Neal’s funnel distribution ergodicity.* Additional measures of ergodicity in Neal’s funnel distribution. According to the Kolmogorov-Smirnov and MMD$^2_a$ statistics, the Riemannian methods with a threshold of $1 \times 10^{-1}$ are competitive with the best-performing Euclidean methods. However, by the sliced Wasserstein metric, all Riemannian methods have better ergodicity compared to the Euclidean variants. In any event, all metrics agree that ergodicity performance is essentially constant for thresholds less than $1 \times 10^{-2}$. We show convergence of independent Markov chains in the marginal distribution of $v$ which is Normal$(0, 3^2)$. We see that convergence of Riemannian methods is faster than for HMC; however, when employing the largest threshold, one can observe the bias in the stationary distribution. In computing the MMD statistic, we use a kernel bandwidth of 8.4.

Figure 5.4: *Ruppert averaging in Neal’s funnel distribution.* The use of Ruppert averaging in Neal’s funnel distribution to adaptively set the convergence threshold to achieve six decimal digits of similarity compared to a transition kernel with a threshold of $1 \times 10^{-10}$. We show a Monte Carlo approximation to $B(\delta)$, which appears smooth and monotonically increasing. The value of $\delta$ satisfying $B(\delta) = -6$ is approximately $\delta = 1 \times 10^{-6.5}$, which is somewhat greater than the value of $1 \times 10^{-7}$ produced by Ruppert averaging.

offers negligible benefits in terms of ergodicity.

We apply the Ruppert averaging procedure in Neal’s funnel distribution in order to identify a threshold that produces, on average, six ($\kappa = 6$) decimal digits of similarity with a numerical integrator whose convergence threshold is $1 \times 10^{-10}$. We show the results of
Figure 5.5: Visualization of the posterior mean of \(v_t\) (in blue) and \(r_t\) (in orange) in the Fitzhugh-Nagumo posterior distribution. The posterior means are visually indistinguishable irrespective of the convergence tolerance used in RMHMC.

this procedure in fig. 5.4. The sequence of \(\bar{L}_n\) stabilizes at zero by iteration 100. The sequence of \(\bar{\delta}_n\) has converged by iteration 500.

Neal’s funnel distribution offers one of the most convincing examples of the benefit of a Riemannian approach to MCMC. We illustrate this phenomenon in fig. 5.2, which compares HMC with variable step-sizes against RMHMC with variable thresholds. For both the variables \((x_1, \ldots, x_{10})\) and the hierarchical variance \(v\), the ESS of the Riemannian methods are orders of magnitude larger than the MCMC procedures without preconditioning.

5.5.2 Fitzhugh-Nagumo Differential Equation Posterior

The Fitzhugh-Nagumo differential equation is a two-dimensional ordinary differential equation of the form,

\[
\dot{v}_t = c \left( v_t - \frac{v_t^3}{3} + r_t \right) \tag{5.25}
\]

\[
\dot{r}_t = - \left( \frac{v_t - a + br_t}{c} \right), \tag{5.26}
\]

where \((a, b, c)\) are parameters of the system. Given the initial condition \(v_0 = -1\) and \(r_0 = +1\), consider observing \(\tilde{v}_{tk} = v_{tk} + \epsilon_{v,k}\) and \(\tilde{r}_{tk} = r_{tk} + \epsilon_{r,k}\) where \(\epsilon_{v,k}, \epsilon_{r,k} \overset{i.i.d.}{\sim} \text{Normal}(0, \sigma^2)\) for \(k = 1, \ldots, 200\) and where \((t_1, \ldots, t_{200})\) are 200 equally-spaced points in the interval \([0, 10)\). The Bayesian inference task at hand is to sample from the poste-
Figure 5.6: *Fitzhugh-Nagumo distribution detailed balance and ergodicity.* Visualization of the errors in reversibility and volume preservation in the Fitzhugh-Nagumo posterior distribution, and the ergodicity of the RMHMC Markov chain as measured by the Kolmogorov-Smirnov statistics over random sub-spaces. In measuring KS, we also compare against i.i.d. samples, for which the average KS statistic (over random one-dimensional subspaces) is also shown.

Collectively denoting the parameters of the Fitzhugh-Nagumo differential equation model by \( q = (a, b, c) \), it follows from the general expression for the Fisher information of a multivariate normal that the Riemannian metric formed by the sum of the Fisher information of the log-likelihood and the negative Hessian of the log-prior assumes the...
Figure 5.7: Fitzhugh-Nagumo distribution sampling metrics. Visualization of the effective sample sizes (ESS) of the variables in the Fitzhugh-Nagumo posterior. Distributions over ESS are computed by splitting a Markov chain of length 100,000 into twenty contiguous sequences of length 5,000. We observe that the effective sample size for RMHMC is in excess of 5,000, indicating super-efficient sampling.

Figure 5.8: Fitzhugh-Nagumo ergodicity metrics. Additional measures of ergodicity in the Fitzhugh-Nagumo differential equation model. In this example, the MMD\(_2\) statistic indicates that a threshold of \(1 \times 10^{-2}\) has the smallest value of the squared maximum mean discrepancy. This is in contrast to the Kolmogorov-Smirnov and sliced Wasserstein metrics, which indicate the ergodicity improves to a threshold of \(1 \times 10^{-3}\). In the MMD statistic we use a kernel bandwidth of 0.088.

The following form,

\[
G_{ij}(q) = \frac{1}{\sigma^2} \sum_{k=1}^{200} \left( \frac{\partial v_{t_k}}{\partial q_i} \frac{\partial v_{t_k}}{\partial q_j} + \frac{\partial r_{t_k}}{\partial q_i} \frac{\partial r_{t_k}}{\partial q_j} \right) + \mathbf{1}\{i = j\}
\]  

(5.27)

We see, therefore, that unlike the other distributions considered thus far, the values of the log-posterior, its gradient, and Riemannian metric of the Fitzhugh-Nagumo model are not available in closed-form. Instead, these quantities are approximated by numerically
Figure 5.9: The use of Ruppert averaging in the Fitzhugh-Nagumo posterior distribution to adaptively set the convergence threshold to achieve four decimal digits of similarity compared to a transition kernel with a threshold of $1 \times 10^{-10}$. We show a Monte Carlo approximation to $B(\delta)$, which appears smooth and monotonically increasing. The value of $\delta$ satisfying $B(\delta) = -4$ is approximately $\delta = 1 \times 10^{-5.3}$, which is approximately the value produced by Ruppert averaging.

solving initial value problems involving sensitivities of the required quantities. In our implementation, we solve these initial value problems using SciPy’s odeint function with its default parameters. In total there are twenty initial value problems to be solved in the FitzHugh-Nagumo posterior: the two equations for the FitzHugh-Nagumo dynamics in eqs. (5.25) and (5.26), six equations for the first-order sensitivities of $v_t$ and $r_t$ with respect to the parameters, and twelve equations for the second order sensitivities. As an interesting consequence of this approximation, the computed “derivatives” of the Hamiltonian are no longer exact up to machine precision due to accumulating errors associated to the numerical solution of the initial value problems. As demonstrated in section 5.A.2, the proof that the generalized leapfrog integrator conserves volume is predicated on the symmetry of the partial derivatives of the Hamiltonian, which may be violated when analytical expressions for derivatives are supplanted by numerical solutions to differential equations.

We can apply the Ruppert averaging method in order to find a threshold in the Fitzhugh-Nagumo differential equation posterior that produces a numerical integrator with four decimal digits of similarity. Figure 5.9 shows convergence of the sequences $\bar{L}_n$ and $\bar{\delta}_n$. The sequence $\bar{\delta}_n$ by iteration 100 and the sequence $\bar{L}_n$ has converged by the five-hundredth iteration.
We follow Girolami and Calderhead [2011] and set a number of integration steps equal to six and use an integration step-size of 0.5 in our experiments. For the case of HMC, we follow Tripuraneni et al. [2017] and use ten integration steps and a step-size of 0.015, which produces an acceptance rate of around ninety-percent. We see in fig. 5.8c that a threshold of $1 \times 10^{-2}$ is sufficient to obtain around 2.5 or 3 decimal digits of similarity relative to a transition kernel with threshold $1 \times 10^{-10}$. In terms of ergodicity, we observe that RMHMC with a threshold of $1 \times 10^{-1}$ produces samples that are arguably of lesser quality than a HMC baseline; see fig. 5.6c. However, a threshold of $1 \times 10^{-2}$ appears to produce an improvement over HMC and for thresholds less than $1 \times 10^{-2}$ there is no evident ergodicity advantage. One notes that the Riemannian metric and its gradients constitute an expensive metric to compute, even though the parameter space of the posterior is only three-dimensional. This is because the Riemannian metric and its gradients require computing solutions to initial value problems. Therefore, the computational burden of computing fixed point solutions is significant in the FitzHugh-Nagumo posterior, particularly in the implicit update of the position variable, for which we must recompute the metric at each iteration.

When assessing ergodicity of the RMHMC algorithm in the FitzHugh-Nagumo model, we use rejection sampling to generate 100,000 independent samples as a point of comparison. To apply rejection sampling, we use a uniform distribution in a cube centered at the posterior mode and whose side lengths are ten times the marginal standard deviations computed from a Laplace approximation to the posterior at the mode. Similar to the conclusion of Girolami and Calderhead [2011], we find the Euclidean HMC performs competitively with RMHMC, though RMHMC does exhibit a distribution of Kolmogorov-Smirnov statistics somewhat more tightly concentrated near zero. The RMHMC algorithm also yields samples that are less auto-correlated, producing a larger ESS, as shown in fig. 5.7. In section 5.A.10 we consider ergodicity failures that can occur when transcription errors invalidate the assumption of a unit Jacobian determinant.
Figure 5.10: We illustrate the dependency of the reversibility metric on the procedure used to resolve the implicit updates to momentum and position. In most examples, employing Newton’s method to update both the position and the momentum variables in the generalized leapfrog integrator produced faster convergence to, and better respect for, reversibility of the proposal up to numerical precision. We note that in the banana-shaped distribution, using Newton’s method to update both momentum and position has severe errors in reversibility; see our discussion in the text and section 5.A.9 for an explanation.

5.5.3 Alternatives to Fixed Point Iteration: Newton’s Method

We now proceed to consider an alternative to fixed point iterations to resolve the implicit update to the position and momentum in the generalized leapfrog integrator. In particular, we evaluate using Newton’s method (definition 5.2.23) to compute approximate solutions to fixed point equations. Because Newton’s method has second order convergence, we may expect Newton’s method, if successful, to be more robust to the specification of the convergence threshold.

In the generalized leapfrog integrator, we may choose to apply Newton’s method to solve the implicit update to position (eq. (5.2)) or to position (eq. (5.3)) or both. As discussed in section 5.A.7.6, the implicit updates for position differs fundamentally from the implicit equation to update momentum in the generalized leapfrog integrator: the update
Figure 5.11: We illustrate the dependency of the Jacobian determinant metric on the procedure used to resolve the implicit updates to momentum and position. Employing Newton’s method to update both the position and the momentum variables in the generalized leapfrog integrator yields faster convergence toward a volume-preserving proposal. We note that in the banana-shaped distribution, using Newton’s method to update both momentum and position has severe errors in volume preservation; see our discussion in the text and section 5.A.9 for an explanation.

to momentum does not require recomputing quantities of the posterior. In contrast, the implicit update to the position must recompute the Riemannian metric at each iteration. Therefore, to the extent that posterior quantities are the most expensive computations in RMHMC, the momentum fixed point iteration can be resolved relatively cheaply. If the Jacobian of the fixed point equation can be found, then it is conceivable to solve these implicit relationships via Newton’s method and enjoy a faster order of convergence than fixed point iteration. Hence, we consider three variations of RMHMC. The first uses fixed point iterations in order to resolve both of the implicit updates to the momentum variable. The second method updates the momentum variable using Newton iterations and employs fixed point iterations for the implicit update to position. The third variation uses Newton’s method for both implicit updates. Equations for the required Jacobians for either the
implicit update to the momentum or position are given in eq. (5.A.90) and eq. (5.A.94), respectively.

We now investigate how the application of Newton’s method influences the errors in reversibility and volume preservation in RMHMC. In fig. 5.10 we show the reversibility as a function of threshold for the three variations of the RMHMC procedure we consider. In many cases, Newton’s method produces integrators that much more closely respect reversibility of the proposal operator; this is attributable to the second-order convergence properties of Newton’s method. In terms of reversibility, employing Newton’s method on the momentum update only is virtually indistinguishable from an implementation based on fixed point iterations. Similarly, in fig. 5.11 we show how the three variations of RMHMC respect volume preservation. When employing Newton’s method to resolve the implicit updates to momentum and position, the volume preservation is insensitive to the convergence threshold over many orders of magnitude. We note that RMHMC using only fixed point iterations or with Newton’s method only applied to the momentum update are visually indistinguishable in many cases.

In MCMC, we are principally interested in the ergodicity of the Markov chain: that is, our most important metric is the closeness of the Markov chain samples to the target distribution. Additional experiments showing the minimum ESS per second and comparing the number of iterations required to solve the fixed point equations in RMHMC with Newton’s method can be found in section 5.A.8.8. To investigate ergodicity in the context of the three variations of RMHMC, we show in fig. 5.A.26 the Kolmogorov-Smirnov statistic along randomly sampled one-dimensional sub-spaces for the banana-shaped distribution, the FitzHugh-Nagumo posterior, and the multiscale Student-\(t\) distribution. In the case of the FitzHugh-Nagumo posterior or the multiscale Student-\(t\) distribution, the ergodicity of RMHMC is essentially constant as a function of the threshold; this can be constrained with an implementation based on fixed point iterations, for which the convergence threshold will affect ergodicity. In cases wherein Newton’s method produces an integrator that
satisfies the conditions of reversibility and volume preservation, its use can eliminate much of the sensitivity of the RMHMC algorithm to the selection of the convergence tolerance. Beyond ergodicity, we may characterize the variation of RMHMC that employs Newton’s method in terms of time-normalized ESS. In fig. 5.A.24 we show three examples wherein the use of Newton’s method degrades performance as measured by the effective sample size per second. On the other hand, in fig. 5.A.25 we show three case studies where fixed point iterations and Newton’s method produced comparable effective sample sizes per second.

However, a notable exception to the superiority of Newton’s method for both position and momentum is the banana-shaped distribution, for which we see there is a notable proportion of transitions that utterly fail to respect reversibility. In addition, there are a substantial proportion of transitions whose estimated Jacobian determinant is vastly different from unity when employing Newton’s method to update both position and momentum in the banana-shaped distribution. We investigated this phenomenon and observed that Newton’s method was able to identify unstable solutions of the fixed point equation defining the implicit update to the momentum and position in eqs. (5.2) and (5.3); these unstable solutions appear to produce maps for which volume preservation and reversibility are drawn into question by figs. 5.10a and 5.11a. The failure of assumption 5.2.15 and assumption 5.2.16 implies that stationarity will not hold for this Markov chain transition kernel. This is supported in fig. 5.A.26 which reveals severely degraded ergodicity relative to the competing implementations. In section 5.A.9 we investigate this phenomenon further and propose to augment Newton’s method with a “stability check” to ensure that the solutions of fixed point equations obtained by Newton’s method are stable. As shown in fig. 5.A.27a the stability check is shown to eliminate the pathology of unstable fixed point solutions.

These results demonstrate that Newton’s method can eliminate much of the sensitivity of RMHMC to the convergence threshold, both in terms of detailed balance (figs. 5.10

254
and (5.11) and ergodicity (fig. 5.A.26). The results in the banana-shaped distribution show that unstable fixed point solutions can cause a degradation of volume preservation and reversibility, but that implementing a stability check (see section 5.A.9) can ensure that RMHMC with Newton’s method is well-behaved.

5.6 Discussion and Conclusion

This work has examined the role of convergence thresholds in Riemannian manifold Hamiltonian Monte Carlo. The integrators used in implementations of RMHMC depend on non-zero convergence tolerances, which will affect the volume preservations and reversibility properties of the integrator. Because reversibility and volume preservation are critical ingredients in the proofs of detailed balance of HMC, it is necessary to understand how these properties degrade in the presence of non-zero convergence tolerances. We would like to emphasize the following empirical observations from our case studies:

1. The RMHMC algorithms tend to produce samples with better ergodicity and larger effective sample sizes than Euclidean methods, though the $O(m^3)$ computational effort represents a heavy burden. We find that the question of whether or not RMHMC has a better effective sample size per second depends on the convergence threshold being used. Very small thresholds such as $1 \times 10^{-9}$ tend to be associated to the best reversibility and volume preservation, but also the largest number of fixed point iterations at each step of the generalized leapfrog integrator, and therefore largest computation cost. Our measures of ergodicity indicate that these minuscule thresholds are no better than more modest thresholds with smaller computational costs.

2. We observe that the empirical sample quality, as measured by Kolmogorov-Smirnov statistics along random one-dimensional subspaces and related measures, is sensitive to the choice of convergence threshold only to a point. This implies that there exist
diminishing returns to using a convergence threshold beyond some critical value. While divining the value of this critical threshold would seem challenging without the kind of analysis conducted herein, its existence suggests that proper adaptation of the convergence threshold can produce important improvements in computational efficiency.

3. A method based on Ruppert averaging has been proposed for selecting a convergence tolerance based on specifying a desired number of decimal digits of similarity with a numerical integrator with a stringent convergence threshold. This procedure only requires basic arithmetic operations beyond computing the proposal state and acceptance probability of the RMHMC Markov chain with variable convergence thresholds. Therefore, tuning the convergence threshold in this manner is about as fast as running RMHMC itself. If a practitioner can hypothesize a particular minimal scale of the posterior distribution, then this quantity can guide the selection of the convergence tolerance such that differences (on average) between the numerical integrators are beneath the scale of the posterior. By tuning the convergence tolerance of the method, one can obtain improvements in computational expediency without significant detectable differences in either the ergodicity or effective sample size properties of the RMHMC procedure.

4. We find that Newton’s method, consistent with its stronger position as a second-order solver, enjoys more rapid convergence toward numerical reversibility and volume preservation under our measures. We additionally find that, within the scope of convergence thresholds and experimental designs here, the ergodic properties of RMHMC implemented with Newton’s method are less sensitive to the choice of threshold. However, we also illuminated a circumstance in the banana-shaped posterior where Newton’s method actually produced degraded sampling behavior.

5. Violations of detailed balance can emerge from sources other than the convergence
tolerance. Indeed, the correctness of derivatives assumes a special significance in RMHMC, which we investigated in section 5.A.10. In Euclidean HMC, an incorrectly specified derivative can be expected to slow mixing but not to invalidate the correctness of the MCMC procedure itself. This no longer holds in the case of RMHMC, where detailed balance depends on volume preservation and volume preservation depends on the symmetry of partial derivatives. While the use of certain automatic differentiation (AD) tools may alleviate some of these concerns, we make the additional observations that (i) metrics such as the Fisher information (which is the expectation of the negative Hessian of the log-likelihood) may not be readily obtained from AD, and (ii) implementation errors in the AD rules of even the most popular libraries are corrected on a regular basis. At the present time, we suggest that at least some minimal effort be made to ensure that volume preservation holds in a given application.

We hope that this work raises interest and awareness around this aspect of the computation in RMHMC.
Appendix 5.A

Appendices to Chapter 5
5.A.1 Proofs of Reversibility and Volume Preservation of Hamiltonian Mechanics

Lemma 5.A.1.1. Let $q_t$ and $p_t$ be solutions to Hamilton’s equations of motion given in eqs. (5.8) and (5.9). Then $\frac{d}{dt} H(q_t, p_t) = 0$.

Proof.

\[
\frac{d}{dt} H(q_t, p_t) = \nabla_q H(q_t, p_t) \cdot \nabla_p H(q_t, p_t) - \nabla_p H(q_t, p_t) \cdot \nabla_q H(q_t, p_t) \quad (5.A.1)
\]

\[= 0. \quad (5.A.2)\]

Lemma 5.A.1.2. Let $q_t$ and $p_t$ be solutions to Hamilton’s equations of motion given in eqs. (5.8) and (5.9). Then $\text{div}((\dot{q}_t, \dot{p}_t)) = 0$.

Proof.

\[
\text{div}((\dot{q}_t, \dot{p}_t)) = \nabla_q \cdot \dot{q}_t + \nabla_p \dot{p}_t \quad (5.A.3)
\]

\[= \nabla_q \cdot \nabla_p H(q_t, p_t) - \nabla_p \cdot \nabla_q H(q_t, p_t) \quad (5.A.4)
\]

\[= \sum_{i=1}^{m} \frac{\partial^2 H}{\partial p_i \partial q_i}(q_t, p_t) - \sum_{i=1}^{m} \frac{\partial^2 H}{\partial q_i \partial p_i}(q_t, p_t) \quad (5.A.5)
\]

\[= 0. \quad (5.A.6)\]

Lemma 5.A.1.3. Let $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be a smooth Hamiltonian satisfying assumption 5.2.6 Let $(q_t, p_t)$ be solutions to Hamilton’s equations of motion given in eqs. (5.8) and (5.9). Then, for fixed $\tau \in \mathbb{R}$, the trajectories $\tilde{q}_s = q_{\tau-s}$ and $\tilde{p}_s = -p_{\tau-s}$ are also solutions to Hamilton’s equations of motion.
Proof. By direct calculation,

\[
\frac{d}{ds} \tilde{q}_s = -\dot{q}_{\tau-s} \\
= -\nabla_p H(q_{\tau-s}, p_{\tau-s}) \tag{5.A.7}
\]

\[
= -\nabla_p H(\tilde{q}_s, -\tilde{p}_s) \tag{5.A.8}
\]

\[
= \nabla_p H(\tilde{q}_s, \tilde{p}_s) \tag{5.A.9}
\]

and

\[
\frac{d}{ds} \tilde{p}_s = \dot{p}_{\tau-s} \\
= -\nabla_q H(q_{\tau-s}, p_{\tau-s}) \tag{5.A.11}
\]

\[
= -\nabla_q H(\tilde{q}_s, \tilde{p}_s). \tag{5.A.12}
\]

\[
\nabla_p H(\tilde{q}_s, \tilde{p}_s)
\]
5.A.2 Proofs of Reversibility and Volume Preservation of the Generalized Leapfrog Integrator

Proposition 5.A.2.1. Let $H: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ be a smooth Hamiltonian satisfying assumption 5.2.6. The generalized leapfrog integrator is reversible under negation of the momentum variable.

Proof. Given an initial position in phase-space $(q_n, p_n)$, the generalized leapfrog integrator computes the following series of updates:

\begin{align}
    p_{n+1/2} &= p_n - \frac{\epsilon}{2} \nabla_q H(q_n, p_{n+1/2}) \tag{5.A.14} \\
    q_{n+1} &= q_n + \frac{\epsilon}{2} \left( \nabla_p H(q_n, p_{n+1/2}) + \nabla_p H(q_{n+1}, p_{n+1/2}) \right) \tag{5.A.15} \\
    p_{n+1} &= p_{n+1/2} - \frac{\epsilon}{2} \nabla_q H(q_{n+1}, p_{n+1/2}). \tag{5.A.16}
\end{align}

Now we apply the integrator a second time from the initial position $(q_{n+1}, -p_{n+1})$. This produces the following update to the momentum:

\begin{equation}
    p_{n+1+1/2} = -p_{n+1} - \frac{\epsilon}{2} \nabla_q H(q_{n+1}, p_{n+1+1/2}) \tag{5.A.17}
\end{equation}

But by rearranging eq. (5.A.16) and using the condition $\nabla_q H(q, p) = \nabla_q H(q, -p)$ we have,

\begin{equation}
    -p_{n+1/2} = -p_{n+1} - \frac{\epsilon}{2} \nabla_q H(q_{n+1}, -p_{n+1/2}), \tag{5.A.18}
\end{equation}

so we obtain $p_{n+1+1/2} = -p_{n+1}$. Now we compute the update to the momentum variable
and use the fact that $-\nabla_p H(q, -p) = \nabla_p H(q, p)$ to obtain,

\begin{align}
q_{n+2} &= q_{n+1} + \frac{\epsilon}{2} \left( \nabla_p H(q_{n+1}, -p_{n+1/2}) + \nabla_p H(q_{n+2}, -p_{n+1/2}) \right) \quad (5.A.19) \\
&= q_{n+1} - \frac{\epsilon}{2} \left( \nabla_p H(q_{n+1}, p_{n+1/2}) + \nabla_p H(q_{n+2}, p_{n+1/2}) \right). \quad (5.A.20)
\end{align}

By rearranging eq. (5.A.15), we see that $q_{n+2} = q_n$ solves this implicit relation. Finally we compute the explicit update to the momentum variable,

\begin{align}
p_{n+2} &= -p_{n+1/2} - \frac{\epsilon}{2} \nabla_q H(q_n, -p_{n+1/2}) \quad (5.A.21) \\
&= -p_{n+1/2} - \frac{\epsilon}{2} \nabla_q H(q_n, p_{n+1/2}) \quad (5.A.22) \\
&= -p_n \quad (5.A.23)
\end{align}

by rearranging eq. (5.A.14). \hfill \square

**Proposition 5.A.2.2.** The generalized leapfrog integrator is volume preserving.

**Proof.** One can prove that the generalized leapfrog integrator is volume preserving by employing an analysis of differential forms. We refer the interested reader to [Leimkuhler and Reich 2005](#) for an introduction to this technique. The first step is to compute the coordinate 1-forms of the integrator.

\begin{align}
\text{d}p_{n+1/2} &= \text{d}p_n - \frac{\epsilon}{2} \nabla_q \nabla_q H(q_n, p_{n+1/2}) \text{ d}q_n - \frac{\epsilon}{2} \nabla_p \nabla_q H(q_n, p_{n+1/2}) \text{ d}p_{n+1/2} \quad (5.A.24) \\
\text{d}q_{n+1} &= \text{d}q_n + \frac{\epsilon}{2} \left( \nabla_q \nabla_p H(q_n, p_{n+1/2}) \text{ d}q_n + \nabla_q \nabla_q H(q_{n+1}, p_{n+1/2}) \text{ d}q_{n+1} \right) + (** \text{ d}p_{n+1/2} \quad (5.A.25) \\
\text{d}p_{n+1} &= \text{d}p_{n+1/2} - \frac{\epsilon}{2} \nabla_q \nabla_q H(q_{n+1}, p_{n+1/2}) \text{ d}q_{n+1} - \frac{\epsilon}{2} \nabla_p \nabla_q H(q_{n+1}, p_{n+1/2}) \text{ d}p_{n+1/2} \quad (5.A.26)
\end{align}
Next we use the coordinate 1-forms to compute wedge products:

\[ dq_n \wedge dp_{n+1/2} = dq_n \wedge dp_n - \frac{\epsilon}{2} dq_n \wedge \nabla_p \nabla_q H(q_n, p_{n+1/2}) \, dp_{n+1/2} \]  
\[ dq_{n+1} \wedge dp_{n+1/2} = dq_n \wedge dp_{n+1/2} + \frac{\epsilon}{2} \nabla_q \nabla_p H(q_n, p_{n+1/2}) \, dq_n \wedge dp_{n+1/2} \]  
\[ + \frac{\epsilon}{2} \nabla_q \nabla_p H(q_{n+1}, p_{n+1/2}) \, dq_{n+1} \wedge dp_{n+1/2} \]  
\[ \text{(5.A.27)} \]

\[ dq_{n+1} \wedge dp_{n+1/2} = dq_n \wedge dp_n + \frac{\epsilon}{2} \nabla_q \nabla_p H(q_{n+1}, p_{n+1/2}) \, dq_{n+1} \wedge dp_{n+1/2} \]  
\[ \text{(5.A.28)} \]

\[ dq_{n+1} \wedge dp_{n+1} = dq_{n+1} \wedge dp_{n+1/2} - dq_{n+1} \wedge \frac{\epsilon}{2} \nabla_p \nabla_q H(q_{n+1}, p_{n+1/2}) \, dp_{n+1/2} \]  
\[ \text{(5.A.29)} \]

\[ dq_{n+1} \wedge dp_{n+1} = dq_n \wedge dp_n, \]  
\[ \text{(5.A.30)} \]

which establishes volume preservation. (Actually, this proof establishes an even stronger property known as *symplecticness*, which is important in the theory of Hamiltonian mechanics, but which we have eschewed in our discussion.) The fact that

\[ dq_{n+1} \wedge \nabla_p \nabla_q H(q_{n+1}, p_{n+1/2}) \, dp_{n+1/2} = \nabla_q \nabla_p H(q_{n+1}, p_{n+1/2}) \, dq_{n+1} \wedge dp_{n+1/2} \]  
\[ \text{(5.A.32)} \]

is used in eq. (5.A.31), which is predicated on the symmetry of partial derivatives. One therefore sees that one has required the symmetry of partial derivatives in order to establish the volume-preservation property. One concludes, therefore, that one cannot apply the generalized leapfrog method to *arbitrary* vector fields in position and momentum and expect to obtain a volume-preserving numerical integrator. \( \square \)

We briefly remark on the special case when \( \frac{\partial}{\partial q_i} G(q) = 0 \) for \( i = 1, \ldots, m \) in Hamilton-
nians of the form in eq. (5.1). In this case, the implicit update to the momentum becomes,

\[
p_{n+1/2} = p_n - \frac{\epsilon}{2} \left[ -\nabla \mathcal{L}(q_n) + \frac{1}{2} \text{trace} \left( G^{-1}(q_n) \nabla G(q_n) \right) + \frac{1}{2} p_{n+1/2} G^{-1}(q_n) \nabla G(q_n) G^{-1}(q_n) p_{n+1/2} \right]
\]

(5.A.33)

\[
= p_n + \frac{\epsilon}{2} \nabla \mathcal{L}(q_n)
\]

(5.A.34)

so that the first update to momentum becomes explicit. Similarly, the implicit update to position is,

\[
q_{n+1} = q_n + \frac{\epsilon}{2} \left[ G^{-1}(q_n) p_{n+1/2} + G^{-1}(q_{n+1}) p_{n+1/2} \right]
\]

(5.A.35)

\[
= q_n + \frac{\epsilon}{2} \left[ G^{-1} p_{n+1/2} + G^{-1} p_{n+1/2} \right]
\]

(5.A.36)

\[
= q_n + \epsilon G^{-1} p_{n+1/2}
\]

(5.A.37)

so that the update to position also becomes explicit in this case. In fact, these simplifications show that the generalized leapfrog method reduces to the leapfrog method in algorithm 7 when the metric is constant. This leads us to conclude that when the metric is constant, the convergence tolerance becomes irrelevant. Intuitively, in cases wherein the metric is “close to constant” the convergence tolerance will not have a large effect on the ergodicity of the algorithm.
5.A.3 Establishing Detailed Balance in Hamiltonian Monte Carlo

Theorem 5.A.3.1. Let $\Phi$ be a self-inverse and volume-preserving map from $\mathbb{R}^{2m}$ to itself. Consider a probability distribution whose density is proportional to $\exp(-H(q,p))$ where $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is a smooth function. Consider the Markov chain transition kernel defined as,

$$K((q,p), B) = \min \left\{ 1, \exp(H(q,p) - H(\Phi(q,p))) \right\} \cdot 1 \{ \Phi(q,p) \in B \}$$

$$+ \left( 1 - \min \left\{ 1, \exp(H(q,p) - H(\Phi(q,p))) \right\} \right) \cdot 1 \{ (q,p) \in B \},$$

(5.A.38)

where $B$ is a Borel subset of $\mathbb{R}^{2m}$. Then the stationary distribution of $K$ is the distribution with density $\pi(q,p) \propto \exp(-H(q,p))$.

Proof. The following proof is similar to that in [Brofos and Lederman 2021a]. It suffices to show that the transition kernel satisfies detailed balance with respect to the density $\pi$. Let $A$ and $B$ be Borel subsets of $\mathbb{R}^{2m}$. Let $z$ be a random variable drawn from the
distribution with density $\pi(z)$ and suppose that $z'|z \sim K(z, \cdot)$. Then,

$$\Pr [z \in A \text{ and } z' \in B] = \int_A \pi(z) \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} 1 \{ \Phi(z) \in B \} \, dz$$

$$+ \int_A \pi(z) \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \right) 1 \{ z \in B \} \, dz$$

$$(5.A.39)$$

$$=$$  

$$\int_{A \cap \Phi(B)} \pi(z) \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \, dz$$

$$+ \int_B \pi(z) \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \right) 1 \{ z \in A \} \, dz$$

$$(5.A.40)$$

$$=$$  

$$\int_{\Phi(A) \cap B} \pi(\Phi(z')) \min \left\{ 1, \frac{\pi(z')}{\pi(\Phi(z'))} \right\} \, dz'$$

$$+ \int_B \pi(z) \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \right) 1 \{ z \in A \} \, dz$$

$$(5.A.41)$$

$$=$$  

$$\int_{\Phi(A) \cap B} \pi(z') \min \left\{ 1, \frac{\pi(\Phi(z'))}{\pi(z')} \right\} \, dz'$$

$$+ \int_B \pi(z) \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \right) 1 \{ z \in A \} \, dz$$

$$(5.A.42)$$

$$=$$  

$$\int_B 1 \{ z' \in \Phi(A) \} \pi(z') \min \left\{ 1, \frac{\pi(\Phi(z'))}{\pi(z')} \right\} \, dz'$$

$$+ \int_B \pi(z) \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \right) 1 \{ z \in A \} \, dz$$

$$(5.A.43)$$

$$=$$  

$$\Pr [z \in B \text{ and } z' \in A]$$

$$(5.A.44)$$

The fact that distribution whose density is $\pi$ is the stationary distribution of such a Markov chain then follows immediately from the choice $A = \mathbb{R}^{2m}$; mathematically,

$$\Pr [z \in \mathbb{R}^{2m} \text{ and } z' \in B] = \Pr [z' \in B] = \Pr [z \in B].$$

$$(5.A.45)$$
The fact that $\Phi$ is a self-inverse function with unit Jacobian determinant is used in the change-of-variables in eq. (5.A.41). Thus, we see how these two properties play a critical role in establishing detailed balance of the HMC algorithm.
5.A.4 Computational Complexity of Implicit Updates

The generalized leapfrog integrator involves two fixed point iterations. The first of these, given in eq. (5.2), provides a half-step update to the momentum variable. A principle advantage of the generalized leapfrog method is that it facilitates caching of reusable quantities from iteration to iteration. In the context of the first momentum update, the computation of the gradient of the log-posterior, the Riemannian metric and its inverse, and the Jacobian of the Riemannian metric can be cached from the explicit update to the momentum in eq. (5.4); this averts recomputing these quantities within each fixed-point iteration. However, at iteration $n$, if we define $v_{n+1/2} = G(q_n)^{-1}p_{n+1/2}$ then the quantity $v_{n+1/2}^T \frac{\partial G(q_n)}{\partial q_i} v_{n+1/2}$ incurs a computational cost $O(m^2)$. Since this computation is replicated for $i = 1, \ldots, m$, the total computational complexity to update the momentum is $O(m^3)$.

The second fixed-point iteration provides, given in eq. (5.3), a full-step update to the position variable. At iteration $n$, $v_{n+1/2} = G(q_n)^{-1}p_{n+1/2}$ may be cached but $G(q_{n+1})^{-1}p_{n+1/2}$ requires that the Riemannian metric, and its inverse, be recomputed at each iteration. Because matrix inversion scales as $O(m^3)$, this position update shares the same overall computational complexity with the fixed point update to the momentum. Although by this analysis both implicitly defined updates scale as $O(m^3)$, the update to position differs fundamentally in the sense that it must compute new quantities from the posterior at each fixed point iteration; namely, the Riemannian metric must be recomputed for every new candidate solution to the fixed point equation to update the position variable. As a practical matter, the computational effort required to compute the Riemannian metric may be more substantial than that required to invert it, such as in the case of section 5.5.2 wherein the Riemannian metric takes values in $\mathbb{R}^{3 \times 3}$ but whose computation requires the solution to an eight-dimensional initial value problem at two-hundred predetermined locations in
time. Therefore, in these cases we expect the implicit update to the position to be the more costly of the two.

The final step of the generalized leapfrog integrator requires us to compute the gradient of the log-posterior, the Riemannian metric and its inverse, and the Jacobian of the Riemannian metric in order to give an explicit half-step update to the momentum variable. We measure the total computational complexity of the generalized leapfrog integrator by calculating the number of fixed point iterations used in updating the momentum and position variables. The number of fixed point iterations depends on the convergence threshold, thereby revealing that the computational efficiency of the RMHMC algorithm will depend on the threshold.
Algorithm 7 The leapfrog integrator, which is a volume-preserving, reversible, second-order accurate numerical integrator for separable Hamiltonians.

Input: An initial state \((q_n, p_n) \in \mathbb{R}^m \times \mathbb{R}^m\), a smooth, separable Hamiltonian \(H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}\), a step-size \(\epsilon \in \mathbb{R}\).

Compute an explicit half-step in the momentum variable,

\[
P_{n+1/2} = p_n - \frac{\epsilon}{2} \nabla_q H(q_n, p_n). \tag{5.A.46}
\]

Compute an explicit full step in the position variable,

\[
q_{n+1} = q_n + \epsilon \nabla_p H(q_n, p_{n+1/2}). \tag{5.A.47}
\]

Apply an explicit half-step in the momentum variable,

\[
P_{n+1} = P_{n+1/2} - \frac{\epsilon}{2} \nabla_q H(q_{n+1}, p_{n+1/2}). \tag{5.A.48}
\]

Return: \((q_{n+1}, p_{n+1})\), the next position in phase-space along the integrated trajectory.

5.A.5 Algorithms

In this appendix, we give algorithmic implementations of some of the techniques we have described. Algorithm 7 shows the idealized generalized leapfrog method (Definition 5.2.2). A practical implementation of this method will require a subroutine to find solutions to the fixed point equations defining this integrator. Therefore, an algorithm implementing fixed point iterations (Definition 5.2.22) is provided in Algorithm 8. Similarly, an algorithm for Newton’s method (Definition 5.2.23) is given in Algorithm 9. With these methods of solving fixed point equations, we present an algorithm for implementing the generalized leapfrog integrator (Definition 5.2.4) in Algorithm 10. Using the generalized leapfrog integrator, we provide an algorithm for Riemannian manifold Hamiltonian Monte Carlo in Algorithm 11. In Algorithm 12 we provide an algorithm implementing the approach of Section 5.4.2 for identifying a threshold that produces a desired number of decimal digits of similarity between two integrators.
Algorithm 8 Algorithmic implementation of employing fixed point iterations to solve a fixed point equation.

**Input:** A function \( g : \mathbb{R}^m \to \mathbb{R}^m \) whose fixed point is sought, a convergence tolerance \( \delta > 0 \), an initial guess for the fixed point solution \( z_0 \).

Set \( \Delta \leftarrow \infty \).

Set \( n \leftarrow 0 \).

**while** \( \Delta > \delta \) **do**

- Iterate the fixed point and set \( z_{n+1} \leftarrow g(z_n) \).
- Measure convergence and set \( \Delta \leftarrow \| z_{n+1} - z_n \|_\infty \).
- Update iteration \( n \leftarrow n + 1 \).

**end while**

Set \( z_* \leftarrow z_n \).

**Return:** The approximate fixed point solution \( z_* \).

Algorithm 9 Algorithmic implementation of employing Newton’s method to solve a fixed point equation.

**Input:** A smooth function \( g : \mathbb{R}^m \to \mathbb{R}^m \) whose fixed point is sought, a convergence tolerance \( \delta > 0 \), an initial guess for the fixed point solution \( z_0 \).

Set \( \Delta \leftarrow \infty \).

Set \( n \leftarrow 0 \).

**while** \( \Delta > \delta \) **do**

- Apply the Newton update and set \( z_{n+1} \leftarrow z_n - (\text{Id} + \nabla g(z_n))^{-1}g(z_n) \).
- Measure convergence and set \( \Delta \leftarrow \| z_{n+1} - z_n \|_\infty \).
- Update iteration \( n \leftarrow n + 1 \).

**end while**

Set \( z_* \leftarrow z_n \).

**Return:** The approximate fixed point solution \( z_* \).

Algorithm 10 Implementation of the generalized leapfrog integrator using either fixed point iteration or Newton’s method to resolve the fixed point equations.

**Input:** An initial state \( (q, p) \in \mathbb{R}^m \times \mathbb{R}^m \), a smooth Hamiltonian \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \), a step-size \( \epsilon \in \mathbb{R} \), a convergence threshold \( \delta > 0 \).

Define the function \( g_p(\tilde{p}) = p - \frac{\epsilon}{2} \nabla_q H(q, \tilde{p}) \).

Solve the fixed point equation \( \tilde{p} = g_p(\tilde{p}) \) using algorithm 8 or algorithm 9 to tolerance \( \delta \) with initial guess \( \tilde{p}_0 = p - \frac{\epsilon}{2} \nabla_q H(q, p) \). Denote the approximate solution by \( \tilde{p}_* \).

Define the function \( g_q(\tilde{q}) = q + \frac{\epsilon}{2}(\nabla_p H(q, \tilde{p}_*) + \nabla_p H(\tilde{q}, \tilde{p}_*)) \).

Solve the fixed point equation \( \tilde{q} = g_q(\tilde{q}) \) using algorithm 8 or algorithm 9 to tolerance \( \delta \) with initial guess \( \tilde{q}_0 = q + \epsilon \nabla_p H(q, \tilde{p}_*) \). Denote the approximate solution by \( \tilde{q}_* \).

Compute \( \tilde{p}_* \) using eq. (5.7).

**Return:** \( (\tilde{q}_*, \tilde{p}_*) \in \mathbb{R}^m \times \mathbb{R}^m \), the next position in phase-space along the integrated trajectory.
Algorithm 11 Algorithm implementing the Riemannian manifold Hamiltonian Monte Carlo (RMHMC) Markov chain.

**Input:** The state $q_n \in \mathbb{R}^m$ of the Markov chain at the $n$-th step, a smooth log-density $\mathcal{L} : \mathbb{R}^m \to \mathbb{R}$, a Riemannian metric $G : \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m$, an integration step-size $\epsilon \in \mathbb{R}$ and a number of integration steps $k \in \mathbb{N}$, a convergence tolerance $\delta > 0$.

Sample $p_n \sim \text{Normal}(0, G(q_n))$.

Define the function $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ as given in eq. (5.1).

Set $q \leftarrow q_n$ and $p \leftarrow p_n$.

for $i = 1, \ldots, k$ do

Apply algorithm 10 with initial state $(q, p)$, smooth Hamiltonian $H$, step-size $\epsilon$, and convergence tolerance $\delta$. Denote the output by $(\tilde{q}, \tilde{p})$.

Set $q \leftarrow \tilde{q}$ and $p \leftarrow \tilde{p}$.

end for

Set $\tilde{q}_{n+1} \leftarrow q$ and $\tilde{p}_{n+1} \leftarrow p$.

Apply the momentum flip operator to produce a self-inverse proposal and set $\tilde{p}_{n+1} \leftarrow -\tilde{p}_{n+1}$.

Set $M \leftarrow \min \{1, H(q_n, p_n) - H(\tilde{q}_{n+1}, \tilde{p}_{n+1})\}$.

Sample $u \sim \text{Uniform}(0, 1)$.

if $u < M$ then

Accept the proposal and set $(q_{n+1}, p_{n+1}) \leftarrow (\tilde{q}_{n+1}, \tilde{p}_{n+1})$.

else

Reject the proposal and set $(q_{n+1}, p_{n+1}) \leftarrow (q_n, p_n)$.

end if

Return: The next state of the Markov chain $q_{n+1}$.

5.A.6 Posterior Summary
Algorithm 12 Implementation of Ruppert averaging to identify a threshold that produces a desired number of decimal digits of similarity between two numerical integrators on average.

**Input:** Constants $D \in \mathbb{R}_+$ and $\omega \in (1/2, 1)$, initial log$_{10}$-threshold $\xi_0$, comparison convergence threshold $\delta'$, initial Markov chain state $q_0 \in \mathbb{R}^m$, a smooth log-density $\mathcal{L} : \mathbb{R}^m \to \mathbb{R}$, a Riemannian metric $G : \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m$, an integration step-size $\epsilon \in \mathbb{R}$ and a number of integration steps $k \in \mathbb{N}$, desired negative number of decimal digits of similarity $\kappa$, maximum number of iterations $n \in \mathbb{N}$.

Define the function $B(\xi)$ as in eq. (5.24).

Set $\xi_0 \leftarrow 0$.

for $k = 1, \ldots, n$ do

Apply algorithm 11 with state $q_{k-1}$, smooth log-density $\mathcal{L}$, Riemannian metric $G$, integration step-size $\epsilon \in \mathbb{R}$ and number of integration steps $k \in \mathbb{N}$, and convergence tolerance $\delta'$. Denote the next state of the Markov chain $q_k$.

Sample $p_k \sim \text{Normal}(0, G(q_k))$.

Set $\gamma_k \leftarrow Dk^{-\omega}$.

Compute $L_\xi(q_k, p_k; 10^{\xi_{k-1}}, \delta')$ using eq. (5.23).

Compute $\xi_k$ using eq. (5.20).

Compute $\bar{\xi}_k$ using eq. (5.21).

end for

Set $\delta^* \leftarrow 10^{\bar{\xi}_n}$

**Returns:** The threshold $\delta^*$ that produces the desired number of decimal digits of similarity.

<table>
<thead>
<tr>
<th>Posterior</th>
<th># Dimensions</th>
<th># Samples</th>
<th>Metric</th>
<th>Hierarchical</th>
<th>Exact Gradients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banana</td>
<td>2</td>
<td>1,000,000</td>
<td>Fisher Information + Hessian</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>14 + 1</td>
<td>100,000</td>
<td>Fisher Information + Hessian</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Neal’s Funnel</td>
<td>11</td>
<td>1,000,000</td>
<td>SoftAbs Metric</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Stochastic Volatility</td>
<td>1000 + 3</td>
<td>100,000</td>
<td>Fisher Information + Hessian</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Log-Gaussian Cox-Poisson</td>
<td>1024 + 2</td>
<td>5,000</td>
<td>Fisher Information + Hessian</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Fitzhugh-Nagumo</td>
<td>3</td>
<td>100,000</td>
<td>Fisher Information + Hessian</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Multivariate Student-t</td>
<td>20</td>
<td>1,000,000</td>
<td>Positive Definite Part of Hessian</td>
<td>✗</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 5.A.1: Summary of the posterior distributions we examine in our experimental results. For the hierarchical models, we employ an alternating Metropolis-within-Gibbs-like strategy; we indicate the dimensionality of each alternating component by separating them with a plus-sign. The Neal Funnel distribution has a hierarchical structure but it is sampled in a non-hierarchical manner. As discussed in detail later, the Fitzhugh-Nagumo model computes approximate gradients by solving initial value problems.
5.A.7 Detailed Measures and Metrics

Here we describe our procedures for measuring quantities of interest in assessing the dependence of the RMHMC algorithm on the threshold. We develop metrics for measuring reversibility and volume-preservation of the proposal operator, ergodicity, similarity between transition kernels, and the computational effort expended by the proposal operator. In measuring reversibility and volume-preservation, we modify the technique of Brofos and Lederman [2021a]. We emphasize that by measuring reversibility and volume-preservation, we are not measuring detailed balance. Instead, we are measuring two properties which, taken together, imply detailed balance and therefore stationarity of the RMHMC Markov chain. We carry out all computation in 64-bit precision.

5.A.7.1 Measuring Reversibility

As described previously, the easiest mechanism to ensure reversibility of the RMHMC proposal operator is by negating the momentum variable following integration. Therefore, to assess numerical reversibility, we define the momentum flip operator \( F(q, p) = (q, -p) \).

Under reversibility, it follows that \( F \circ \Phi_k^\epsilon \circ F \circ \Phi_k^\epsilon = \text{Id} \) for any \( k \in \mathbb{N} \) and \( \epsilon \in \mathbb{R} \). To measure the degree of reversibility of the numerical integrator with fixed point convergence tolerance \( \delta \), we set, \( (q_f, p_f) = F \circ \Phi_k^\epsilon(q, p; \delta) \) and \( (q_r, p_r) = F \circ \Phi_k^\epsilon(q_f, p_f; \delta) \) and compute the absolute reversibility error (ARE) by,

\[
\text{ARE} = \sqrt{\|q - q_r\|^2 + \|p - p_r\|^2}. \tag{5.A.49}
\]
We additionally consider a normalized version of this absolute error by dividing by $\sqrt{\|q\|^2 + \|p\|^2}$: the relative reversibility error (RRE) is

$$RRE = \frac{ARE}{\sqrt{\|q\|^2 + \|p\|^2}}. \quad (5.A.50)$$

The results of the relative error analysis are shown in section 5.A.12.

5.A.7.2 Measuring Volume Preservation

The (non-generalized) leapfrog integrator is always a volume-preserving transformation because its three steps consist only of translations of the position (resp. momentum) by quantities depending only on the momentum (resp. position). For volume-preservation property of the generalized leapfrog method is complicated by the implicit relations used to define the integrator. Nevertheless, viewing the concatenation of the momentum flip operator and the generalized leapfrog integrator as a map from $\mathbb{R}^{2m}$ to itself, we recall that a necessary and sufficient condition for a smooth map to be volume-preserving is that it has unit Jacobian determinant. Writing $z = (q, p)$ and letting $e_i$ be the $i^{th}$ standard basis vector of $\mathbb{R}^{2m}$ we approximate the Jacobian of the generalized leapfrog integrator by the central difference formula whose $i^{th}$ column is given by,

$$\tilde{J}_i = \frac{\Phi_k(z + \omega e_i/2; \delta) - \Phi_k(z - \omega e_i/2; \delta)}{\omega}. \quad (5.A.51)$$

where $\omega > 0$ is the finite difference perturbation size. Forming the approximate Jacobian $\tilde{J}$ in this way, we compute its determinant and compare its absolute difference against unity in order to obtain a measure of volume preservation: the absolute volume preservation error (AVPE) is

$$AVPE = \left| \det(\tilde{J}) - 1 \right|. \quad (5.A.52)$$
The use of a finite difference approximation in computing the Jacobian will produce round-off and truncation errors that prevent perfect estimation of the true Jacobian. Therefore, we measure the true violation of volume preservation with error. In our experiments we search over values of \( \omega \) from the set \( \{1 \times 10^{-8}, 1 \times 10^{-7}, 1 \times 10^{-6}, 1 \times 10^{-5}, 1 \times 10^{-4}, 1 \times 10^{-3}\} \) and report the volume preservation results for the value of \( \omega \) that produces estimates of the Jacobian determinant that are closest to zero when using a convergence of \( 1 \times 10^{-9} \). Further details on the sensitivity of these estimates are given in section 5.A.11.

One could also choose to examine the error in volume preservation in a relative sense. However, since the analytical Jacobian determinant is equal to one, the relative error in volume preservation is eq. (5.A.52) divided by one. Hence we see that, in the case of the volume preservation metric, the absolute error and relative error coincide.

5.A.7.3 Measuring Ergodicity

Ergodicity of a Markov chain refers to the property that, from any initial condition, the state of the Markov chain is, asymptotically, distributed according to the target distribution. In practice, of course, we cannot actually assess this asymptotic behavior because the chain must be stopped at some finite, but large, number of steps. However, in cases where it is possible to generate i.i.d. samples from the posterior, we can assess the similarity of the Markov chain samples and analytic samples drawn from the target distribution. Moreover, in these cases we can initialize the Markov chain in stationary distribution by drawing an initial state from the target distribution. Therefore, in order to assess the ergodicity of RMHMC, we propose to compare the samples produced by RMHMC against a collection of i.i.d. samples drawn from the target distribution. In the case of the banana-shaped distribution and the Fitzhugh-Nagumo model, which are of dimension two and three, respectively, it is feasible to generate i.i.d. samples via rejection sampling. For Neal’s funnel distribution and for the multi-scale Student-\( t \) distribution, we can generate
i.i.d. samples analytically. Formally, let \((q_1, \ldots, q_n)\) be a collection of \(m\)-dimensional samples produced by a Markov chain and let \((q'_1, \ldots, q'_{n'})\) be a collection of i.i.d. samples from an \(m\)-dimensional target distribution.

**Cramér-Wold Methods** To assess the ergodicity of the RMHMC samplers of varying thresholds, we draw on the Cramér-Wold theorem [Billingsley, 1986] and the Kolmogorov-Smirnov statistic for inspiration.

**Theorem 5.A.7.1** (Cramér-Wold). A density function in \(\mathbb{R}^m\) is determined by the its projection onto all the one-dimensional subspaces of \(\mathbb{R}^m\).

**Definition 5.A.7.2** (Kolmogorov-Smirnov Statistic). Given two sets of data \((\omega_1, \ldots, \omega_n)\) and \((\omega'_1, \ldots, \omega'_{n'})\) in \(\mathbb{R}\), denote their empirical cumulative distribution functions by \(\hat{F}\) and \(\hat{F}'\), respectively. The Kolmogorov-Smirnov statistic is

\[
KS(\hat{F}, \hat{F}') = \sup_{x \in \mathbb{R}} |\hat{F}(x) - \hat{F}'(x)|.
\]

Therefore, we propose the measure the ergodicity of the sampler by comparing the Markov chain samples to the analytic samples across many random one-dimensional projections. For as many iterations as desired, compute the following:

1. Sample \(u \sim \text{Uniform}(S^{m-1})\).

2. Compute the orthogonal projections onto the vector space spanned by \(u\); that is, \(\omega_i = u^\top q_i \in \mathbb{R}\) for \(i = 1, \ldots, n\) and \(\omega'_i = u^\top q'_i \in \mathbb{R}\) for \(i = 1, \ldots, n'\).

3. Compute the Kolmogorov-Smirnov statistic (the maximal absolute difference in the cumulative distribution functions) between \((\omega_1, \ldots, \omega_n)\) and \((\omega'_1, \ldots, \omega'_{n'})\).

A procedure similar to the one described above was previously advocated in Cuesta-Albertos et al. [2006] for the purposes of crafting a two-sample test for equality of distributions. By constructing a histogram of these Kolmogorov-Smirnov statistics across
numerous random directions, one obtains a quantitative measure of the closeness of the distribution of the Markov chain samples and the i.i.d. samples. We note that we do not attempt to compute a $p$-value associated to these Kolmogorov-Smirnov statistics due to the serial auto-correlation of the Markov chain samples, which would invalidate any independence assumption.

**Maximum Mean Discrepancy** We consider the method of maximum mean discrepancy developed in [Gretton et al.]\(^{[2012]}\). Given two random variables $x \sim \pi$ and $y \sim \tilde{\pi}$, the maximum mean discrepancy is defined by,

$$\text{MMD}[^{\mathcal{F}, \pi, \tilde{\pi}}] \overset{\text{def}}{=} \sup_{\phi \in \mathcal{F}} \left( \mathbb{E}_{x \sim \pi} \phi(x) - \mathbb{E}_{y \sim \tilde{\pi}} \phi(y) \right), \quad (5.A.54)$$

where $\mathcal{F}$ is a prescribed set of $\mathbb{R}$-valued functions. Given a reproducing kernel Hilbert space (RKHS) $\mathcal{H}$ with kernel $k$, [Gretton et al.]\(^{[2012]}\) showed that the squared maximum mean discrepancy enjoys the following characterization,

$$\text{MMD}^2[^{\mathcal{F}, \pi, \tilde{\pi}}] = \mathbb{E}_{x,x' \sim \pi} k(x, x') + \mathbb{E}_{y,y' \sim \tilde{\pi}} k(y, y') - 2 \mathbb{E}_{x \sim \pi, y \sim \tilde{\pi}} k(x, y). \quad (5.A.55)$$

when $\mathcal{F}$ is the set of functions in the unit ball of the RKHS: $\mathcal{F} = \{ \phi \in \mathcal{H} : \| \phi \|_H \leq 1 \}$. Under suitable conditions on the RKHS, it has been shown that $\pi = \tilde{\pi} \iff \text{MMD}[^{\mathcal{F}, \pi, \tilde{\pi}}] = 0$. We use the following unbiased estimator of the squared maximum mean discrepancy as a measure of similarity between the Monte Carlo and analytical samples,

$$\text{MMD}^2_a[^{\mathcal{F}, \pi, \tilde{\pi}}] = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i}^{n} k(q_i, q_j) + \frac{1}{n'(n'-1)} \sum_{i'=1}^{n'} \sum_{j \neq i'}^{n'} k(q'_{i'}, q'_{j'}) - \frac{2}{nn'} \sum_{i=1}^{n} \sum_{j=1}^{n'} k(q_i, q'_{j'}). \quad (5.A.56)$$

As a measure of ergodicity, when sampling has been effective we expect that MMD will be close to zero; since the estimator we employ is unbiased, this can produce a nega-
tive quantity. In our experiments, we report $|\text{MMD}_u^2[\mathcal{F}, \pi, \tilde{\pi}]|$ so as to avoid presenting negative values of what ought to be a non-negative quantity. We use the squared exponential kernel $k(q, q') = \exp(-\|q - q'\|^2_2/h^2)$ where $h$ is the kernel bandwidth. We use the median distance heuristic among the i.i.d. samples to select the bandwidth $h$ so that $h = \text{median}(\{|\|q_i' - q_j'\|\}_{i \neq j})$. The computational complexity of the MMD estimator is $O(n^2 + (n')^2)$.

**Sliced Wasserstein Distances** A somewhat similar method to the Cramér-Wold procedure described above can be formulated based on Wasserstein distances rather than Kolmogorov-Smirnov statistics along one-dimensional projections. Given two probability measures $\Pi$ and $\tilde{\Pi}$ on $\mathbb{R}$ with cumulative distribution functions $F_\Pi$ and $F_{\tilde{\Pi}}$, respectively, the 1-Wasserstein distance between $\Pi$ and $\tilde{\Pi}$ is [Ramdas et al., 2017],

$$W_1(\Pi, \tilde{\Pi}) = \int_{0}^{1} \left| F_{\Pi}^{-1}(t) - F_{\tilde{\Pi}}^{-1}(t) \right| \, dt.$$ 

(5.A.57)

Using this one-dimensional characterization of the 1-Wasserstein distance, a distance – called the sliced Wasserstein distance – on probability measures on $\mathbb{R}^m$. Let $\Pi$ now be a probability measure on $\mathbb{R}^m$ may be constructed. Then the sliced Wasserstein distance is defined by

$$SW_1(\Pi, \tilde{\Pi}) = \int_{\mathbb{S}^{m-1}} W_1(\mathcal{R}_\theta(\Pi), \mathcal{R}_\theta(\tilde{\Pi})) \, d\text{Vol}_{\mathbb{S}^{m-1}}(\theta),$$

(5.A.58)

where $\mathcal{R}_\theta(\Pi)$ is the probability measure of the random variable $x^\top \theta$ when $x \sim \Pi$. Nadjahi et al. [2021] computes a Monte Carlo approximation of eq. (5.A.58) by sampling unit vectors uniformly over $\mathbb{S}^{m-1}$ and treating $(\omega_1, \ldots, \omega_n)$ and $(\omega_1', \ldots, \omega_n')$ as the locations of Dirac measures by which to define discrete probability measures. (The quantities $\omega_i$ and $\omega_i'$ were defined in step two in our discussion of Cramér-Wold Methods.) In the special
case that $n = n'$, the one-dimensional Wasserstein distance assumes the simple form,

$$W_1(\Pi, \hat{\Pi}) = \frac{1}{n} \sum_{i=1}^{n} |\omega(i) - \omega'(i)|,$$  \hspace{1cm} (5.A.59)

where $(\omega(1), \ldots, \omega(n))$ and $(\omega'(1), \ldots, \omega'(n))$ are, respectively, $(\omega_1, \ldots, \omega_n)$ and $(\omega'_1, \ldots, \omega'_n)$ sorted in ascending order.

**Discretized Differences in Probability** In section [5.A.8.1](#), the low-dimensionality of the posterior distribution permits us to employ the method of [Biau and Gyorfi][2007] which considers the $L_1$ distance between discretized probability densities. Given a partition $\mathcal{P}$ of $\mathbb{R}^m$, we compute the statistic,

$$DL_1(\{q_i\}_{i=1}^{n}, \{q'_i\}_{i=1}^{n'}) = \sum_{r \in \mathcal{P}} \text{Vol}(r) \cdot \left| \frac{1}{n} \sum_{i=1}^{n} 1 \{q_i \in r\} - \frac{1}{n'} \sum_{i=1}^{n'} 1 \{q'_i \in r\} \right|. \hspace{1cm} (5.A.60)$$

This can be interpreted as an approximation of the $L_1$ distance between probability densities, where the quality of the approximation depends on the number of samples $n$ and $n'$ and the fineness of the partition. In section [5.A.8.1](#) we partition $[-30, 10] \times [-10, 10] \subset \mathbb{R}^2$ (which contains virtually all of the probability mass of the posterior in that example) into 2,500 equally sized rectangles and compute the $DL_1$ statistic.

**Methods Based on Multiple Chains** Let $q_0$ be a fixed initial position variable. Consider $r$ independent (RM)HMC Markov chains starting from this shared initial position. Let $q_{ij}$ denote the position variable at step $j$ of the $i$-th Markov chain. On the $j$-th step, how close is the distribution of the set $\{q_{1j}, \ldots, q_{rj}\}$ to the target distribution? Note that, given the fixed initial position $q_0$, $q_{ij}$ is independent of $q_{kj}$. Therefore, the question as posed is clearly distinct from the question, “How close is the distribution $\{q_{i1}, \ldots, q_{in}\}$ to the target distribution?” Answering this question can give an indication of the convergence speed of the HMC and RMHMC procedures and, in the latter case, its sensitivity to the convergence
threshold.

Computing the previously described metrics would be prohibitively expensive to compute for every step of the Markov chain. However, in sections 5.A.8.7 and 5.5.1 there are singular dimensions of the posterior that are particularly challenging to sample due to the multiscale structure of the target distribution, but which nonetheless have known marginal distributions. By exclusively considering these single dimensions, we may assess convergence with respect to the most challenging dimension of the posterior. Let \( l \) denote the index of this challenging dimension of the posterior so that \( q_{ijl} \) is the \( l \)-th dimension of the state at the \( j \)-th step in the \( i \)-th chain. In our experiments, we compute the single sample Kolmogorov-Smirnov statistic comparing the distribution of \( \{q_{1jl}, \ldots, q_{rjl}\} \) against the known marginal. We set \( r = 1,000 \) and consider \( j = 1, \ldots, 10,000 \).

5.A.7.4 Measuring Sample Independence

Ergodicity measures how close the iterates of the Markov chain are to the target distribution of interest. However, the samples generated by Markov chains exhibit serial autocorrelation and therefore are not independent. The degree of dependence between samples with effect the determine the precision of the Monte Carlo approximation of posterior expectations. A standard measure in the MCMC literature is the effective number of independent samples that a set of Markov chain samples represents. The effective sample size is the equivalent number of independent samples that would produce an estimator with the same variance as the auto-correlated samples produced by the Markov chain. Formally, following Gelman et al. [2004], let \( (q_1, \ldots, q_n) \) be a sequence of univariate Markov chain samples and let \( \rho_t \) be the auto-correlation of \( (q_1, \ldots, q_n) \) with lag \( t \). The effective sample size (ESS) is the quantity,

\[
\text{ESS} = \frac{n}{1 + 2 \sum_{t=1}^{\infty} \rho_t}.
\]  

(5.A.61)
In practice, $\rho_t$ is not known and must be estimated from the sequence $(q_1, \ldots, q_n)$ itself. We utilize the procedure of Kumar et al. [2019] to compute the ESS in our experiments. As a practical matter, one is concerned not only with the effective sample size in absolute terms, but also with the effective sample size per unit of computation. To measure this quantity, we divide the ESS by the running time (in seconds) of the algorithm. In distributions with multiple dimensions, we may consider the mean ESS, which is the average ESS among each dimension of the Markov chain. Similarly, the minimum ESS is the minimum ESS among each dimension of the Markov chain.

5.A.7.5 Measuring Transition Kernel Similarity

Given two RMHMC transition kernels $K((q, p), \cdot; \epsilon, k, \delta)$ and $K((q, p), \cdot; \epsilon, k, \delta')$ with the same step-size $\epsilon$ and number of integration steps $k$, but with differing fixed point convergence thresholds $\delta$ and $\delta'$, how can we measure their similarity? For Bayesian inference tasks, we are primarily concerned with their similarity in the $q$-dimensions, since the $p$-dimensions are auxiliary variables that serve only to facilitate the construction of a phase-space. Therefore, let $\text{Proj}_q : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ denote the projection to the $q$-dimensions alone. We propose to measure the similarity of $K((q, p), \cdot; \epsilon, k, \delta)$ and $K((q, p), \cdot; \epsilon, k, \delta')$ as,

$$\mathbb{E}_{q \sim (\cdot)} \mathbb{E}_{p \sim \text{Normal}(0, \mathcal{G}(q))} \mathbb{E}_{u \sim \text{Uniform}(0, 1)} \| \text{Proj}_q(\Psi_\epsilon^k((q, p), u; \delta)) - \text{Proj}_q(\Psi_\epsilon^k((q, p), u; \delta')) \|_2,$$

(5.A.62)

where we average over a suitable distribution over the position and momentum variables. This is the expected difference in the samples generated by the transition kernels $K_\delta$ and $K_{\delta'}$ when ensuring that both transitions are computed using the same integration step-size, the same number of integration steps, the same current position in phase-space, and the same uniform random number used in applying the Metropolis-Hastings accept-reject
criterion. The distribution over $q$, the random position variable, is arbitrary and one therefore has this degree of freedom when measuring the similarity of transition kernels. In our experiments, we will use either i.i.d. samples from the target distribution, or samples from another Markov chain to approximate the expectation over the $q$-variables. Note that when both transition kernels reject the proposal computed by the RMHMC integrator, the expected difference is zero; in our visualizations of this metric, we show a distribution of differences in the cases where at least one of the two transition kernels did not reject its proposal and the probability that both transition kernels reject, which we call the “rejection agreement.” In our experiments, we compare transition kernels with thresholds $\delta \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 1 \times 10^{-7}, 10^{-8}, 10^{-9}\}$ against the transition kernel with $\delta' = 1 \times 10^{-10}$.

5.A.7.6 Measuring Computational Effort

As a measure of computational effort, we report the number of fixed point iterations required to resolve the implicit updates to the momentum and position.
Figure 5.A.1: Banana-shaped posterior detailed balance and similarity measures. Visualization of the error in reversibility (see fig. 5.A.1a), error in volume-preservation (see fig. 5.A.1b), and the number of decimal digits of similarity in transition kernels (see fig. 5.A.1c) for variable thresholds in the banana posterior distribution.

Figure 5.A.2: Visualization of the computational effort required to sample with RMHMC from the banana posterior. We show the number of fixed point iterations required to compute the two implicit steps of the generalized leapfrog integrator.
5.A.3 Banana-shaped distribution Kolmogorov-Smirnov. Visualization of the distribution of Kolmogorov-Smirnov statistics over variable thresholds in the banana-shaped posterior. The maximum Kolmogorov-Smirnov statistic of Euclidean HMC with a step-size of 0.1 is shown as a dashed blue line. In measuring KS, we also compare against i.i.d. samples, for which the average KS statistic (over random one-dimensional subspaces) is also shown.

5.A.4 Banana-shaped distribution ergodicity measures. Additional measures of ergodicity in the banana-shaped distribution. For thresholds smaller than $1 \times 10^{-2}$, the Riemannian methods produce a modest improvement in the ergodicity of the samples produced by the Markov chain. In computing the MMD metric, we employ a bandwidth of 1.727.

5.A.8 Additional Details on Experimentation

5.A.8.1 Banana-Shaped Distribution

Consider the following generative model:

$$\theta_1, \theta_2 \overset{i.i.d.}{\sim} \text{Normal}(0, \sigma_\theta^2)$$  \hspace{1cm} (5.A.63)

$$y_i | \theta_1, \theta_2 \overset{i.i.d.}{\sim} \text{Normal}(\frac{\theta_1}{289} + \frac{\theta_2}{\sigma_y^2}) \text{ for } i = 1, \ldots, 100.$$  \hspace{1cm} (5.A.64)
Figure 5.A.5: *Banana-shaped distribution minimum ESS per second.* Visualization of the effective sample size per second of RMHMC compared to HMC on the banana-shaped distribution. Notice in particular the differentiation of the RMHMC with the generalized leapfrog integrator compared to the implicit midpoint method.

Figure 5.A.6: The use of Ruppert averaging in the banana-shaped distribution to adaptively set the convergence threshold to achieve eight decimal digits of similarity compared to a transition kernel with a threshold of $1 \times 10^{-10}$. We show a Monte Carlo approximation to $B(\delta)$, which appears smooth, monotonically increasing. The value of $\delta$ satisfying $B(\delta) = -8$ is approximately $\delta = 1 \times 10^{-8}$, which is approximately the value of produced by Ruppert averaging.

As can be readily observed, the likelihood function of $(\theta_1, \theta_2)$ given $(y_1, \ldots, y_{100})$ is non-identifiable since for any real number $c \geq \theta_1$, there are two values of $\theta_2$ for which $c = \theta_1 + \theta_2^2$; namely, $\theta_2 = \pm \sqrt{c - \theta_1}$. The purpose of of the banana-shaped distribution, suggested by Bornn and Cornebise in their rejoinder to Girolami and Calderhead [2011], is therefore to give a representative example of a non-identifiable likelihood and the eponymous banana-shaped posterior that it produces. In our experiments we set $\sigma_y = 2$ and...
$\sigma_\theta = 2$. When sampling observations $(y_1, \ldots, y_{100})$ we set $\theta_1 = 1/2$ and $\theta_2 = \sqrt{1 - 1/4}$. The sum of the Fisher information and the negative Hessian of the log-prior associated to the banana-shaped distribution is given by,

$$G(\theta_1, \theta_2) = \begin{pmatrix}
\frac{1}{\sigma_\theta^2} + \frac{n}{\sigma_\theta^2} & \frac{2n\theta_2}{\sigma_\theta^2} \\
\frac{2n\theta_2}{\sigma_\theta^2} & \frac{1}{\sigma_\theta^2} + \frac{4n\theta_2^2}{\sigma_\theta^2}
\end{pmatrix}. \quad (5.A.65)$$

We now turn our discussion to inference in this statistical model. We consider Euclidean HMC with a step-size of 0.1 and 20 integration steps. For RMHMC with the generalized leapfrog integrator, we consider a step-size of 0.04 and 20 integration steps. As described by Bornn and Cornebise, RMHMC requires a small integration step-size for the banana-shaped distribution because of divergences in the fixed point iterations for the momentum variable in eqs. (5.2) and (5.3). The banana-shaped distribution places a non-trivial amount of probability mass in long, thin “tails” that extend in the $\theta_1$-variable and which are symmetric for negations of the $\theta_2$-variable, whose exploration is inhibited by a small step-size. We also implement RMHMC using the implicit midpoint integrator with a step-size of 0.1 and 20 integration steps, which does not suffer the same divergences as the generalized leapfrog integrator due to its superior stability. We also observe the implicit midpoint integrator has an acceptance probability of ninety-five percent, even greater than the generalized leapfrog integrator with a step-size less than half that used by the implicit midpoint method. The implicit midpoint method produces RMHMC with the best ergodicity properties, dominating both Euclidean HMC and RMHMC with the implicit midpoint integrator when 1,000,000 samples are drawn. In terms of ergodicity, a threshold of $1 \times 10^{-3}$ appears to be sufficient for inference in the banana-shaped distribution. However, the computational expediency of Euclidean HMC causes it to enjoy the best ESS per second, which must be balanced against the superior ergodicity of the the Riemannian variants. Notably, however, a threshold of $1 \times 10^{-3}$ is by no means optimal
in terms of producing the best reversibility and volume preservation among the implicitly
defined integrators, indicating that computational benefits may be obtained via principled
selection of this convergence parameter.

We also consider the use of the Ruppert averaging procedure in the banana-shaped
posterior in order to identify a threshold that has, on average, eight \((\kappa = 8)\) decimal digits
of similarity with a numerical integrator whose convergence threshold is \(1 \times 10^{-10}\). Results
are shown in fig. 5.A.6. We see that convergence of \(\bar{L}_n\) to zero is rapid; the sequence
\((\bar{\delta}_1, \bar{\delta}_2, \ldots)\) converges in, approximately, one-hundred fifty iterations. In fig. 5.A.6c we
show a Monte Carlo approximation to \(B(\delta)\) for the banana-shaped distribution and, in
addition, ten random samples of \(L(\delta)\) for randomly selected values of the position and
momentum. One observes that individual samples are not monotonically increasing, nor
do they appear to be particularly smooth. However, the average function, shown in blue,
does appear monotonically increasing and smooth. These observations will be replicated
in our other experiments.

### 5.A.8.2 Bayesian Logistic Regression

We consider a hierarchical Bayesian logistic regression defined by the following generative
model

\[
\alpha \sim \text{Gamma}(k, \theta) \tag{5.A.66}
\]

\[
\beta_i|\alpha \overset{i.i.d.}{\sim} \text{Normal}(0, \alpha^{-1}) \text{ for } i = 1, \ldots, p \tag{5.A.67}
\]

\[
y_i|\beta, x_i \overset{i.i.d.}{\sim} \text{Bernoulli}(\sigma(x_i^\top \beta)) \text{ for } i = 1, \ldots, n, \tag{5.A.68}
\]

where \(\sigma : \mathbb{R} \rightarrow (0,1)\) is the sigmoid function. In our experiments we analyze the
heart dataset from Girolami and Calderhead \([2011]\). This dataset consists of \(p = 14\) regression coefficients and \(n = 270\) observations; including the latent prior preci-
Figure 5.A.7: Visualization of the time-normalized effective sample sizes (ESS) of the variables in the logistic regression posterior. Distributions over ESS are computed by splitting a Markov chain of length 100,000 into twenty contiguous sequences of length 5,000.

Figure 5.A.8: Visualization of the error in reversibility (see fig. 5.A.8a), error in volume-preservation (see fig. 5.A.8b), and the number of decimal digits of similarity in transition kernels (see fig. 5.A.8c) for variable thresholds in the logistic posterior distribution.

This produces a 15-dimensional posterior distribution. In our experiments, we set \( k = 1 \) and \( \theta = 2 \). We employ a Metropolis-within-Gibbs sampling strategy whereby we alternate between sampling \( \beta \mid ((x_1, y_1), \ldots, (x_n, y_n)) \), \( \alpha \) using RMHMC and sampling \( \alpha \mid \beta, ((x_1, y_1), \ldots, (x_n, y_n)) \) analytically since

\[
\alpha \mid \beta, ((x_1, y_1), \ldots, (x_n, y_n)) \sim \text{Gamma} \left( k + \frac{p}{2}, \frac{1}{\beta^\top \beta + \frac{1}{\theta}} \right). \quad (5.A.69)
\]
The Fisher information metric for sampling the former of these distributions depends on the value of $\beta$, thereby necessitating the use of the generalized leapfrog integrator. In particular, the sum of the Fisher information and the negative Hessian of the log-prior can be shown to be $G(\beta) = X^\top \Lambda X + \alpha \text{Id}$ where $\Lambda$ is a diagonal matrix with entries $\Lambda_{i,i} = \sigma(x_i^\top \beta)(1 - \sigma(x_i^\top \beta))$ and $X \in \mathbb{R}^{n \times p}$ is the row-wise concatenation of $x_1, \ldots, x_n$.

We compare the ESS of the RMHMC algorithm in fig. 5.A.7. We compute the ESS by taking a single chain of length 100,000 and splitting it into twenty contiguous arrays of length 5,000. Within each contiguous sample, we compute the ESS and report the minimum ESS among the linear coefficients and the ESS of the precision variable. In this experiment, the minimum ESS is effectively constant as a function of the threshold, relative to the ESS per second exhibited by HMC. Indeed, this example provided a circumstance wherein no threshold employed in RMHMC was able to produce a time normalized minimum ESS which was competitive with Euclidean HMC. It is worth noting, however, that RMHMC tends to outperform HMC when time is not accounted for, as demonstrated in the top panel of fig. 5.A.7. We employ RMHMC with a step-size of 0.2 and twenty integration steps. For reference, we also report these ESS statistics for HMC with a step-sizes of 0.1 and 0.2 and twenty integration steps.

We visualize the violation of reversibility and volume preservation over varying thresholds in fig. 5.A.8. This analysis reveals that, in the worst case, the proposal operator enjoys approximately two decimal digits of reversibility and volume preservation. Moreover, for a threshold of $1 \times 10^{-3}$, the transition kernels are in agreement to the fourth decimal place.

5.A.8.3 Stochastic Volatility Model

Stochastic volatility models are random processes which are characterized by the randomness inherent in their variance (or “volatility”). We consider the following generative
Figure 5.A.9: Visualization of the posterior distribution over the stochastic volatilities wherein the process hyperparameters \((\sigma^2, \phi, \beta)\) are sampled by RMHMC with variable threshold. Visually, the posterior distributions are indistinguishable.

- \(x_{t+1} | x_t, \phi, \sigma^2 \sim \text{Normal}(\phi x_t, \sigma^2)\) for \(t = 1, \ldots, T - 1\)  \hspace{1cm} (5.A.71)
- \(y_t | \beta, x_t \sim \text{Normal}(0, \beta^2 \exp(x_t))\) for \(t = 1, \ldots, T\)  \hspace{1cm} (5.A.72)

Figure 5.A.10: Stochastic volatility errors in detailed balance. Visualization of the error in reversibility (see fig. 5.A.10a), error in volume-preservation (see fig. 5.A.10b), and the number of decimal digits of similarity in transition kernels (see fig. 5.A.10c) for variable thresholds in the stochastic volatility posterior.

Figure 5.A.11: Visualization of the ESS per second for the three latent parameters of the stochastic volatility model.
with priors $\sigma^2 \sim \text{Inv-}\chi^2(10, 1/20)$, $(\phi + 1)/2 \sim \text{Beta}(20, 3/2)$ (so that $\phi \in (-1, +1)$), and the prior density over $\beta$ is proportional to $1/\beta$. The Bayesian inference task is to infer the posterior distribution of $(x_1, \ldots, x_T, \sigma^2, \phi, \beta)$ given observations $(y_1, \ldots, y_T)$. In our experiments we set $T = 1,000$, producing a posterior of dimensionality 1,003. Following [Girolami and Calderhead] (2011), we employ a Metropolis-within-Gibbs-like strategy by alternating between sampling the distribution $(x_1, \ldots, x_T)|\sigma^2, \phi, \beta, (y_1, \ldots, y_T)$ and sampling the distribution $(\phi, \beta, \sigma^2)|(x_1, \ldots, x_T), (y_1, \ldots, y_T)$. In the former case, the Fisher information metric is constant with respect to $(x_1, \ldots, x_T)$, thereby allowing us to use the usual leapfrog integrator to produce samples; moreover, the metric has a special tri-diagonal structure, facilitating the use of specialized numerical linear algebra routines. However, the Fisher information metric of the distribution $(\phi, \beta, \sigma^2)|(x_1, \ldots, x_T), (y_1, \ldots, y_T)$ depends on $(\phi, \beta, \sigma^2)$, thereby producing a non-separable Hamiltonian and necessitating the use of the generalized leapfrog integrator. In order to respect the constraints $\sigma^2 > 0$ and $\phi \in (-1, +1)$, we define auxiliary variables $(\gamma, \alpha)$ and employ the smooth, invertible transformations $\sigma^2 = \exp(\gamma)$ and $\phi = \tanh(\alpha)$. The Fisher information metric for the

Figure 5.A.12: Visualization of the number of fixed point iterations required to compute the implicit updates to position and momentum required by the generalized leapfrog integrator in the stochastic volatility model.
transformed variables is,

\[
G(\gamma, \alpha, \beta) = \begin{pmatrix}
2T & 2\phi & 0 \\
2\phi & 2\phi^2 + (T - 1)(1 - \phi^2) & 0 \\
0 & 0 & 2T/\beta^2
\end{pmatrix}.
\]  

(5.A.73)

In sampling the posterior of \((\sigma^2, \phi)\), we use six integration steps with a step-size of 0.5. We seek to draw 100,000 observations from the posterior.

We visualize the posterior over the stochastic volatilities for variable thresholds in fig. 5.A.9. Visually, the posterior distributions are indistinguishable; this conclusion is reinforced by the close similarity of the posterior marginals of over \((\phi, \beta, \sigma^2)\), wherein only the largest threshold \(1 \times 10^{-1}\) shows any dissimilarity, which is nonetheless minor. This similarity is quantified in our assessment of the similarity of the Markov chain transition kernel, wherein we see that the threshold \(1 \times 10^{-1}\) enjoys nearly five decimal digits of similarity relative to the transition kernel with threshold \(1 \times 10^{-10}\). In figs. 5.A.12a and 5.A.12b we visualize the number of fixed point iterations required by the generalized leapfrog method by convergence tolerance. We observe that for a threshold of \(1 \times 10^{-1}\), only one or two fixed point iterations are required to resolve the implicit updates of the position and momentum variables, respectively, which compares favorably to the six or seven required by the generalized leapfrog method with a convergence tolerance of \(1 \times 10^{-9}\).

In fig. 5.A.11 we evaluate the effective sample size per second of the Riemannian and non-Riemannian HMC variants. Only in the case of the variable \(\phi\) does RMHMC offer any benefits, while sampling efficiency is degraded by using RMHMC on the variables \(\beta\) and \(\sigma^2\). This occurs due to the relative computational burden of RMHMC, which cannot always be compensated for, in terms of time-normalized metrics, by the geometric advantages.
5.A.8.4 Log-Gaussian Cox-Poisson Model

The log-Gaussian Cox-Poisson model allows us to model count data within a spatial grid. In particular, consider a $N \times N$ grid on the unit square. Within the $(i, j)$ sub-region, the counts of some quantity of interest are denoted $y_{i,j}$. We model these count observations as following a Poisson distribution whose rate is spatially correlated according to a Gaussian process. Formally, we consider the following generative procedure:

\[
\begin{align*}
\beta & \sim \text{Normal}(2, 1/2) \quad (5.A.74) \\
\sigma^2 & \sim \text{Normal}(2, 1/2) \quad (5.A.75) \\
\Sigma_{(i,j),(i',j')} & \sim \sigma^2 \exp \left( -\frac{\sqrt{(i-i')^2 + (j-j')^2}}{N\beta} \right) \quad (5.A.76) \\
\text{vec}(x)|\Sigma & \sim \text{MultivariateNormal}(\mu_1, \Sigma) \quad (5.A.77) \\
y_{i,j}|x_{i,j} & \sim \text{Poisson} \left( \exp(x_{i,j})/N^2 \right) \quad (5.A.78)
\end{align*}
\]

where $\text{vec} : \mathbb{R}^{N\times N} \to \mathbb{R}^{N^2}$ by concatenating columns into a single vector. The Bayesian inference task is to infer both the Gaussian process posterior $x$ and the posterior distribu-
Figure 5.A.14: (Cox-Poisson Model detailed balance and sampling metrics.) Visualization of the error in reversibility (see fig. 5.A.14a) and error in volume-preservation (see fig. 5.A.14b) in the log-Gaussian Cox-Poisson posterior. We also illustrate the ESS per second for sampling the hyperparameters of the Cox-Poisson model and a comparison against a Euclidean HMC method.

tion of $(\beta, \sigma^2)$ given observations of the Poisson process. Following Girolami and Calderhead [2011], we employ a Metropolis-within-Gibbs-like strategy and alternate between sampling the posterior of $x|{(y_{i,j})}_{i,j=1}^N, \sigma^2, \beta$ and the posterior of $(\sigma^2, \beta)|x$. In generating data from this model, we set $\sigma^2 = 1.91$, $\beta = 1/33$, and $\mu = \log 126 - 1.91/2$. In our experiments we set $N = 32$ so that the total dimensionality of the posterior is $32^2 + 2 = 1026$. Similarly to the situation in section 5.A.8.3, the $x|{(y_{i,j})}_{i,j=1}^N, \sigma^2, \beta$ has a constant Fisher information metric; however, the conditional distribution of the hyperparameters $(\sigma^2, \beta)|x$
depends on $\sigma^2$ and $\beta$, thereby necessitating the use of the generalized leapfrog integrator. Once again, to respect the constraints $\sigma^2 > 0$ and $\beta > 0$, we employ the transformation 

$$\sigma^2 = \exp(\phi_1)$$ and $$\beta = \exp(\phi_2).$$

Since the conditional distribution of $\text{vec}(x)$ given $\phi_1$ and $\phi_2$ is simply a multivariate Gaussian, the Fisher information is

$$G(\phi_1, \phi_2) = \frac{1}{2} \begin{pmatrix} N^2 & \Omega(\phi_1, \phi_2) \\ \Omega(\phi_1, \phi_2) & \Lambda(\phi_1, \phi_2) \end{pmatrix} \quad (5.A.79)$$

$$\Omega(\phi_1, \phi_2) = \text{trace} \left( \Sigma^{-1}(\phi_1, \phi_2) \frac{\partial}{\partial \phi_2} \Sigma(\phi_1, \phi_2) \right) \quad (5.A.80)$$

$$\Lambda(\phi_1, \phi_2) = \text{trace} \left( \Sigma^{-1}(\phi_1, \phi_2) \frac{\partial}{\partial \phi_2} \Sigma(\phi_1, \phi_2) \Sigma^{-1}(\phi_1, \phi_2) \frac{\partial}{\partial \phi_2} \Sigma(\phi_1, \phi_2) \right). \quad (5.A.81)$$

In sampling from the posterior distribution of $(\sigma^2, \beta)$ we use the generalized leapfrog integrator with six integration steps and a step-size of 0.5. We seek to draw 5,000 samples after an initial burn-in period of 1,000 samples.

We visualize the posterior mean and standard deviation of the log-Gaussian Cox-Poisson process in fig. 5.A.13. As in the case of section 5.A.8.3, we observe few detectable differences within the first two moments of the posterior. When assessing the degree to which reversibility and volume preservation are violated, we observe that the worst-case
behavior of the generalized leapfrog integrator still exhibits error only in the third or fourth
decimal digit. The similarity of transition kernels also reveals that the worst case difference
among transitions still maintains approximately three digits of similarity with a transition
kernel whose threshold is $1 \times 10^{-10}$. When contextualized in terms of computational ef-
fort, one sees that a threshold of $1 \times 10^{-1}$ requires substantially fewer fixed point iteration
on average for both the position and momentum variables, compared with nearly ten re-
quired by a threshold of $1 \times 10^{-9}$. Since computing the Riemannian metric requires the
inverse of $\Sigma$, which is a $N^2 \times N^2$ matrix, one infers that the computational complexity of
computing the Riemannian metric for the Cox-Poisson model scales as $O(N^6)$; therefore,
one desires few fixed point iterations, particularly in the implicit update to position, for
which the Riemannian metric must be recomputed at each iteration.

We evaluate the ESS per second for the Log-Gaussian Cox-Poisson model in figs. 5.A.14c
and 5.A.14d. Of the two hyperparameters in the Cox-Poisson model, $\beta$ is the more chal-
lenging, having the smallest time-normalized ESS. Even with the most conservative con-
vergence threshold, RMHMC outperforms the Euclidean algorithm; as a result, RMHMC
has superior minimum ESS per second relative to Euclidean HMC.

5.A.8.5 Details for Neal’s Funnel Distribution

The SoftAbs metric is constructed as follows. Let $H(x, v)$ denote the Hessian of the
joint density of $(x, v)$. Let $H(x, v) = QAQ^T$ be an eigen-decomposition of $H$, where
$Q \equiv Q(x, v)$ is the matrix of eigen-vectors and $\Lambda \equiv \Lambda(x, v)$ is a diagonal matrix of
eigen-values. The Hessian of Neal’s funnel distribution is not positive definite and it
is therefore inadmissible as a Riemannian metric. The SoftAbs metric is constructed
from the Hessian by smoothly transforming the eigen-values of the Hessian according
to $\tilde{\lambda}_i(x, v) = \alpha \coth(\alpha \lambda_i(x, v))$ and $\tilde{\Lambda}(x, v) = \text{diag}(\tilde{\lambda}_1(x, v), \ldots, \tilde{\lambda}_{11}(x, v))$, where $\alpha > 0$
is a tunable parameter controlling the smoothness of the SoftAbs transformation. Indeed,
Figure 5.A.16: Number of fixed point iterations in Neal’s funnel distribution. Visualization of the computational effort required to sample with RMHMC from Neal’s funnel distribution. We show the number of fixed point iterations required to compute the two implicit steps of the generalized leapfrog integrator.

the transformation $\lambda \mapsto \text{coth}(\alpha \lambda)$ is a smooth approximation to the absolute value function. The SoftAbs metric is then $G(x, v) = Q(x, v)\tilde{\Lambda}(x, v)Q(x, v)^\top$. In our experiments we set $\alpha = 10^4$.

5.A.8.6 Details for Fitzhugh-Nagumo

We give some additional details about how the metric employed in the Fitzhugh-Nagumo experiments is computed in practice. One first observes that there is no immediate closed-form expression for the partial derivatives $\partial v_{tk}/\partial q_i$ or $\partial r_{tk}/\partial q_i$ based on the model description. However, we may leverage implicit differentiation of eqs. (5.25) and (5.26), in order to deduce sensitivity equations [Arriola and Hyman 2009] for these quantities. For instance,

$$\frac{d}{dt} \frac{\partial v_t}{\partial a} = \frac{\partial}{\partial a} \frac{dv_t}{dt} = c \left( \frac{\partial r_t}{\partial a} - (v_t^2 - 1) \frac{\partial v_t}{\partial a} \right)$$

(5.A.82)
In order to obtain initial conditions for these sensitivity dynamics, one observes that the initial condition of the Fitzhugh-Nagumo system \((v_0, r_0) = (-1, +1)\) does not depend on \(a, b,\) or \(c\). The RMHMC algorithm requires the derivatives of the Riemannian metric; applying the chain rule to eq. (5.27) yields the following equations,

\[
\frac{\partial}{\partial q_l} G_{ij}(q) = \frac{1}{\sigma^2} \sum_{k=1}^{200} \frac{\partial^2 v_{t_k}}{\partial q_i \partial q_j} \frac{\partial v_{t_k}}{\partial q_l} + \frac{\partial^2 v_{t_k}}{\partial q_i \partial q_l} \frac{\partial v_{t_k}}{\partial q_j} + \frac{\partial^2 r_{t_k}}{\partial q_i \partial q_j} \frac{\partial r_{t_k}}{\partial q_l} + \frac{\partial r_{t_k}}{\partial q_i} \frac{\partial^2 r_{t_k}}{\partial q_j \partial q_l} \quad (5.83)
\]

To compute these derivatives requires additional sensitivity equations: the second partial derivatives of the states. However, this presents no additional conceptual difficulty since we may apply implicit differentiation a second time to obtain the required system of dif-

(a) Number of fixed point iterations for position variable.

(b) Number of fixed point iterations for momentum variable.

Figure 5.A.17: Visualization of the error in reversibility (see fig. 5.6a) and error in volume-preservation (see fig. 5.6b) for variable thresholds in the Fitzhugh-Nagumo posterior distribution.

In order to obtain initial conditions for these sensitivity dynamics, one observes that the initial condition of the Fitzhugh-Nagumo system \((v_0, r_0) = (-1, +1)\) does not depend on \(a, b,\) or \(c\). The RMHMC algorithm requires the derivatives of the Riemannian metric; applying the chain rule to eq. (5.27) yields the following equations,
### 5.A.8.7 Multi-Scale Phenomena in a Multivariate Student-t Distribution

Following [Lan et al. 2015], we next consider sampling an $m$-dimensional Student-$t$ distribution with $\nu$ degrees-of-freedom whose log-density function is, $L(q) = -\frac{\nu + m}{2} \log \left(1 + \frac{1}{\nu} q^\top \Sigma^{-1} q\right)$, where $\Sigma \in \mathbb{R}^{m \times m}$ is a positive definite matrix. The Hessian of the log-density of the multivariate Student-$t$ density is given by

\begin{align*}
    \frac{d}{dt} \frac{\partial^2 v_t}{\partial a^2} &= c \left( \frac{\partial^2 r_t}{\partial a^2} - \left(v_t^2 - 1\right) \frac{\partial^2 v_t}{\partial a^2} - 2v_t \left( \frac{\partial v_t}{\partial a} \right)^2 \right) \quad (5.A.84) \\
    \frac{d}{dt} \frac{\partial^2 v_t}{\partial b \partial a} &= c \left( \frac{\partial^2 r_t}{\partial a \partial b} - 2v_t \frac{\partial v_t}{\partial a} \frac{\partial v_t}{\partial b} - \left(v_t^2 - 1\right) \frac{\partial^2 v_t}{\partial a \partial b} \right) \quad (5.A.85) \\
    \frac{d}{dt} \frac{\partial^2 v_t}{\partial c \partial a} &= c \left( \frac{\partial^2 r_t}{\partial a \partial c} - \left(v_t^2 - 1\right) \frac{\partial^2 v_t}{\partial a \partial c} \right) + \frac{\partial r_t}{\partial a} - 2c \frac{\partial v_t}{\partial c} \frac{\partial v_t}{\partial a} v_t + \left(v_t^2 - 1\right) \frac{\partial}{\partial a} v_t. \quad (5.A.86)
\end{align*}

Figure 5.A.18: Student-$t$ distribution multiscale phenomena. Visualization of how multiscale properties of a target multivariate Student-$t$ distribution affect the ergodicity and computational efficiency of the Euclidean HMC and RMHMC algorithms. In the presence of severe multiscale attributes of the posterior, Euclidean methods of HMC exhibit poor sampling performance by these metrics.
Figure 5.A.19: Visualization of the error in reversibility (see fig. 5.A.19a), error in volume-preservation (see fig. 5.A.19b), and the number of decimal digits of similarity in transition kernels (see fig. 5.A.19c) for variable thresholds in a multivariate Student-t distribution with $\sigma^2 = 10,000$.

The negative Hessian of the log-density is not positive definite everywhere; however, instead of employing the SoftAbs procedure as in section 5.5.1 to produce a positive definite metric, we instead consider the metric $G(q) = \frac{\nu + m}{\nu(1 + q^\top \Sigma^{-1} q / \nu)} \Sigma^{-1}$. This metric can be motivated by computing the negative Hessian of the multivariate Student-t log-density and keeping only the term that is guaranteed to be positive definite. The covariance of the Student-t distribution is computed from the matrix $\Sigma$ and the degrees-of-freedom. In particular, when $q$ is drawn from a multivariate Student-t distribution, $\text{Cov}(q) = \frac{\nu}{(\nu - 2)}\Sigma$. In our experiments we set $\nu = 5$, $m = 20$, and $\Sigma = \text{diag}(1, \ldots, 1, \sigma^2)$; we consider values of $\sigma^2 \in \{1 \times 10^0, 1 \times 10^1, 1 \times 10^2, 1 \times 10^3, 1 \times 10^4\}$. By examining variable values of $\sigma^2$, we force $\Sigma$ to incorporate increasingly severe multiscale phenomena; this can be related to the condition number of the matrix $\Sigma$, which is the ratio of the largest eigenvalue to the smallest eigenvalue. We expect Euclidean variants of HMC to struggle in the presence of multiscale distributions [Pourzanjani and Petzold, 2019, Lan et al., 2015, Betancourt]
Figure 5.A.20: Student-\(t\) distribution sampling metrics and ergodicity. Visualization of the sample quality of the variables in a multivariate Student-\(t\) distribution as measured by the ESS per second (fig. 5.A.20a) and the distribution of Kolmogorov-Smirnov (KS) statistics over random one-dimensional subspaces (fig. 5.A.20b). In measuring KS, we also compare against i.i.d. samples, for which the average KS statistic (over random one-dimensional subspaces) is also shown. Distributions over ESS are computed by splitting a Markov chain of length 1,000,000 into twenty contiguous sequences of length 50,000.

For the Euclidean methods we consider twenty integration steps and a step-size in \(\{0.1, 0.5, 0.8\}\). For Riemannian methods we also consider twenty integration steps and a fixed step-size of 0.3. In fig. 5.A.18 we visualize the effect of the condition number on the ergodicity and effective sample size per second metrics. We show results for a Riemannian
Figure 5.A.21: Additional measures of ergodicity in the multivariate Student-t distribution. In the presence of the multiscale phenomena, the Euclidean HMC algorithms cannot achieve the same level of ergodicity as the Riemannian algorithms except in the case wherein a very large convergence tolerance is used in RMHMC. When $\sigma^2 = 1 \times 10^4$, we also show convergence of the dimension with largest variation, whose marginal is $t(5, \sigma^2)$. The Riemannian methods exhibit faster convergence compared to HMC. For $\sigma^2 = 1 \times 10^4$, the MMD statistic uses a bandwidth of 110.72 in the squared exponential kernel.

Figure 5.A.22: Ruppert averaging in multiscale Student-t distribution. The use of Ruppert averaging in the multiscale Student-t distribution to adaptively set the convergence threshold to achieve eight decimal digits of similarity compared to a transition kernel with a threshold of $1 \times 10^{-10}$. We show a Monte Carlo approximation to $B(\delta)$, which appears smooth, monotonically increasing. The value of $\delta$ satisfying $B(\delta) = -8$ is approximately $\delta = 1 \times 10^{-7.83}$, which is somewhat greater than the value of $1 \times 10^{-8}$ produced by Ruppert averaging. We note that the function $B(\delta)$ is approximately constant over the interval $[-3, -2]$; this could present difficulty in optimizing thresholds in this range since it violates the fourth assumption for convergence of the Robbins-Monro algorithm.

method with threshold fixed at $1 \times 10^{-5}$ and report results for the best performing Euclidean algorithm among the step-sizes considered. For condition numbers of 100, 1,000, and 10,000, we see that the Euclidean algorithms exhibit degraded performance. Whereas the effective sample size per second of the Riemannian methods is effectively constant as a function of the condition number, Euclidean methods are able to generate fewer effective
samples per second as the covariance matrix becomes increasingly ill-conditioned. This is also reflected in our measure of ergodicity, wherein large condition numbers result in worse sampling from the target distribution.

In the sequel, we fix $\sigma^2 = 1 \times 10^4$. Fig. 5.A.20a shows the time-normalized ESS; we observe that RMHMC, irrespective of convergence threshold, has superior sampling efficiency. In fig. 5.A.20b we show the sensitivity of the ergodicity of RMHMC to the choice of threshold and compare it against the Euclidean algorithms with variable step-size. In sampling from this distribution, we observe that a threshold of $1 \times 10^{-2}$ is sufficient to achieve a measure of ergodicity that is comparable to a threshold of $1 \times 10^{-9}$. We further note that the Riemannian algorithm with the largest step-size achieves a measure of ergodicity comparable to the best Euclidean HMC implementation. In fig. 5.A.19a we visualize the reversibility of the proposal: we observe that a threshold of $1 \times 10^{-2}$ exhibits around one decimal digit of error in reversibility. In fig. 5.A.20a we repeat this analysis for the effective sample size per second. Since smaller convergence thresholds require strictly greater computational effort, we observe that, for thresholds less than or equal to $1 \times 10^{-2}$, the effective sample size per second is a decreasing function of threshold.

### 5.A.8.8 Details for Newton’s Method

Recall that the fixed equation for the momentum variable in the generalized leapfrog is

$$\tilde{p} = p_n - \frac{\epsilon}{2} \nabla_q H(q_n, \tilde{p})$$

(5.A.88)

which we can represent as a root-finding problem (in the variable $\tilde{p}$) by rearranging as,

$$g(\tilde{p}) = \tilde{p} - p_n + \frac{\epsilon}{2} \nabla_q H(q_n, \tilde{p}) = 0.$$  

(5.A.89)
Figure 5.A.23: We compute the number of iterations required to resolve the implicit updates to the momentum variable when using either fixed point iteration or Newton’s method. We report the mean number of iterations with error bars showing dispersion about the mean. In each circumstance, Newton’s method converges faster than fixed point iterations, reflecting its superior order of convergence. Only around three iterations of Newton’s method are required on average in each example. We observe that in many cases, Newton’s method appears to be less sensitive to the convergence tolerance.

The Jacobian is therefore

$$\nabla g(\bar{p}) = \text{Id} + \frac{\epsilon}{2} \nabla_p \nabla_q H(q_0, \bar{p}).$$  

(5.A.90)

For Hamiltonians of the form eq. (5.1), we have

$$\nabla_p \nabla_q H(q, p) = -G^{-1}(q)\nabla_q G(q)G^{-1}(q)\bar{p},$$  

(5.A.91)

which immediately gives a computable Jacobian to employ in Newton’s method. Although Newton’s method has a superior order of convergence than fixed point iteration, we observe that the two procedures do not have the same computational complexity. In particular, Newton’s method requires that we solve the linear system, $\nabla g(\bar{p})s = g(\bar{p})$ at each step.
for the vector $s \in \mathbb{R}^m$; in general, the matrix $\nabla g(\bar{p})$ does not have any special structure, so solving this linear system has cubic complexity. We expect, therefore, that resolving the implicit update to momentum via Newton’s method must balance the faster order of convergence with the greater computational burden.

We may also consider Newton iterations to resolve the implicit update to the position variable in RMHMC with the generalized leapfrog integrator. The implicit update to position in eq. (5.3) and for Hamiltonians of the form eq. (5.1) is,

$$q_{n+1} = q_n + \frac{\epsilon}{2} \left( G^{-1}(q_{n+1})\bar{p} + G^{-1}(q_n)\bar{p} \right).$$  (5.92)

This can be represented as a root-finding problem by defining the function,

$$g(q_{n+1}) = q_{n+1} - q_n - \frac{\epsilon}{2} G^{-1}(q_{n+1})\bar{p} - \frac{\epsilon}{2} G^{-1}(q_n)\bar{p},$$  (5.93)

whose Jacobian is,

$$\nabla g(q_{n+1}) = \text{Id} - \frac{\epsilon}{2} G^{-1}(q_{n+1}) \nabla G(q_{n+1}) G^{-1}(q_{n+1})\bar{p}.$$  (5.94)
Figure 5.A.24: Three circumstances are shown wherein the use of Newton’s method to resolve the implicit update to the momentum resulted in degraded sampling efficiency due to the higher computational burden of Newton’s method compared to fixed point iteration. In the banana-shaped distribution, although the posterior is only two-dimensional, significant computation is wasted while searching for solutions to the momentum update that do not exist.
Figure 5.A.25: Three circumstances in which Newton’s method produced performance comparable to fixed point iteration when resolving the implicit update to the momentum. In these cases, Newton’s method may actually be the preferred method, since it is somewhat less sensitive to the convergence tolerance. In the first row we show results for the stochastic volatility model (SV); the second row shows the log-Gauss Cox-Poisson model (CP); the third row shows results for the Fitzhugh-Nagumo differential equation model (FN).
Figure 5.A.26: We visualize the ergodicity properties for the three implementations of RMHMC. In the case where Newton’s method is employed for resolving the implicit updates to momentum and position, the ergodicity of the method is essentially constant with respect to the convergence tolerance. However, this is not desirable in the case of the banana-shaped posterior distribution, wherein ergodicity failures result from severe violations of reversibility and volume preservation in some cases.
Figure 5.A.27: Comparison of fixed point iterations, Newton’s method (to update both position and momentum), and Newton’s method with a stability check. We observe that implementing a stability check for Newton’s method immensely improves the ergodicity of the method. We also show the difference in Hamiltonian energies along trajectories that have at least one unstable fixed point. These results show that even unstable fixed points can lead to proposals that are accepted, which is why it is important to incorporate a stability check.

5.A.9 Stability Checks in Newton Methods

As noted in our discussion of the banana-shaped distribution when employing Newton’s method to update both the momentum and position, we observed degraded ergodicity metrics as computed by our Kolmogorov-Smirnov metric. The degradation is visible in fig. 5.A.26a. Upon investigation, we observed two phenomena. First, that numerical assessment of volume preservation and reversibility suggested that employing Newton’s method to update momentum and position could, on occasion, produce proposals which appeared to severely violate both assumptions. Second, that these instances of violated reversibility and volume preservation were associated to Newton’s method identifying unstable fixed points, which we now define.

Definition 5.A.9.1. Let \( g : \mathbb{R}^m \to \mathbb{R}^m \) and let \( z^* \) be the solution to the fixed point equation \( z = g(z) \). We say that \( z^* \) is stable if \( \| \nabla g(z^*) \|_{op} < 1 \). Otherwise we say that \( z^* \) is unstable.
As seen from theorem 5.2.25, a stable fixed point means that there is some non-zero radius of convergence for fixed point iterations. We propose, therefore, to verify that the approximate solution $z^*$ obtained by Newton’s method is stable by computing $\|\nabla g(z^*)\|_{op}$ and comparing this value against the theoretical threshold of one. If at any step of the generalized leapfrog integrator the solutions obtained by Newton’s method are unstable, we ensure that the proposal is rejected. We refer to this procedure as Newton’s method with a stability check. We note that, in relation to prior work, Cornebise and Julien, in their discussion of Girolami and Calderhead [2011], compute the probability of the implicit update to the momentum having a solution by checking that $\|\nabla g_p(p^*_s)\|_{op} < 1.2$ where $g_p$ is defined in eq. (5.5); however, in their case $p^*_s$ is obtained via fixed point iteration.

Results comparing the generalized leapfrog integrator with fixed point iterations, and Newton’s method (with and without the stability check) to update both the position and momentum variables are shown in fig. 5.A.27a. We observe that incorporating the stability check improves the ergodicity of the method as measured by our Kolmorogov-Smirnov statistic, while simultaneously decreasing the sensitivity to the convergence threshold.

It may surprising unstable solutions can still produce proposal maps (consisting of $k$ integration steps) $(q_0, p_0) \mapsto (q_k, p_k)$ for which the proposal is accepted. Recall that the acceptance probability of RMHMC has the following form:

$$\min \{1, \exp (H(q_0, p_0) - H(q_k, p_k))\}.$$ (5.A.95)

This motivates us to inspect the differences in the Hamiltonian energy among trajectories containing an unstable fixed point. In fig. 5.A.27b we identified 10,000 initial conditions $(q_0, p_0)$ which produced a trajectory containing at least one unstable fixed point; fig. 5.A.27b then shows the distribution of the $\log_{10}$ difference in Hamiltonians. We computed the difference in the initial Hamiltonian $H(q_0, p_0)$ and the Hamiltonian at the terminal position $H(q_k, p_k)$. We found that 36.5% of these had $H(q_k, p_k) < H(q_0, p_0)$, which
are guaranteed to be accepted by the Markov chain.
Figure 5.A.28: Visualization of how a failure to respect the symmetry of partial derivatives yields severely degraded sample quality; indeed, the target distribution may be warped. Although reversibility can be reduced by a decreasing threshold, the property of volume-preservation cannot be recovered irrespective of threshold. As a consequence, the RMHMC algorithm is no longer ergodic for the target distribution, which is reflected in the degraded ergodicity measure.

5.A.10 Non-Volume-Preserving Errors in the Fitzhugh-Nagumo Posterior

Violations of volume preservation may appear in other, more subtle ways, which would raise no immediate indication of error unless one knew to look for it. As a concrete example, we consider a transcription error in the specification of the forward sensitivity equa-
tions necessary to compute the derivatives of the Riemannian metric used in the Fitzhugh-Nagumo model. Specifically, we consider replacing the expression in eq. (5.A.86) as follows:

\[
\frac{d}{dt} \frac{\partial^2 v_t}{\partial c \partial a} = c \left( \frac{\partial^2 r_t}{\partial a \partial c} - (v_t^2 - 1) \frac{\partial^2 v_t}{\partial a \partial c} \right) + \frac{\partial r_t}{\partial a} - 2c \frac{\partial v_t}{\partial c} \frac{\partial v_t}{\partial a} v_t + 1 - v_t^2 \frac{\partial v_t}{\partial a}, \quad (5.A.96)
\]

and we incorrectly set,

\[
\frac{d}{dt} \frac{\partial^2 v_t}{\partial c \partial b} = -2cv_t \frac{\partial v_t}{\partial c} \frac{\partial v_t}{\partial a} + 1 - v_t^2 \frac{\partial v_t}{\partial b} - c (v_t^2 - 1) \frac{\partial^2 v_t}{\partial c \partial b} + \frac{\partial r_t}{\partial b} + \frac{\partial r_t}{\partial c \partial b}, \quad (5.A.97)
\]
\[
\frac{d}{dt} \frac{\partial^2 v_t}{\partial^2 c} = -2cv_t \frac{\partial v_t}{\partial c} \frac{\partial v_t}{\partial c} + 2 - v_t^2 \frac{\partial v_t}{\partial c} - c (v_t^2 - 1) \frac{\partial^2 v_t}{\partial^2 c} - v_t^2 \frac{\partial v_t}{\partial c} + 2 \frac{\partial r_t}{\partial c} + c \frac{\partial^2 r_t}{\partial^2 c} \quad (5.A.98)
\]

The result of this modification is that the symmetry of partial derivatives is violated; therefore, no convergence threshold can be used in the generalized leapfrog integrator to produce a volume preserving proposal. In section 5.A.13 we give a theoretical treatment of what occurs when the Metropolis-Hastings correction is applied without properly accounting for the change in volume due to the proposal. We analyze the effect of this modification in fig. 5.A.28. We see that although reversibility may be reduced via a diminished threshold, it is not possible to produce a volume-preserving proposal in the presence of mis-specified sensitivity equations that break the symmetry of partial derivatives. In terms of ergodicity, the failure to account for substantial changes in volume has destroyed the stationarity property of the desired target distribution and the ergodicity metric we employ reveals degraded performance relative to a correct implementation of the Fitzhugh-Nagumo sensitivity equations. It is worth observing that despite the incorrectly specified derivatives, volume preservation is still preserved to around one decimal digit, which explains why samples produced by this incorrect procedure, while certainly degraded, are not absolutely awful, as shown in fig. 5.A.28c (around 1.5 decimal digits of similarity.
with the target posterior as measured by Kolmogorov-Smirnov statistics along random one-dimensional sub-spaces).
Figure 5.A.29: We show how the estimated volume preservation of the generalized leapfrog integrator is affected by the choice of finite difference perturbation size. We report the degree of volume-preservation for the generalized leapfrog integrator with a convergence threshold of $\delta = 1 \times 10^{-9}$.

5.A.11 Sensitivity of Jacobian Determinant Estimates to Perturbation Size

We evaluate the sensitivity of our estimates of volume preservation to the finite difference perturbation size. In fig. 5.A.29 we show the distribution of Jacobian determinants for all of our experiments for perturbations in the set

$$\{1 \times 10^{-8}, 1 \times 10^{-7}, 1 \times 10^{-6}, 1 \times 10^{-5}, 1 \times 10^{-4}, 1 \times 10^{-3}\}.$$ (5.A.99)
Our criterion to select a perturbation size is to identify which perturbation produces an estimated *median* Jacobian determinant that is closest to zero when using a convergence threshold of $1 \times 10^{-9}$. For the banana-shaped distribution in fig. 5.A.29a, the selected perturbation is $1 \times 10^{-5}$. For the Bayesian logistic regression model in fig. 5.A.29b, the selected perturbation is $1 \times 10^{-5}$. For Neal’s funnel distribution in fig. 5.A.29c, the selected perturbation is $1 \times 10^{-5}$. For the stochastic volatility model in fig. 5.A.29d, the selected perturbation is $1 \times 10^{-4}$. For the Cox-Poisson model in fig. 5.A.29e, the selected perturbation is $1 \times 10^{-3}$. For the Fitzhugh-Nagumo posterior in fig. 5.A.29f, the selected perturbation is $1 \times 10^{-5}$. For the multiscale Student-$t$ distribution with $\sigma^2 = 1 \times 10^4$ in fig. 5.A.29g, the selected perturbation is $1 \times 10^{-5}$. 
Figure 5.A.30: We measure the relative error in reversibility by normalizing the absolute reversibility error by the norm of initial condition. This measure contextualizes the reversibility error in terms of the scale of the distribution.

5.A.12 Relative Error of Reversibility

One may also choose to measure errors in reversibility and volume preservation in a relative sense. As discussed in section 5.A.7.2, absolute error and relative error coincide when measuring errors in volume preservation. However, one can measure relative error in the case of reversibility, as discussed in section 5.A.7.1. The use of relative error in measuring reversibility permits us to contextualize the absolute error in reversibility in terms of the norm of the initial position in phase space; this is made precise in eq. (5.A.50). We show in fig. 5.A.30 the relative error in reversibility for each of the posterior distributions con-
sidered in the main text. Qualitatively the relative error is similar to the absolute error in
that, as expected, it is a decreasing function of the convergence threshold.
5.A.13 The Stationary Distribution of a Reversible Proposal with Non-Unit Jacobian Determinant

Proposition 5.A.13.1. Let \( \Phi : \mathbb{R}^m \to \mathbb{R}^m \) be a smooth function that is self-inverse; that is, \( \Phi = \Phi^{-1} \). Denote the Jacobian determinant of \( \Phi \) at \( z \in \mathbb{R}^m \) by \( J_\Phi(z) \). Then

\[
J_\Phi(\Phi(z)) = \frac{1}{J_\Phi(z)}.
\] (5.A.100)

Proof. By the inverse function theorem

\[
\text{Id} = \nabla \Phi^{-1}(\Phi(z)) \cdot \nabla \Phi(z).
\] (5.A.101)

Taking the determinant on both sides gives,

\[
1 = J_{\Phi^{-1}}(\Phi(z)) J_\Phi(z)
\] (5.A.102)

\[
\implies J_{\Phi^{-1}}(\Phi(z)) = \frac{1}{J_\Phi(z)}.
\] (5.A.103)

The fact that \( \Phi = \Phi^{-1} \) then yields the desired conclusion. \( \square \)

Proposition 5.A.13.2. Let \( \pi : \mathbb{R}^m \to \mathbb{R}_+ \) be a density and let \( \Phi : \mathbb{R}^m \to \mathbb{R}^m \) be a smooth, self-inverse function. Consider a Markov chain at state \( z \in \mathbb{R}^m \) and transitions to state \( z' = \Phi(z) \) with probability \( \min \{1, \pi(\Phi(z))/\pi(z)\} \). Denote this Markov chain transition operator by

\[
\Psi(z) = \begin{cases} 
\Phi(z) & \text{w.p. } \frac{\pi(\Phi(z))}{\pi(z)} \\
z & \text{otherwise.}
\end{cases}
\] (5.A.104)
Then the stationary distribution of the chain is

$$\bar{\pi}(z) \propto \pi(z) \cdot \sqrt{|J_\Phi(z)|}. \quad (5.A.105)$$

**Proof.** The detailed balance condition states that for all regions $A, B \subset \mathbb{R}^m$ we have

$$\Pr_{z \sim \bar{\pi}}[z \in A \text{ and } \Phi(z) \in B] = \Pr_{z \sim \bar{\pi}}[z \in B \text{ and } \Phi(z) \in A]. \quad (5.A.106)$$

Let $Z = \int_{\mathbb{R}^m} \pi(z) \cdot \sqrt{|J_\Phi(z)|} \, dz$ be the normalizing constant of $\bar{\pi}(z)$. We have,

$$\int_{\mathbb{R}^m} \mathbf{1}_{\{z \in A\}} \cdot \mathbf{1}_{\{\Phi(z) \in B\}} \cdot \bar{\pi}(z) \cdot \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \, dz \quad (5.A.107)$$

$$= \frac{1}{Z} \int_{\mathbb{R}^m} \mathbf{1}_{\{z \in A\}} \cdot \mathbf{1}_{\{\Phi(z) \in B\}} \cdot \pi(z) \cdot \sqrt{|J_\Phi(z)|} \cdot \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \, dz \quad (5.A.108)$$

$$= \frac{1}{Z} \int_{\mathbb{R}^m} \mathbf{1}_{\{\Phi(z') \in A\}} \cdot \mathbf{1}_{\{z' \in B\}} \cdot \pi(\Phi(z')) \cdot \sqrt{|J_\Phi(\Phi(z'))|} \cdot \min \left\{ 1, \frac{\pi(z')}{\pi(\Phi(z'))} \right\} \cdot |J_\Phi(z')| \, dz' \quad (5.A.109)$$

$$= \frac{1}{Z} \int_{\mathbb{R}^m} \mathbf{1}_{\{\Phi(z') \in A\}} \cdot \mathbf{1}_{\{z' \in B\}} \cdot \pi(z') \cdot \sqrt{\frac{1}{|J_\Phi(z')|}} \cdot \min \left\{ 1, \frac{\pi(z')}{\pi(\Phi(z'))} \right\} \cdot |J_\Phi(z')| \, dz' \quad (5.A.110)$$

$$= \frac{1}{Z} \int_{\mathbb{R}^m} \mathbf{1}_{\{\Phi(z') \in A\}} \cdot \mathbf{1}_{\{z' \in B\}} \cdot \pi(z') \cdot \sqrt{|J_\Phi(z')|} \cdot \min \left\{ 1, \frac{\pi(z')}{\pi(z)} \right\} \, dz' \quad (5.A.111)$$

and similarly,

$$\int_{\mathbb{R}^m} \mathbf{1}_{\{z \in A\}} \cdot \mathbf{1}_{\{z \in B\}} \cdot \bar{\pi}(z) \cdot \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \right) \, dz = \int_{\mathbb{R}^m} \mathbf{1}_{\{z \in B\}} \cdot \mathbf{1}_{\{z \in A\}} \cdot \bar{\pi}(z) \cdot \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \right) \, dz. \quad (5.A.113)$$
These two equations verify the symmetry of $A$ and $B$ in the Metropolis-Hastings accept-reject decision. Therefore,

\[
\Pr_{z \sim \pi} [z \in A \text{ and } \Phi(z) \in B] = \int_{\mathbb{R}^m} 1 \{z \in A\} \cdot 1 \{\Phi(z) \in B\} \cdot \tilde{\pi}(z) \cdot \min \left\{1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \, dz \\
+ \int_{\mathbb{R}^m} 1 \{z \in A\} \cdot 1 \{z \in B\} \cdot \tilde{\pi}(z) \cdot \left(1 - \min \left\{1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \right) \, dz
\]

(5.A.114)

\[
= \int_{\mathbb{R}^m} 1 \{z \in B\} \cdot 1 \{\Phi(z) \in A\} \cdot \tilde{\pi}(z) \cdot \min \left\{1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \, dz \\
+ \int_{\mathbb{R}^m} 1 \{z \in B\} \cdot 1 \{z \in A\} \cdot \tilde{\pi}(z) \cdot \left(1 - \min \left\{1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \right) \, dz
\]

(5.A.115)

\[
= \Pr_{z \sim \pi} [z \in B \text{ and } \Phi(z) \in A]
\]

(5.A.116)

Since $\tilde{\pi}$ satisfies detailed balance with respect to the transition operator, it follows that $\tilde{\pi}$ is the stationary distribution of the Markov chain with transition operator $\Psi$. \hfill \Box

This result demonstrates what can go wrong when the Metropolis-Hastings accept-reject rule does not account for the change-of-volume of the function $\Phi$. We can interpret $\tilde{\pi}(z)$ as proportional to a perturbation of $\pi(z)$ according to the square-root of the Jacobian determinant of the transformation $\Phi$ at $z$. More formally,

**Proposition 5.A.13.3.**

\[
\text{KL}(\pi \parallel \tilde{\pi}) = \log \mathbb{E}_{z \sim \pi} \sqrt{|J_{\Phi}(z)|} - \mathbb{E}_{z \sim \pi} \log \sqrt{|J_{\Phi}(z)|}
\]

(5.A.117)

\[
\geq 0.
\]

(5.A.118)
Proof.

\[
\text{KL}(\pi \| \bar{\pi}) = \mathbb{E}_{z \sim \pi} \log \frac{\pi(z)}{\bar{\pi}(z)} \tag{5.A.119}
\]

\[
= \mathbb{E}_{z \sim \pi} \left[ \log \pi(z) - \log \pi(z) - \log \sqrt{J_\Phi(z)} + \log Z \right] \tag{5.A.120}
\]

\[
= \log Z - \mathbb{E}_{z \sim \pi} \log \sqrt{J_\Phi(z)} \tag{5.A.121}
\]

Now recall

\[
Z = \int_{\mathbb{R}^m} \pi(z) \cdot \sqrt{|J_\Phi(z)|} \, dz \tag{5.A.122}
\]

\[
= \mathbb{E}_{z \sim \pi} \sqrt{|J_\Phi(z)|}, \tag{5.A.123}
\]

which proves equality. Non-negativity then follows from Jensen’s inequality. \qed
5.A.14  Ergodicity Issues in Fixed Point Iteration

We now turn our attention to developing a theory for the pathologies that can occur when fixed point iterations fail to converge.

**Lemma 5.A.14.1.** Let \( \Phi : \mathbb{R}^m \to \mathbb{R}^m \) be an involution and define the set

\[
S = \{ x \in \mathbb{R}^m : \Phi(x) \neq x \}.
\]

(5.A.124)

Let \( x \in S \). Then, if \( y = \Phi(x) \), we have \( y \in S \).

**Proof.** We first establish that \( \Phi(y) \) is non-trivial. We have,

\[
\Phi(y) = \Phi(\Phi(x)) \quad (5.A.125)
\]

\[
= x \quad (5.A.126)
\]

\[
\neq \Phi(x) \quad (5.A.127)
\]

\[
= y. \quad (5.A.128)
\]

Hence, \( \Phi(y) \neq y \).

**Lemma 5.A.14.2.** Let \( \pi : \mathbb{R}^m \to \mathbb{R}_+ \) be a probability density on \( \mathbb{R}^m \) with respect to a measure \( \mu : \mathcal{B}(\mathbb{R}^m) \to \mathbb{R}_+ \). Let \( \Phi \) and \( S \) be as in lemma 5.A.14.1 and assume further that \( \Phi \) is a diffeomorphism except on a \( \mu \)-negligible set. Then involutive Monte Carlo, with involution \( \Phi \), satisfies detailed balance with respect to \( \pi \).

**Proof.** For \( \mu \)-almost all \( x \), we may take the involutive Monte Carlo transition kernel to be,

\[
K(x, A) = \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} 1 \{ \Phi(x) \in A \}
\]

\[
+ \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} \right) 1 \{ x \in A \}. \quad (5.A.129)
\]
Let $B \in \mathfrak{B}(\mathbb{R}^m)$. Then,

\[
\int_B \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} \mathbf{1} \{ \Phi(x) \in A \} \pi(x) \mu(dx) \quad (5.A.130)
\]

\[
= \int_{B \cap \Phi(A)} \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} \pi(x) \mu(dx) \quad (5.A.131)
\]

\[
= \int_{B \cap \Phi(A)} \min \{ \pi(x), \pi(\Phi(x)) |\det(\nabla \Phi(x))| \} \mu(dx) \quad (5.A.132)
\]

\[
= \int_{B \cap \Phi(A)} \min \{ \pi(x) |\det(\nabla \Phi(\Phi(x)))|, \pi(\Phi(x)) \} |\det(\nabla \Phi(x))| \mu(dx) \quad (5.A.133)
\]

\[
= \int_{\Phi(B) \setminus A} \min \{ \pi(\Phi(x)) |\det(\nabla \Phi(x))|, \pi(x) \} \mu(dx) \quad (5.A.134)
\]

\[
= \int_A \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} \mathbf{1} \{ \Phi(x) \in B \} \pi(x) \mu(dx). \quad (5.A.135)
\]

Moreover,

\[
\int_B \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} \right) \mathbf{1} \{ x \in A \} \pi(x) \mu(dx) \quad (5.A.136)
\]

\[
= \int_A \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\} \right) \mathbf{1} \{ x \in B \} \pi(x) \mu(dx). \quad (5.A.137)
\]

These two equations together imply detailed balance.

However, the division of $\mathbb{R}^m$ into $S$ and $\mathbb{R}^m \setminus S$ suggests a failure of ergodicity since $S$ is closed under $\Phi$ from lemma 5.A.14.1. Therefore:

**Corollary 5.A.14.3.** Let $\pi : \mathbb{R}^m \to \mathbb{R}_+$ be a probability density on $\mathbb{R}^m$ with respect to a measure $\mu : \mathfrak{B}(\mathbb{R}^m) \to \mathbb{R}_+$. Let $\Phi$ and $S$ be as in lemma 5.A.14.1 and assume further that $\Phi$ is a diffeomorphism except on a $\mu$-negligible set. Then involutive Monte Carlo, with involution $\Phi$, satisfies detailed balance on $S$ with respect to

\[
\tilde{\pi}(x) = \frac{\pi(x)}{\text{Pr}(x \in S)} \mathbf{1} \{ x \in S \}. \quad (5.A.138)
\]

**Proof.** It suffices to observe that for any $A \subseteq S$ we have that $\Phi(A) \subseteq S$ and that, for
\[ x \in S, \]
\[
\min \left\{ 1, \frac{\pi(\Phi(x))}{\hat{\pi}(x)} |\det(\nabla \Phi(x))| \right\} = \min \left\{ 1, \frac{\hat{\pi}(\Phi(x))}{\pi(x)} |\det(\nabla \Phi(x))| \right\}. \quad (5.A.139)
\]

A calculation nearly identical to the proof of lemma 5.A.14.2 then yields the result. □

**Example 36.** Let us now see an example of how this theory may be applied. For certain, pathological vector fields, the generalized leapfrog integrator may not have a solution to one of its implicit updates, as we saw in the case of the banana-shaped distribution in section 5.A.9. In the case wherein these fixed point iterations do not converge, one may decide to simply remain at the current state. Therefore, there is a subset \( S \) of the phase space in which the involution is set as the identity. The above calculations show the danger of this circumstance: Not only may the non-existence of solutions to the implicit updates cause slower mixing (by remaining at the current state), but they may also yield an ergodicity failure if reducibility in the sense of corollary 5.A.14.3 occurs.}
Chapter 6

Several Remarks on the Numerical Integrator in Lagrangian Monte Carlo

This chapter is adapted from [Brofos and Lederman 2022]. This research represents my investigations into an alternative sampling procedure based on the Lagrangian formalism rather than the Hamiltonian one. This chapter makes several contributions towards the understanding of numerical integrators based on Lagrangian mechanics, with particular attention given to their role in Markov chain Monte Carlo. In particular, this chapter identifies and clarifies a misunderstanding of the numerical order of these methods, and proposes mechanisms by which to accelerate the numerical method. Composition and experimentation are original to me.

Abstract. Riemannian manifold Hamiltonian Monte Carlo (RMHMC) is a powerful method of Bayesian inference that exploits underlying geometric information of the posterior distribution in order to efficiently traverse the parameter space. However, the form of the Hamiltonian necessitates complicated numerical integrators, such as the generalized leapfrog method, that preserve the detailed balance condition. The distinguishing feature of these numerical integrators is that they involve solutions to implicitly defined equations. Lagrangian Monte Carlo (LMC) proposes to eliminate the fixed point iterations by transitioning from the Hamiltonian formalism to Lagrangian dynamics, wherein a fully explicit...
integrator is available. This work makes several contributions regarding the numerical integrator used in LMC. First, it has been claimed in the literature that the integrator is only first-order accurate for the Lagrangian equations of motion; to the contrary, we show that the LMC integrator enjoys second order accuracy. Second, the current conception of LMC requires four determinant computations in every step in order to maintain detailed balance; we propose a simple modification to the integration procedure in LMC in order to reduce the number of determinant computations from four to two while still retaining a fully explicit numerical integration scheme. Third, we demonstrate that the LMC integrator enjoys a certain robustness to human error that is not shared with the generalized leapfrog integrator, which can invalidate detailed balance in the latter case. We discuss these contributions within the context of several benchmark Bayesian inference tasks.

6.1 Introduction

Let $L : \mathbb{R}^m \to \mathbb{R}$ be the log-density of a smooth distribution, known up to an additive constant. A critical problem in Bayesian inference is the design of procedures that generate samples from a target density $\pi(q) \propto \exp(L(q))$. Hamiltonian Monte Carlo (HMC) [Duane et al., 1987, Betancourt 2017] is a Markov chain Monte Carlo (MCMC) method for sampling from arbitrary differentiable probability distributions based on numerical solutions to Hamilton’s equations of motion. Unlike random walk Metropolis or the Metropolis-adjusted Langevin algorithm, HMC can propose distant states, thereby dramatically decreasing the autocorrelation between states of the chain and increasing sampling efficiency.

The efficiency of HMC can be further improved by incorporating geometric concepts into the proposal mechanism. Information geometry [Amari 2016] provides a framework for representing sets of probability densities as a Riemannian manifold whose metric can be chosen as the Fisher information matrix. When $L$ decomposes into the sum of a log-
likelihood and a log-prior terms, Girolami and Calderhead [2011] proposed Riemannian manifold Hamiltonian Monte Carlo (RMHMC), which sought to precondition Hamiltonian dynamics with the inverse of the sum of Fisher information of the log-likelihood and the negative Hessian of the log-prior. In general, both the Fisher information and the Hessian of the log-prior depend on the parameter, denoted here by $q$, to be sampled; this dependency produces a complicated Hamiltonian. In order to maintain the detailed balance condition of the RMHMC Markov chain, elaborate numerical integrators such as the generalized leapfrog or implicit midpoint methods must be used [Pourzanjani and Petzold, 2019, Girolami and Calderhead, 2011, Brofos and Lederman, 2021a]. These numerical integrators involve implicitly defined updates, wherein updates are defined as the solution to a fixed point equation, which are typically resolved to a prescribed convergence tolerance by fixed point iteration.

Lan et al. [2015] introduced Lagrangian Monte Carlo (LMC) as an alternative to RMHMC. Unlike RMHMC, which is based on numerical solutions to Hamiltonian mechanics, LMC is instead inspired by the Lagrangian formalism of classical mechanics. Although Lagrangian and Hamiltonian mechanics are formally equivalent from a physical perspective, methods of numerical integration assume a simplified form when expressed as Lagrangian mechanics. Indeed, the simplification is so significant that Lan et al. [2015] was able to devise a fully explicit numerical integrator for which detailed balance could be maintained when used as a proposal operator in MCMC. However, the elimination of fixed point iterations was replaced by the requirement that four Jacobian determinants be computed in a single step of the explicit integrator. Moreover, it was stated that the explicit method had only first-order accuracy as a numerical integrator, supposedly producing less accurate solutions than the generalized leapfrog method, whose accuracy is second-order.

The purpose of the present work is three-fold. First, we propose a simple mechanism by which to reduce the number of Jacobian determinant evaluations from four to two in a single step of the integrator. This is achieved by inverting the sequence in which position
and velocity are integrated. Second, we will clarify that the order of the explicit integrator, with or without inversion, actually has second-order accuracy, the same as those integrators which are commonly used in RMHMC. Third, we discuss how LMC enjoys a greater robustness to human error than RMHMC. The outline of the remainder of this paper is as follows. In section 6.2 we discuss preliminary material on Hamiltonian and Lagrangian mechanics, numerical integrators, and MCMC methods based off of these physical models. In section 6.3 we review some related work in the literature on integration methods for RMHMC. In section 6.4 we proceed to our analysis of the LMC integrator, where we describe how one may reduce the number of Jacobian determinant computations while maintaining a fully explicit integration method and give a proof that the integrator has second-order accuracy. In section 6.5 we turn to the evaluation of the proposed modifications to LMC; we evaluate performance on several benchmark Bayesian inference tasks and give numerical evidence to support the claim of second-order accuracy.

### 6.2 Preliminaries

In this section we review the necessary background for Hamiltonian and Lagrangian Monte Carlo. Section 6.2.2 reviews the most important concepts from RMHMC and LMC, giving perspective on the varieties of integrators and how they are employed in a Markov chain Monte Carlo procedure. We then proceed in section 6.A.2.1 to recall the Hamiltonian and Lagrangian formalisms from physics. Section 6.A.2.2 then treats the matter of numerical integration of the Hamiltonian and Lagrangian mechanics. In section 6.A.2.3 we then review Hamiltonian and Lagrangian Monte Carlo using the framework of involutive Monte Carlo, wherein we give special attention to the Jacobian determinant computations that are necessitated in the Lagrangian construction.
6.2.1 Notation

In the context of classical mechanics, we denote by $q$ the position variable, $v$ the velocity variable, $p$ the momentum variable, and $a$ the acceleration variable, all elements of $\mathbb{R}^m$. We adopt the notation $q^{(k)}$ to refer to the $k$-th element of $q$, with similar conventions being employed for $v$ and $p$. Denoting $z = (q, p)$; we call $z$ a point in phase space. We denote by $\text{PD}(m)$ the set of $m \times m$ positive definite matrices. Given a map $\Phi : \mathbb{R}^m \to \mathbb{R}^m$, we use the notation $\Phi^k$ to mean the composition $\Phi \circ \cdots \circ \Phi$ ($k$ times). We write $\text{Id}_m$ to denote the identity matrix of size $m \times m$. We denote the Borel $\sigma$-algebra on $\mathbb{R}^m$ by $\mathcal{B}(\mathbb{R}^m)$.

6.2.2 Lagrangian Monte Carlo: The Essentials

Lagrangian Monte Carlo (LMC) is a geometric method of Bayesian inference that seeks to incorporate second-order information about the posterior in order to produce effective proposals, similar to Riemannian manifold Hamiltonian Monte Carlo (RMHMC). In this section we review the fundamentals of these methods. For applications in HMC, an important class of Hamiltonians have the following form:

**Definition 6.2.1. The Riemannian Hamiltonian**

\[
H(q, p) = U(q) + K(q, p)
\]  

(6.1)

where $U : \mathbb{R}^m \to \mathbb{R}$ is called the potential energy function and $K : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is the kinetic energy function, having the form $K(q, p) = \frac{1}{2} p^\top G^{-1}(q)p$, where $G : \mathbb{R}^m \to \text{PD}(m)$ where $G$ is called the metric.

The Riemannian metric defines Christoffel symbols which convey information about the curvature and shape of $\mathbb{R}^m$ imbued with the metric $G$. 
Definition 6.2.2. The Christoffel symbols are the $m^3$ functions defined by

$$\Gamma^k_{ij}(q) = \frac{1}{2} \sum_{l=1}^{m} G^{-1}_{kl}(q) \left( \frac{\partial}{\partial q^{(i)}} G_{lj}(q) + \frac{\partial}{\partial q^{(j)}} G_{li}(q) - \frac{\partial}{\partial q^{(l)}} G_{ij}(q) \right). \quad (6.2)$$

We note that the Christoffel symbols are symmetric in their lower indices (i.e. $\Gamma^k_{ij}(q) = \Gamma^k_{ji}(q)$). The Christoffel symbols play a prominent role in the development of Lagrangian mechanics. As a notational convenience, we will define the matrix-valued function $\Omega : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^{m \times m}$ whose $(i, j)$-th entry is

$$\Omega_{ij}(\epsilon, q, v) = \frac{\epsilon}{2} \sum_{k=1}^{m} \Gamma^i_{kj}(q) v^{(k)}. \quad (6.3)$$

The Riemannian Hamiltonian in definition 6.2.1 produces equations of motion (see section 6.A.2.1 for details), which do not have closed-form solutions. This necessitates the use of numerical integrators. We now review two integrators that form the basis of our evaluations: the generalized leapfrog integrator, which is a reversible, volume-preserving, and second-order accurate, and the Lagrangian leapfrog method of Lan et al. [2015].

Definition 6.2.3. The **generalized leapfrog integrator** for the Hamiltonian equations of motion in eqs. (6.A.1) and (6.A.2) is a map $(q, p) \mapsto (\tilde{q}, \tilde{p})$ defined by,

$$\dot{p}^{(k)} = p^{(k)} - \frac{\epsilon}{2} \left( -\frac{1}{2} \frac{v^T}{2} G^{-1}(q) \left( \frac{\partial G}{\partial q^{(k)}}(q) \right) G^{-1}(q) \dot{p} + \frac{\partial U}{\partial q^{(k)}}(q) \right) \quad (6.4)$$

$$\tilde{q} = q + \frac{\epsilon}{2} \left( G^{-1}(q) \dot{p} + G^{-1}(\tilde{q}) \tilde{p} \right) \quad (6.5)$$

$$\tilde{p}^{(k)} = \dot{p}^{(k)} - \frac{\epsilon}{2} \left( -\frac{1}{2} \frac{v^T}{2} G^{-1}(q) \left( \frac{\partial G}{\partial q^{(k)}}(q) \right) G^{-1}(q) \dot{p} + \frac{\partial U}{\partial q^{(k)}}(\tilde{q}) \right). \quad (6.6)$$

Pseudo-code implementing the generalized leapfrog algorithm is given in algorithm 13 in section 6.A.1.

Definition 6.2.4. The **Lagrangian leapfrog integrator** for the Lagrangian equations of mo-
The function given in eq. (6.A.7) is a map \((q, v) \mapsto (\tilde{q}, \tilde{v})\) defined by,

\[
\tilde{v} = \left[ \text{Id}_m + \Omega(\epsilon, q, v) \right]^{-1} \left[ v - \frac{\epsilon}{2} G^{-1}(q) \nabla U(q) \right] \quad (6.7)
\]

\[
\tilde{q} = q + \epsilon \tilde{v} \quad (6.8)
\]

\[
\tilde{v} = \left[ \text{Id}_m + \Omega(\epsilon, \tilde{q}, \tilde{v}) \right]^{-1} \left[ \tilde{v} - \frac{\epsilon}{2} G^{-1}(\tilde{q}) \nabla U(\tilde{q}) \right]. \quad (6.9)
\]

Pseudo-code implementing the Lagrangian leapfrog algorithm is given in algorithm 14 in section 6.A.1. Unlike the generalized leapfrog integrator (definition 6.2.3), which is a symplectic transformation and therefore necessarily volume-preserving, the Lagrangian leapfrog (definition 6.2.4) is not volume-preserving. Its Jacobian determinant is computed in eq. (6.A.18). A thorough treatment of numerical integrators is provided in section 6.A.2.2. In the context of Monte Carlo, this means that Markov chains constructed from repeated applications of the Lagrangian leapfrog integrator will require a Jacobian determinant computation, whereas methods based on the generalized leapfrog will not (its Jacobian determinant is one). Such a Markov chain is the subject of the following example.

**Example 37.** Let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be as in definition 6.2.1 and define a probability density \( \pi(q, p) \propto \exp(-H(q, p)) \). Let \( q \in \mathbb{R}^m \) be given; a single Markov chain step is constructed as follows. Sample \( p \sim \text{Normal}(0, G(q)) \). Fix \( k \in \mathbb{N} \). We consider RMHMC and LMC separately:

**LMC** Let \( \tilde{\Phi}_\epsilon \) denote the Lagrangian leapfrog (definition 6.2.4). Compute the proposal \((\tilde{q}, \tilde{v}) = \tilde{\Phi}(q, G^{-1}(q)p)\) and set \( \tilde{p} = G(\tilde{q})\tilde{v} \). Compute the Jacobian determinant \( J \) of the map \((q, p) \mapsto (\tilde{q}, \tilde{p})\) using eq. (6.A.18).

**RMHMC** Let \( \hat{\Phi}_\epsilon \) denote the generalized leapfrog integrator (definition 6.2.3). Compute the proposal \((\tilde{q}, \tilde{p}) = \hat{\Phi}(q, p)\) and set \( J = 1 \).

Accept the proposal state \( \tilde{q} \) with probability \( \alpha((q, p), (\tilde{q}, \tilde{p}), J) = \min \left\{ 1, \frac{\pi(\tilde{q}, \tilde{p})}{\pi(q, p)} \cdot J \right\} \); otherwise remain at the current state \( q \).  

333
A more rigorous treatment of the LMC and RMHMC Markov chains is given in section 6.A.2.3 using the framework of diffeomorphism Monte Carlo. Pseudo-code is provided in algorithm 15 in section 6.A.1.

6.3 Related Work

The focus of the present work is to investigate the numerical methods of integration that were proposed in [Lan et al. 2015]. Brofos and Lederman [2021a] gave an evaluation of the implicit midpoint integrator for RMHMC with special attention paid to the errors in reversibility and volume preservation that were produced by the implicit midpoint algorithm compared to the generalized leapfrog method, as well as the energy conservation properties enjoyed by the implicit midpoint integrator. Other mechanisms of explicit integration have been considered with applications to RMHMC foremost in mind, such as Cobb et al. [2019] which produced a reversible, volume-preserving numerical method in an expanded phase-space. Due to the expansion of phase-space, this integrator cannot be used to produce a Markov chain satisfying detailed balance. The work of Zhang and Sutton [2014] explored alternating blockwise Metropolis-within-Gibbs-like strategies with Riemannian metrics chosen to produce separable Hamiltonians within each block; each block can then be integrated using the standard leapfrog integrator. LMC has previously been criticized in the literature for having unfavorable performance in high dimensions; this failure of LMC in relation to RMHMC is discussed in Betancourt et al. [2014] and we will see evidence of this degradation in section 6.5.4.

6.4 Analytical Apparatus

In this section we describe an algorithmic recommendation for the numerical integrator used in Lagrangian Monte Carlo and we clarify certain statements around the order of
this numerical method. Specifically, we show how to reduce the number of determinant computations from four to two, and that the integrator of Lagrangian dynamics has third-order local error, comparable to the error of the (generalized) leapfrog method used in HMC.

### 6.4.1 Inversion of the Integration Sequence

A disadvantage of the Lagrangian integrator is that it involves four Jacobian determinant computations at each step of the integrator. In general, computing the Jacobian determinant of an $m \times m$ matrix incurs a computational cost like $O(m^3)$. Therefore, it seems worthwhile to investigate mechanisms by which to reduce the number of these calculations that are required. [Lan et al. 2015] proposed one method that has only two Jacobian determinant computations, but necessitates the return to implicit methods of integration.

To retain the advantages of explicit integration, we propose a method that computes two Jacobian determinants in every step and consists only of explicit integration steps. To achieve this, we propose a conceptually simple procedure: invert the sequence in which position and velocity are updated in the Lagrangian integrator so that position is updated twice at the beginning and end of each step and velocity is updated once in between each update to position. Formally:

**Definition 6.4.1.** The *inverted Lagrangian leapfrog integrator* for the Lagrangian equations of motion given in eq. (6.A.7) is a map $(q, v) \mapsto (\tilde{q}, \tilde{v})$ defined by,

\[
\begin{align*}
\tilde{q} &= q + \frac{\epsilon}{2} v \\
\tilde{v} &= [\text{Id}_m + 2\Omega(\epsilon, \tilde{q}, v)]^{-1} \left[ v - \epsilon G^{-1}(\tilde{q}) \nabla U(\tilde{q}) \right] \\
\tilde{q} &= \tilde{q} + \frac{\epsilon}{2} \tilde{v}.
\end{align*}
\]

Pseudo-code for this procedure is provided in algorithm 16 in section 6.A.1. The basic
modification requires only two Jacobian determinant computations per step since the two updates to position in eqs. (6.10) and (6.12), being shear transformations, are volume-preserving in \((q, v)\)-space [Modi et al., 2021]. The required change in volume due to the mapping \((q, p) \mapsto (\tilde{q}, \tilde{p})\) is readily obtained as

\[
\det \left( \frac{\partial (\tilde{q}, \tilde{p})}{\partial (q, p)} \right) = \frac{\det(G(\tilde{q})) \det(\text{Id}_m - \Omega(\epsilon, \dot{q}, \tilde{v}))}{\det(G(q)) \det(\text{Id}_m + \Omega(\epsilon, \dot{q}, v))}. \tag{6.13}
\]

Although this Jacobian determinant differs from that produced by algorithm 14 in section 6.A.1, we still have the following important property.

**Lemma 6.4.2.** The inverted Lagrangian leapfrog integrator is self-adjoint and has at least first-order local error.

A proof is given in section 6.A.4.

**Corollary 6.4.3.** The inverted Lagrangian leapfrog integrator has at least second-order local error.

**Proof.** This follows as an immediate corollary of lemma 6.4.2. \(\square\)

**Definition 6.4.4.** Let \(\Phi_\epsilon\) be the inverted Lagrangian leapfrog integrator with step-size \(\epsilon \in \mathbb{R}\) (definition 6.4.1). Let \(k \in \mathbb{N}\) be the number of integration steps. The inverted Lagrangian Monte Carlo (ILMC) is an instance involutive Monte Carlo (definition 6.A.2.16) with involution \(F \circ \Phi^k_\epsilon\) where \(F\) is the momentum flip operator (definition 6.A.2.18).

Let us denote by \(\hat{\Phi}_\epsilon\) the Lagrangian leapfrog integrator (definition 6.2.4) and \(\check{\Phi}_\epsilon\) the inverted Lagrangian leapfrog (definition 6.4.1). Let \(\Phi_\epsilon\) be the exact time \(\epsilon\) solution of eq. (6.A.7). Because both methods are second-order accurate, it follows that

\[
\| \hat{\Phi}_\epsilon - \check{\Phi}_\epsilon \| \leq \| \hat{\Phi}_\epsilon - \Phi_\epsilon \| + \| \check{\Phi}_\epsilon - \Phi_\epsilon \|
\]

\[
= \mathcal{O}(\epsilon^3) \tag{6.15}
\]
Figure 6.1: Visualization of the difference in expected distance when integrating a trajectory of a separable, quadratic Hamiltonian for a varying number of steps and integration step-sizes. This difference is always non-negative, indicating that one expects a greater distance between position variables when integrating with the leapfrog, rather than inverted leapfrog, algorithm.

Intuitively, in the limit of small step-sizes, the difference in proposals generated by the Lagrangian leapfrog and the inverted Lagrangian leapfrog will be minimal, but ILMC (definition 6.4.4) involves half the number of Jacobian determinant computations compared to LMC (definition 6.A.2.20) and should therefore be preferred. On the other hand, for large step-sizes, the situation is less clear.

### 6.4.2 Aversions to Inverting the Integration Sequence

In the case when $G = \text{Id}_m$, the Lagrangian integrator devolve into the standard leapfrog integrator (definition 6.A.5.1) that is ubiquitous in HMC. There are good reasons why implementations of HMC integrate in the order of an initial half-step in momentum, a full step in position, and then a second half-step of momentum. This is described visually by Bou-Rabee and Sanz-Serna [2018] for the Hamiltonian $H(q, p) = q^2/2 + p^2/2$, who make the argument that the leapfrog integrator produces high acceptance probabilities along the $q$-axis, which are desirable. By contrast, inverted leapfrog (definition 6.A.5.2) produces large acceptance probabilities along the $p$-axis, which are not valuable in HMC. Here we
wish to expand on this example to consider what happens when HMC Markov chains based on the standard leapfrog and inverted leapfrog are initialized at stationarity. Based on the symmetric roles of \( q \) and \( p \) in the Hamiltonian, it is tempting to conclude that these Markov chains would exhibit similar performance profiles; surprisingly, this is not the case as shown in the following example.

**Example 38.** Consider a Hamiltonian of the form \( H(q, p) = \omega^2 q^2 / 2 + p^2 / 2 \). This Hamiltonian corresponds to the distributions \( q \sim \text{Normal}(0, 1/\omega^2) \) and \( p \sim \text{Normal}(0, 1) \). Let \( \hat{\Phi}_\epsilon : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m \times \mathbb{R}^m \) and \( \tilde{\Phi}_\epsilon : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m \times \mathbb{R}^m \) be the single step leapfrog (definition 6.A.5.1) and inverted leapfrog (definition 6.A.5.2) methods with step-size \( \epsilon \), respectively. Let \( \text{Proj}_q(q, p) = q \) be the projection onto the \( q \)-variables. Then,

\[
\mathbb{E} \left[ (\text{Proj}_q \circ \hat{\Phi}_\epsilon(q, p) - q)^2 \right] = \frac{\epsilon^4 \omega^2}{4} + \epsilon^2 \quad (6.16)
\]

\[
\mathbb{E} \left[ (\text{Proj}_q \circ \tilde{\Phi}_\epsilon(q, p) - q)^2 \right] = \frac{\epsilon^4 \omega^2}{4} + \epsilon^2 \left( 1 - \frac{\epsilon^2 \omega^2}{4} \right). \quad (6.17)
\]

Both the leapfrog and inverted leapfrog methods are only numerically stable when \( \epsilon^2 \omega^2 < 4 \) [Leimkuhler and Reich, 2005]. Hence, \( \mathbb{E} \left[ (\text{Proj}_q \circ \hat{\Phi}_\epsilon(q, p) - q)^2 \right] < \mathbb{E} \left[ (\text{Proj}_q \circ \tilde{\Phi}_\epsilon(q, p) - q)^2 \right] \), which we interpret to mean that, in a single step of the integrator, the expected squared distance between initial and terminal position is greater for the leapfrog integrator than for the inverted leapfrog. By deriving the propagator matrices (see section 6.A.5) of the leapfrog and inverted leapfrog integrators, we may deduce exact formulas for the \( k \)-step quantities \( \mathbb{E} \left[ (\text{Proj}_q \circ \hat{\Phi}_\epsilon^k(q, p) - q)^2 \right] \) and \( \mathbb{E} \left[ (\text{Proj}_q \circ \tilde{\Phi}_\epsilon^k(q, p) - q)^2 \right] \). In fig. 6.1 we show the difference of these expected squared distances as a function of the number of steps and for several step-sizes. Notably, this difference is always non-negative, indicating that one expects the leapfrog to produce more distant proposals than the inverted leapfrog in the Gaussian case.

Because autocorrelation is related to the distance to subsequent samples, one expects
the inverted leapfrog method to exhibit fewer effective samples. This will be true even if LMC Markov chain is initialized in the stationary distribution. This illustrates an important short-coming of inverting the integration sequence, which must be balanced against computational savings obtained by the reduction in Jacobian determinant computations in the LMC setting.

6.4.3 The Order of the Explicit Integrators

In [Lan et al. [2015]], the authors showed that the local error rate of the Lagrangian integrator is at least $O(\epsilon^2)$. We now build on this foundation in order to deduce that the local error rate of the Lagrangian leapfrog is at least $O(\epsilon^3)$. This means that the order of the LMC integrator matches the local and global error rates of the integrators used in HMC and RMHMC. Proofs of lemmas 6.4.5 and 6.4.6 may be found in section 6.A.4.

Lemma 6.4.5. The explicit integrator of the Lagrangian dynamics has at least first-order local error.

Lemma 6.4.6. The explicit integrator of Lagrangian dynamics is self-adjoint.

Proposition 6.4.7. The explicit integrator of the Lagrangian dynamics has at least third-order local error.

Proof. From lemma 6.4.5 we know that $r \geq 1$. Suppose $r = 1$. From lemma 6.4.6 we know that the integrator is self-adjoint. From theorem 6.A.2.12 we know that the order of a self-adjoint method must be even; hence $r$ cannot be odd. But $r = 1$ by assumption, a contradiction. Therefore, it must be that $r \geq 2$ so that the explicit integrator has, at least, third-order local error.

Denote by $\hat{\Phi}_\epsilon$ the second-order integrator of Lagrangian dynamics. If $\hat{\Phi}_\epsilon(q_0, v_0) = (\hat{q}_\epsilon, \hat{v}_\epsilon)$, one wonders if $(\hat{q}_\epsilon, \hat{p}_\epsilon)$ is a second-order approximation of the Hamiltonian dy-
namics, where \( \dot{p}_e = G(\dot{q}_e)\dot{v}_e \). Indeed, this is true and follows as an immediate consequence of proposition 6.A.2.13 with diffeomorphism \((q, v) \mapsto (q, Gv)\).

### 6.4.4 Efficient Computation of the Jacobian Determinant and the Update to Velocity

The update to the velocity in eqs. (6.7), (6.9) and (6.11) and the associated change-in-volume in eq. (6.13) involve manipulations of a matrix of the form \( \text{Id} + \Omega(\epsilon, q, v) \): in the former case, we must solve a linear system involving this matrix while in the latter case we must compute the absolute value of its Jacobian determinant. An efficient procedure by which to achieve both of these objectives is to compute the PLU decomposition of \( \text{Id} + \Omega(\epsilon, q, v) = \text{PLU} \) where \( P \) is a permutation matrix, \( L \) is a lower-triangular matrix with unit diagonal, and \( U \) is an upper-triangular matrix. The computational cost of this decomposition is \( \frac{2}{3}m^3 + O(m^2) \). Linear systems can be solved using the PLU decomposition by applying the permutation and solving the triangular systems via forward-backward substitution. Moreover, the required Jacobian determinant is simply \( \prod_{i=1}^{m} U_{ii} \), since the determinant of the permutation matrix has unit magnitude and the lower-triangular matrix \( L \) has unit Jacobian determinant since all of its diagonal elements are equal to one.

### 6.4.5 Built-In Robustness of the Lagrangian Integrator

As shown in Brofos and Lederman [2021b], the volume-preservation property (i.e. \( J = 1 \) in algorithm 13 in section 6.A.1) of the generalized leapfrog integrator is predicated on the symmetry of partial derivatives \( \nabla_q^T \nabla_p H(q, p) = \nabla_p \nabla_q H(q, p) \). In an implementation of the generalized leapfrog integrator we may suppose that we have functions \( g_k(q) \) representing \( \frac{\partial G}{\partial q_{(k)}}(q) \). Substituting this function into the definition of the generalized leapfrog
integrator (definition 6.2.3) yields the following map \((q, p) \mapsto (\tilde{q}, \tilde{p})\)

\[
\tilde{p}^{(k)} = p^{(k)} - \frac{\epsilon}{2} \delta_k(q, \tilde{p}) \tag{6.18}
\]

\[
\tilde{q} = q + \frac{\epsilon}{2} \left( \Delta(q, \tilde{p}) + \Delta(\tilde{q}, \tilde{p}) \right) \tag{6.19}
\]

\[
\tilde{p}^{(k)} = \tilde{p}^{(k)} - \frac{\epsilon}{2} \delta_k(\tilde{q}, \tilde{p}), \tag{6.20}
\]

where

\[
\delta_k(q, p) = -\frac{1}{2} p^\top G^{-1}(q) g_k(q) G^{-1}(q) p + \frac{\partial U}{\partial q^{(k)}}(q) \tag{6.21}
\]

\[
\Delta(q, p) = G^{-1}(q) p \tag{6.22}
\]

When \( g_k(q) = \frac{\partial G}{\partial q^{(k)}}(q) \), the resulting map is necessarily volume-preserving. However, we may then ask the question, “What happens when \( g_k(q) \) is incorrectly implemented so that, in fact, \( g_k(q) \neq \frac{\partial G}{\partial q^{(k)}}(q) \)?” The symmetry of partial derivatives has therefore been violated since

\[
\frac{\partial \Delta}{\partial q^{(k)}}(q, p) = \frac{\partial G^{-1}}{\partial q^{(k)}}(q) p \neq -G^{-1}(q) g_k(q) G^{-1}(q) p = \frac{\partial \delta_k}{\partial p}(q, p). \tag{6.23}
\]

In the case of the RMHMC Markov chain (definition 6.A.2.19), detailed balance is no longer satisfied and there is no expectation that the RMHMC will converge to the target distribution.

The situation is different in the case of the LMC Markov chain (definition 6.A.2.20). The fundamental difference is that LMC expects the transformation \((q, v) \mapsto (\tilde{q}, \tilde{v})\) to be non-volume-preserving, hence necessitating the Jacobian determinant correction in eq. (6.A.15). To see that the change-in-volume is still correctly computed even when \( g_k(q) \neq \frac{\partial G}{\partial q^{(k)}}(q) \), we observe that the Lagrangian leapfrog’s (definition 6.2.4) update to
velocity in eq. (6.7) is a special case of the following map:

\[
\dot{v} = (\text{Id}_m + A(q, v))^{-1} (v - b(q))
\]  

(6.24)

where \(A : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^{m \times m}\) and \(b : \mathbb{R}^m \to \mathbb{R}^m\). Under the assumption that \(A(q, v)\dot{v} = A(q, \tilde{v})v\) (which holds for the correctly implemented LMC with \(A = \Omega(\epsilon, q, v)\) by proposition 6.A.2.8), the Jacobian determinant of the map \((q, v) \mapsto (q, \dot{v})\) is

\[
\left| \det \frac{\partial (q, \dot{v})}{\partial (q, v)} \right| = \left| \frac{\det(\text{Id}_m + A(q, v))}{\det(\text{Id}_m + A(q, \dot{v}))} \right|. 
\]  

(6.25)

Hence, if we adopt the notation \(g_{k,ij}(q)\) as the \((i, j)\)-th element of \(g_k(q)\), and substitute

\[
\tilde{\Gamma}_{ij}^k(q) = \frac{1}{2} \sum_{l=1}^m \Gamma_{kl}^{-1}(q) (g_{i,ij}(q) + g_{j,li}(q) - g_{l,ij}(q))
\]  

(6.26)

for eq. (6.2) and define \(\tilde{\Omega}_{ij}(\epsilon, q, v) = \frac{1}{2} \sum_{k=1}^n \tilde{\Gamma}_{kl}^i(q) v^{(k)}\) then we still have \(\tilde{\Omega}(\epsilon, q, v)\dot{v} = \tilde{\Omega}(\epsilon, q, \dot{v})v\) using the fact that \(\tilde{\Gamma}_{ij}^k(q) = \tilde{\Gamma}_{ji}^k(q)\). Hence, with \(A(q, v) = \tilde{\Omega}(\epsilon, q, v)\), eq. (6.25) applies to compute the Jacobian determinant.

### 6.5 Experimentation

We now turn our attention to the empirical evaluation of the numerical integrator of Lagrangian dynamics in terms of its numerical order, its inverted variant, and its robustness to misspecification of the derivatives of the metric. We begin in section 6.5.1 by numerically validating the second-order behavior of the numerical integrator. In the subsequent material, we evaluate the integrator with and without inversion in a banana-shaped distribution, in Bayesian logistic regression, in a multiscale Student-\(t\) distribution, and in a stochastic volatility model. As baselines, we consider RMHMC and HMC. Code to reproduce these experiments may be found at [https://github.com/JamesBrofos/](https://github.com/JamesBrofos/)
We consider three metrics by which to assess the convergence of the Markov chain produced by ILMC and the baselines. First, we consider the expected squared jump distance (ESJD) as described in Gelman and Pasarica [2007]; this measures the expected squared distance between the current state and the next state, where the expectation is computed over the acceptance probability. The larger the ESJD, the less serial autocorrelation in the Markov chain samples. We also consider the effective sample size (ESS) normalized by time elapsed, which gives an indication of the sampling efficiency of each method. We use the implementation of ESS as given in Kumar et al. [2019]. We also consider the method of Brofos and Lederman [2021b] for measuring the ergodicity of the Markov chain given i.i.d. samples. Under this procedure, we project the i.i.d. samples and the Markov chain samples along one-hundred random directions and measure the average value of the Kolmogorov-Smirnov statistics of these one-dimensional distributions. By the Cramér-Wold theorem, the closer these Kolmogorov-Smirnov statistics are concentrated toward zero, the higher the fidelity between the Markov chain samples and the i.i.d. samples. In implementing the generalized leapfrog integrator (definition 6.2.3), we resolve the fixed point equations eqs. (6.4) and (6.5) using fixed point iteration to a convergence tolerance of $1 \times 10^{-6}$, with convergence measured in $\| \cdot \|_\infty$.

6.5.1 Demonstration of Second-Order Error

We consider the following non-separable Hamiltonian $H(q,p) = \frac{q^2 p^2}{2}$, which describes geodesic motion in $\mathbb{R}$ when equipped with the metric $G(q) = 1/q^2$. After converting from momentum to velocity $v_t = \frac{p_t}{q_t}$, we obtain the second-order differential equation $a_t = \frac{v_t^2}{q_t}$. Given initial conditions $q_0$ and $p_0$ (in the Hamiltonian formalism), the exact solution to this differential equation is $q_t = q_0 \exp(q_0 p_0 t)$ and $v_t = q_0^2 p_0 \exp(q_0 p_0 t)$. If, as claimed, the Lagrangian integrator is indeed second-order, then it should exhibit third-order local error.
We see that the observed local error decreases on a log-log scale with a slope of three, corresponding to the claimed third-order local error. By contrast, we also show a line with a slope of two, indicating that the integrator exhibits accuracy better than first-order.

according to definition 6.A.2.10. We evaluate this by examining the squared error between the output of the numerical integrator and the analytical solution for a decreasing sequence of step-sizes $\epsilon$; that is, denoting the output of a single step of the Lagrangian integrator by $(\hat{q}_\epsilon, \hat{v}_\epsilon)$, we measure $\|\hat{q}_\epsilon - q_\epsilon\|^2_2 + \|\hat{v}_\epsilon - v_\epsilon\|^2_2$. We observe in fig. 6.2 that this error decreases linearly on a log-log scale and, critically, the slope of this linear relation is three. This demonstrates numerically the third-order local error of the Lagrangian integrator and gives support to the claim that the method is of second-order accuracy.

### 6.5.2 Banana-Shaped Posterior Distribution

The banana-shaped distribution was proposed in [Girolami and Calderhead, 2011] by Cornebise and Julien as an example of Bayesian inference in non-identifiable models. In this example, a non-identifiable likelihood function in Bayesian linear regression is regularized according to a normal prior, the effect of which is to produce a density with symmetric,
Figure 6.3: Bananna-Shaped Distribution metrics. We show the ESJD, the minimum ESS per second, and the distribution of Kolmogorov-Smirnov statistics for sampling from the banana-shaped distribution using RMHMC, LMC, and ILMC. Surprisingly, we do not observe degradation of the ESJD when employing ILMC. Therefore, combined with its faster sampling iteration, ILMC enjoys the best ESS per second and a distribution of Kolmogorov-Smirnov statistics that is comparable to LMC.

The generative model of this distribution is as follows:

\[ (\theta_1, \theta_2) \overset{i.i.d.}{\sim} \text{Normal}(0, \sigma_\theta^2) \]  
\[ y_i | \theta_1, \theta_2 \overset{i.i.d.}{\sim} \text{Normal}(\theta_1 + \theta_2^2, \sigma_y^2) \] \text{for } i = 1, \ldots, n. \] 

This distribution also illustrates a short-coming of the generalized leapfrog method. For large step-sizes, the implicit update to the momentum variable will not have a solution; therefore, the generalized leapfrog integrator is compelled to adopt a significantly smaller step-size than can be used even by the standard leapfrog method. Indeed, an advantage of explicit numerical integrators is that one does not need to fret that constituent update steps in the integrator will not have solutions. We seek to draw 1,000,000 samples from this posterior.

In our experiments we set \( \sigma_\theta^2 = \sigma_y^2 = 2 \), \( n = 100 \), and set parameter values \( \theta_1 = 1/2 \) and \( \theta_2 = \sqrt{1 - 1/2} \) for generating synthetic data. For HMC we use a step-size of 0.1 and ten integration steps, which was found to produce an acceptance probability between eighty and ninety percent. As for the Riemannian metric, we adopt the sum of the Fisher information of the log-likelihood and the negative Hessian of the log-prior. For RMHMC,
we use a step-size of 0.04 and twenty integration steps, which produces an acceptance probability of around ninety percent. For LMC and ILMC, we use twenty integration steps with a step-size of 0.1, which yields an acceptance probability of around ninety percent. These parameter configurations were found to produce reasonable Markov chains based on hand-tuning. Results showing the effective sample size (ESS) per second are provided in fig. 6.3; we see that RMHMC struggles in this distribution, being even worse than ordinary HMC due to the requirement to use a small step-size. By contrast, LMC and ILMC do significantly better, with ILMC having the best ESS per second due to its elimination of two Jacobian determinant computations. We also show the distribution of this average value over ten trials for each sampling method. We find that the geometric methods based on the Lagrangian formalism perform similarly under this ergodicity measure and outperform competing methods.

6.5.3 Bayesian Logistic Regression

We consider Bayesian logistic regression on a breast cancer and a thyroid cancer dataset. The breast cancer dataset has 277 observations and ten covariates while the thyroid cancer dataset has 215 observations and six covariates. The Bayesian generative model is
We show the ESJD and the minimum and mean ESS per second for the thyroid cancer dataset using RMHMC, LMC, and ILMC. We observe that LMC and ILMC have significantly degraded movement through the sample space as computed by the ESJD; however, this is compensated for by their superior computational efficiency, ultimately yielding more effectively independent samples per second and RMHMC.

assumed to have the following form:

\[
\beta \sim \text{Normal}(0, \alpha^{-1} \text{Id})
\]

\[
y_i | x_i, \beta \sim \text{Bernoulli}(\sigma(x_i^\top \beta)) \quad \text{for} \quad i = 1, \ldots, n,
\]

where \(\sigma : \mathbb{R} \rightarrow (0, 1)\) denotes the sigmoid function. As the Riemannian metric, we adopt the sum of the Fisher information of the log-likelihood and the negative Hessian of the log-prior. We show in figs. 6.4 and 6.5 the minimum ESS per second, where we have also considered varying the convergence threshold used to solve fixed point iterations in RMHMC; as discussed in detail by Brofos and Lederman [2021b], implementations of RMHMC require that the implicit updates to momentum and position be resolved using an iterative procedure such as fixed point iteration or Newton’s method. The tolerance in these numerical methods directly controls the degree to which reversibility and volume preservation are violated by the implementation of the generalized leapfrog method: For small values of the threshold, these theoretical properties are closely respected by the numerical method, while for large thresholds, violations occur. Reversibility and volume preservation imply detailed balance in HMC, and violations imply that the detailed balance
Figure 6.6: Multi-Scale Student-$t$ $\nu = 5 \times 10^0$ metrics. We show the ESJD, the minimum ESS per second, and the distribution of Kolmogorov-Smirnov statistics for sampling from the multiscale Student-$t$ distribution with $\nu = 5 \times 10^0$ using RMHMC, LMC, and ILMC. We observe that RMHMC enjoys the best ESJD, time-normalized ESS, and Kolmogorov-Smirnov statistics. This illustrates an important limitation of the Lagrangian procedure in that its performance may actually be unable to match RMHMC in certain circumstances.

may not hold in an implementation of RMHMC with large thresholds. By contrast, LMC and ILMC are fully explicit and detailed balance is respected to machine precision for either method. We observe that both sampling methods based on Lagrangian mechanics exhibit higher ESS per second than their counterpart based on the Hamiltonian formalism except for the largest thresholds (with the greatest bias); we refer the interested reader to [Brofos and Lederman 2021b] for a detailed discussion on the effects of the convergence threshold on the bias of the RMHMC Markov chain. Moreover, ILMC outperforms LMC on this metric. When comparing the ESJD, we observe that, consistent with our understanding in the Gaussian case, the ILMC method has the smallest expected distance traveled. However, this is offset by a faster sampling iteration due to only requiring two Jacobian determinant calculations instead of four.

6.5.4 Multiscale Student’s $t$-Distribution

Multiscale distributions can be challenging for HMC since trajectories integrated by the leapfrog method will exhibit significant oscillations along directions of the distributions possessing the smallest spatial scale. To investigate this phenomenon, we consider sampling from a multivariate Student-$t$ distribution with a multiscale covariance; in particular
we consider a distribution with density function,

\[ \pi(x) \propto \left[ 1 + \frac{1}{\eta} x^\top \Sigma^{-1} x \right]^{-(\eta+m)/2}, \tag{6.31} \]

where \( x \in \mathbb{R}^m \), \( \eta > 2 \) is the degrees-of-freedom, and \( \Sigma \) is the scale matrix. We consider scale matrices of the form \( \Sigma = \text{diag}(1, \ldots, 1, \sigma^2) \in \mathbb{R}^{m \times m} \). In our experiments we set \( m = 20 \) and \( \eta \in \{5 \times 10^0, 5 \times 10^3\} \) and consider multiscale distributions for \( \sigma^2 \in \{1 \times 10^1, 1 \times 10^2, 1 \times 10^3, 1 \times 10^4\} \). We choose these two values of the degrees-of-freedom to demonstrate two distinct sampling behaviors. For the Riemannian methods, we consider a step-size of 0.7 and twenty integration steps. As the Riemannian metric, we use the positive definite term in the negative Hessian of the log-density of the distribution. For \( \nu = 5 \times 10^9 \), we observe that the ESJD is largest for RMHMC; this occurs because RMHMC enjoys a far superior acceptance probability in this scenario at 95% whereas LMC and ILMC have acceptance probabilities of 40% and 62%, respectively. This degraded performance of the Lagrangian methods is then reflected in the time-normalized ESS, which shows RMHMC dominating LMC, ILMC, and HMC. For the case of \( \eta = 5 \times 10^3 \), circumstances are more favorable to the Lagrangian methods,
Figure 6.8: *Stochastic Volatility Model metrics.* We show the ESJD and the minimum and mean ESS per second for the stochastic volatility model using HMC, RMHMC, LMC, and ILMC. In this example, ILMC moves less efficiently through the same space as measured by ESJD, and ultimately produces fewer effective samples per second relative to LMC. However, it is intriguing to note that ILMC nonetheless produces more effective transitions than RMHMC and HMC.

with ILMC exhibiting the best performance in terms of time-normalized ESS, with the LMC method giving the second best results. In terms of ESJD, we observe that ILMC moves less far in sample space than LMC, but that this is offset by the faster sampling.

As in the case of the banana-shaped distribution, we may sample from this target density analytically in order to assess the ergodicity properties of the samplers. We find that the ergodicity of the geometric methods is essentially constant with respect to the multiscale parameter, whereas the performance of Euclidean HMC noticeably degenerates. For both $\nu = 5 \times 10^3$ and $\nu = 5 \times 10^0$, one observes that the performance of the geometric methods is essentially constant over the multiple scales of the target distribution, demonstrating the beneficial effect of capturing the geometry of the target.
6.5.5 Stochastic Volatility Model

We consider Bayesian inference in a stochastic volatility model. We consider the following generative model:

\[ x_t | x_{t-1}, \phi, \sigma^2 \sim \text{Normal}(\phi x_{t-1}, \sigma^2) \]  
\[ y_t | \beta, x_t \sim \text{Normal}(0, \beta^2 \exp(x_t)) \]

for \( t = 2, \ldots, T \) in eq. (6.32) and \( t = 1, \ldots, T \) in eq. (6.33) and with priors \( x_1 \sim \text{Normal}(0, \sigma^2/(1 - \phi^2)) \), \((\phi + 1)/2 \sim \text{Beta}(20, 3/2)\), \(1/\sigma^2 \sim \xi^2(10, 1/20)\), and the prior over \( \beta \) being proportional to \( 1/\beta^2 \). Given \((y_1, \ldots, y_T)\), we seek to sample the posterior of \((x_1, \ldots, x_T, \phi, \beta, \sigma^2)\). We follow [Girolami and Calderhead 2011] and employ a Metropolis-within-Gibbs-like alternating procedure for sampling the posteriors of \((x_1, \ldots, x_T)\) and \((\phi, \beta, \sigma^2)\). In our experiments we set \( T = 1,000 \) and use values of \( \phi = 0.98, \beta = 0.65, \) and \( \sigma^2 = 0.15^2 \). For HMC, we use a step-size of 0.01 and fifty integration steps when sampling \((\phi, \beta, \sigma^2)\); for the geometric methods, we use a step-size of 0.5 and six integration steps. As for the Riemannian metric, we adopt the sum of the Fisher information of the log-likelihood and the negative Hessian of the log-prior. We compare the average ESS per second among the three latent variables \((\phi, \beta, \sigma^2)\) with results reported in fig. 6.8. We find that LMC and ILMC are the strongest performing methods, with LMC having better ESS per second due to its more efficient traversal of the sample space. Indeed, ILMC has degraded performance in this example, owing to its greater autocorrelation, but nevertheless outperforms HMC and RMHMC.
6.5.6 Fitzhugh-Nagumo Model

We now investigate the LMC and RMHMC in the Fitzhugh-Nagumo posterior distribution.

The Fitzhugh-Nagumo posterior distribution has the following generative model,

\[(a, b, c)^{\text{i.i.d.}} \sim \text{Normal}(0, 1)\]  \hfill (6.34)

\[\hat{r}_{tk}^{\text{i.i.d.}} \sim \text{Normal}(r_{tk}, \sigma^2)\]  \hfill (6.35)

\[\hat{v}_{tk}^{\text{i.i.d.}} \sim \text{Normal}(v_{tk}, \sigma^2) \quad \text{for } k = 1, \ldots, n,\]  \hfill (6.36)

where \(t_1, \ldots, t_n\) are evenly spaced time points in \([0, 10]\) and \(r_t\) and \(v_t\) obey the differential equations,

\[
\dot{r}_t = \frac{-v_t - a + br_t}{c} \hfill (6.37)
\]

\[
\dot{v}_t = cv_t - \frac{cv_t^3}{3} + cr_t. \hfill (6.38)
\]
In our experiments we give initial conditions $v_0 = 1$, $r_0 = -1$, $\sigma^2 = 1/4$, and $n = 200$. The objective is to infer the posterior of $q = (a, b, c)$ given observations $\{(\hat{v}_{tk}, \hat{r}_{tk})\}_{k=1}^{200}$.

Here we emphasize a different aspect of the computation: its robustness to misspecification. As described in section 6.4.5, the proof that the generalized leapfrog integrator is a volume-preserving transformation when used to integrate Hamiltonian vector fields is the symmetry of partial derivatives. We therefore consider the repercussions of an error in the implementation of partial derivatives that invalidates this requirement. To give further context to this experiment, we quote from Dahlquist and Björck [2003]: “In all numerical work, one must expect that clerical errors, errors in hand calculation, and misunderstandings will occur. [...] Most of the errors depend on the so-called human factor. [...] We take up these sources of error in order to emphasize that both the person who carries out a calculation and the person who guides the work of others can plan so that such sources of error are not damaging.” It seems to us, therefore, that the robustness of a Markov chain procedure to human misspecification is a most desirable circumstance.

As described in Girolami and Calderhead [2011], computing the gradient of the log-posterior, the Riemannian metric, and the Jacobian of the Riemannian metric requires us to compute sensitivity equations of the form

$$\frac{d}{dt} \frac{\partial r_t}{\partial q_i}, \frac{d}{dt} \frac{\partial v_t}{\partial q_i}, \frac{d}{dt} \frac{\partial^2 r_t}{\partial q_i \partial q_j}, \text{ and } \frac{d}{dt} \frac{\partial^2 v_t}{\partial q_i \partial q_j} \text{ for } i = 1, 2, 3 \text{ and } j = 1, 2, 3.$$ 

We employ a Riemannian metric of the form,

$$G_{ij}(q) = \frac{1}{\sigma^2} \sum_{k=1}^{n} \left( \frac{\partial v_{tk}}{\partial q_i} \frac{\partial v_{tk}}{\partial q_j} + \frac{\partial r_{tk}}{\partial q_i} \frac{\partial r_{tk}}{\partial q_j} \right) + \mathbf{1} \{i = j\}. \tag{6.39}$$

If any of these sensitivity equations are misspecified, then we will destroy the symmetry of partial derivatives required by the leapfrog integrator to preserve volume. However, the integrator used in LMC and ILMC is non-volume-preserving by design, and the appropriate volume correction is readily computable during sampling. We expect, therefore, that the volume correction employed in LMC and ILMC will provide robustness against these incorrectly computed quantities. We note that, in this case, there is no obvious mechanism
by which to compute the required Jacobian determinant when the generalized leapfrog integrator is no longer a symplectic transformation.

In fig. 6.9 we show the minimum ESS per second and Kolmogorov-Smirnov statistics for sampling from the Fitzhugh-Nagumo posterior. We observe that LMC outperforms ILMC in this example, but that both Lagrangian methods outperform HMC and RMHMC. In terms of ergodicity, all methods perform similarly. When we introduce changes to the sensitivity equations, we observe that RMHMC’s ergodicity severely degrades. For HMC, for which higher-order sensitivities are not required (because of the Euclidean metric), ergodicity is identical to the previous case. However, for LMC and ILMC, in which the higher-order sensitivities are required to compute the Christoffel symbols, the fact that the sensitivities have been incorrectly specified has not noticeably degraded ergodicity. This is a virtue of the LMC and ILMC Markov chains that they are more robust to human errors of this variety.

6.6 Conclusion

This work has examined the numerical integrator for Lagrangian Monte Carlo (LMC). Motivated by the observation that LMC requires four Jacobian determinant evaluations, mechanisms by which this number may be reduced were examined. By inverting the sequence of integration so that position, rather than velocity, is updated twice, the number of Jacobian determinant evaluations in each step was reduced from four to two while still maintaining a fully explicit method. Empirical evaluations of this method were provided to show several situations in which the proposed integration strategy enjoys the best time-normalized performance among several alternatives. Moreover, it has been demonstrated in this work that the local error of the Lagrangian leapfrog (and its inverted counterpart) is third order, which improves the previously known order. Additionally, an important robustness property was characterized that LMC possesses and an illustration wherein
human error will invalidate stationarity in RMHMC but not in LMC was given.

Methods of Bayesian inference that incorporate geometric understanding exhibit a pleasing aesthetic, yet they are burdened by numerical considerations – among which are fixed points and cubic complexity – that have limited their adoption. We hope that this research sparks renewed interest in mechanisms of improving these geometric methods.
Algorithm 13 Pseudo-code implementing the generalized leapfrog integrator for computing approximate solutions to Hamilton’s equations of motion. We assume a Hamiltonian in the form of eq. (6.1).

1: **Input**: An integration step-size $\epsilon \in \mathbb{R}$, an initial position in phase space $(q, p) \in \mathbb{R}^m \times \mathbb{R}^m$.
2: Compute an implicit half-step update to the momentum variable to obtain $\dot{p}$ using eq. (6.4).
3: Compute an implicit full-step to the position variable to obtain $\tilde{q}$ using eq. (6.5).
4: Compute an explicit half-step update to the momentum variable to obtain $\hat{p}$ using eq. (6.6).
5: Compute the Jacobian determinant $J = 1$.
6: **Output**: The updated position and momentum $(\tilde{q}, \hat{p})$ and the Jacobian determinant $J$.

Appendix 6.A

Appendices to Chapter 6

6.A.1 Algorithms

In this appendix we give pseudo-code implementations of algorithms featured in this work. In algorithm 13 we show the generalized leapfrog, which contrasts with the Lagrangian leapfrog in algorithm 14 in the presence of implicitly-defined integration steps. Algorithm 16 shows the Lagrangian integrator but with an inverted sequence of integration (updating position before velocity). In algorithm 15 we show an algorithmic implementation of a single-step of the Hamiltonian Monte Carlo Markov chain, which can be carried out using any of the integrators considered in this work.
Algorithm 14 Pseudo-code implementing the Lagrangian leapfrog integrator for computing approximate solutions to the Lagrangian equations of motion. We assume a Lagrangian in the form of eq. (6.A.3). The conversion from momentum to velocity and back again is included for notational simplicity; in practice, when applying multiple steps of this integrator, these conversions need only be computed for the first conversion from momentum to velocity and the last conversion from velocity to momentum.

1: **Input**: An integration step-size \( \epsilon \in \mathbb{R} \), an initial position in phase space \((q, p) \in \mathbb{R}^m \times \mathbb{R}^m\).
2: Use the inverse Legendre transform (definition 6.A.2.5) to convert from momentum to velocity \(v = G^{-1}(q)p\).
3: Compute an explicit half-step update to the velocity variable to obtain \(\tilde{v}\) using eq. (6.7).
4: Compute an explicit full-step to the position variable to obtain \(\tilde{q}\) using eq. (6.8).
5: Compute an explicit half-step update to the velocity variable to obtain \(\tilde{v}\) using eq. (6.9).
6: Use the Legendre transform (definition 6.A.2.4) to convert from velocity to momentum \(\tilde{p} = G(\tilde{q})\tilde{v}\).
7: Compute the Jacobian determinant \(J = \frac{\partial(q, p)}{\partial(\tilde{q}, \tilde{p})}\) using eq. (6.A.18).
8: **Output**: The updated position and momentum \((\tilde{q}, \tilde{p})\) and the Jacobian determinant \(J\).

Algorithm 15 Single step of the Hamiltonian / Lagrangian Monte Carlo transition kernel.

1: **Input**: The current state of the Markov chain \((q_n, p_n)\), an integration step-size \(\epsilon \in \mathbb{R}\), a number of integration steps \(k \in \mathbb{N}\), a numerical integrator \(\Phi\) (as described in any of algorithms 13, 14 and 16).
2: Resample the momentum \(p_n|q_n \sim \text{Normal}(0, G(q_n))\).
3: Initialize the Jacobian determinant \(J = 1\) and initial integration state \((\tilde{q}, \tilde{p}) = (q_n, p_n)\).
4: for \(i = 1, \ldots, k\) do
5: Compute the proposal by applying the numerical integrator: \(((\tilde{q}, \tilde{p}), \tilde{J}) = \Phi_\epsilon(\tilde{q}, \tilde{p})\).
6: Update the Jacobian determinant \(J = J \times \tilde{J}\)
7: end for
8: Apply the momentum flip operator (definition 6.A.2.18) to obtain \((\tilde{q}, \tilde{p}) = F(\tilde{q}, \tilde{p})\).
9: Sample \(u \sim \text{Uniform}(0, 1)\) and compute the acceptance probability \(a = \alpha((q_n, p_n), (\tilde{q}, \tilde{p}), J)\) using eq. (6.A.15).
10: if \(u < a\) then
11: Accept the proposal \((q_{n+1}, p_{n+1}) = (\tilde{q}, \tilde{p})\).
12: else
13: Reject the proposal \((q_{n+1}, p_{n+1}) = (q_n, p_n)\).
14: end if
15: **Output**: The next state of the Markov chain \((q_{n+1}, p_{n+1})\).
Algorithm 16 Pseudo-code implementing the inverted Lagrangian leapfrog integrator for computing approximate solutions to the Lagrangian equations of motion. We assume a Lagrangian in the form of eq. (6.A.3). This version of leapfrog integration for Lagrangian dynamics requires only two Jacobian determinant computations in each step, as opposed to the four required by algorithm 14.

1: **Input**: An integration step-size $\epsilon \in \mathbb{R}$, an initial position in phase space $(q, p) \in \mathbb{R}^m \times \mathbb{R}^m$.

2: Use the inverse Legendre transform (definition 6.A.2.5) to convert from momentum to velocity $v = G^{-1}(q)p$.

3: Compute an explicit half-step to the position variable to obtain $\dot{\tilde{v}}$ using eq. (6.10).

4: Compute an explicit full-step update to the velocity variable to obtain $\tilde{v}$ using eq. (6.11).

5: Compute an explicit half-step to the position variable to obtain $\tilde{q}$ using eq. (6.12).

6: Use the Legendre transform (definition 6.A.2.4) to convert from velocity to momentum $\tilde{p} = G(\tilde{q})\tilde{v}$.

7: Compute the Jacobian determinant $J = \frac{\partial (\tilde{q}, \tilde{p})}{\partial (q, p)}$ using eq. (6.A.18).

8: **Output**: The updated position and momentum $(\tilde{q}, \tilde{p})$ and the Jacobian determinant $J$.

6.A.2 Expanded Preliminaries

6.A.2.1 Hamiltonian and Lagrangian Mechanics

*Definition 6.A.2.1.* Let $H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ be a smooth function, which we call the Hamiltonian. Hamilton’s equations of motion are defined as the solutions to the initial value problem

\[ v_{t} = \nabla_p H(q_t, p_t) \]  \hspace{1cm} (6.A.1)

\[ \dot{p}_t = -\nabla_q H(q_t, p_t), \]  \hspace{1cm} (6.A.2)

with $(q_0, p_0)$ a given initial position in phase space.

Definition 6.A.2.1 gives us a system of coupled first-order differential equations. Hamilton’s equations of motion exhibit several key properties [Marsden and Ratiu 2010], which we summarize.
**Theorem 6.A.2.2.** The Hamiltonian mechanics given in definition 6.A.2.1 possess the following three properties:

1. They preserve the Hamiltonian \( \frac{d}{dt} H(q_t, p_t) = 0. \)

2. Denoting \( \dot{z}_t = (v_t, \dot{p}_t) \) the equations of motion in phase space, Hamiltonian mechanics preserve volume in phase space: \( \text{div}(\dot{z}_t) = 0. \)

3. Under the conditions that \( \nabla_q H(q, -p) = \nabla_q H(q, p) \) and \( -\nabla_p H(q, p) = \nabla_p H(q, -p) \), the equations of motion are reversible via negation of the momentum variable.

A proof is given in section 6.A.3. In the special case when \( U(q) = 0 \) in eq. (6.1) for all \( q \in \mathbb{R}^m \), Hamilton’s equations of motion produce geodesic motion on the Riemannian manifold \((\mathbb{R}^m, G)\), where \( G \) assumes the role of the Riemannian metric [Calin and Chang 2004].

**Definition 6.A.2.3.** Given a Riemannian Hamiltonian (definition 6.2.1), define the associated Lagrangian function by,

\[
L(q, v) = \tilde{K}(q, v) - U(q), \tag{6.A.3}
\]

where \( \tilde{K}(q, v) = \frac{1}{2} v^T G(q)v. \)

Hamiltonian mechanics are related to Lagrangian dynamics by the Legendre transform which converts between momentum and velocity [Marsden and West 2001].

**Definition 6.A.2.4.** Let \( L : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be a Lagrangian of the form in eq. (6.A.3). The Legendre transform relates the momentum and velocity according to \( p = \frac{\partial L}{\partial v}(q, v) = G(q)v. \)

**Definition 6.A.2.5.** Let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be a Hamiltonian of the form in eq. (6.1). The inverse Legendre transform relates velocity and momentum according to \( v = \frac{\partial H}{\partial p}(q, p) = G^{-1}(q)p. \)
The Lagrangian then determines equations of motion in accordance with the following physical principle.

Definition 6.A.2.6. Hamilton’s principle states that the equations of motion \((q_t, v_t)\) over an interval of time \([a, b]\), with known boundary conditions \(q_a\) and \(q_b\), are solutions of the variational equation

\[
\frac{\delta S[q_t]}{\delta q_t} = 0 \tag{6.A.4}
\]

where

\[
S[q_t] = \int_a^b L(q_t, v_t) \, dt \tag{6.A.5}
\]

where \(q_t\) and \(v_t\) are related by \(v_t = \frac{d}{dt} q_t\).

Hamilton’s principle states that equations of motion (as specified by \(q\) and \(v\)) should extremize the Lagrangian subject to the boundary conditions on \(q_a\) and \(q_b\).

Theorem 6.A.2.7. Under Hamilton’s principle (definition 6.A.2.6), the equations of motion \((q_t, v_t)\) must be solutions of the Euler-Lagrange equation

\[
\nabla_q L(q_t, v_t) = \frac{d}{dt} \nabla_v L(q_t, v_t). \tag{6.A.6}
\]

The equations of motion from theorem 6.A.2.7 (or definition 6.A.2.6) are, in fact, equivalent to the motion produced under Hamiltonian mechanics in definition 6.A.2.1 [Marsden and Ratiu 2010]. For Lagrangians of the form in eq. (6.A.3), the \(k\)-th element of the acceleration is,

\[
a_t^{(k)} = - \sum_{i=1}^m \sum_{j=1}^m \Gamma_{ij}^k(q_t) v_t^{(i)} v_t^{(j)} - \sum_{l=1}^m G_{kl}^{-1}(q_t) \frac{\partial U}{\partial q_t^{(l)}}(q_t), \tag{6.A.7}
\]
where $\Gamma^k_{ij}$ are the Christoffel symbols (definition 6.2.2). We conclude this section by noting that $\Omega$ in eq. (6.3) enjoys the following properties:

**Proposition 6.A.2.8.** The function $\Omega : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m \times m$ in eq. (6.3) satisfies

\[
\Omega(\epsilon, q, v)\dot{v} = \Omega(\epsilon, q, \dot{v})v \quad (6.A.8)
\]

\[
\frac{\partial}{\partial v}(\Omega(\epsilon, q, v)\dot{v}) = \Omega(\epsilon, q, \dot{v}). \quad (6.A.9)
\]

**Proof.** The result in eq. (6.A.8) can be seen from

\[
(\Omega(\epsilon, q, v)\dot{v})_i = \frac{\epsilon}{2} \sum_{j=1}^{m} \sum_{k=1}^{m} \Gamma^i_{kj}(q) v^{(k)} \ddot{v}^{(j)} \quad (6.A.10)
\]

\[
= \frac{\epsilon}{2} \sum_{j=1}^{m} \sum_{k=1}^{m} \Gamma^i_{jk}(q) v^{(k)} \ddot{v}^{(j)} \quad (6.A.11)
\]

\[
= \frac{\epsilon}{2} \sum_{k=1}^{m} \sum_{j=1}^{m} \Gamma^i_{jk}(q) v^{(k)} \ddot{v}^{(j)} \quad (6.A.12)
\]

\[
= (\Omega(\epsilon, q, \dot{v})v)_i. \quad (6.A.13)
\]

Equation (6.A.9) then follows as an immediate corollary.\[\square\]

Like Hamiltonian mechanics, Lagrangian dynamics conserve the Hamiltonian (when $p_t = \mathbf{G}(q_t)v_t$) and are reversible. However, Lagrangian dynamics do not conserve volume in $(q, v)$-space.

**6.A.2.2 Background on Numerical Integrators**

The Hamiltonian equations of motion in eqs. (6.A.1) and (6.A.2) and the Lagrangian motion described in eq. (6.A.7) rarely have closed-form solutions. Therefore, it is necessary to investigate methods of numerical integration to produce approximate solutions to these initial value problems. In the following section, we review key ideas from Hairer et al.
Definition 6.A.2.9. Let \( g : \mathbb{R}^m \to \mathbb{R}^m \). The solution to an initial value problem is a function \( z(\cdot) : \mathbb{R} \to \mathbb{R}^m \) for which \( \frac{d}{dt} z_t = g(z_t) \) and for which the initial value \( z_0 \in \mathbb{R}^m \) is known. In this case \( g \) is called a \((\text{time-homogenous})\) vector field.

Definition 6.A.2.10. Let \( z_t \) be the solution to an initial value problem (definition 6.A.2.9) with initial value \( z_0 \). A numerical integrator with step-size \( \epsilon \) is a mapping \( \Phi_\epsilon : \mathbb{R}^m \to \mathbb{R}^m \) and is said to have order \( p \) if \( \| \Phi_\epsilon(z_0) - z_\epsilon \| = \mathcal{O}(\epsilon^{p+1}) \).

Definition 6.A.2.11. The adjoint of a numerical method \( \Phi_\epsilon : \mathbb{R}^m \to \mathbb{R}^m \) is defined by the relation \( \Phi_\epsilon^*(z) = \Phi_\epsilon^{-1}(z) \). The numerical method \( \Phi_\epsilon \) is said to be self-adjoint if \( \Phi_\epsilon^* = \Phi_\epsilon \).

Theorem 6.A.2.12. Consider an initial value problem \( \dot{z}_t = g(z_t) \) with initial condition \( z_0 \in \mathbb{R}^m \). Let \( \Phi_\epsilon \) be a one-step numerical integrator of (maximal) order \( r \in \mathbb{N} \). If \( \Phi_\epsilon \) is self-adjoint, then \( r \) is even.

Proposition 6.A.2.13. Let \( g : \mathbb{R}^m \to \mathbb{R}^m \) be a time-homogenous vector field. Let \( \Xi : \mathbb{R}^m \to \mathbb{R}^m \) be a diffeomorphism. Let \( z_t \) be the solution to the initial value problem \( \dot{z}_t = g(z_t) \) given \( z_0 \). Let \( \Phi_\epsilon : \mathbb{R}^m \to \mathbb{R}^m \) be the flow map of \( z_\epsilon \) and suppose that \( \hat{\Phi}_\epsilon \) is a \( p \)-th order approximation of \( \Phi_\epsilon \). Then \( \Xi \circ \hat{\Phi}_\epsilon \) is a \( p \)-th order approximation of \( \Xi \circ \Phi_\epsilon \).

See proposition 2.2.18 for an identical statement and proof.


6.A.2.3 Hamiltonian and Lagrangian Monte Carlo

Our objective in Bayesian inference is to draw samples from the distribution whose density is \( \pi(q) \propto \exp(\mathcal{L}(q)) \). We now review basic concepts from Markov chain Monte Carlo.

Definition 6.A.2.15. A Markov chain transition kernel is a map \( K : \mathbb{R}^m \times \mathcal{B}(\mathbb{R}^m) \to [0, 1] \) satisfying (i) for every \( x \in \mathbb{R}^m \) the map \( A \mapsto K(x, A) \) is a probability measure and (ii) for
every $A \in \mathcal{B}(\mathbb{R}^m)$, the map $x \mapsto K(x, A)$ is measurable. Given a Markov chain transition kernel $K$, a Markov chain is defined inductively by $x_{n+1} | x_n \sim K(x_n, \cdot)$.

HMC accomplishes this by artificially expanding the distribution to incorporate a momentum variable. Defining $U(q) = -\mathcal{L}(q) + \frac{1}{2} \log \det(G(q))$, let $H(q, p)$ be the Hamiltonian given in eq. (6.1) and observe that the density $\pi(q, p) \propto \exp(-H(q, p))$ has $\pi(q) = \int_{\mathbb{R}^m} \pi(q, p) \, dp$ and $p|q \sim \text{Normal}(0, G(q))$. In order to unify methods of Bayesian inference based on either the generalized leapfrog integrator (definition 6.2.3) or the Lagrangian leapfrog (definition 6.2.4) under one framework, we now introduce the Markov chain transition kernel based on smooth involutions.

**Definition 6.A.2.16** (Neklyudov et al. [2020]). Let $\Phi : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m$ be a smooth involution (i.e. $\Phi = \Phi^{-1}$). Let $\pi : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}_+$ be a probability density with respect to Lebesgue measure. Then we define the Markov chain transition kernel of involutive Monte Carlo by

$$K((q, p), (A, B)) = \alpha((q, p), (\tilde{q}, \tilde{p}), J) \mathbf{1}\{(\tilde{q}, \tilde{p}) \in (A, B)\} \tag{6.A.14}$$

$$+ \left(1 - \alpha((q, p), (\tilde{q}, \tilde{p}), J)\right) \mathbf{1}\{(q, p) \in (A, B)\}$$

$$\alpha((q, p), (\tilde{q}, \tilde{p}), J) = \min \left\{1, \frac{\exp(-H(\tilde{q}, \tilde{p}))}{\exp(-H(q, p))} \cdot J\right\} \tag{6.A.15}$$

$$J = \left|\det\left(\frac{\partial(\tilde{q}, \tilde{p})}{\partial(q, p)}\right)\right|, \tag{6.A.16}$$

where $(A, B) \in \mathcal{B}(\mathbb{R}^m \times \mathbb{R}^m)$ and $(\tilde{q}, \tilde{p}) = \Phi(q, p)$.

**Proposition 6.A.2.17.** The Markov chain transition kernel of involutive Monte Carlo satisfies detailed balance with respect to the distribution whose density is $\pi(q, p) \propto \exp(-H(q, p))$.

A proof is provided in section 6.A.6. Central to the construction of involutions of interest to us is the momentum flip operator, defined as follows.

**Definition 6.A.2.18.** The momentum flip operator is the map $F(q, p) = (q, -p)$. 

363
The fact that numerical integrators can be combined with the momentum flip operator in order to produce involutions is covered in section 6.A.6 in the case of the Lagrangian leapfrog; other integrators are handled similarly. We provide pseudo-code implementing a single step of involutive Monte Carlo in algorithm 15 with involutions provided by the generalized leapfrog or Lagrangian leapfrog integrator. Given an initial point \((q_0, p_0)\) in phase space drawn from the distribution \(\pi(q, p)\), the sequence of states \(((q_1, p_1), (q_2, p_2), \ldots)\) computed by algorithm 15 are guaranteed to have \(\pi(q, p)\) as their marginal distributions. Under the additional conditions that the HMC Markov chain is irreducible and aperiodic, HMC also produces an ergodic chain.

The generalized leapfrog integrator is volume preserving. However, the integrator of Lagrangian dynamics is not. The required change-of-volume can be deduced as follows. First, observe that the update eq. (6.8) is immediately volume-preserving since it is merely the translation of the position variable by a quantity. The change of volume incurred in eq. (6.7) has a Jacobian determinant given by,

\[
\begin{vmatrix}
\det \left( \frac{\partial (q, \tilde{v})}{\partial (q, v)} \right)
\end{vmatrix} = \frac{\det(\text{Id}_m - \Omega(\epsilon, q, \tilde{v}))}{\det(\text{Id}_m + \Omega(\epsilon, q, v))}.
\]

(6.A.17)

The update in eq. (6.9) incurs a similar change of volume. Thus, when employing algorithm 14, the Jacobian determinant of the transformation \((q, p) \mapsto (\tilde{q}, \tilde{p})\) is,

\[
\begin{vmatrix}
\det \left( \frac{\partial (\tilde{q}, \tilde{p})}{\partial (q, p)} \right)
\end{vmatrix} = \det(G^{-1}(q))\det(G(\tilde{q})) \times \frac{\det(\text{Id}_m - \Omega(\epsilon, q, \tilde{v}))\det(\text{Id}_m - \Omega(\epsilon, q, v))}{\det(\text{Id}_m + \Omega(\epsilon, q, \tilde{v}))\det(\text{Id}_m + \Omega(\epsilon, q, v))}.
\]

(6.A.18)

**Definition 6.A.2.19.** Let \(\Phi_\epsilon\) be the generalized leapfrog integrator with step-size \(\epsilon \in \mathbb{R}\) (definition 6.2.3). Let \(k \in \mathbb{N}\) be a number of integration steps. The Riemannian manifold Hamiltonian Monte Carlo (RMHMC) Markov chain is an instance involutive Monte Carlo (definition 6.A.2.16) with involution \(F \circ \Phi_\epsilon^k\).
**Definition 6.A.2.20.** Let \( \Phi \) be the Lagrangian leapfrog integrator with step-size \( \epsilon \in \mathbb{R} \) (definition 6.2.4). Let \( k \in \mathbb{N} \) be a number of integration steps. The *Lagrangian Monte Carlo (LMC) Markov chain* is an instance involutive Monte Carlo (definition 6.A.2.16) with involution \( F \circ \Phi^k \).

**6.A.3 Proofs Concerning Mechanics Systems**

We give a proof of theorem 6.A.2.2.

**Proof.** We first show that the Hamiltonian energy is conserved.

\[
\frac{d}{dt} H(q_t, p_t) = \nabla_q H(q_t, p_t) \cdot v_t + \nabla_p H(q_t, p_t) \cdot \dot{p}_t \\
= \nabla_q H(q_t, p_t) \cdot \nabla_p H(q_t, p_t) - \nabla_p H(q_t, p_t) \cdot \nabla_q H(q_t, p_t) \\
= 0.
\]

(6.A.19)

(6.A.20)

(6.A.21)

Next we show that Hamiltonian mechanics conserve volume in \((q, p)\)-space. This is equivalent to the vector field \((v_t, \dot{p}_t)\) having zero divergence, which we now verify.

\[
\text{div}(v_t, \dot{p}_t) = \nabla_q \cdot v_t + \nabla_p \cdot \dot{p}_t \\
= \nabla_q \cdot \nabla_p H(q_t, p_t) - \nabla_p \cdot \nabla_q H(q_t, p_t) \\
= \sum_{i=1}^{m} \frac{\partial^2 H}{\partial q_i \partial p_i} H(q_t, p_t) - \sum_{i=1}^{m} \frac{\partial^2 H}{\partial p_i \partial q_i} H(q_t, p_t) \\
= 0,
\]

(6.A.22)

(6.A.23)

(6.A.24)

(6.A.25)

by symmetry of partial derivatives. Finally we show that under the conditions that \( \nabla_q H(q, -p) = \nabla_q H(q, p) \) and \( \nabla_p H(q, -p) = -\nabla_p H(q, p) \) that the equations of motion are reversible under negation of the momentum variable. To see this, fix \( \tau \in \mathbb{R}_+ \) and consider \((q_t, p_t)\) satisfying eqs. (6.A.1) and (6.A.2) for \( t \in [0, \tau] \). Let \( \tilde{p}_t = -p_{\tau-t} \) and \( \tilde{q}_t = q_{\tau-t} \). We find...
that $\tilde{q}_t$ and $\tilde{p}_t$ obey the following equations of motion:

\[
\frac{d}{dt} \tilde{q}_t = -v_{\tau-t} \\
= -\nabla_p H(q_{\tau-t}, p_{\tau-t}) \\
= \nabla_p H(\tilde{q}_t, -p_{\tau-t}) \\
= \nabla_p H(\tilde{q}_t, \tilde{p}_t) \\
= \nabla_p H(\tilde{q}_t, \tilde{p}_t) \\
= \nabla_p H(\tilde{q}_t, \tilde{p}_t).
\]

\[
\frac{d}{dt} \tilde{p}_t = \dot{p}_{\tau-t} \\
= -\nabla_q H(q_{\tau-t}, p_{\tau-t}) \\
= -\nabla_q H(q_{\tau-t}, -p_{\tau-t}) \\
= -\nabla_q H(\tilde{q}_t, \tilde{p}_t).
\]

Thus we see that $(\tilde{q}_t, \tilde{p}_t)$ are also solutions to Hamilton’s equations of motion and satisfy $\tilde{q}_r = q_0$ and $\tilde{p}_r = -p_0$, demonstrating reversibility.
6.A.4 Proofs Concerning the Numerical Order of the Lagrangian Integrators

The claim of first order accuracy was derived in [Lan et al. 2015]; we have included a proof for completeness.

**Proof of lemma 6.4.5** Let $q_t$ and $v_t$ be solutions to the initial value problem,

\begin{align*}
\dot{q}_t &= v_t \quad (6.A.34) \\
\dot{v}_t &= -\Omega(2, q_t, v_t)v_t - G^{-1}(q_t)\nabla U(q_t). \quad (6.A.35)
\end{align*}

Given the initial value $q_0 = q$ and $v_0 = v$, we can expand the solution in a Taylor series about $t = 0$.

\begin{align*}
q_t &= q_0 + t\dot{q}_0 + \mathcal{O}(t^2) \quad (6.A.36) \\
&= q + tv + \mathcal{O}(t^2) \quad (6.A.37) \\
v_t &= v_0 - t \left( \Omega(2, q_0, v_0)v_0 - G^{-1}(q_0)\nabla U(q_0) \right) + \mathcal{O}(t^2) \quad (6.A.38) \\
&= v - \Omega(2t, q, v)v - tG^{-1}(q)\nabla U(q) + \mathcal{O}(t^2). \quad (6.A.39)
\end{align*}

Letting $t = \epsilon$ gives the following approximations,

\begin{align*}
q_\epsilon &= q + \epsilon v + \mathcal{O}(\epsilon^2) \quad (6.A.40) \\
v_\epsilon &= v - \Omega(2\epsilon, q, v)v - \epsilon G^{-1}(q)\nabla U(q) + \mathcal{O}(\epsilon^2) \quad (6.A.41)
\end{align*}

Now we expand the steps of the explicit numerical integrator. We begin with the first
Now we expand the update to position.

\[ \tilde{q} = q + \epsilon \tilde{v} \]  
\[ = q + \epsilon v + \mathcal{O}(\epsilon^2) \]  

We conclude by expanding the second update to velocity in terms of the initial conditions.

\[ \ddot{v} = (I_m + \Omega(\epsilon, \tilde{q}, \tilde{v}))^{-1} \left( \dot{v} - \frac{\epsilon}{2} G^{-1}(\tilde{q}) \nabla U(\tilde{q}) \right) \]  
\[ = \dot{v} - \Omega(\epsilon, \tilde{q}, \tilde{v}) \dot{v} - \frac{\epsilon}{2} G^{-1}(\tilde{q}) \nabla U(\tilde{q}) + \mathcal{O}(\epsilon^2) \]  
\[ = \dot{v} - \Omega(\epsilon, q, v) v - \frac{\epsilon}{2} G^{-1}(q) \nabla U(q) + \mathcal{O}(\epsilon^2) \]  
\[ = v - \Omega(2\epsilon, q, v) v - \epsilon G^{-1}(q) \nabla U(q) + \mathcal{O}(\epsilon^2). \]  

Therefore, a single step of the numerical integrator with step-size \( \epsilon \) agrees with the analytical solution to the initial value problem (from the same initial condition) to at least first order in \( \epsilon \).

Proof of Lemma 6.4.2 To demonstrate that the inverted integrator has at least first-order
error, we expand the steps of the integrator as follows. First,

\[ \dot{q} = q + \frac{\epsilon}{2} v \quad (6.A.51) \]

\[ \ddot{v} = (\text{Id}_m + \Omega(2\epsilon, \dot{q}, v))^{-1} \left( v - \epsilon G^{-1}(\dot{q}) \nabla U(\dot{q}) \right) \quad (6.A.52) \]

\[ = (\text{Id}_m + \Omega(2\epsilon, \dot{q}, v)) \left( v - \epsilon G^{-1}(\dot{q}) \nabla U(\dot{q}) \right) + \mathcal{O}(\epsilon^2) \quad (6.A.53) \]

\[ = v - \Omega(2\epsilon, q, v)v - \epsilon G^{-1}(q) \nabla U(q) + \mathcal{O}(\epsilon^2) \quad (6.A.54) \]

\[ \tilde{q} = \ddot{q} + \frac{\epsilon}{2} \ddot{v} \quad (6.A.55) \]

\[ = q + \frac{\epsilon}{2} v + \frac{\epsilon}{2} v + \mathcal{O}(\epsilon^2) \quad (6.A.56) \]

\[ = q + \epsilon v. \quad (6.A.57) \]

This verifies that the inverted Lagrangian leapfrog has at least first order accuracy.

In order to show that the inverted leapfrog integrator is symmetric, we proceed as follows. Recall that the three steps of the inverted Lagrangian leapfrog are

\[ \dot{q} = q + \frac{\epsilon}{2} v \quad (6.A.58) \]

\[ \ddot{v} = (\text{Id}_m + \Omega(2\epsilon, \dot{q}, v))^{-1} \left( v - \epsilon G^{-1}(\dot{q}) \nabla U(\dot{q}) \right) \quad (6.A.59) \]

\[ \tilde{q} = \ddot{q} + \frac{\epsilon}{2} \ddot{v}. \quad (6.A.60) \]

Therefore, we consider integrating from initial position \((\tilde{q}, \tilde{v})\) with a negated step-size \(-\epsilon\) as follows:

\[ \dot{q}' = \tilde{q} - \frac{\epsilon}{2} \tilde{v} \quad (6.A.61) \]

\[ = \ddot{q} \quad (6.A.62) \]
For the velocity we have,

\[ \tilde{v}' = (\text{Id}_m - \Omega(2\epsilon, \tilde{q}, \tilde{v}))^{-1} \left( \tilde{v} + \epsilon \mathbf{G}^{-1}(\tilde{q}) \nabla U(\tilde{q}) \right) \]  
\( (6.A.63) \)

\[ \Rightarrow (\text{Id}_m - \Omega(2\epsilon, \tilde{q}, \tilde{v})) \tilde{v}' = \tilde{v} + \epsilon \mathbf{G}^{-1}(\tilde{q}) \nabla U(\tilde{q}) \]  
\( (6.A.64) \)

\[ \Rightarrow \tilde{v} + \Omega(2\epsilon, \tilde{q}, \tilde{v}') \tilde{v} = \tilde{v}' - \epsilon \mathbf{G}^{-1}(\tilde{q}) \nabla U(\tilde{q}) \]  
\( (6.A.65) \)

\[ \Rightarrow \tilde{v} = (\text{Id}_m + \Omega(2\epsilon, \tilde{q}, \tilde{v}' \tilde{v}))^{-1} \left( \tilde{v}' - \epsilon \mathbf{G}^{-1}(\tilde{q}) \nabla U(\tilde{q}) \right) \]  
\( (6.A.66) \)

\[ \Rightarrow \tilde{v}' = v \]  
\( (6.A.67) \)

Finally, the last update to the position is,

\[ \tilde{q}' = \tilde{q} - \frac{\epsilon}{2} v \]  
\( (6.A.68) \)

\[ = q. \]  
\( (6.A.69) \)

Hence we see that the inverted Lagrangian leapfrog is also self-adjoint. As noted in the main text, the combination of self-adjointness and at least first-order accuracy immediately imply second-order accuracy. \( \square \)

**Proof of lemma 6.4.6** Self-adjointness of a numerical integrator follows immediately from the condition \( \Phi_{-\epsilon} \circ \Phi_{\epsilon} = \text{Id} \). Therefore, to demonstrate that a numerical method is self-adjoint it suffices to establish this condition. Consider the first update to the velocity:

\[ \ddot{v} = (\text{Id}_m + \Omega(\epsilon, q, v))^{-1} \left( v - \frac{\epsilon}{2} \mathbf{G}(q)^{-1} \nabla U(q) \right) \]  
\( (6.A.70) \)

\[ \iff \ddot{v} + \Omega(\epsilon, q, v) \ddot{v} = v - \frac{\epsilon}{2} \mathbf{G}(q)^{-1} \nabla U(q) \]  
\( (6.A.71) \)

\[ \iff \ddot{v} + \frac{\epsilon}{2} \mathbf{G}(q)^{-1} \nabla U(q) = v - \Omega(\epsilon, q, \ddot{v}) v \]  
\( (6.A.72) \)

\[ \iff \ddot{v} + \frac{\epsilon}{2} \mathbf{G}(q)^{-1} \nabla U(q) = (\text{Id}_m - \Omega(\epsilon, q, \ddot{v})) v \]  
\( (6.A.73) \)

\[ \iff v = (\text{Id}_m - \Omega(\epsilon, q, \ddot{v}))^{-1} \left( \ddot{v} + \frac{\epsilon}{2} \mathbf{G}^{-1}(q) \nabla U(q) \right). \]  
\( (6.A.74) \)
An identical series of computations reveals,

$$\dot{v} = (\text{Id}_m - \Omega(\epsilon, \tilde{q}, \tilde{v}))^{-1} \left( \tilde{v} + \frac{\epsilon}{2} G^{-1}(\tilde{q}) \nabla U(\tilde{q}) \right). \tag{6.A.75}$$

Hence, applying $\Phi_{-\epsilon}$ to $(\tilde{q}, \tilde{v})$ yields the following series of updates,

$$\dot{v}' = (\text{Id}_m - \Omega(\epsilon, \tilde{q}, \tilde{v}))^{-1} \left( \tilde{v} + \frac{\epsilon}{2} G^{-1}(\tilde{q}) \nabla U(\tilde{q}) \right) \tag{6.A.76}$$

$$= \dot{v} \tag{6.A.77}$$

$$q' = \tilde{q} - \epsilon \dot{v}' \tag{6.A.78}$$

$$= \tilde{q} - \epsilon \dot{v} \tag{6.A.79}$$

$$= q \tag{6.A.80}$$

$$v' = (\text{Id}_m - \Omega(\epsilon, q, \tilde{v}))^{-1} \left( \tilde{v} + \frac{\epsilon}{2} G^{-1}(q) \nabla U(q) \right) \tag{6.A.81}$$

$$= v. \tag{6.A.82}$$

Hence we return to the initial condition $(q, v)$. This verifies that the explicit integrator employed in LMC is self-adjoint. \qed
6.A.5 Propagator Matrices for the Leapfrog and Inverted Leapfrog

**Definition 6.A.5.1.** The generalized leapfrog integrator for the Hamiltonian equations of motion in eqs. (6.A.1) and (6.A.2) with Hamiltonian \( H(q, p) = U(q) + \frac{p^\top p}{2} \) is a map \( (q, p) \mapsto (\tilde{q}, \tilde{p}) \) defined by,

\[
\begin{align*}
\tilde{p} &= p - \frac{\epsilon}{2} \nabla U(q) \quad (6.A.83) \\
\tilde{q} &= q + \epsilon \tilde{p} \quad (6.A.84) \\
\tilde{p} &= \tilde{p} - \frac{\epsilon}{2} \nabla U(\tilde{q}). \quad (6.A.85)
\end{align*}
\]

**Definition 6.A.5.2.** The generalized leapfrog integrator for the Hamiltonian equations of motion in eqs. (6.A.1) and (6.A.2) with Hamiltonian \( H(q, p) = U(q) + \frac{p^\top p}{2} \) is a map \( (q, p) \mapsto (\tilde{q}, \tilde{p}) \) defined by,

\[
\begin{align*}
\tilde{q} &= q + \frac{\epsilon}{2} p \quad (6.A.86) \\
\tilde{p} &= p - \epsilon \nabla U(q) \quad (6.A.87) \\
\tilde{q} &= \tilde{q} + \frac{\epsilon}{2} \tilde{p}. \quad (6.A.88)
\end{align*}
\]

For Hamiltonians of the form \( H(q, p) = \omega^2 q^2/2 + p^2/2 \), the action of the leapfrog and inverted leapfrog integrators are linear. This means that there are matrices, called “propagator matrices,” which, when acting on the vector \( (q, p) \in \mathbb{R}^2 \), produce the same position in phase space as the integrators themselves. Computing integer matrix powers of these matrices can then produce the multi-step output of the integrators. In the case of
the leapfrog integrator, the propagator matrix is $[\text{Leimkuhler and Reich}, 2005]$, 

$$
\begin{pmatrix}
\tilde{q} \\
\tilde{p}
\end{pmatrix} = \begin{pmatrix}
1 - \frac{\epsilon^2 \omega^2}{2} & \epsilon \\
\epsilon \omega^2 \left(1 - \frac{\epsilon^2 \omega^2}{4}\right) & 1 - \frac{\epsilon^2 \omega^2}{2}
\end{pmatrix}
\begin{pmatrix}
q \\
p
\end{pmatrix}.
$$

(6.A.89)

The propagator matrix for the inverted leapfrog is,

$$
\begin{pmatrix}
\tilde{q} \\
\tilde{p}
\end{pmatrix} = \begin{pmatrix}
1 - \frac{\epsilon^2 \omega^2}{2} & \epsilon \left(1 - \frac{\epsilon^2 \omega^2}{4}\right) \\
-\epsilon \omega^2 & 1 - \frac{\epsilon^2 \omega^2}{2}
\end{pmatrix}
\begin{pmatrix}
q \\
p
\end{pmatrix}.
$$

(6.A.90)

At stationarity,

$$
\begin{pmatrix}
q \\
p
\end{pmatrix} \sim \text{Normal}\left(\begin{pmatrix}
0 \\
0
\end{pmatrix}, \begin{pmatrix}
1/\omega^2 & 0 \\
0 & 1
\end{pmatrix}\right).
$$

(6.A.91)

Therefore, for an integrator with propagator matrix $R$, the distribution of the $k$-step transition is,

$$
\text{Normal}\left(\begin{pmatrix}
0 \\
0
\end{pmatrix}, R^k \begin{pmatrix}
1/\omega^2 & 0 \\
0 & 1
\end{pmatrix} (R^k)^\top\right),
$$

(6.A.92)
6.A.6 Involutive Monte Carlo

Proof of Proposition 6.A.2.17: The detailed balance condition states that for any Borel sets \((A, B), (C, D) \in \mathcal{B}(\mathbb{R}^m \times \mathbb{R}^m)\) we have,

\[
\int_{(A,B)} K((q, p), (C, D)) \, \pi(q, p) \, dq \, dp = \int_{(C,D)} K((q, p), (A, B)) \, \pi(q, p) \, dq \, dp.
\]

(6.A.93)
Let $\pi(q, p)$ be the probability density on $\mathbb{R}^m \times \mathbb{R}^m$ that is proportional to $\exp(-H(q, p))$.

Using the fact that $\Phi$ is an involution, it suffices to verify

\[
\int_{(A, B)} \alpha((q, p), \Phi(q, p), |\det (\nabla \Phi(x))|) \mathbf{1}\{(\tilde{q}, \tilde{p}) \in (C, D)\} \pi(q, p) \, dq \, dp \quad (6.\text{A}.94)
\]

\[
= \int_{(A, B)} \min \left\{ 1, \frac{\pi(\Phi(q, p))}{\pi(q, p)} |\det (\nabla \Phi(x))| \right\} \mathbf{1}\{\Phi(q, p) \in (C, D)\} \pi(q, p) \, dq \, dp \quad (6.\text{A}.95)
\]

\[
= \int_{(A, B)} \min \{\pi(q, p), \pi(\Phi(q, p))|\det (\nabla \Phi(x))|\} \mathbf{1}\{\Phi(q, p) \in (C, D)\} \, dq \, dp \quad (6.\text{A}.96)
\]

\[
= \int_{(A, B)} \min \left\{ \frac{\pi(q, p)}{\pi(\Phi(q, p)) |\det (\nabla \Phi(x))|}, 1 \right\} \mathbf{1}\{\Phi(q, p) \in (C, D)\} \pi(\Phi(q, p)) |\det (\nabla \Phi(x))| \, dq \, dp \quad (6.\text{A}.97)
\]

\[
= \int_{(A, B)} \min \left\{ \frac{\pi(q, p)}{\pi(\Phi(q, p)) |\det (\nabla \Phi(x))|}, 1 \right\} \mathbf{1}\{\Phi(q, p) \in (C, D)\} \pi(q, p) |\det (\nabla \Phi(x))| \, dq \, dp \quad (6.\text{A}.98)
\]

\[
= \int_{\Phi(A, B)} \min \left\{ 1, \frac{\pi(\Phi(\tilde{q}, \tilde{p}))}{\pi(\tilde{q}, \tilde{p}) |\det (\nabla \Phi(\tilde{q}, \tilde{p}))|} \right\} \mathbf{1}\{(\tilde{q}, \tilde{p}) \in (C, D)\} \pi(\tilde{q}, \tilde{p}) \, d\tilde{q} \, d\tilde{p} \quad (6.\text{A}.99)
\]

\[
= \int_{(C, D)} \alpha((q, p), \Phi(q, p), |\det (\nabla \Phi(q, p))|) \mathbf{1}\{(q, p) \in \Phi(A, B)\} \pi(q, p) \, dq \, dp \quad (6.\text{A}.100)
\]

\[
= \int_{(C, D)} \alpha((q, p), \Phi(q, p), |\det (\nabla \Phi(q, p))|) \mathbf{1}\{\Phi(q, p) \in (A, B)\} \pi(q, p) \, dq \, dp \quad (6.\text{A}.101)
\]
Moreover, integrating over the rejection components of the transition kernel already have symmetry in \((A, B)\) and \((C, D)\):

\[
\int_{(A,B)} (1 - \alpha((q, p), (\tilde{q}, \tilde{p}), |\det(\nabla \Phi(x))|)) \mathbf{1} \{(q, p) \in (C, D)\} \pi(q, p) \, dq \, dp
\]

(6.A.102)

\[
= \int_{(C,D)} (1 - \alpha((q, p), (\tilde{q}, \tilde{p}), |\det(\nabla \Phi(x))|)) \mathbf{1} \{(q, p) \in (A, B)\} \pi(q, p) \, dq \, dp.
\]

(6.A.103)

Detailed balance follows as a consequence.

\[\square\]

**Lemma 6.A.6.1.** Let \(\Phi_\epsilon : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m\) be the Lagrangian leapfrog integrator with step-size \(\epsilon \in \mathbb{R}\). Then \(F \circ \Phi_\epsilon\) is an involution.

**Proof.** Let \((\tilde{q}, \tilde{v}) = \Phi_\epsilon(q, v)\). Now we apply the Lagrangian leapfrog to \((\tilde{q}, -\tilde{v})\).

\[
\tilde{v}'' = (\mathbf{Id}_m + \Omega(\epsilon, \tilde{q}, -\tilde{v}))^{-1} \left( -\tilde{v} - \frac{\epsilon}{2} \mathbf{G}^{-1}(\tilde{q}) \nabla U(\tilde{q}) \right)
\]

(6.A.104)

\[
= -(\mathbf{Id}_m - \Omega(\epsilon, \tilde{q}, \tilde{v}))^{-1} \left( \tilde{v} + \frac{\epsilon}{2} \mathbf{G}^{-1}(\tilde{q}) \nabla U(\tilde{q}) \right)
\]

(6.A.105)

\[
= -\tilde{v}
\]

(6.A.106)

from eq. (6.A.76). Moreover,

\[
\tilde{q}'' = \tilde{q} + \epsilon \tilde{v}''
\]

(6.A.107)

\[
= \tilde{q} - \epsilon \tilde{v}
\]

(6.A.108)

\[
= q.
\]

(6.A.109)
Finally,

\[ \tilde{v}'' = (\text{Id}_m + \Omega(\epsilon, q, \tilde{v}))^{-1} \left( -\tilde{v} - \frac{\epsilon}{2} \mathbf{G}^{-1}(q) \nabla U(q) \right) \]  

(6.A.110)

\[ = -(\text{Id}_m - \Omega(\epsilon, q, \tilde{v}))^{-1} \left( \tilde{v} + \frac{\epsilon}{2} \mathbf{G}^{-1}(q) \nabla U(q) \right) \]  

(6.A.111)

\[ = -v \]  

(6.A.112)

from eq. (6.A.81).
Chapter 7

Geometric Ergodicity in Modified Variations of Riemannian Manifold and Lagrangian Monte Carlo

This chapter presents my investigations into equipping geometric Markov chain methods with an ergodicity theory of some form. This takes the role of modifying RMHMC and LMC such that their single-step implementation is substituted for the manifold Metropolis-adjusted Langevin algorithm (MMALA), whereby the modified versions of RMHMC and LMC “inherit” geometric ergodicity in cases that MMALA is geometrically ergodic. This chapter is joint work with Roy Lederman and Vivek Roy; composition and experimentation are original to me.

Abstract. Riemannian manifold Hamiltonian (RMHMC) and Lagrangian Monte Carlo (LMC) have emerged as powerful methods of Bayesian inference. Unlike Euclidean Hamiltonian Monte Carlo (EHMC) and the Metropolis-adjusted Langevin algorithm (MALA), the geometric ergodicity of these Riemannian algorithms has not been extensively studied. On the other hand, the manifold Metropolis-adjusted Langevin algorithm (MMALA) has recently been shown to exhibit geometric ergodicity under certain conditions. This work investigates the mixture of the LMC and RMHMC transition kernels with MMALA in order to equip the resulting method with an “inherited” geometric ergodicity theory. We
motivate this mixture kernel based on an analogy between single-step HMC and MALA. We then proceed to evaluate the original and modified transition kernels on several benchmark Bayesian inference tasks.

7.1 Introduction

Bayesian inference seeks to combine subjective sources of information with observational data. By specifying one’s prior beliefs and correctly capturing sources of uncertainty in a stochastic system, one may employ the Bayesian approach in order to capture and reason about uncertainty under the posterior distribution. Formally, the posterior distribution is a probability distribution \( \Pi : \mathcal{B}(\mathbb{R}^m) \to [0, 1] \) with density \( \pi : \mathbb{R}^m \to \mathbb{R}_+ \). As a practical matter in Bayesian inference, one is concerned with the computation of expectations with respect to \( \Pi \) of integrable functions \( h : \mathbb{R}^m \to \mathbb{R} \); that is, we wish to compute \( \mathbb{E}_{q \sim \Pi} [h(q)] \). Since \( \pi \) is generally intractable, these expectations are not available in closed form and Monte Carlo methods based on samples from \( \pi \) are often used to approximate these means. For instance, one may compute the mean and variance of the posterior by employing appropriate choices of expectation. Therefore, a central problem in Bayesian inference is the generation of samples from the posterior distribution.

One popular Monte Carlo method is based on approximate samples from the posterior distribution generated via the technique of Markov chain Monte Carlo (MCMC). In MCMC, given some initial state of the chain, \( q^0 \), a sequence of random variables \((q^1, q^2, \ldots)\) is generated inductively: the conditional distribution \( q^n | q^{n-1}, \ldots, q^0 \) obeys the Markov property. Under certain desirable regularity conditions, the random variables \( q^n \) will have a distribution which, asymptotically, converges to \( \Pi \). Establishing the rate of convergence, or at least an upper bound on that rate, serves to quantify how quickly the chain \((q^1, q^2, \ldots)\) mixes toward the target distribution. In the case where the rate of convergence is geometrically fast, the Markov chain enjoys a strong stability property in that
the estimator $n^{-1} \sum_{i=1}^{n} h(q^i)$ of $\mathbb{E}_{q \sim \Pi}[h(q)]$ can be equipped with a central limit theorem under additional technical assumptions. This central limit theorem has important practical implications in that it allows the use of asymptotically valid standard errors of MCMC estimates, which, in turn, can be used to decide how long to run the Markov chains [Roy 2020].

One of the most popular MCMC methods is Euclidean Hamiltonian Monte Carlo (EHMC) [Neal 2010b]. In EHMC, one computes approximate solutions to Hamilton’s equations of motion using numerical integrators, taking care to ensure that the resulting Markov chain satisfies detailed balance in a phase-space consisting of the position variables $q \in \mathbb{R}^m$ and auxiliary momentum variables $p \in \mathbb{R}^m$. The rate of convergence of HMC was established by Livingstone et al. [2016] in the presence of conditions on the log-density [see also Durmus et al. 2020]. Traditional EHMC, however, struggles in distributions that exhibit multiple spatial scales [Pourzanjani and Petzold 2019, Betancourt 2012]. This observation led to the development of geometric methods of HMC, specifically the Riemannian manifold Hamiltonian Monte Carlo (RMHMC) method of Girolami and Calderhead [2011].

Unlike EHMC, RMHMC adapts to the second-order structure of the posterior, which allows it to align its proposals in the direction of the posterior that exhibits the greatest local variation. However, the sophisticated form of the Hamiltonian employed in RMHMC necessitates the use of complex numerical integrators that are significantly more expensive than the numerical integrator employed in EHMC. This concern was partially alleviated in the introduction of Lagrangian Monte Carlo (LMC) in Lan et al. [2015], which was able to construct a MCMC method with a more efficient numerical integration procedure, while simultaneously continuing to take advantage of second-order geometric knowledge in order to efficiently traverse the posterior. Both RMHMC and LMC can be viewed as generalizations of the Euclidean Hamiltonian Monte Carlo algorithm, which incorporate second-order geometric information about the posterior into the Markov chain tran-
sition kernel. Incorporating second-order information allows these geometric methods of MCMC to explore the typical region of the target distribution more efficiently.

However, neither RMHMC nor LMC have, as of yet, been equipped with a geometric ergodicity theory. The focus in this work is in establishing the geometric ergodicity of modified versions of RMHMC and LMC. In implementations of RMHMC and LMC it is common to randomize the number of integration steps, and to take one-step with positive probability. In the case of Euclidean HMC, the equivalence between single-step HMC and the Metropolis-adjusted Langevin algorithm (MALA) is critical for establishing geometric ergodicity of HMC in Livingstone et al. [2016]. Our approach is two-fold. First, we propose a simple modification to both RMHMC and LMC which is motivated by a unique correspondence between single-step Euclidean HMC and MALA. In particular, we propose that instead of applying the transition kernel corresponding to RMHMC or LMC with a single integration step, one instead applies the transition kernel of the manifold Metropolis-adjusted Langevin algorithm (MMALA). Recent work in Roy and Zhang [2021] gave conditions under which MMALA is geometrically ergodic. We may imbue these modified variations of RMHMC and LMC with a geometric ergodicity theory via inheritance once one establishes that the RMHMC and LMC transition kernels (with a fixed, non-random number of integration steps) are reversible in the required sense. This construction is described in section 7.3. In section 7.4 we proceed to a numerical evaluation of the proposed modification of RMHMC and LMC, with special attention given to the probability of applying the MMALA transition kernel. We begin, however, in section 7.2 with an overview of required mathematical concepts from numerical integration and Markov chains.
7.2 Preliminaries

In section 7.2.1 we review the generalized leapfrog and Lagrangian leapfrog employed in geometric MCMC methods for performing Bayesian inference. Section 7.2.2 covers Markov chains based on involutions as well as those based on discretizations of Langevin diffusions. We also discuss geometric ergodicity, mixture transition kernels, and the equivalence between single-step HMC and MALA in the Euclidean regime.

7.2.1 Numerical Integrators

Hamiltonian and Lagrangian mechanics are equivalent, with the momentum and velocity being related by the Legendre transform [Marsden and Ratiu 2010]. In this section, we will consider numerical integrators of these equations of motion. Throughout our discussion, we will consider a Hamiltonian function of the following form.

**Definition 7.2.1.** Let $\pi : \mathbb{R}^m \to \mathbb{R}_+$ be a smooth probability density and let $G : \mathbb{R}^m \to \text{PD}(m)$. The Riemannian Hamiltonian is the function $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ defined by

$$H(q, p) = -\log \pi(q) + \frac{1}{2} \log \det(G(q)) + \frac{1}{2} p^\top G^{-1}(q)p.$$ (7.1)

Moreover, assuming $\int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \exp(-H(q, p)) \, dq \, dp < +\infty$, the Riemannian density is the probability density $\pi(q, p) \propto \exp(-H(q, p))$.

**Remark.** The term $\frac{1}{2} \log \det(G(q)) + \frac{1}{2} p^\top G^{-1}(q)p$ appearing in the definition of the function $H$ in eq. (7.1) is chosen such that the conditional density $\pi(q, p)$ satisfies

$$\pi(p|q) = \text{Normal}(0, G(q)).$$ (7.2)

It is easily seen that the $q$-marginal density of $\pi(q, p)$ obtained by marginalizing out $p$ is $\pi(q)$.
Hamiltonian functions $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ produce Hamilton’s equations of motion which, by definition, are solutions to the coupled differential equations

$$\dot{q}_t = \nabla_p H(q_t, p_t) \quad (7.3)$$
$$\dot{p}_t = -\nabla_q H(q_t, p_t). \quad (7.4)$$

Except in special cases, a closed-form solution for the map $t \mapsto (q_t, p_t) \in \mathbb{R}^m \times \mathbb{R}^m$ does not exist; this necessitates the use of numerical integrators in order to approximate the equations of motion obeying eqs. (7.3) and (7.4). One such example is the generalized leapfrog method, which we now define.

**Definition 7.2.2.** The **generalized leapfrog integrator** with step-size $\epsilon \in \mathbb{R}$ applied to the Riemannian Hamiltonian in eq. (7.1) is the map $(q, p) \mapsto \hat{\Phi}_{\epsilon}(q, p)$ defined by,

$$\ddot{p} = p - \frac{\epsilon}{2} \nabla_q H(q, \dot{p}) \quad (7.5)$$
$$\ddot{q} = q + \frac{\epsilon}{2} (\nabla_p H(q, \dot{p}) + \nabla_p H(\dot{q}, \ddot{p})) \quad (7.6)$$
$$= q + \frac{\epsilon}{2} (G^{-1}(q) + G^{-1}(\dot{q})) \ddot{p} \quad (7.7)$$
$$\ddot{\tilde{p}} = \ddot{p} - \frac{\epsilon}{2} \nabla_q H(\ddot{q}, \ddot{p}) \quad (7.8)$$
$$= p - \frac{\epsilon}{2} \nabla_q H(q, \ddot{p}) - \frac{\epsilon}{2} \nabla_q H(\ddot{q}, \ddot{p}), \quad (7.9)$$

where $\hat{\Phi}_{\epsilon}(q, p) = (\ddot{q}, \ddot{p})$.

**Remark.** One can prove [Leimkuhler and Reich, 2005, Hairer et al., 2006] that the generalized leapfrog integrator is a symplectic transformation and that therefore $|\det(\nabla \Phi(q, p))| = 1$.

**Remark.** We must now give a remark about notation. Numerical integration plays a central role in our analysis, and we give particular attention to multiple steps of integration. At the same time, we will also discuss Markov chains, which also consist of multiple steps.
As an attempt to differentiate these two notions of step, we will use a lower index to refer to steps of numerical integration, whereas we will use upper indices to denote Markov chain steps. Given initial data \((q_0, p_0) \in \mathbb{R}^m \times \mathbb{R}^m\), we denote \(k\)-steps of the generalized leapfrog integrator by \((q_k, p_k) = \hat{\Phi}_k(q_0, p_0) = \hat{\Phi}_\epsilon(q_{k-1}, p_{k-1})\). Similarly, we let \(\tilde{\beta}_k = p_{k-1} - \frac{\epsilon}{2} \nabla_q H(q_{k-1}, \tilde{\beta}_k)\), which we call the intermediate momentum at the \(k\)-th step.

A Hamiltonian of the form in eq. (7.1) may be transformed into a Lagrangian as

\[
L(q, \dot{q}) = -\log \pi(q) + \frac{1}{2} \log \det(G(q)) - \frac{1}{2} \dot{q}^\top G(q) \dot{q}.
\] (7.10)

As noted at the beginning of this subsection, Lagrangian and Hamiltonian mechanics are formally equivalent, with the momentum \(p \in \mathbb{R}^m\) being related to the velocity \(\dot{q} \in \mathbb{R}^m\) according to the Legendre transformation \(p = G(q) \dot{q}\). The Lagrangian produces equivalent equations of motion called the Euler-Lagrange equations as

\[
\nabla_q L(q_t, \dot{q}_t) = \frac{d}{dt} \nabla_{\dot{q}_t} L(q_t, \dot{q}_t).
\] (7.11)

One advantage possessed by the Lagrangian formalism over the Hamiltonian approach is that one may identify explicit numerical integrators [Lan et al., 2015] of the equations of motion, such as the Lagrangian leapfrog.

Definition 7.2.3. The Lagrangian leapfrog integrator with step-size \(\epsilon \in \mathbb{R}\) applied to the
Riemannian Hamiltonian in eq. (7.1) is the map \((q, p) \mapsto \tilde{\Phi}_\epsilon(q, p)\) defined by,

\[
v = G^{-1}(q)p
\]

\[
\tilde{v} = v - \frac{\epsilon}{2} G^{-1}(q) \nabla_q U(q) - \frac{\epsilon}{2} \Omega(q, v) \tilde{v}
\]

\[
\tilde{q} = q + \epsilon \tilde{v}
\]

\[
\tilde{v} = \tilde{v} - \frac{\epsilon}{2} G^{-1}(\tilde{q}) \nabla_q U(\tilde{q}) - \frac{\epsilon}{2} \Omega(\tilde{q}, \tilde{v}) \tilde{v}
\]

\[
= v - \frac{\epsilon}{2} G^{-1}(q) \nabla_q U(q) - \frac{\epsilon}{2} \Omega(q, v) \tilde{v} - \frac{\epsilon}{2} G^{-1}(\tilde{q}) \nabla_q U(\tilde{q}) - \frac{\epsilon}{2} \Omega(\tilde{q}, \tilde{v}) \tilde{v}
\]

\[
\tilde{p} = G^{-1}(\tilde{q}) \tilde{v}
\]

where \((\tilde{q}, \tilde{p}) = \tilde{\Phi}_\epsilon(q, p)\) and

\[
U(q) = -\log \pi(q) + \frac{1}{2} \log \det(G(q))
\]

\[
\Omega_{ij}(q, v) = \sum_{k=1}^m \Gamma^i_{kj}(q)v_k
\]

\[
\Gamma^i_{kj}(q) = \frac{1}{2} \sum_{l=1}^m G^{-1}_{kl}(q) \left( \frac{\partial}{\partial q_i} G_{lj}(q) + \frac{\partial}{\partial q_j} G_{il}(q) - \frac{\partial}{\partial q_l} G_{ij}(q) \right).
\]

The Jacobian determinant of the transformation \(\tilde{\Phi}_\epsilon\) (defined in definition 7.2.3) is,

\[
\left| \det(\nabla \tilde{\Phi}_\epsilon(q, p)) \right| = \det(G^{-1}(q)) \det(G(\tilde{q})) \left| \frac{\det(\text{Id} + \epsilon \Omega(q, v)/2) \det(\text{Id} - \epsilon \Omega(\tilde{q}, \tilde{v})/2)}{\det(\text{Id} + \epsilon \Omega(q, v)/2) \det(\text{Id} - \epsilon \Omega(q, v)/2)} \right|
\]

\[
(7.21)
\]

**Remark.** Given initial data \((q_0, p_0) \in \mathbb{R}^m \times \mathbb{R}^m\), we denote \(k\)-steps of the generalized (or Lagrangian) leapfrog integrator by \((q_k, p_k) = \tilde{\Phi}_\epsilon^{k}(q_0, p_0) = \tilde{\Phi}_\epsilon(q_{k-1}, p_{k-1})\).

In the case where \(G\) is a constant function of \(q\), the generalized leapfrog integrator reduces to the standard leapfrog integrator, which is defined as follows.

**Definition 7.2.4.** Let \(\pi : \mathbb{R}^m \rightarrow \mathbb{R}_+\) be a smooth probability density. A Euclidean Hamil-
tonian is a smooth map \( H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \) of the form,

\[
H(q, p) = -\log \pi(q) + \frac{1}{2} p^\top G^{-1} p,
\]

(7.22)

where \( G \in \text{PD}(m) \).

**Definition 7.2.5.** Consider a Euclidean Hamiltonian as defined in definition 7.2.4. The **Euclidean leapfrog integrator** is the map \((q, p) \mapsto \tilde{\Phi}_\epsilon(q, p)\) defined by

\[
\begin{align*}
\tilde{p} &= p + \frac{\epsilon}{2} \nabla \log \pi(q) \\
\tilde{q} &= q + \epsilon G^{-1} \tilde{p} \\
\tilde{p} &= \tilde{p} + \frac{\epsilon}{2} \nabla \log \pi(\tilde{q}),
\end{align*}
\]

(7.23) (7.24) (7.25)

where \( \tilde{\Phi}_\epsilon(q, p) = (\tilde{q}, \tilde{p}) \).

The Euclidean leapfrog integrator is the *de-facto* standard numerical integrator employed in EHMC. This is because of its accuracy and computational efficiency; in contrast to the generalized leapfrog integrator in definition 7.2.2, the Euclidean leapfrog is fully explicit. Nevertheless, the generalized leapfrog and Euclidean leapfrog are exactly equivalent when applied to a Hamiltonian in the form of definition 7.2.4; this is made precise in the following result.

**Proposition 7.2.6.** When the generalized leapfrog (definition 7.2.2) or Lagrangian leapfrog (definition 7.2.3) is applied to a Euclidean Hamiltonian in the form of definition 7.2.4 (in the sense that we take \( G(q) \equiv G \) for every \( q \in \mathbb{R}^m \)), the resulting map is equivalent to the Euclidean leapfrog method in definition 7.2.5.

**Proof.** In the case of the generalized leapfrog integrator, this is immediate from the fact that \( \nabla_q H(q, p) = -\nabla \log \pi(q) \) which does not depend on \( p \) and \( \nabla_p H(q, p) = Gp \) which does not depend on \( q \).
In the case of the Lagrangian leapfrog, the form of the Euclidean Hamiltonian means that \( \Omega(q, v) \) is uniformly zero (because all of the derivatives of the metric vanish). The Lagrangian leapfrog integrator may be re-expressed as

\[
v = G^{-1} \dot{p} \tag{7.26}
\]

\[
\ddot{v} = v + \frac{\epsilon}{2} G^{-1} \nabla \log \pi(q) \tag{7.27}
\]

\[
= G^{-1} \ddot{p} \tag{7.28}
\]

\[
\ddot{q} = q + \epsilon \ddot{v} \tag{7.29}
\]

\[
= q + \epsilon G^{-1} \ddot{p} \tag{7.30}
\]

\[
\ddot{v} = \ddot{v} + \frac{\epsilon}{2} G^{-1} \nabla \log \pi(\ddot{q}) \tag{7.31}
\]

\[
= G^{-1} \ddot{p}, \tag{7.32}
\]

where \( \ddot{p} \) and \( \ddot{p} \) are as in eqs. (7.23) and (7.25).

### 7.2.2 Markov Chains

The purpose of this section is to review fundamentals of Markov chains, including their construction, convergence properties, and central limit theorems. Markov chains are stochastic processes that may be inductively defined by their transition kernel. Since the transition kernel yields a probability measure depending only on the current state, one sees by inspection that the Markov chain satisfies the Markov property.

**Definition 7.2.7.** A Markov chain transition kernel on \( \mathbb{R}^m \) is a function \( K : \mathbb{R}^m \times \mathcal{B}(\mathbb{R}^m) \to [0, 1] \) such that (i) for \( q \in \mathbb{R}^m \), \( K(q, \cdot) \) is a probability measure and (ii) for fixed \( A \in \mathcal{B}(\mathbb{R}^m) \) the function \( q \mapsto K(q, A) \) is measurable.

Within the context of simulation-based inference and mainly Bayesian inference, one is interested in constructing a Markov chain which converges to a specified target probability
distribution; as noted in the introduction, this target distribution is typically known by its density up to a normalizing constant. We now define two notions of convergence.

**Definition 7.2.8.** Given a Markov chain transition kernel \( K \), a Markov chain is a sequence of random variables defined inductively by \( q^{n+1} \mid q^n \sim K(q^n, \cdot) \).

**Definition 7.2.9.** The *total variation distance* between two probability measures \( \Pi : \mathcal{B}(\mathbb{R}^m) \rightarrow [0, 1] \) and \( \Xi : \mathcal{B}(\mathbb{R}^m) \rightarrow [0, 1] \) is defined by

\[
\|\Pi - \Xi\|_{TV} = 2 \sup_{A \in \mathcal{B}(\mathbb{R}^m)} |\Pi(A) - \Xi(A)|.
\] (7.33)

**Definition 7.2.10.** A Markov chain with transition kernel \( K \) is said to be ergodic if

\[
\lim_{n \to \infty} \|K^n(q, \cdot) - \Pi(\cdot)\|_{TV} = 0,
\] (7.34)

where for \( A \in \mathcal{B}(\mathbb{R}^m) \), \( K^n(q, A) = \Pr(q^n \in A \mid q^0 = q) \) is the \( n \)-step Markov transition probability.

It will turn out that one can establish ergodicity of a Markov chain if one can establish three separate properties: irreducibility, aperiodicity, and stationarity. We now define each of these concepts.

**Definition 7.2.11.** A Markov chain transition kernel \( K \) is said to be \( \Pi \)-irreducible if for every set \( A \in \mathcal{B}(\mathbb{R}^m) \) with \( \Pi(A) > 0 \) there exists \( n \in \mathbb{N} \) for which \( K^n(q, A) > 0 \) for all \( q \in \mathbb{R}^m \).

**Definition 7.2.12.** Given a Markov chain transition kernel \( K : \mathbb{R}^m \times \mathcal{B}(\mathbb{R}^m) \rightarrow [0, 1] \), a set \( C \subseteq \mathbb{R}^m \) is called small if there exists a \( n \in \mathbb{N} \), a \( \delta > 0 \), and a probability measure \( \nu : \mathcal{B}(\mathbb{R}^m) \rightarrow [0, 1] \) such that for any \( q \in C \) and \( A \in \mathcal{B}(\mathbb{R}^m) \) we have \( K^n(q, A) \geq \delta \cdot \nu(A) \).

We may therefore say that the set \( C \) is \((n, \delta, \nu)\)-small.

**Definition 7.2.13.** A Markov chain transition kernel \( K \) is called aperiodic if there exists a
small set $C$ for which the greatest common divisor of the set

$$\{n' \in \mathbb{N} : \exists \delta > 0 \text{ and probability measure } \nu \text{ such that } C \text{ is } (n', \delta, \nu)\text{-small}\} \quad (7.35)$$

is one.

**Definition 7.2.14.** A Markov chain with transition kernel $K$ is said to be stationary for the probability measure $\Pi$ if for any Borel set $A \in \mathcal{B}(\mathbb{R}^m)$ we have $\mathbb{E}_{q \sim \Pi} K(q, A) = \Pi(A)$.

**Definition 7.2.15.** A Markov chain with transition kernel $K$ is called geometrically ergodic if there exists $\rho \in (0, 1)$ and function $V : \mathbb{R}^m \to \mathbb{R}_+$ such that

$$\|K^n(q, \cdot) - \Pi(\cdot)\|_{TV} \leq V(q)\rho^n. \quad (7.36)$$

It is clear from the definitions that geometric ergodicity is a stronger form of convergence than mere ergodicity. In the latter case, however, simple conditions under which a Markov chain is ergodic can be provided.

**Theorem 7.2.16** (Tierney [1994]). Suppose that $K$ is a Markov chain transition kernel that is $\Pi$-irreducible, aperiodic, and for which $\Pi$ is the stationary probability measure. Then $K$ produces an ergodic Markov chain.

However, in addition to giving an upper bound on the mixing of the Markov chain into the target probability measure, geometric ergodicity is also useful for establishing a central limit theorem for expectations computed from the sequence $(q^1, q^2, \ldots)$. We make this precise as follows.

**Theorem 7.2.17** (Meyn and Tweedie [1993]). Let $h : \mathbb{R}^m \to \mathbb{R}$ be a function for which $\mathbb{E}_{q \sim \Pi} |h(q)|^{2+\gamma} < \infty$ for some $\gamma > 0$. Let $K$ be a Markov chain transition kernel that is aperiodic, $\Pi$-irreducible, and for which $\Pi$ is the unique stationary distribution. Assume further that $K$ converges geometrically to $\Pi$. Define $S_K[h] = n^{-1} \sum_{i=1}^n h(q^i)$ where
\(q^{i+1}|q^i \sim K(q^i, \cdot)\). Then, as \(n \to \infty\),

\[
\sqrt{n} \left( S_K[h] - \mathbb{E}_{q \sim \Pi}[h(q)] \right) \xrightarrow{d} \text{Normal}(0, \tau_h^2) \tag{7.37}
\]

where \(\tau_h^2\) is a constant, less than infinity, that depends on \(h\) (and \(K\)). The quantity \(\tau_h^2\) has a closed-form given by

\[
\tau_h^2 = \text{Var}_{q \sim \Pi}(h(q)) + 2 \sum_{k=1}^{\infty} \text{Cov}_{q \sim \Pi}(h(q), h(q^k)). \tag{7.38}
\]

Establishing that a Markov chain has \(\Pi\) as a stationary distribution is often easily achieved by showing that the chain satisfies detailed balance with respect to \(\Pi\). As we require detailed balance when discussing the marginal transition kernels of RMHMC and LMC in section 7.3, we define this notion now.

**Definition 7.2.18.** A Markov chain transition kernel is said to satisfy detailed balance with respect to the probability distribution \(\Pi\) (equivalently, the Markov chain is called reversible with respect to \(\Pi\)) if for any sets \(Q, Q' \in \mathcal{B}(\mathbb{R}^m)\),

\[
\int_Q K(q, Q') \, \Pi(dq) = \int_{Q'} K(q, Q) \, \Pi(dq). \tag{7.39}
\]

**Remark.** When detailed balance holds for a Markov chain transition kernel \(K\), it follows that \(\Pi\) is the stationary probability measure (definition 7.2.14) of the Markov transition kernel \(K\).

**Metropolis-Hastings Kernels and Generalized Langevin Algorithms**

In this section we review Metropolis-Hastings Markov chains. We begin by formally defining these objects before proceeding to a special case of Metropolis-Hastings method based on Langevin diffusion, which recalls the constructions considered in [Roy and Zhang](unnamed).
Definition 7.2.19 (Metropolis-Hastings Algorithm). Let \( \pi : \mathbb{R}^m \to \mathbb{R}_+ \) be a probability density and, for each \( \tilde{q} \in \mathbb{R}^m \), let \( \tilde{\pi}(\cdot|\tilde{q}) : \mathbb{R}^m \to \mathbb{R}_+ \) be a probability density depending on \( \tilde{q} \). The Metropolis-Hastings transition kernel with proposal density \( \tilde{\pi}(\cdot|\tilde{q}) \) is,

\[
K(q, A) = \int_{\tilde{q}} \min \left\{ 1, \frac{\pi(\tilde{q}) \tilde{\pi}(q|\tilde{q})}{\tilde{\pi}(\tilde{q}|q)} \right\} \tilde{\pi}(q|\tilde{q}) \, d\tilde{q} + \left( 1 - \int_{\mathbb{R}^m} \min \left\{ 1, \frac{\pi(\tilde{q}) \tilde{\pi}(q|\tilde{q})}{\tilde{\pi}(\tilde{q}|q)} \right\} \tilde{\pi}(q|\tilde{q}) \, d\tilde{q} \right) 1 \{ q \in A \},
\]

for \( q \in \mathbb{R}^m \) and \( A \in \mathcal{B}(\mathbb{R}^m) \).

Definition 7.2.20 (Generalized Metropolis-Adjusted Langevin Algorithm). Fix \( \epsilon \in \mathbb{R} \backslash \{0\} \). Let \( c_\epsilon : \mathbb{R}^m \to \mathbb{R}^m \) and let \( A : \mathbb{R}^m \to \text{PD}(m) \). The generalized Metropolis-adjusted Langevin algorithm is an instance of the Metropolis-Hastings algorithm with proposal density

\[
\tilde{\pi}(\tilde{q}|q) = \text{Normal}(\tilde{q}; c_\epsilon(q), \epsilon^2 A(q)).
\]

We denote the Markov chain transition kernel of the generalized Metropolis-adjusted Langevin algorithm with step-size \( \epsilon \) by \( J_\epsilon : \mathbb{R}^m \times \mathcal{B}(\mathbb{R}^m) \to [0, 1] \).

Definition 7.2.21 (Riemannian Manifold Metropolis-Adjusted Langevin Algorithm). In the special case where

\[
c_\epsilon(q) = \frac{\epsilon^2}{2} A(q) \nabla \log \pi(q) + \frac{\epsilon^2}{2} \Gamma(q)
\]

\[
\Gamma_i(q) = \sum_{j=1}^m \frac{\partial}{\partial q_j} A_{ij}(q)
\]

we call the resulting method the Riemannian manifold Metropolis-adjusted Langevin algorithm (MMALA).

Remark. The form of the proposal in eq. (7.42) is based on an Euler-Maruyama discretiza-
tion of a Langevin diffusion on a manifold with Riemannian metric $G(q) = A^{-1}(q)$; see Xifara et al. [2014] for details. In summary, eq. (7.42) is the drift component of the Euler-Maruyama discretization (with step-size $\epsilon^2$) applied to the following stochastic differential equation:

$$dX_t = \frac{1}{2} G^{-1}(X_t) \nabla \log \pi(X_t) \, dt + \Gamma(X_t) \, dt + G^{-1/2}(X_t) \, dB_t,$$  

(7.44)

where $B_t$ is Euclidean Brownian motion at time $t \in \mathbb{R}$. When $G(q)$ is the Fisher information matrix, the term $G^{-1}(X_t) \nabla \log \pi(X_t)$ can be identified as the natural gradient of the function $\log \pi$ under the geometry generated by the Fisher metric [Amari and Nagaoka 2000] at position $X_t$. On the other hand, the term $\Gamma(X_t) \, dt + G^{-1/2}(X_t) \, dB_t$ corresponds to manifold Brownian motion, since its infinitesimal generator is the Laplace-Beltrami operator on the manifold [Hsu 2002]. Therefore, as the stochastic differential equation in eq. (7.44) consists of a term comprising the manifold gradient of a log-density and another term comprising manifold Brownian motion, it is called the Riemannian Langevin equation in correspondence with the Euclidean case.

**Definition 7.2.22 (Simplified Riemannian Manifold Metropolis-Adjusted Langevin Algorithm).** Computing the function $\Gamma$ in definition 7.2.21 may be inconvenient to evaluate. It can be ignored while still yielding a Markov chain transition kernel, giving the special case,

$$c_\epsilon(q) = \frac{\epsilon^2}{2} A(q) \nabla \log \pi(q).$$  

(7.45)

The resulting method is called the simplified Riemannian manifold Metropolis-adjusted Langevin algorithm, which was considered by Girolami and Calderhead [2011].

We adopt the abbreviation SMALA to refer to the simplified Metropolis-adjusted Langevin algorithm.
Remark. There are three common choices of the function $A$. The first is that $A$ is a constant function, in which case we simply write $A \in \text{PD}(m)$. A second option is that $A$ is chosen as the inverse of the sum of the Fisher information matrix and the negative Hessian of the log-prior, which can capture second-order geometry of both the likelihood and the prior; the use of the inverse of the sum of the Fisher information and the negative Hessian of the log-prior as a preconditioner is the approach advocated by [Girolami and Calderhead 2011]. A third option is that $A$ is the inverse of the SoftAbs metric, which is a smooth transformation of the Hessian of the log-density of the target distribution; for details see [Betancourt 2012].

Geometric Methods of Bayesian Inference

Definition 7.2.23 (Involutive Monte Carlo [Neklyudov et al., 2020]). Let $\Phi : \mathbb{R}^m \to \mathbb{R}^m$ be a smooth involution and let $\pi : \mathbb{R}^m \to \mathbb{R}_+$ be a probability density on $\mathbb{R}^m$. The Markov chain transition kernel of involutive Monte Carlo with target density $\pi$ is

$$K(q, A) = \min \left\{ 1, \frac{\pi(\Phi(q))}{\pi(q)} \left| \text{det}(\nabla \Phi(q)) \right| \right\} 1 \{ \Phi(q) \in A \}
+ \left( 1 - \min \left\{ 1, \frac{\pi(\Phi(q))}{\pi(q)} \left| \text{det}(\nabla \Phi(q)) \right| \right\} \right) 1 \{ q \in A \},$$

(7.46)

for $q \in \mathbb{R}^m$ and $A \in \mathcal{B}(\mathbb{R}^m)$.

It is easily verified that the transition kernel of involutive Monte Carlo satisfies detailed balance with respect to the density $\pi$. Involutive Monte Carlo gives rise to two special transition kernels corresponding to RMHMC and LMC.

Definition 7.2.24. Let $\hat{\Phi}_\epsilon$ be as in definition 7.2.2 and let $H$ be as in definition 7.2.1. The involution of Riemannian manifold Hamiltonian Monte Carlo with step-size $\epsilon$ and $k$
integration steps is,

$$
\Phi = F \circ \tilde{\Phi}_\epsilon \circ \cdots \circ \tilde{\Phi}_\epsilon, \quad (7.47)
$$

where $F(q, p) = (q, -p)$. The target density of Riemannian manifold Hamiltonian Monte Carlo is $\pi(q, p) \propto \exp(-H(q, p))$. The transition kernel of involutive Monte Carlo (definition 7.2.23) with involution $\Phi$ given in eq. (7.47) is called the transition kernel of Riemannian manifold Hamiltonian Monte Carlo with step-size $\epsilon$ and $k$ integration steps.

**Remark.** It can be shown that if $\Phi$ is an invertible function and if $F \circ \Phi$ is an involution that $F \circ \Phi \circ \cdots \Phi$ is also an involution. This explains why eq. (7.47) is also an involution. For further details, see proposition 7.A.1.1.

In the special case where $G$ is a constant function of $q$, the resulting Markov chain is called Euclidean Hamiltonian Monte Carlo.

**Definition 7.2.25.** Let $\tilde{\Phi}_\epsilon$ be as in definition 7.2.5 and let $H$ be as in definition 7.2.4. The involution of Euclidean Hamiltonian Monte Carlo (EHMC) with step-size $\epsilon$ and $k$ integration steps is,

$$
\Phi = F \circ \tilde{\Phi}_\epsilon \circ \cdots \circ \tilde{\Phi}_\epsilon, \quad (7.48)
$$

where $F(q, p) = (q, -p)$. The target density of Euclidean Hamiltonian Monte Carlo is $\pi(q, p) \propto \exp(-H(q, p))$. The transition kernel of involutive Monte Carlo (definition 7.2.23) with involution $\Phi$ given in eq. (7.48) is called the transition kernel of Euclidean Hamiltonian Monte Carlo with step-size $\epsilon$ and $k$ integration steps.

**Definition 7.2.26.** Let $\tilde{\Phi}_\epsilon$ be as in definition 7.2.3 and let $H$ be as in definition 7.2.1. The
involution of Lagrangian Monte Carlo with step-size $\epsilon$ and $k$ integration steps is,

$$
\Phi = F \circ \tilde{\Phi}_\epsilon \circ \cdots \circ \tilde{\Phi}_\epsilon
$$

(7.49)

with $F(q, p) = (q, -p)$. The target density of Lagrangian Hamiltonian Monte Carlo is $\pi(q, p) \propto \exp(-H(q, p))$. We call this the transition kernel of Lagrangian Monte Carlo with step-size $\epsilon$ and $k$ integration steps.

Remark. Unlike the transition kernel employed in Riemannian manifold Hamiltonian Monte Carlo, the Jacobian determinant of the involution appearing in Lagrangian Monte Carlo must be computed through $k$ applications of eq. (7.21), where, as before, $k$ is the number of integration steps.

In the special case where $G$ is a constant function of $q$, Riemannian manifold Hamiltonian Monte Carlo and Lagrangian Monte Carlo produces a transition kernel that is equivalent to Euclidean Hamiltonian Monte Carlo with constant mass matrix $G$.

**Proposition 7.2.27.** The transition kernels of RMHMC (or LMC) with step-size $\epsilon$ and $k$ integration steps when applied to the Euclidean Hamiltonian in definition 7.2.4 (in the sense that we take $G(q) \equiv G$ for every $q \in \mathbb{R}^m$) are both equivalent to the EHMC transition kernel.

**Proof.** It follows from proposition 7.2.6 that the proposals produced by RMHMC and LMC are identical to EHMC. It remains to be verified that the acceptance decisions are identical, too. This can be verified by observing that the Riemannian Hamiltonian in eq. (7.1) is the Euclidean Hamiltonian in eq. (7.22) up to an additive constant (since $G$ does not depend on position). Hence, the Euclidean and Riemannian densities are proportional to one another, and the acceptance probabilities will be identical. □

We now elaborate on the connection between EHMC with a single-step and the Metropolis-adjusted Langevin algorithm: in particular, these two methods can be constructed to be
exactly equivalent in $q$-space, producing equal proposals and identical acceptance probabilities.

**Proposition 7.2.28** (Neal [2010a]). Let $A \in \mathbb{R}^{m \times m}$ be a fixed positive definite matrix. Consider EHMC with the Hamiltonian $H(q, p) = -\log \pi(q) + \frac{1}{2} p^\top A^{-1} p$ and MALA with proposal distribution $\tilde{q}|q \sim \text{Normal}(q + \frac{\epsilon^2}{2} A^{-1} \nabla \log \pi(q), \epsilon^2 A^{-1})$. The marginal transition kernel of EHMC with step-size $\epsilon$ and a single integration step can be constructed to be exactly equivalent to the transition kernel of MALA.

A proof is provided in section 7.A.1.1. The analysis of mixture transition kernels will be central to our analysis. In standard implementations of Riemannian manifold Hamiltonian and Lagrangian Monte Carlo, it is typical to randomize the number of integration steps. This is done to avert any Markov chain pathologies (such as irreducibility failures) that may result from using a fixed number of integration steps. We therefore consider Markov chain transition kernels which are mixtures.

**Definition 7.2.29.** Let $K_{\epsilon,k} : (\mathbb{R}^m \times \mathbb{R}^m) \times \mathcal{B}(\mathbb{R}^m \times \mathbb{R}^m) \to \mathbb{R}_+$ be the Markov chain transition kernel of RMHMC (or LMC) with step-size $\epsilon$ and $k$ integration steps. Let $\tilde{K}_{\epsilon,k}$ be the marginal transition kernel (defined in lemma 7.2.30) of $K_{\epsilon,k}$. Define the mixture transition kernel of RMHMC (or LMC) to be

$$\tilde{K}_\epsilon = \sum_{k=1}^{\infty} \alpha_k \tilde{K}_{\epsilon,k}, \tag{7.50}$$

where $(\alpha_1, \alpha_2, \ldots)$ is a probability vector.

The fact that EHMC, RMHMC, and LMC all satisfy detailed balance in $(q, p)$-space causes us to examine the progression of $q$-states alone in between Gibbs resampling steps of the momentum. This leads to marginal Markov chain transition kernels of the following form:
Lemma 7.2.30. Let $K : (\mathbb{R}^m \times \mathbb{R}^m) \times \mathfrak{B}(\mathbb{R}^m \times \mathbb{R}^m) \to \mathbb{R}$ be a Markov chain transition kernel. Suppose that $K$ satisfies detailed balance with respect to the density $\pi(q, p)$ with $q$-marginal distribution $\pi(q)$ and conditional density $\pi(p|q)$; i.e. $\pi(q, p) = \pi(q) \pi(p|q)$. Consider the marginal chain constructed as follows. Given $q^n = q$, sample $p^n \sim \pi(p|q)$, sample $(q^{n+1}, p^{n+1}) \sim K((q^n, p^n), \cdot)$ and discard both momenta. The transition kernel of the marginal chain satisfies

$$
\tilde{K}(q, Q) = \int_{\mathbb{R}^m} K((q, p), (Q, \mathbb{R}^m)) \pi(p|q) \, dp
$$

(7.51)

where $Q \in \mathfrak{B}(\mathbb{R}^m)$.

A proof is given in section 7.A.1.2

Proposition 7.2.31. Let $\tilde{K}$ be the marginal transition kernel described in lemma 7.2.30. The marginal chain satisfies detailed balance with respect to the distribution whose density is $\pi(q)$.

A proof is provided in section 7.A.1.3

Ergodicity Theorems

The geometric ergodicity of MALA, MMALA, and SMALA have been examined in the Markov chain literature. In section 7.3 we will see how the geometric ergodicity of a single component of a mixture Markov chain transition kernel implies the geometric ergodicity of the mixture itself. With this destination in mind, we now recall two key results in this direction.

Theorem 7.2.32 (Roberts and Tweedie [1996]). Consider the transition kernel $K$ of the Metropolis-adjusted Langevin algorithm with $A = \text{Id}$ and $c_e(q) = x + \frac{c^2}{2} \nabla \log \pi(q)$. Then $K$ is geometrically ergodic under the following conditions:
1. We have

$$\liminf_{\|q\|\to\infty} \left( \|q\| - \|q + \frac{\epsilon^2}{2} \nabla \log \pi(q)\| \right) > 0. \quad (7.52)$$

2. We have

$$\lim_{\|q\|\to\infty} \int_{(A(q) \cup I(q)) \cap (A(q) \cap I(q))} \tilde{\pi}(\tilde{q}|q) \, d\tilde{q} = 0, \quad (7.53)$$

where

$$A(q) = \{ \tilde{q} \in \mathcal{X} : \pi(q) \tilde{\pi}(\tilde{q}|q) \leq \pi(y) \tilde{\pi}(\tilde{q}|\tilde{q}) \} \quad (7.54)$$

$$I(q) = \{ \tilde{q} \in \mathcal{X} : \|\tilde{q}\| \leq \|q\| \}. \quad (7.55)$$

**Theorem 7.2.33** (Roy and Zhang [2021]). Consider the transition kernel $K$ of the generalized Metropolis-adjusted Langevin algorithm. Then $K$ is geometrically ergodic under the following conditions:

1. There exist matrices $A_l \in \text{PD}(m)$ and $A_u \in \text{PD}(m)$ such that $A_l \leq A(q) \leq A_u$.

2. When $A \subset \mathbb{R}^m$ is bounded, the function $c_\epsilon : \mathbb{R}^m \to \mathbb{R}^m$ is bounded on $A$.

3. There is a quantity

$$C = \limsup_{\|q\|\to\infty} \int_{R(q)} \left( 1 - \tilde{\pi}(\tilde{q}|q) \min \left\{ 1, \frac{\pi(\tilde{q}) \tilde{\pi}(\tilde{q}|\tilde{q})}{\pi(q) \tilde{\pi}(\tilde{q}|q)} \right\} \right) \, dy \quad (7.56)$$

which is strictly less than one, where

$$R(q) = \{ \tilde{q} \in \mathbb{R}^m : \pi(q) \tilde{\pi}(\tilde{q}|x) > \pi(q) \tilde{\pi}(\tilde{q}|q) \}. \quad (7.57)$$
4. There exists \( s > 0 \) such that,

\[
\liminf_{\|q\| \to \infty} (\|A_{u}^{-1/2}(q)x\| - \|A_{u}^{-1/2}(q)c_{\epsilon}(q)\|) > \frac{\log(D(s)) - \log(1 - C)}{s},
\]

where,

\[
D(s) = \epsilon^{-m/2} \left( \frac{\pi}{2} \right)^{m/2-1} \left( \frac{\det(A_{u})}{\det(A_{l})} \right)^{1/2} \exp(\epsilon s^2 / 2) \int_{0}^{\infty} \exp\left( -(r - \epsilon s)^2 / 2\epsilon \right) r^{m-1} \, dr
\]

In section 7.3 we will examine mixture transition kernels and their geometric ergodicity. In the case that one transition kernel is geometrically ergodic and all transition kernels are reversible, the mixture transition kernel is geometrically ergodic, too, from the following result.

**Theorem 7.2.34** (Lee and Łatuszyński [2014]). Let \((\alpha_1, \alpha_2, \ldots)\) be a sequence satisfying \( \alpha_i \geq 0 \) and \( \sum_{i=1}^{\infty} \alpha_i = 1 \). Let \( \tilde{K} = \sum_{i=1}^{\infty} \alpha_i K_i \) be a mixture of reversible transition kernels with invariant distribution \( \Pi \). Suppose that \( K_1 \) satisfies the properties: (i) \( \Pi \) is the unique invariant distribution, and (ii) \( K_1 \) is geometrically ergodic. Then \( \tilde{K} \) is geometrically ergodic.

We term the derived geometric ergodicity of \( \tilde{K} \) from \( K_1 \) as “inherited” geometric ergodicity.

### 7.3 Inherited Geometric Ergodicity

In this section we discuss the notion of “inherited” geometric ergodicity, wherein we modify RMHMC and LMC to be geometrically ergodic when MMALA is. In the Euclidean case, by proposition 7.2.28, in cases where the Metropolis-adjusted Langevin algorithm is geometrically ergodic, so is HMC by invoking theorem 7.2.34. The situation in the ge-
ometric setting is more complicated, since there is no analogue of proposition 7.2.28 to show that the marginal transition kernel of a single-step of RMHMC or LMC is exactly equivalent to the Markov chain transition kernel of the generalized Metropolis-adjusted Langevin algorithm for a suitable choice of mean function. This leads us to propose the following Markov chain transition kernel.

**Definition 7.3.1.** Let $K_{\epsilon,k} : (\mathbb{R}^m \times \mathbb{R}^m) \times \mathcal{B}(\mathbb{R}^m \times \mathbb{R}^m) \to \mathbb{R}_+$ be the Markov chain transition kernel of RMHMC (or LMC) with step-size $\epsilon$ and $k$ integration steps. Let $\tilde{K}_{\epsilon,k}$ be the marginal transition kernel (defined in lemma 7.2.30). Let $(\alpha_1, \alpha_2, \ldots)$ be a sequence satisfying $\alpha_i \geq 0$ and $\sum_{i=1}^{\infty} \alpha_i = 1$. Let $J_\epsilon$ be the Markov chain transition kernel of the generalized Metropolis-adjusted Langevin algorithm (defined in definition 7.2.20). The **Langevin mixture transition kernel of RMHMC (or LMC)**, which we abbreviate by $\text{LMRMHMC (or LMLMC)}$, is defined by

$$\tilde{K}_\epsilon = \alpha_1 J_\epsilon + \sum_{k=2}^{\infty} \alpha_k \tilde{K}_{\epsilon,k}.$$  

(7.60)

We call $\alpha_1$ the MMALA mixture weight.

The modified transition kernel simply replaces a single-step of RMHMC (or LMC) by the transition kernel of the generalized Metropolis-adjusted Langevin algorithm. In order to apply theorem 7.2.34, it is necessary to verify that the marginal transition kernels of RMHMC and LMC are reversible. Fortunately, this is readily shown as follows.

**Corollary 7.3.2.** For fixed $\epsilon \in \mathbb{R}$ and $k \in \mathbb{N}$, the RMHMC and LMC transition kernels with step-size $\epsilon$ and $k$ integration steps satisfy detailed balance in $(q,p)$-space, it follows that their marginal chains are reversible with respect to the distribution $\Pi$ by proposition 7.2.31.

**Proposition 7.3.3.** Suppose $\alpha_1 > 0$. Under the conditions of theorem 7.2.33 the Markov chain transition kernel of LMRMHMC (or LMLMC) in eq. (7.60) is geometrically ergodic.
Proof. This follows as an immediate consequence of theorem 7.2.34 using the fact that the marginal transition kernels or RMHMC (or LMC) are reversible by corollary 7.3.2.

Remark. As a practical matter, we choose the mixture probabilities \((\alpha_1, \alpha_2, \ldots)\) in the following way. For a particular target distribution, we will choose a maximal number of integration steps \(k_{\text{max}}\) for which \(\alpha_k = 0\) for \(k > k_{\text{max}}\). Then, given a particular selection of \(\alpha_1 \in (0, 1]\), we split the remaining probability mass equally for each \(k \in \{2, \ldots, k_{\text{max}}\}\); that is, the fraction can be expressed as \(\alpha_k = \frac{1 - \alpha_1}{k_{\text{max}} - 1}\) for \(k = 2, \ldots, k_{\text{max}}\). In our experiments, we will consider variable choices for the mixing parameter \(\alpha_1\). The special case \(\alpha_1 = 0\), will correspond by convention to the unmodified RMHMC and LMC transition kernels with a single-step computed using the prescribed involution, as described in definition 7.2.29.

Mixing with the MMALA transition kernel also immediately establishes irreducibility, aperiodicity, and the smallness of all compact sets, as the following result reveals.

Lemma 7.3.4. Let \(\pi(q) \propto \exp(\mathcal{L}(q))\) be continuous and bounded on compact sets and denote by \(\Pi\) the probability measure with density \(\pi\). Suppose \(\alpha_1 > 0\). The marginal transition kernel of LMRMHMC (or LMLMC) is \(\Pi\)-irreducible, aperiodic, and all non-negligible compact sets are small.

Proof. With probability \(\alpha_1\), the Markov chain transitions according to a Metropolis-Hastings accept-reject decision with a normal proposal distribution. Hence, for any set \(A \subset \mathcal{B}(\mathbb{R}^m)\) for which \(\int_A \pi(q) \, dq > 0\), we have,

\[
\tilde{K}_\epsilon(q, A) \geq \alpha_1 J_\epsilon(q, A)
\]

\[
> 0
\]

since a normal proposal distribution is non-vanishing everywhere on \(\mathbb{R}^m\). The fact that \(\tilde{K}_\epsilon\) is aperiodic and that all non-negligible compact sets are small follows as an immediate
consequence of Lemma 1.2 from Mengersen and Tweedie [1996].

Corollary 7.3.5. From lemma 7.3.4 we have that the modified Markov chain of RMHMC (or LMC) is \(\Pi\)-irreducible, aperiodic, and from proposition 7.2.31 \(\Pi\) is the stationary distribution. Hence, it follows from theorem 7.2.16 that the modified Markov chain of RMHMC (or LMC) produces an ergodic Markov chain.

7.4 Experimentation

We turn now to the investigation of the proposed modified variations of RMHMC and LMC. These examples are chosen to represent a wide class of posterior distributions. In computing the convergence of the Markov chain under the maximum mean discrepancy metric, we measure similarity between 10,000 i.i.d. samples of the target distribution, and an equal number of independent Markov chains; in this case, we compute the unbiased estimator of the maximum mean discrepancy over the course of the first one-hundred sampling steps. For the expected squared jump distance (ESJD) and the effective sample size (ESS) metrics, we consider a ten replicates of a long Markov chain consisting of 1,000,000 samples, except in the case of the Fitzhugh-Nagumo posterior where we sample only 100,000 times. In each experiment, we consider mixing LMC and RMHMC with the transition kernel described in definition 7.2.21 except in the case of Neal’s funnel distribution, wherein we consider a mixture with definition 7.2.22. We begin in section 7.4.1 to describe measures and metrics by which we may assess the performance of the Markov chains corresponding to LMRHMC and LMLMC.

Code for reproducing these experimental results may be found at [https://tinyurl.com/29kz7krz](https://tinyurl.com/29kz7krz).
7.4.1 Measures and Metrics

We now give details of the evaluation metrics by which we compare Markov chains. We first recall the method of maximum mean discrepancy due to [Gretton et al. 2012].

**Definition 7.4.1.** Let \( k : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \) be a positive definite function that is symmetric in its arguments. Let \( \Pi \) and \( \Pi' \) be two probability distributions on \( \mathbb{R}^m \). The squared maximum mean discrepancy between \( \Pi \) and \( \Pi' \) is defined by,

\[
\text{MMD}^2(k, \Pi, \Pi') = \mathbb{E}_{q,q' \sim \Pi} k(q, q') + \mathbb{E}_{q,q' \sim \Pi'} k(q, q') - 2 \mathbb{E}_{q \sim \Pi, q' \sim \Pi'} k(q, q').
\]

(7.63)

Let \( (q^1_s, \ldots, q^r_s) \sim \Pi \) and \( (q^1, \ldots, q^s) \sim \Pi' \). An unbiased estimator of the squared maximum mean discrepancy is,

\[
\text{MMD}^2_u(k, \{q^i_s\}_{i=1}^r, \{q^j\}_{i=1}^s) = \frac{1}{r(r-1)} \sum_{i=1}^r \sum_{j \neq i}^r k(q^i_s, q^j_s) + \frac{1}{s(s-1)} \sum_{i=1}^s \sum_{j \neq i}^s k(q^i, q^j)
- \frac{2}{rs} \sum_{i=1}^r \sum_{j=1}^s k(q^i_s, q^j).
\]

(7.64)

In our evaluations we adopt a squared exponential positive definite kernel \( k(q, q') = \exp(-\|q - q'\|^2/2h) \), where \( h \in \mathbb{R}_+ \) is a parameter called the *kernel bandwidth*. In our experiments we set \( h \) to be the median distance between i.i.d. samples from the target distribution. Recall the definition of geometric ergodicity given in definition 7.2.15. The measure of convergence is the total variation norm, which we can discuss theoretically but cannot evaluate in a computational setting. Instead, we can measure convergence to the target distribution \( \Pi \) as a function of \( n \), the number of steps, by means of eq. (7.63), for which we can obtain an unbiased estimate via eq. (7.64) if we have samples from \( K^n(q^0, \cdot) \) and samples from \( \Pi \). Indeed, as in definition 7.4.1 let \( (q^1_s, \ldots, q^r_s) \sim \Pi \) and, for a fixed initial position \( q^0 \) and number of steps \( n \) let \( (q^1, \ldots, q^s) \sim K^n(q^0, \cdot) \) and
we can compute an unbiased estimate of the squared maximum mean discrepancy. In the latter case, \((q^1, \ldots, q^s)\) can be obtained by running \(s\) independent Markov chains for \(n\) steps; independent samples from the target distribution \(\Pi\) may be available in certain benchmark cases. In our experiments, i.i.d. samples from the target distribution may be generated from the banana-shaped posterior, Neal’s funnel distribution, the Fitzhugh-Nagumo posterior, and the multi-scale Student-\(t\) distribution.

**Remark.** Recall the definition of geometric ergodicity given in definition \[7.2.15\] Taking logarithms reveals

\[
\log \| K_n(q, \cdot) - \Pi(\cdot) \|_{TV} \leq n \log \rho + \log V(q). \quad (7.65)
\]

Therefore, one may claim to see evidence of geometric ergodicity if, as a function of \(n\), there is a linear decrease in the total variation distance on a logarithmic scale. Of course, we cannot directly compute the total variation distance, but we may look for a similar negative linear trend when \(|\text{MMD}^2_u|\) is plotted on a logarithmic scale.

As an additional measure of ergodicity, we consider comparing Markov chain samples against i.i.d. samples via random projection onto one-dimensional sub-spaces. Let \((q^1, \ldots, q^s) \sim \Pi\) and let \(q^n \sim K^n(q^0, \cdot)\) for \(n = 1, \ldots, s\). Let \(u\) be a random unit vector. We compute the Kolmogorov-Smirnov (KS) statistic for the projections \((u^\top q^1, \ldots, u^\top q^s)\) and \((u^\top q^1, \ldots, u^\top q^s)\). Repeating this process for one-hundred randomly generated unit vectors yields a distribution over Kolmogorov-Smirnov statistics. The more tightly concentrated this distribution is near zero, the closer the distribution of Markov chain iterates \((q^1, \ldots, q^s)\) is to the collection of i.i.d. samples \((q^1, \ldots, q^s)\) from the target distribution.

We also consider the expected squared jump distance (ESJD) \[Gelman and Pasarica 2007\], which measures the dissimilarity between subsequent states of the Markov chain. Intuitively, a Markov chain that moves more efficiently through the sample space (higher ESJD) will exhibit smaller sample auto-correlation. Formally, the ESJD is

\[
\mathbb{E}_{q \sim \Pi} \| q' - q \|^2
\]
where \( q' \mid q \sim K(q, \cdot) \). Note that the choice of norm \( \| \cdot \| \) is left to the practitioner and may be selected to capture geometric properties of the target distribution. This expectation is typically approximated as follows: let \( q^k \) be the state of the Markov chain at step \( k \) and at step \( k + 1 \) a candidate state is generated, denoted \( q^{k+1}_{\text{prop}} \), the proposal state is accepted with probability \( \alpha_k \); the following empirical mean is then taken as our approximation to the ESJD

\[
\text{ESJD}(\{(q^k, q^{k+1}_{\text{prop}}, \alpha_k)\}_{k=1}^n) = \frac{1}{n} \sum_{k=1}^n \alpha_k \| q^{k+1}_{\text{prop}} - q^k \|^2. \tag{7.66}
\]

In our experiments in section 7.4.8, we observe that the ESJD can be misleading as a measure. Therefore, we also introduce the median squared jump distance (MSJD) which we define as

\[
\text{MSJD}(\{(q^k, q^{k+1}_{\text{prop}}, \alpha_k)\}_{k=1}^n) = \text{Median} \left( \{ \alpha_k \| q^{k+1}_{\text{prop}} - q^k \|^2 \}_{k=1}^n \right). \tag{7.67}
\]

As the median, we expect the MSJD to exhibit less sensitivity to outliers than the ESJD.

We additionally consider the effective sample size (ESS) as a metric for our Markov chain procedures. We compute ESS using the technique of Kumar et al. [2019], who describe the method succinctly as follows. For \( i = 1, \ldots, p \), let \( (q_{i,1}^1, q_{i,2}^1, \ldots, q_{i,n}^i) \) be a sequence of \( \mathbb{R}^m \)-valued parameters. The integer \( p \) is called the number of chains. The effective sample size of the \( j \)-th parameter is computed according to,

\[
\text{ESS} \left( \{ q_{ji,1}, \ldots, q_{ji,n} \}_{i=1}^p \right) = \frac{pn}{\hat{\tau}} \tag{7.68}
\]

\[
\hat{\tau} = \left( 2 \sum_{k=1}^r \hat{\rho}_{2k} + \hat{\rho}_{2k+1} \right) - 1, \tag{7.69}
\]

where \( \hat{\rho}_k \) is an estimate, based on all of the \( p \) sequences, of the autocorrelation of the \( j \)-th parameter with a \( k \)-step lag (for details see Vehtari et al. [2021]) and \( r \) is the smallest
Figure 7.1: We examine the ergodicity of the Markov chains as a function of the sampling step in the banana-shaped posterior distribution. We observe that the use of a smaller step-size degrades the ergodicity of the RMHMC, particularly when compared against the LMC algorithm. In both RMHMC and LMC, we observe that aggressively mixing with MMALA causes the Markov chain to mix more slowly; on the other hand, a relatively small mixing probability leads to ergodicity that is nearly indistinguishable from the unmodified implementations of RMHMC and LMC.

Figure 7.2: We show the ESJD, the minimum ESS, and the time-normalized minimum ESS for inference in the banana-shaped posterior distribution. We observe that for both RMHMC and LMC, aggressively mixing with MMALA causes the ESJD distance to decrease; this correspondingly produces a decrease in the minimum ESS. For LMC, the time-normalized minimum ESS is a decreasing function of the mixing probability.

integer for which $\hat{\rho}_{2(r+1)} + \hat{\rho}_{2(r+1)+1} \leq 0$. In our experiments, we set $p = 2$ by taking a single long Markov chain and splitting it in half at the middle. As a practical matter, we will also report the minimum, over all parameters of the posterior, ESS per second in order to represent the computational efficiency of the method.
7.4.2 Banana-Shaped Distribution

Our first example considers the following generative model, which represents an example of non-identifiable parameters.

\[ \theta_1, \theta_2 \sim \text{i.i.d.} \text{Normal}(0, \sigma^2) \]  

\[ y_i|\theta_1, \theta_2 \sim \text{i.i.d.} \text{Normal}(\theta_1 + \theta_2^2, \sigma^2_y) \text{ for } i = 1, \ldots, N. \]

Given observations \( \{y_i\}_{i=1}^N \), we wish to sample the posterior distribution of \((\theta_1, \theta_2)\), a two-dimensional posterior. In our experiments, we consider \( n = 100 \). We employ Euclidean HMC with a step-size of \( \epsilon = 0.1 \); RMHMC is implemented with a step-size of 0.04; in LMC we use an integration step-size of \( \epsilon = 0.1 \). In each case, we set \( k_{\text{max}} = 10 \). For the geometric methods, the sum of the Fisher information and the negative Hessian of the log-prior is used as a Riemannian metric. The reason that a smaller integration step-size is employed in RMHMC is that eq. (7.5) in the defining involution will fail to have a solution for \( \dot{p} \) for large step-sizes; this necessitates the use of a smaller integration step.

In fig. 7.1, we visualize the statistic \( |\text{MMD}_u^2| \) over one-hundred steps of the Markov chain for euclidean HMC (EHMC), RMHMC and LMC; we color the RMHMC and LMC Markov chains according to how aggressively they mix with MMALA. Neither EHMC nor RMHMC exhibit clear evidence of geometric ergodicity in this target distribution; on the other hand, there exist a broad range of mixing probabilities for which LMC exhibits a linear decrease, as a function of \( n \), in \( |\text{MMD}_u^2| \) on a logarithmic scale. In fig. 7.2 we visualize the ESJD, the minimum ESS, and the minimum ESS per second for the three MCMC algorithms. We observe that LMC exhibits by far the strongest performance on these metrics, while RMHMC languishes due to its small step-size and the pathologies of the generalized leapfrog integrator applied to this posterior.
Figure 7.3: We show the ESJD, the minimum ESS, and the time-normalized minimum ESS for inference in the hierarchical Bayesian logistic regression posterior distribution. We observe that both of the geometric methods do not perform as well as standard HMC in this inference task. For each mixing weight with MMALA, we observe that RMHMC enjoys a greater ESJD and minimum ESS when compared against LMC; however, as LMC is a more computationally expedient procedure, LMC has a greater time-normalized ESS. In all of these metrics, however, EHMC dominates the geometric inference algorithms.

7.4.3 Hierarchical Bayesian Logistic Regression

We consider sampling from the hierarchical Bayesian logistic regression model

\[
\alpha \sim \text{Gamma}(\omega, \theta) \quad (7.72)
\]

\[
\beta_i | \alpha \sim \text{i.i.d. Normal}(0, \alpha^{-1}) \text{ for } i = 1, \ldots, m \quad (7.73)
\]

\[
y_i | x_i, \beta \sim \text{Bernoulli} \left( \frac{1}{1 + \exp(-x_i^T \beta)} \right) \text{ for } i = 1, \ldots, N. \quad (7.74)
\]

We consider a logistic regression dataset with \( N = 270 \) observations and \( m = 14 \) covariates. We set \( \omega = 10 \) and \( \theta = 2 \) in our experiments. We employ a Metropolis-within-Gibbs sampling procedure wherein we alternate between sampling the posterior distributions \( \alpha | \beta, k, \theta \) and \( \beta | \alpha, \{(x_i, y_i)\}_{i=1}^n \); sampling the former can be performed analytically, whereas we employ EHMC, RMHMC, and LMC to sample the latter. In Euclidean HMC, we set \( \epsilon = 0.1 \); in RMHMC and LMC we set \( \epsilon = 0.8 \). We set \( k_{\text{max}} = 10 \) as the upper bound on the number of integration steps in each case. For implementing both LMRMHHMC and LMLMC, we use the sum of the Fisher information and the negative Hessian of the log-prior as a metric.
Figure 7.4: We examine the ergodicity of the Markov chains as a function of the sampling step in Neal’s funnel distribution. We observe that EHMC struggles in this distribution due to the multiple spatial scales inherent in the funnel-shaped distribution. Moreover, we observe that aggressively mixing RMHMC with SMALA causes convergence in MMD to become increasingly slow, with convergence of the pure SMALA algorithm being slower than EHMC. On the other hand, a modest amount of mixing with SMALA produces ergodicity that is nearly indistinguishable from the unmodified implementation of RMHMC.

In fig. 7.3 we show the ESJD, the minimum ESS, and the minimum ESS per second for the hierarchical Bayesian logistic regression posterior. Neither of the geometric methods perform well in this posterior, consistently under-performing EHMC. A criticism of LMC is that its performance degrades significantly in higher dimensions [Betancourt et al., 2014]; one sees evidence of this phenomenon in the smaller ESS generated by LMC; however, the computational savings due to eliminating the fixed point iterations still allow LMC to edge out a stronger time-normalized ESS compared to RMHMC.

### 7.4.4 Neal’s Funnel Distribution

Neal’s funnel distribution [Neal, 2003] is a density defined in the following hierarchical manner.

\[ v \sim \text{Normal}(0, 9) \]  \hspace{1cm} (7.75)

\[ x_i | v \sim \text{i.i.d. Normal}(0, \exp(-v)) \text{ for } i = 1, \ldots, N. \]  \hspace{1cm} (7.76)
Figure 7.5: We show the ESJD, the distribution of Kolmogorov-Smirnov statistics, and the time-normalized minimum ESS for inference in Neal’s funnel distribution. Although RMHMC and EHMC have comparable time-normalized ESS, it is clear from the distribution of KS statistics that RMHMC produces samples that are closer to the target distribution in EHMC.

This distribution is shaped like a funnel, in which the thickness of the “neck” is being controlled by the random variable $v$. This model is reflective of posteriors encountered in hierarchical models with sparse data. The objective in this task is to jointly sample $(v, x_1, \ldots, x_N)$, producing a $(N + 1)$-dimensional target distribution. In Euclidean HMC we employ an integration step-size of $\epsilon = 0.1$ and $k_{\text{max}} = 10$. Our implementation of RMHMC uses $k_{\text{max}} = 20$ integration steps with a step-size of $\epsilon = 0.1$ with the SoftAbs metric. We do not consider LMC in this task since we found it non-obvious how the SoftAbs structure could be extended into the LMC framework while preserving the cubic computational cost at each step. In this experiment, we use the SoftAbs metric.

In fig. 7.4 we visualize $|\text{MMD}^2_u|$ for both EHMC and RMHMC. As expected, EHMC struggles to sample from Neal’s funnel distribution due to the multiscale phenomena. On the other hand, RMHMC exhibits much stronger convergence properties, having a linear decrease over several possible mixing probabilities with SMALA. We find that aggressively mixing with SMALA can be counter-productive, however, due to the less efficient traversal of the target distribution by single-step methods. In fig. 7.5 we show the ESJD, the Kolmogorov-Smirnov statistics, and the minimum ESS per second. On all of these metrics, RMHMC clearly outperforms EHMC.
Figure 7.6: We show the ESJD, the minimum ESS, and the time-normalized minimum ESS for inference in the posterior distribution of the stochastic volatility model. We observe that both of the geometric methods with a modest mixing probability with MMALA produce time-normalized ESS that are competitive with EHMC. For each mixing weight with MMALA, we observe that LMC enjoys a greater ESJD when compared against RMHMC.

7.4.5 Stochastic Volatility Model

We consider a stochastic volatility model with the following generative model.

\[
\frac{\phi + 1}{2} \sim \text{Beta}(20, 3/2) \tag{7.77}
\]

\[
\sigma^2 \sim \text{InverseChiSquared}(10, 1/20) \tag{7.78}
\]

\[
x_{1|\phi, \sigma^2} \sim \text{Normal}(0, \sigma^2/(1 - \phi^2)) \tag{7.79}
\]

\[
x_{t+1|x_t, \phi, \sigma^2} \sim \text{Normal}(\phi x_t, \sigma^2) \text{ for } t = 2, \ldots, T - 1 \tag{7.80}
\]

\[
y_{t|\beta, x_t} \sim \text{Normal}(0, \beta^2 \exp(x_t)) \text{ for } t = 1, \ldots, T. \tag{7.81}
\]

Additionally, the parameter \( \beta \) is equipped with an improper prior proportional to \( \beta^{-1} \). In this example, we seek to generate samples from the joint distribution \( x_1, \ldots, x_T, \beta, \phi, \sigma^2 \) given observations \( y_1, \ldots, y_T \). We employ Metropolis-within-Gibbs-like strategy wherein we alternate between sampling \( (x_1, \ldots, x_T) | (y_1, \ldots, y_T) \phi, \sigma^2, \beta \) and \( (\phi, \sigma^2, \beta) | \{(x_i, y_i)\}_{i=1}^T \); in the former case we employ Euclidean HMC whereas in the latter case we compare Euclidean HMC, RMHMC, and LMC. In sampling either distribution, the Riemannian metric is chosen as the sum of the Fisher information and the negative Hessian of the log-prior.
Figure 7.7: We show the ESJD, the minimum ESS, and the time-normalized minimum ESS for inference in the posterior distribution of the log-Gaussian Cox-Poisson model. For any mixing probability, the geometric methods enjoy larger time-normalized effective sample sizes than EHMC, with LMC outperforming RMHMC on these metrics. For nearly every mixing probability, we additionally observe that the ESJD is greater for LMC and RMHMC than for EHMC.

In our experiments we set $T = 1,000$. When using Euclidean HMC to sample $(\phi, \sigma^2, \beta)$ we use $k_{\text{max}} = 50$ integration steps and a step-size of $\epsilon = 0.01$; in the case of RMHMC and LMC we use $k_{\text{max}} = 6$ integration steps and a step-size of $\epsilon = 0.5$.

In fig. 7.6 we show the ESJD, the minimum ESS, and the minimum ESS per second. We observe that employing a modest mixture probability with MMALA produces a Markov chain that is marginally better than EHMC.

### 7.4.6 Log-Gaussian Cox-Poisson Process

We consider inference in a log-Gaussian Cox-Poisson model with the following generative model:

\[
\Sigma_{(i,j),(i',j')} | \sigma^2, \beta = \sigma^2 \exp \left( -\sqrt{(i - i')^2 + (j - j')^2} / (N \beta) \right) \tag{7.82}
\]

\[
\text{vec}(x) | \Sigma \sim \text{MultivariateNormal}(\mu 1, \Sigma) \tag{7.83}
\]

\[
y_{ij} | x_{ij} \sim \text{Poisson}(\exp(x_{ij}) / N^2), \tag{7.84}
\]

with priors $\beta \sim \text{Gamma}(2, 1/2)$ and $\sigma^2 \sim \text{Gamma}(2, 1/2)$. In this example, the objective is to sample the joint distribution $(\beta, \sigma^2; \{x_{ij}\}_{i,j=1}^N)$ given observations $\{y_{ij}\}_{i,j=1}^N$. As in
Figure 7.8: We examine the ergodicity of the Markov chains as a function of the sampling step in the Fitzhugh-Nagumo posterior distribution. We observe that when a modest mixing probability is employed, the modified RMHMC and LMC transition kernels mix more efficiently than the EHMC transition kernel, being nearly indistinguishable from the unmodified RMHMC and LMC transition kernels. When one uses a large mixing probability, we see that the mixing rate is decreased.

the case of stochastic volatility model, we alternatively sample between $\{x_{ij}\}_{i,j=1}^N$ given $\{y_{ij}\}_{i,j=1}^N$, $\sigma^2$, and $\beta$, and $(\beta, \sigma^2)$ given $\{x_{ij}\}_{i,j=1}^N$. In each case, the metric is given by the sum of the Fisher information and the negative Hessian of the log-prior. In our experiments we set $N = 16$. In the former case we employ Euclidean HMC, whereas in the latter case we compare Euclidean HMC, RMHMC and LMC. When implementing Euclidean HMC we employ $k_{\text{max}} = 50$ integration steps and a step-size of $\epsilon = 0.01$; in RMHMC and LMC we use $k_{\text{max}} = 6$ integration steps and a step-size of $\epsilon = 0.5$.

In fig. 7.7 we show the ESJD, the minimum ESS, and the minimum ESS per second. We observe that there is a range of mixture probabilities for which mixing with the MMALA in LMRMHMC and LMLMC exhibit superior performance compared to EHMC. Both RMHMC and LMC exhibit similar minimum ESS metrics, but due to its computational advantage, LMC produces a larger minimum ESS per second.
Figure 7.9: We show the ESJD, the minimum ESS, and the time-normalized minimum ESS for inference in the Fitzhugh-Nagumo posterior distribution. LMC dominates both RMHMC and EHMC in this inference task under the time-normalized ESS metrics with its greater computational expediency. When timing is not accounted for, we observe that RMHMC and LMC enjoy similar ESJD and minimum ESS metrics. We observe that for both RMHMC and LMC, aggressively mixing with MMALA causes the ESJD distance to decrease; this correspondingly produces a decrease in the minimum ESS.

7.4.7 Fitzhugh-Nagumo Model

Given $\mathbb{R}$-valued parameters $a$, $b$, and $c$, the Fitzhugh-Nagumo differential equations are defined by,

\begin{align}
\dot{v}_t &= c \left( v_t - \frac{v_t^3}{3} + r_t \right) \\
\dot{r}_t &= -\left( \frac{v_t - a + br_t}{c} \right). 
\end{align}

(7.85) \hspace{1cm} (7.86)

Given initial conditions $v_0$ and $r_0$, we consider the following generative model:

\begin{align}
(a, b, c) &\sim \text{i.i.d.} \ Normal(0, 1) \quad (7.87) \\
\hat{r}_{tk} | a, b, c, r_{tk}, \sigma^2 &\sim \text{Normal}(r_{tk}, \sigma^2) \text{ for } k = 1, \ldots, n \quad (7.88) \\
\hat{v}_{tk} | a, b, c, v_{tk}, \sigma^2 &\sim \text{Normal}(v_{tk}, \sigma^2) \text{ for } k = 1, \ldots, n, \quad (7.89)
\end{align}

where $t_1, \ldots, t_n$ are equally spaced points between $[0, T]$. In our experiments, we set $v_0 = 1$, $r_0 = -1$, $\sigma^2 = 1/4$, $T = 10$, and $n = 200$. Our metric is given by the sum of the Fisher information and negative Hessian of the log-prior. In our implementations, we use
Figure 7.10: We examine the ergodicity of the Markov chains as a function of the sampling step in the multiscale Student-t distribution. Here we see that, in the case of LMC, it is beneficial from an ergodicity perspective to mix heavily with MMALA, whereas for RMHMC, only modest mixing probabilities can produce a modified Markov chain that is competitive with the unmodified version. The Euclidean HMC struggles in this posterior distribution due to the multiple spatial scales.

Euclidean HMC with a step-size of $\epsilon = 0.01$ and $k_{\text{max}} = 10$ integration steps. In RMHMC and LMC, we employ an integration step-size of $\epsilon = 0.5$ and $k_{\text{max}} = 6$ integration steps.

Figure 7.8 shows the ergodicity measures for EHMC, RMHMC, and LMC. We observe that each of these MCMC algorithms exhibit a linear decrease in $MMD_u^2$ on a logarithmic scale. In fig. 7.9 we visualize the ESJD, the minimum ESS, and the minimum ESS per second for the three MCMC algorithms. We observe that LMC exhibits the strongest performance in terms of the minimum ESS per second, whereas RMHMC only approaches the time-normalized performance of EHMC due to its complexity.

7.4.8 Multi-Scale Student Distribution

Fix $m \in \mathbb{N}$ and let $\Sigma \in \mathbb{PD}(m)$ and $\nu > 2$. The density function of the multivariate Student-t distribution is,

$$
\pi(q) \propto \left(1 + \frac{1}{\nu} x^\top \Sigma^{-1} x\right)^{-(m+\nu)/2}.
$$

(7.90)

In our experiments we consider $m = 20$, $\nu = 5$, and $\Sigma = \text{diag}(1, \ldots, 1, 10^4) \in \mathbb{R}^{m \times m}$. The presence of severely differing spatial scales in the multivariate Student-t distribution
Figure 7.11: We show the ESJD and the MSJD for inference in the multiscale Student-$t$ distribution. Curiously, EHMC enjoys the largest ESJD, but this does not correspond to more efficient sampling, as shown in the inferior distribution of KS statistics. To examine this property further, we also measured the MSJD, which reveals that EHMC is less effective in traversing the parameter space compared to the RMHMC geometric method.

Figure 7.12: We show the distribution of Kolmogorov-Smirnov statistics and the time-normalized minimum ESS for inference in the multiscale Student-$t$ distribution. For the geometric methods, we observe that RMHMC has much larger ESJD than LMC; however, in terms of time-normalized performance, it is best to fully mix either method with MMALA. Indeed, we find that MMALA on its own produces a distribution of KS statistics that is competitive with any of the mixture kernels.

will cause Euclidean HMC with identity mass matrix to exhibit highly oscillatory behavior [Pourzanjani and Petzold, 2019], which will limit the efficiency of the method both in terms of ergodicity and the effective sample size. We employ Euclidean HMC with a step-size of $\epsilon = 0.8$, and RMHMC and LMC with a step-size of $\epsilon = 0.7$; in each case we
employ \( k_{\text{max}} = 20 \) integration steps. In the case of RMHMC and LMC, we consider a metric given by the positive definite term in the Hessian of the log-density of the multivariate Student-\( t \) distribution.

In fig. 7.10 we show \( |\text{MMD}_u^2| \) as a function of the number of Markov chain steps. All of the MCMC algorithms produce a linear decrease in the estimate of the maximum mean discrepancy on a logarithmic scale; however, the methods differ drastically in terms of the slope of this linear relationship. We see that RMHMC exhibits by far the fastest convergence. Notably, LMC exhibits slower convergence on this target distribution than MMALA; this is due to the dimensionality of the posterior, in which LMC struggles to maintain the Hamiltonian energy required to accept proposals. In fig. 7.12 we show the ESJD, KS, and the minimum ESS per second. EHMC produces the largest ESJD but this does not translate into a large ESS due to the oscillatory behavior of the EHMC proposal mechanism, with RMHMC producing the largest minimum ESS per second despite its computational complexity.

### 7.5 Conclusion

This work has considered methods by which to equip RMHMC and LMC with a geometric ergodicity theory. The fundamental technique we adopt is to replace the RMHMC (or LMC) transition kernel consisting of a single integration step with the MMALA transition kernel. This modification is inspired by the Euclidean case, in which single-step HMC and MALA can be constructed to be exactly equivalent. By establishing reversibility of the marginal transition kernels, geometric ergodicity can be inherited from MMALA. We evaluated the modified variations of RMHMC and LMC, called LMRMHMC and LMLMC, respectively, on a suite of Bayesian inference tasks. We found that aggressively mixing with MMALA transition kernel can be detrimental for the performance of the Markov chain on a variety of metrics, but that more modest mixing can produce behaviors compet-
itive with, or exceeding, the original RMHMC or LMC methods while still imbuing the methods with a supporting theory of geometric ergodicity.
Appendix 7.A

Appendices to Chapter 7

7.A.1 Proofs

7.A.1.1 Proof of Proposition 7.2.28

Proof. The proposal distribution of MALA is

\[ \tilde{q}|q \sim \text{Normal} \left( q + \frac{\epsilon^2}{2} A^{-1} \nabla \log \pi(q), \epsilon^2 A^{-1} \right) \] (7.A.1)

\[ \tilde{\pi}(\tilde{q}|q) \propto \exp \left( -\frac{1}{2\epsilon^2} \left( \tilde{q} - q - \frac{\epsilon^2}{2} A^{-1} \nabla \log \pi(q) \right)^\top A \left( \tilde{q} - q - \frac{\epsilon^2}{2} A^{-1} \nabla \log \pi(q) \right) \right) \] (7.A.2)

\[ = \exp \left( -\frac{1}{2\epsilon^2} \left( A\tilde{q} - Aq - \frac{\epsilon^2}{2} \nabla \log \pi(q) \right)^\top A^{-1} \left( A\tilde{q} - Aq - \frac{\epsilon^2}{2} \nabla \log \pi(q) \right) \right). \] (7.A.3)
Therefore, the acceptance probability of MALA is,

\[
\min \left\{ 1, \frac{\pi(\tilde{q}) \pi(q|\tilde{q})}{\pi(q) \pi(\tilde{q}|q)} \right\} = \min \left\{ 1, \frac{\exp \left( -\frac{1}{2\epsilon^2} (\mathbf{A}q - \mathbf{A}\tilde{q} - \frac{\epsilon^2}{2} \nabla \log \pi(q)) \right)}{\exp \left( -\frac{1}{2\epsilon^2} (\mathbf{A}\tilde{q} - \mathbf{A}q - \frac{\epsilon^2}{2} \nabla \log \pi(q)) \right)} \right\}.
\]

(7.A.4)

The proposal \( \tilde{q} \) given \( q \) can be sampled by generating \( z \sim \text{Normal}(0, \text{Id}) \) and setting

\[
\tilde{q} = q + \frac{\epsilon}{2} \mathbf{A}^{-1} \nabla \log \pi(q) + \epsilon \sqrt{\mathbf{A}^{-1}} z.
\]

(7.A.5)

In HMC, the leapfrog integrator is applied to the Hamiltonian \( H(q, p) = -\log \pi(q) + \frac{1}{2} p^\top \mathbf{A}^{-1} p \). The sequence of updates is,

\[
\begin{align*}
\bar{p} &= p + \frac{\epsilon}{2} \nabla \log \pi(q) \quad \text{(7.A.6)} \\
\tilde{q}' &= q + \epsilon \mathbf{A}^{-1} \bar{p} \quad \text{(7.A.7)} \\
&= q + \frac{\epsilon^2}{2} \mathbf{A}^{-1} \nabla \log \pi(q) + \epsilon \mathbf{A}^{-1} p \quad \text{(7.A.8)} \\
\tilde{p}' &= \bar{p} + \frac{\epsilon}{2} \nabla \log \pi(\tilde{q}') \quad \text{(7.A.9)} \\
&= p + \frac{\epsilon}{2} \nabla \log \pi(q) + \frac{\epsilon}{2} \nabla \log \pi(\tilde{q}'). \quad \text{(7.A.10)}
\end{align*}
\]

When sampling from the distribution with density proportional to \( \exp(-H(q, p)) \), one sees by inspection that \( p \) is independent of \( q \) and that \( p \sim \text{Normal}(0, \mathbf{A}) \). Therefore, using the same \( z \) as in eq. (7.A.5), we can sample \( p \) by setting \( p = \sqrt{\mathbf{A}} z \) so that \( \epsilon \mathbf{A}^{-1} p = \epsilon \sqrt{\mathbf{A}^{-1}} z \); we therefore see that, in this case, HMC and MALA produce exactly the same proposal.
(i.e. $\tilde{q}' = \tilde{q}$). Recall that the HMC acceptance probability is,

$$\min \left\{ 1, \frac{\pi(\tilde{q}) \cdot \exp \left(-\frac{1}{2}(\tilde{p}')^\top A^{-1}(\tilde{p}')\right)}{\pi(q) \cdot \exp \left(-\frac{1}{2}p^\top A^{-1}p\right)} \right\}. \quad (7.A.11)$$

By rearranging eq. (7.A.7) we find,

$$p = \frac{1}{\epsilon} \left( A\tilde{q} - Aq - \frac{\epsilon^2}{2} \nabla \log \pi(q) \right) \quad (7.A.12)$$

$$\tilde{p}' = \frac{1}{\epsilon} \left( A\tilde{q} - Aq + \frac{\epsilon^2}{2} \nabla \log \pi(\tilde{q}) \right) \quad (7.A.13)$$

$$= -\frac{1}{\epsilon} \left( Aq - A\tilde{q} - \frac{\epsilon^2}{2} \nabla \log \pi(\tilde{q}) \right). \quad (7.A.14)$$

Substituting these into eq. (7.A.11) and comparing to eq. (7.A.4) shows that not only are $\tilde{q}$ and $\tilde{q}'$ identical but that the acceptance probabilities are also identical. Therefore, the marginal chain of single-step HMC is exactly equivalent to the MALA chain. □

7.A.1.2 Proof of Lemma 7.2.30

Proof.

$$\Pr(q^{n+1} \in Q | q^n = q) = \Pr(q^{n+1} \in Q \text{ and } p^{n+1} \in \mathbb{R}^m | q^n = q) \quad (7.A.15)$$

$$= \int_{\mathbb{R}^m} \Pr(q^{n+1} \in Q \text{ and } p^{n+1} \in \mathbb{R}^m | q^n = q, p^n = p) \pi(p|q) \, dp \quad (7.A.16)$$

$$= \int_{\mathbb{R}^m} K((q, p), (Q, \mathbb{R}^m)) \pi(p|q) \, dp. \quad (7.A.17)$$

□
7.A.1.3 Proof of Proposition 7.2.31

Proof. Given \( Q, Q' \in \mathcal{B}(\mathbb{R}^m) \) we have

\[
\int_Q \tilde{K}(q, Q') \pi(q) \, dq = \int_Q \int_{\mathbb{R}^m} K((q, p), (Q', \mathbb{R}^m)) \pi(p|q) \, dp \, \pi(q) \, dq = \int_Q \int_{\mathbb{R}^m} K((q, p), (Q, \mathbb{R}^m)) \pi(p|q) \, dp \, \pi(q) \, dq = \int_Q \tilde{K}(q, Q) \pi(q) \, dq, \tag{7.A.18}
\]

where in eq. \((7.A.20)\) we have used the fact that the phase space chain satisfies detailed balance with respect to \( \pi(q, p) \). This verifies that the marginal chain satisfies detailed balance. \( \square \)

7.A.1.4 Proof of Involution Composition

Proposition 7.A.1.1. Suppose that \( \Phi : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m \) is an invertible function and let \( F \) be the momentum flip function given in definition 7.2.24. Suppose that \( F \circ \Phi \) is an involution, then \( F \circ \Phi^k \) is also an involution.

Proof. This will be proved by induction with the base case established by assumption. As the inductive hypothesis, assume that \( F \circ \Phi^k \) is an involution. Using the fact that \( F \circ \Phi \) is an involution, we immediately obtain that \( \Phi^{-1} \circ F = F \circ \Phi \). Using the inductive hypothesis,
one also has $\Phi^{-k} \circ F = F \circ \Phi^k$. Therefore, we obtain,

\[
\Phi^{-1} \circ F = F \circ \Phi \quad (7.A.23)
\]

\[
\implies \Phi^{-1} \circ F \circ \Phi^k = F \circ \Phi^{k+1} \quad (7.A.24)
\]

\[
\implies \Phi^{-1} \circ \Phi^{-k} \circ F = F \circ \Phi^{k+1} \quad (7.A.25)
\]

\[
\implies \Phi^{-k-1} \circ F = F \circ \Phi^{k+1} \quad (7.A.26)
\]

\[
\implies F \circ \Phi^{k+1} \circ F \circ \Phi^{k+1} = \text{Id.} \quad (7.A.27)
\]

This verifies that $F \circ \Phi^{k+1}$ is also an involution. \qed
Chapter 8

Correcting an Explicit Integrator with Lagrange Multipliers

This research represents my first investigations into numerical integration and the importance of reversibility and volume preservation for producing a Markov chain that samples correctly from a target distribution. This research presented the opportunity to correct an integrator that had been proposed in the literature by means of techniques from differential geometry so that it possesses the required reversibility and volume preservation properties. Composition and experimentation are original to me.

Abstract. Non-separable Hamiltonians are challenging for numerical integrators because they often require the use of implicitly-defined integration steps. This is unfortunate because there exist powerful methods of Bayesian inference, such as Hamiltonian Monte Carlo (HMC) that are most effective when combined with non-separable Hamiltonians. Recently, an explicit integrator was proposed for HMC with non-separable Hamiltonians, but the method lacked a theoretical foundation upon which to establish symmetry and symplecticness. As a consequence, the integrator cannot provably produce a Markov chain satisfying detailed balance. This article proposes a modification to this explicit integration methodology so that it is symplectic and symmetric when restricted to a particular submanifold of phase-space. This modification allows us to construct a Markov chain transi-
tion operator which provably satisfies the detailed balance condition. We give an empirical evaluation of this modified method against implicitly-defined integrators of non-separable Hamiltonians and motivate an explicit approximation. We discuss the empirical trade-off between conservation of the Hamiltonian, symmetry of the integrator, and symplecticness of the integrator. We stress that this modification to the explicit integrator requires one to identify Lagrange multipliers, which can be equally costly, or more so, compared to solving implicitly-defined integration steps; therefore, our principle contribution is discussing how this integrator may be corrected. In other words, our contribution is to make this method correct, not correct and fast.

8.1 Introduction

Hamiltonian Monte Carlo (HMC) is a popular method for Bayesian inference in differentiable posterior distributions. The foundation of HMC is Hamiltonian mechanics, and therefore the solutions to Hamilton’s equations of motion play an integral role in the HMC sampling procedure. However, closed-form solutions are generally unavailable for Hamilton’s equations of motion, necessitating the use of methods of numerical integration. In order to be incorporated into a Markov chain satisfying the detailed balance condition, an integrator must possess certain key properties of reversibility and volume preservation. For mathematical details on HMC, see Neal [2010a], Betancourt [2017] or section 2.7. For a historical treatment of HMC, see Robert and Casella [2011]. For a thorough introduction to classical mechanics, including the Hamiltonian formalism, see Marsden and Ratiu [2010] or sections 2.2 and 2.3. For a treatise on methods of integration for Hamiltonian dynamics, see Leimkuhler and Reich [2005].

Methods of Hamiltonian Monte Carlo that seek to incorporate second-order geometric information commonly utilize integrators with implicitly defined updates [Brubaker et al., 2012; Girolami and Calderhead, 2011]. In these algorithms, implicit integration is neces-
sary to handle the non-separable form of the Hamiltonian defining the mechanics. Implicit integration, in which steps of the integrator are defined by solutions to fixed-point equations and which are typically resolved via fixed point iteration, are slower than explicit integration, hampering such methods in timing comparisons and time-normalized metrics (such as the effective sample size per second). Recently, Cobb et al. [2019] proposed an explicit integrator for non-separable Hamiltonians. However, this method may not be satisfactory because it is neither symmetric nor symplectic in a manner that is compatible with HMC. The failure of these properties to hold in the explicit integration scheme invalidates the standard proof of detailed balance for HMC.

The purpose of this work is to modify the explicit integration procedure so that it is symmetric and symplectic such that the standard proof of detailed balance in HMC applies. The approach is based on Lagrange multipliers, which forces the explicitly defined numerical integrator to remain on an appropriately constructed manifold. On this manifold, a symmetric proposal operator for HMC can be constructed. Moreover, the explicit integrator augmented with Lagrange multipliers will define a symplectic transformation on this manifold. We emphasize that this work does not attempt to argue that the method of Lagrange multipliers are in any manner more computationally advantageous than fixed point iteration. Rather, our objective is to characterize the integrator of Cobb et al. [2019] within a framework in which detailed balance may be established.

8.2 Preliminaries

We begin our discussion by introducing key concepts from geometric integration. We will recall precise definitions of symmetry and symplecteness on embedded manifolds; this is carried out in section 8.2.1. Subsequently, in section 8.2.2, we proceed to recall Hamiltonian mechanics on embedded manifolds. This then leads to a discussion of numerical integrators on embedded manifolds in section 8.2.3 where we review techniques based
on Lagrange multipliers. Our objective is to develop these notions for use in section 8.3 where we prove that we may repair the integrator of [Cobb et al. 2019] via augmentation with Lagrange multipliers on a suitably defined manifold, thereby producing a symmetric and symplectic (and hence volume preserving) numerical method. As a consequence, the resulting numerical method will be compatible with diffeomorphism Monte Carlo as discussed in section 8.2.4. Throughout this work we will freely use the isomorphism of $\mathbb{R}^m \times \mathbb{R}^m$ and $\mathbb{R}^{2m}$ without further comment. We adopt the notation $\text{PD}(m)$ to be the set of $m \times m$ positive definite matrices.

8.2.1 Implicitly Defined Manifolds

Embedded manifolds may be constructed as the zero level-set of a function. In order to ensure that a manifold so-defined is a proper embedding, it is necessary to assert some structure over a defining function. This is made precise in definition 8.2.1.

**Definition 8.2.1.** Let $g : \mathbb{R}^m \to \mathbb{R}^k$ be a smooth map for which $\nabla g(q) \in \mathbb{R}^{k \times m}$ is a matrix of full-rank for every $q \in \mathbb{R}^m$. Define the set $M = \{q \in \mathbb{R}^m : g(q) = 0\}$. In this case, $M$ is a manifold implicitly defined by $g$.

**Definition 8.2.2.** The tangent space of an implicitly defined manifold $M$ at position $q \in M$ is defined by $T_qM = \{v \in \mathbb{R}^m : \nabla g(q)v = 0\}$.

**Proposition 8.2.3.** Let $M$ be an implicitly defined manifold. The tangent space $T_qM$ is also an implicitly defined manifold with defining function $f(v) = \nabla g(q)v$.

A proof is given in section 8.A.1.1 Thus, in this case, we see that $T_qM$ is a linear subspace of $\mathbb{R}^m$ defined as the kernel of the matrix $\nabla g(q)$. Shortly, we will construct equations of motion on $M$. However, lemma 8.2.4 reveals that equations of motion must produce velocities that lie in the tangent space.
Lemma 8.2.4. Let \( g : \mathbb{R}^m \to \mathbb{R}^k \) and \( M \subseteq \mathbb{R}^m \) be the manifold implicitly defined by \( g \) as in definition 8.2.1. Let \( q(t) : \mathbb{R} \to \mathbb{R}^m \) be a function satisfying \( g(q_t) = 0 \) for every \( t \in \mathbb{R} \) so that \( q_t \in M \) for every \( t \in \mathbb{R} \). Then \( \frac{d}{dt} q_t \in T_{q_t}M \).

A proof is given in section 8.A.1.2.

Definition 8.2.5. The tangent bundle of an implicitly defined manifold \( M \) is defined to be
\[ T^\ast M = \{(q, v) \in \mathbb{R}^m \times \mathbb{R}^m : q \in M \text{ and } v \in T_q^\ast M\} \]

Proposition 8.2.6. Let \( M \) be an implicitly defined manifold. The tangent bundle \( T^\ast M \) is also an implicitly defined manifold with defining function \( h(q, v) = (g(q), \nabla g(q)v)^\top \) viewed as a map from \( \mathbb{R}^{2m} \) to \( \mathbb{R}^{2k} \).

A proof is given in section 8.A.1.3.

When proceeding to develop Hamiltonian mechanics on an implicitly defined manifold, we must introduce the notion of a cotangent space.

Definition 8.2.7. Let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be a smooth function for which \( \nabla_p H(q, p) \) is a function of the form \( p \mapsto G^{-1}(q)p \) where \( G(q) \in \text{PD}(m) \). Let \( M \) be an implicitly defined manifold and let \( q \in M \). The cotangent space at \( q \) is defined as \( T_q^\ast M = \{p \in \mathbb{R}^m : \nabla g(q)\nabla_p H(q, p) = 0\} = \{G(q)v : v \in T_q^\ast M\} \).

The assumption on the structural form of \( \nabla_p H(q, p) \) in definition 8.2.7 means that the cotangent space is also a vector space. Moreover, as a linear transformation of the tangent space, the cotangent space is also an embedded manifold.

Definition 8.2.8. The cotangent bundle is the set
\[ T^\ast M = \{(q, p) \in \mathbb{R}^{2m} : q \in M \text{ and } p \in T_q^\ast M\} \]

Definition 8.2.9. Let \( (q, p) \in T^\ast M \). The space \( T_{(q, p)}T^\ast M \) is the tangent space of the cotangent bundle.

The cotangent bundle is an implicitly defined manifold with defining function \( f(q, p) = (g(q), \nabla g(q)\nabla_p H(q, p)) \), which may be verified via a proof similar to proposition 8.2.6.
Hence we may view the cotangent bundle as an embedded sub-manifold of $\mathbb{R}^{2m}$. Similarly, from proposition 8.2.3, the tangent space of the cotangent bundle (definition 8.2.9) is a linear subspace of $\mathbb{R}^{2m}$. The advantage of viewing $T^*M$ as an embedded sub-manifold is that elements $(q, p) \in T^*M$ may also be viewed as elements of $\mathbb{R}^{2m}$; we make use of this fact liberally in the sequel.

In order to have a rigorous definition of volume on a manifold, we must discuss the notion of differential forms and maps that are constructed from differential forms.

**Definition 8.2.10.** Let $M$ be an implicitly defined manifold. The coordinate differential one-forms $dq_i : T_qM \to \mathbb{R}$ are defined by $d\hat{z}_i(\delta) = \delta_i$ where $\delta \in T_qM$.

It is possible to combine the coordinate differential one-forms to construct more elaborate maps. For instance, one might adopt the notation $dq_i \otimes dq_j : T_qM \times T_qM \to \mathbb{R}$ as the map defined by $(dq_i \otimes dq_j)(\delta, \delta') = \delta_i \delta'_j$. Definition 8.2.11 gives one of the most important constructions for differential geometry.

**Definition 8.2.11.** The wedge product of coordinate differential one-forms $dq_i$ and $dq_j$, denoted $dq_i \wedge dq_j : T_qM \times T_qM \to \mathbb{R}$, is the map defined as

$$
(dq_i \wedge dq_j)(\delta, \delta') \overset{\text{def.}}{=} (dq_i \otimes dq_j)(\delta, \delta') - (dq_i \otimes dq_j)(\delta', \delta) = \delta_i \delta'_j - \delta'_i \delta_j.
$$

The following definition provides a simplified version of an object known more generally as the pullback of differential forms.

**Definition 8.2.12.** Let $M$ be an implicitly defined manifold and let $\Phi : M \to M$ be a map that can be smoothly extended to $\mathbb{R}^m$. Let $z \in M$. If $\hat{z} = \Phi(z)$ then define the coordinate differential one-forms $d\hat{z}_i : T_{\hat{z}}M \to \mathbb{R}$ by

$$
d\hat{z}_i = \sum_{j=1}^m \frac{\partial \hat{z}_i}{\partial z_j} dz_j.
$$
where $\hat{z} = (\hat{z}_1, \ldots, \hat{z}_m) = (\Phi_1(z), \ldots, \Phi_m(z)) \in \mathbb{R}^m$.

A more general definition of the wedge product of differential forms is given in definition 2.3.13. The construction in section 2.3 eventually give rise to the canonical Liouville volume form on $T^*M$ which is defined in definition 2.3.21. The fact that symplectic transformations (defined in definitions 8.2.17 and 8.2.23) also preserve volume is the subject of section 2.3.8.

### 8.2.2 Hamiltonian Mechanics

In this discussion we will recall Hamilton’s equations of motion. Traditionally the position variables in Hamiltonian mechanics are denoted $q$ while the momentum variables are denoted $p$. Here we will eschew this notation and instead denote the position and momentum variables by $q$ and $p$, respectively. We adopt this notation because our development later will require integrators defined on the traditional $(q, p)$-phase space and on an expanded $(q, x, p, y)$-phase space with auxiliary position and momentum variables $x$ and $y$, respectively. In the latter case, we adopt the notation $\alpha = (q, x)$ and $\beta = (p, y)$ to denote the expanded position and momentum variables in the $(q, x, p, y)$-phase space. However, the development here is agnostic to whether the Hamiltonian mechanics are formulated on the traditional $(q, p)$-phase space or the expanded $(q, x, p, y)$-phase space; therefore, we adopt the separate $(q, p)$ notation for this construction. The following discussion will draw heavily on methods of numerical integration on Euclidean spaces and in more general embedded manifolds; these topics are reiterated in detail in sections 2.2 and 2.3. We first formulate Hamiltonian mechanics on $\mathbb{R}^m$ before proceeding to the case of mechanics on an implicitly defined manifold.
In Euclidean Space

Definition 8.2.13. Let \( q_0 \in \mathbb{R}^m \) and \( p_0 \in \mathbb{R}^m \). Given a smooth function \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) called the Hamiltonian, Hamilton’s equations of motion are,

\[
\dot{q}_t \overset{\text{def.}}{=} \nabla_p H(q_t, p_t) \quad (8.4)
\]
\[
\dot{p}_t \overset{\text{def.}}{=} -\nabla_q H(q_t, p_t) \quad (8.5)
\]

where \( q(\cdot) \) and \( p(\cdot) \) are maps from \( \mathbb{R} \) to \( \mathbb{R}^m \), written as \( t \mapsto q_t \) and \( t \mapsto p_t \), which respect the initial conditions and the time-evolution specified in eqs. (8.4) and (8.5).

Definition 8.2.14. A function \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) is called separable if it can be written in the form \( H(q, p) = U(q) + K(p) \) where \( U \) and \( K \) are smooth maps from \( \mathbb{R}^m \) to \( \mathbb{R} \). Otherwise the Hamiltonian is called non-separable.

Definition 8.2.15. Given a smooth function \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \), the Hamiltonian flow map \( \Phi : \mathbb{R} \times \mathbb{R}^{2m} \to \mathbb{R}^{2m} \) of Hamilton’s equations of motion from eqs. (8.4) and (8.5) is defined by \( \Phi_t(q_0, p_0) = (q_t, p_t) \) where \( q_t \) and \( p_t \) are defined in definition 8.2.13. As a notational convenience we will often write \( \Phi_t : \mathbb{R}^{2m} \to \mathbb{R}^{2m} \) as the map \( (q_0, p_0) \mapsto \Phi_t(q_0, p_0) \).

Definition 8.2.16. Let \( \delta, \delta' \in \mathbb{R}^{2m} \) and let \( z \in \mathbb{R}^{2m} \). Write \( z = (q, p) \) where \( q, p \in \mathbb{R}^m \). The canonical Euclidean symplectic structure \( \Omega : \mathbb{R}^{2m} \times \mathbb{R}^{2m} \to \mathbb{R} \) is the map defined by

\[
\Omega(\delta, \delta') \overset{\text{def.}}{=} \sum_{i=1}^{m}(d\delta_i \wedge d\delta_{m+i})(\delta, \delta') \quad (8.6)
\]
\[
= \sum_{i=1}^{m}(d\delta_i \wedge dp_i)(\delta, \delta') \quad (8.7)
\]
\[
= \sum_{i=1}^{m} \delta_i \delta'_{m+i} - \delta_{m+i} \delta'_i \quad (8.8)
\]

Notice that eq. (8.8) follows by applying definition 8.2.11 and the fact that we have partitioned \( \mathbb{R}^{2m} \) into two sets of coordinates: the first half of which is denoted by \( q \in \mathbb{R}^m \).
and the second half denoted by \( p \in \mathbb{R}^m \). Often we adopt the notational short-hand \( dq = (dq_1, \ldots, dq_m) \) and \( dp = (dp_1, \ldots, dp_m) \) and write
\[
\Omega = \sum_{i=1}^{m} dq_i \wedge dp_i = dq \wedge dp. \tag{8.9}
\]

**Definition 8.2.17.** Let \( (q, p) \in \mathbb{R}^{2m} \). A diffeomorphism \( \Phi : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m} \) is called Euclidean symplectic if, under the change-of-variables \( (\hat{q}, \hat{p}) = \Phi(q, p) \), we have
\[
\sum_{i=1}^{m} d\hat{q}_i \wedge d\hat{p}_i = \sum_{i=1}^{m} dq_i \wedge dp_i \tag{8.10}
\]
where \( d\hat{q}_i \) and \( d\hat{p}_i \) are computed according to definition 8.2.12.

Alternatively, using the notational short-hand of eq. (8.9), one may express eq. (8.10) as \( d\hat{q} \wedge d\hat{p} = dq \wedge dp \).

**On Implicitly Defined Manifolds**

**Definition 8.2.18.** Let \( M \subset \mathbb{R}^m \) be a manifold implicitly defined by a function \( g : \mathbb{R}^m \rightarrow \mathbb{R}^k \) as in definition 8.2.1. Write the Jacobian of \( g \) by \( \nabla g : \mathbb{R}^m \rightarrow \mathbb{R}^{k \times m} \). Let \( H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \) be smooth. Let \( q_0 \in M \) and \( p_0 \in T^*_q M \) (definition 8.2.7). Hamilton’s equations of motion on \( T^* M \) (definition 8.2.8) are defined as,
\[
\dot{q}_t \overset{\text{def.}}{=} \nabla_p H(q_t, p_t) \tag{8.11}
\]
\[
\dot{p}_t \overset{\text{def.}}{=} -\nabla_q H(q_t, p_t) + \nabla g(q_t)^\top \lambda(q_t, p_t) \tag{8.12}
\]
\[
0 = g(q_t) \tag{8.13}
\]
where \( \lambda : T^* M \rightarrow \mathbb{R}^k \) and \( q(\cdot) : \mathbb{R} \rightarrow M \) and \( p(\cdot) : \mathbb{R} \rightarrow T^*_q(\cdot) \).

**Lemma 8.2.19.** The momentum function \( p(\cdot) : \mathbb{R} \rightarrow T^*_q(\cdot) \) satisfies the hidden constraint \( \nabla g(q_t)\nabla_p H(q_t, p_t) = 0 \).
A proof is given in section 8.A.1.4.

**Lemma 8.2.20** (Leimkuhler and Reich [2005]). *The function* $\lambda : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^k$ *is uniquely defined if the matrix* $\nabla g(q)\nabla^2 p H(q, p) \nabla g(q)^T$ *is invertible.*

A proof is given in section 8.A.1.5.

**Definition 8.2.21.** Let $M$ be a manifold implicitly defined by a function $g$ as in definition 8.2.1. Let $H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ be smooth and let $T^* M$ denote the associated cotangent bundle. A map $\Phi : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m}$ is said to be manifold-constrained if $\Phi(z) \in T^* M$ whenever $z \in T^* M$.

Equivalently, a manifold-constrained map is a map $\Phi : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m}$ such that if $(q, p) \in \mathbb{R}^{2m}$ satisfies $q \in M$ and $p \in T^*_q M$ and if we write $(\hat{q}, \hat{p}) = \Phi(q, p)$, then $\hat{q} \in M$, $\hat{p} \in T^*_q M$.

**Definition 8.2.22.** Let $M$ be a manifold implicitly defined by a function $g$ as in definition 8.2.1. Let $H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ be smooth and let $T^* M$ denote the associated cotangent bundle viewed as an embedded sub-manifold of $\mathbb{R}^{2m}$. Let $TT^* M$ denote the tangent space of the cotangent bundle, also viewed as an embedded sub-manifold of $\mathbb{R}^{2m}$. Let $(q, p) \in T^* M$ and let $dq \wedge dp$ be the canonical Euclidean symplectic structure (definition 8.2.16 with $(q, p)$ viewed as an element of $\mathbb{R}^{2m}$). The *canonical manifold symplectic structure* is the restriction of $dq \wedge dp$ to $T(q, p) T^* M \times T(q, p) T^* M \subset \mathbb{R}^{2m} \times \mathbb{R}^{2m}$.

As in the Euclidean case, we continue to write the canonical manifold symplectic structure as $dq \wedge dp = \sum_{i=1}^m dq_i \wedge dp_i$ using the short-hand notation of eq. (8.9).

**Definition 8.2.23.** Let $M$ be a manifold implicitly defined by a function $g$ as in definition 8.2.1. Let $H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ be smooth and let $T^* M$ denote the associated cotangent bundle viewed as an embedded sub-manifold of $\mathbb{R}^{2m}$. Let $(q, p) \in T^* M$ and write $dq \wedge dp$ (using the short-hand notation of eq. (8.9)) for the canonical manifold symplectic structure (definition 8.2.22). We say that $\Phi : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m}$ is a manifold-constrained symplectic
transformation if $\Phi$ is manifold-constrained (definition 8.2.21) and if $dq \wedge dp = dq \wedge dp$ where $dq_i$ and $dp_i$ are computed according to definition 8.2.12.

**Remark.** In either the Euclidean case or the case of an implicitly defined manifold, symplectic transformations are volume-preserving in the sense described in section 2.3.8.

### 8.2.3 Methods of Numerical Integration

As in the case of formulating Hamiltonian mechanics, we present methods of numerical integration both in the case of Euclidean spaces and on implicitly defined manifolds. Central to our later discussion will be the fact that the certain numerical integrators have inverses that can be obtained by negating the integration step-size. Such methods are called symmetric.

**Definition 8.2.24.** Let $M$ be a set. A map $\Phi : \mathbb{R} \times M \to M$ is called symmetric on $M$ if, for every $\epsilon \in \mathbb{R}$, $\Phi_\epsilon \circ \Phi_{-\epsilon} = \text{Id}$. Equivalently, a map $\Phi_\epsilon$ is symmetric if $\Phi_{-\epsilon} = \Phi_{\epsilon}^{-1}$.

Importantly, if for every $\epsilon \in \mathbb{R}$, $\Phi_\epsilon : M \to M$ is symmetric and smooth, then it means that $\Phi_\epsilon$ is a diffeomorphism of $\mathbb{R}^{2m}$ with inverse map $\Phi_{-\epsilon}$.

**In Euclidean Space**

**Definition 8.2.25.** A map $\hat{\Phi} : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ is said to be a $k^{\text{th}}$-order approximation of $\Phi : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ if, for any $z \in \mathbb{R}^m$,

$$\|\hat{\Phi}_t(z) - \Phi_t(z)\| = \mathcal{O}(t^{k+1}).$$

(8.14)

**Definition 8.2.26.** Let $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be a smooth Hamiltonian and consider the equations of motion in eqs. (8.4) and (8.5). The implicit midpoint integrator is the map
\( \Phi : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m \times \mathbb{R}^m \) defined by the following operations.

\[
\begin{pmatrix}
\tilde{q} \\
\tilde{p}
\end{pmatrix} = \begin{pmatrix}
q \\
p
\end{pmatrix} + \epsilon \begin{pmatrix}
\nabla_p H(q, \tilde{p}) \\
-\nabla_q H(q, \tilde{p})
\end{pmatrix}
\]  

(8.15)

where \( \tilde{q} = (q + q)/2 \) and \( \tilde{p} = (p + p)/2 \) and \( (\tilde{q}, \tilde{p}) = \Phi_\epsilon(q, p) \).

**Proposition 8.2.27.** The implicit midpoint integrator is a 2nd-order approximation of the Hamiltonian flow map (definition 8.2.15), is symmetric on \( \mathbb{R}^m \times \mathbb{R}^m \) (definition 8.2.24), and, for fixed \( t, \Phi(t, \cdot) : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m} \) is symplectic (definition 8.2.17).

**Definition 8.2.28.** Let \( H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \) be a smooth Hamiltonian and consider the equations of motion in eqs. (8.4) and (8.5). The generalized leapfrog integrator is the map \( \Phi_\epsilon : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^m \times \mathbb{R}^m \) defined by the following operations.

\[
\begin{align*}
\tilde{p} &= p - \frac{\epsilon}{2} \nabla_q H(q, \tilde{p}) \\
\tilde{q} &= q + \frac{\epsilon}{2} \left[ \nabla_p (q, \tilde{p}) + \nabla_p (\tilde{q}, \tilde{p}) \right] \\
\tilde{p} &= \tilde{p} - \frac{\epsilon}{2} \nabla_q H(\tilde{q}, \tilde{p})
\end{align*}
\]  

(8.16) (8.17) (8.18)

where \( (\tilde{q}, \tilde{p}) = \Phi_\epsilon(q, p) \).

**Proposition 8.2.29.** The generalized leapfrog integrator is a 2nd-order approximation of the Hamiltonian flow map (definition 8.2.15), is symmetric on \( \mathbb{R}^m \times \mathbb{R}^m \) (definition 8.2.24), and, for fixed \( t, \Phi(t, \cdot) : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m} \) is symplectic (definition 8.2.17).

Proofs of propositions 8.2.27 and 8.2.29 results may be found in section 2.2.

**On Implicitly Defined Manifolds**

**Theorem 8.2.30.** Let \( H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \) be smooth. Let \( \Phi \) be the Hamiltonian flow map (definition 8.2.15) for Hamilton’s equations of motion (definition 8.2.13). Let \( \Phi \) be a
2nd-order approximation of $\Phi$ and suppose further that $\hat{\Phi}$ is symmetric on $\mathbb{R}^m \times \mathbb{R}^m$ (definition 8.2.24). Let $M$ be a manifold implicitly defined by a function $g$ as in definition 8.2.1. Let $q_0 \in M$ and suppose $p_0 \in \mathbb{R}^m$ satisfies the hidden constraint from lemma 8.2.19. Then a 2nd-order approximation $\hat{\Psi}_{\epsilon} : T^*M \rightarrow T^*M$ of the constrained Hamiltonian motion in eqs. (8.11) to (8.13) can be constructed from the solution to the following system of equations:

\begin{align}
\bar{p} &= p - \frac{\epsilon}{2} \nabla g(q)^\top \mu \\
(q, \tilde{p}) &= \hat{\Phi}_{\epsilon}(q, \bar{p}) \\
0 &= g(\tilde{q}) \\
\tilde{p} &= \bar{p} - \frac{\epsilon}{2} \nabla g(\tilde{q})^\top \mu' \\
0 &= \nabla g(\tilde{q}) \nabla_p H(\tilde{q}, \tilde{p})
\end{align}

and $\hat{\Psi}_{\epsilon}(q, p) \overset{\text{def.}}{=} (\tilde{q}, \tilde{p})$. Moreover, $\hat{\Psi}$ is symmetric on $T^*M$ (definition 8.2.24) and, for fixed $\epsilon$, $\hat{\Psi}_{\epsilon} : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m}$ is a manifold-constrained symplectic transformation (definition 8.2.23).

A proof is given in section 8.A.1.6. For a thorough treatment of numerical integrators, which contains this result, we also refer the reader to [Leimkuhler and Reich 2005].

Remark. The fact that $\hat{\Psi}_{\epsilon}$ is symmetric on $T^*M$ is important when applied in the context of diffeomorphism Monte Carlo, since definition 8.2.24 implies that $\hat{\Psi}_{\epsilon}^{-1} = \hat{\Psi}_{-\epsilon}$. Later, in section 8.3.3 and in particular in algorithm 22, we will see how the randomization of the sign of the integration step-size allows us to randomly choose between a diffeomorphism and its inverse as required by definition 8.2.32.

Proposition 8.2.31. Let $H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ be smooth and satisfy the condition of lemma 8.2.20. Let $M$ be a manifold implicitly defined by a function $g$ as in definition 8.2.1, let $G$ be the Jacobian of $g$, which is assumed to have full-rank everywhere in $\mathbb{R}^m$. Then,
provided \( \epsilon \in \mathbb{R} \) is sufficiently small, there is a unique \( \mu \) (respectively \( \mu' \)) in eq. (8.19) (resp. eq. (8.22)) in a neighborhood of \( 0 \in \mathbb{R}^k \) satisfying eq. (8.21) (resp. eq. (8.23)).

For a proof of this result, we refer the interested reader to [McLachlan et al., 2014].

8.2.4 Markov Chain Monte Carlo

In this section we recall the basic results of diffeomorphism Monte Carlo, which are detailed in chapter 3.

Definition 8.2.32. Let \( \Pi : \mathcal{B}(\mathbb{R}^{2m}) \to [0, 1] \) be a probability measure on \( \mathbb{R}^{2m} \) with density \( \pi : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}_+ \). Let \( z \in \mathbb{R}^{2m} \) and let \( \Phi : \mathbb{R}^{2m} \to \mathbb{R}^{2m} \) be a diffeomorphism with unit Jacobian determinant (i.e. \( \Phi^{-1} \) is smooth and \( |\det(\nabla_z \Phi(z))| = 1 \)). The Markov chain transition kernel of diffeomorphism Monte Carlo is

\[
K(z, A) = \frac{1}{2} \int_A \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \, dz + \frac{1}{2} \left( 1 - \int_{\mathbb{R}^{2m}} \min \left\{ 1, \frac{\pi(\Phi(z))}{\pi(z)} \right\} \, dz \right) \\
+ \frac{1}{2} \int_A \min \left\{ 1, \frac{\pi(\Phi^{-1}(z))}{\pi(z)} \right\} \, dz + \frac{1}{2} \left( 1 - \int_{\mathbb{R}^{2m}} \min \left\{ 1, \frac{\pi(\Phi^{-1}(z))}{\pi(z)} \right\} \, dz \right),
\]

(8.24)

where \( A \in \mathcal{B}(\mathbb{R}^{2m}) \).

Lemma 8.2.33. Let \( K \) be the transition kernel of diffeomorphism Monte Carlo in definition 8.2.32. Then \( K \) satisfies detailed balance with respect to the distribution \( \Pi \).

Therefore, if we can identify a diffeomorphism \( \Phi \) with unit Jacobian determinant on \( \mathbb{R}^{2m} \), then we may apply this diffeomorphism in definition 8.2.32 to obtain a Markov chain satisfying detailed balance. The use of numerical integrators that satisfy these properties is a central component of the Hamiltonian Monte Carlo algorithm.

Corollary 8.2.34. Let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be a smooth Hamiltonian and let \( \pi(q, p) \propto \exp(-H(q, p)) \) be the probability density of a probability measure \( \Pi : \mathcal{B}(\mathbb{R}^{2m}) \to [0, 1] \).
Let \( \hat{\Phi} \) be a symmetric (definition \[8.2.24\]), symplectic (definition \[8.2.17\]), and 2nd-order (definition \[8.2.25\]) accurate approximation of the Hamiltonian flow map (definition \[8.2.15\]) of the Hamiltonian \( H \). We observe that \( \hat{\Phi}_\epsilon \) is a volume-preserving diffeomorphism with inverse \( \hat{\Phi}_{-\epsilon} \). Hence we may construct a Markov chain transition kernel from \( \hat{\Phi}_\epsilon \), respecting detailed balance with respect to \( \Pi \), using the framework of diffeomorphism Monte Carlo in definition \[8.2.32\].

Remark. The assumption that \( \hat{\Phi}_\epsilon \) is a second-order approximation in corollary \[8.2.34\] is typical throughout the literature on HMC and suggests that the HMC Markov chain should have high acceptance probabilities since we anticipate that \( \pi(\hat{\Phi}_\epsilon(z)) \approx \pi(z) \) (and also \( \pi(\hat{\Phi}_{-\epsilon}(z)) \approx \pi(z) \)). However, the fact that the integrator is second-order is immaterial insofar as detailed balance is concerned.

### 8.3 Analytical Apparatus

In this section we will discuss the explicit integrator (definition \[8.3.10\]) that was proposed by Tao [2016] and which has been suggested for use in Markov chain Monte Carlo by Cobb et al. [2019]. We will propose a formulation of this integrator on a particular manifold – the equality manifold (definition \[8.3.1\]) – which leads to the explicit integrator with Lagrange multipliers (definition \[8.3.12\]). We then discuss methods of resolving the Lagrange multipliers in section \[8.3.2\] and conclude by summarizing the algorithmic approaches we consider in section \[8.3.3\].

#### 8.3.1 Symmetry and Symplecticness via Lagrange Multipliers

Having developed the preliminaries of methods of geometric integration and embedded manifolds, we proceed now to discuss the integration of non-separable Hamiltonian systems. We begin with a problem statement.
**Problem.** Given a non-separable, smooth Hamiltonian $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ (definition 8.2.14), construct a map $\hat{\Phi} : \mathbb{R} \times \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ which is (at least) a 2nd-order approximation (definition 8.2.25) of the corresponding Hamiltonian flow map $\Phi$ (definition 8.2.15), is a symmetric map (definition 8.2.24), and, for fixed $\epsilon \in \mathbb{R}$, the map $\hat{\Phi}_\epsilon : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ is a symplectic transformation (definition 8.2.17).

We have already seen two possible solutions to this problem using the implicit midpoint integrator (definition 8.2.26) or the generalized leapfrog algorithm (definition 8.2.28), as discussed in propositions 8.2.27 and 8.2.29. Both of these integrators are defined implicitly.

**Remark.** As mentioned in section 8.2.2, Hamiltonian mechanics is traditionally developed using the notation $(q, p)$ to refer to a point of phase space in $\mathbb{R}^{2m}$. The implicitly defined integrators in definitions 8.2.26 and 8.2.28 may be applied on $(q, p)$-phase space. However, our development will require an expanded $(q, x, p, y)$-phase space, consisting of auxiliary position variables $x$ and momentum variables $y$. In the expanded phase space, we adopt the notation $\alpha = (q, x)$ to refer to the expanded position variables and $\beta = (p, y)$ to refer to the expanded momentum variables.

**Remark.** In this section we also reserve some special notation for Hamiltonians. We use the notation $H$ to refer exclusively to a Hamiltonian on $\mathbb{R}^m \times \mathbb{R}^m \cong \mathbb{R}^{2m}$; we use the notation $S$ for a Hamiltonian on the space $\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \cong \mathbb{R}^{4m}$; finally, we use the notation $H_\omega$ (defined in definition 8.3.5) to refer to a particular procedure for creating a Hamiltonian on $\mathbb{R}^{4m}$ given a Hamiltonian on $\mathbb{R}^{2m}$.

The purpose of the remainder of this section is to develop an integrator using a manifold construction rather than implicitly defined updates. We begin by introducing the implicitly defined manifold (definition 8.2.1) that will be of particular interest to us.

**Definition 8.3.1.** Let $\alpha = (q, x) \in \mathbb{R}^m \times \mathbb{R}^m \cong \mathbb{R}^{2m}$ where $q, x \in \mathbb{R}^m$. Consider the implicitly defined manifold $E \subset \mathbb{R}^{2m}$ defined by the constraint $e(\alpha) \equiv e(q, x) = q - x$. 

439
That is,

\[ E \overset{\text{def.}}{=} \{ (q, x) \in \mathbb{R}^{2m} : e(q, x) = 0 \} \quad (8.25) \]

We call \( E \) the equality manifold.

**Lemma 8.3.2.** The function \( e(\alpha) \) has the full-rank Jacobian \( \nabla e(\alpha) = (\text{Id}, -\text{Id}) \in \mathbb{R}^{m \times 2m} \).

A proof is given in section 8.A.1.7

**Corollary 8.3.3.** The equality manifold is an implicitly defined manifold in the sense of definition 8.2.1.

**Remark.** Let \( S : \mathbb{R}^{2m} \times \mathbb{R}^{2m} \to \mathbb{R} \) be a smooth Hamiltonian. When formulating Hamiltonian mechanics on \( T^*E \subset \mathbb{R}^{2m} \times \mathbb{R}^{2m} \), we will expand \( (\alpha, \beta) \in T^*E \) as \( \alpha = (q, x) \) and \( \beta = (p, y) \) and will write the Hamiltonian both as a function \( S(\alpha, \beta) \) and as \( S : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) via \( S(q, x, p, y) = S((q, x), (p, y)) \) making use of the isomorphism of \( \mathbb{R}^m \times \mathbb{R}^m \) and \( \mathbb{R}^{2m} \).

**Proposition 8.3.4.** Let \( E \) be defined as in definition 8.3.1 and consider a Hamiltonian \( S : \mathbb{R}^{2m} \times \mathbb{R}^{2m} \to \mathbb{R} \). Let \( \alpha_0 = (q_0, x_0) \in \mathbb{R}^{2m} \) and \( \beta_0 = (p_0, y_0) \in \mathbb{R}^{2m} \) where \( ((q_0, x_0), (p_0, y_0)) \in T^*E \). From definition 8.2.18 the manifold-constrained equations of motion from definition 8.2.18 on \( T^*E \), which are as follows:

\[ \dot{\alpha}_t = \nabla_{\beta} S(\alpha_t, \beta_t) \quad (8.26) \]

\[ \dot{\beta}_t = -\nabla_{\alpha} S(\alpha_t, \beta_t) \quad (8.27) \]

\[ 0 = e(\alpha_t) \quad (8.28) \]

where \( \alpha(\cdot) \) and \( \beta(\cdot) \) are maps from \( \mathbb{R} \) to \( \mathbb{R}^{2m} \) and satisfy the initial conditions. Then \( \beta \)
satisfies the hidden constraint

$$\nabla_p S(q_t, x_t, p_t, y_t) = \nabla_y S(q_t, x_t, p_t, y_t)$$  \hspace{1cm} (8.29)

where $\alpha_t \equiv (q_t, x_t)$, $\beta_t \equiv (p_t, y_t)$, and $S(\alpha, \beta) \equiv S(q, x, p, y)$.

A proof is given in section 8.A.1.8.

**Definition 8.3.5 (Tao [2016]).** Let $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be smooth and fix $\omega \geq 0$. The $\omega$-expanded Hamiltonian $H_\omega : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is defined by,

$$H_\omega(q, x, p, y) = H(q, y) + H(x, p) + \frac{\omega^2}{2} (\|q - x\|^2 + \|p - y\|^2)$$  \hspace{1cm} (8.30)

**Lemma 8.3.6.** Let $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be smooth and fix $\omega \geq 0$. Let $H_\omega$ be the corresponding $\omega$-expanded Hamiltonian from definition 8.3.5. The Hamiltonian equations of motion generated by $H_\omega$ on $T^*E$ are given by

$$\dot{q}_t = \nabla_p H_\omega(q_t, x_t, p_t, y_t) = \nabla_p H(x_t, p_t) + \omega(p_t - y_t)$$  \hspace{1cm} (8.31)

$$\dot{x}_t = \nabla_y H_\omega(q_t, x_t, p_t, y_t) = \nabla_y H(q_t, y_t) + \omega(y_t - p_t)$$  \hspace{1cm} (8.32)

$$\dot{p}_t = -\nabla_q H_\omega(q_t, x_t, p_t, y_t) + \lambda(q_t, x_t, p_t, y_t) = -\nabla_q H(q_t, y_t) - \omega(q_t - x_t) + \lambda(q_t, x_t, p_t, y_t)$$  \hspace{1cm} (8.33)

$$\dot{y}_t = -\nabla_x H_\omega(q_t, x_t, p_t, y_t) = -\nabla_x H(x_t, p_t) - \omega(x_t - q_t) - \lambda(q_t, x_t, p_t, y_t)$$  \hspace{1cm} (8.34)

$$0 = e(q_t, x_t)$$  \hspace{1cm} (8.35)

A proof is given in section 8.A.1.9.
Corollary 8.3.7. Let $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be smooth and fix $\omega \geq 0$. Let $H_\omega$ be the corresponding $\omega$-expanded Hamiltonian from definition 8.3.5. Then eq. (8.29) with $S = H_\omega$ simplifies to

$$\nabla_p H(q_t, p_t) + \omega(p_t - y_t) = \nabla_y H(q_t, y_t) - \omega(p_t - y_t), \quad (8.40)$$

for which the choice $p_t = y_t$ is always a solution.

A proof is given in section 8.A.1.10.

Proposition 8.3.8. Fix $\omega \geq 0$ and let $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be smooth. Let $H_\omega$ be the corresponding $\omega$-expanded Hamiltonian. If $\nabla^2_p H(x, p) + \omega \operatorname{Id}$ is invertible, then the choice $p_t = y_t$ is, locally, the only choice satisfying the hidden constraint condition in eq. (8.40).

A proof is given in section 8.A.1.11.

Example 39. Some particular examples of Hamiltonians $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ will require $p_t = y_t$. For instance, the Hamiltonian used in Riemannian manifold Hamiltonian Monte Carlo is of the form,

$$H(q, p) = U(q) + \frac{1}{2} p^\top G^{-1}(q)p + \frac{1}{2} \log \det((2\pi)^m G(q)), \quad (8.41)$$

where $G(q)$ is the positive-definite matrix of the Riemannian metric at $q \in \mathbb{R}^m$ and $U : \mathbb{R}^m \to \mathbb{R}$. Consider the associated $\omega$-expanded Hamiltonian for the Hamiltonian in eq. (8.41). Then the condition in eq. (8.40) becomes,

$$G^{-1}(q_t)p_t + \omega(p_t - y_t) = G^{-1}(x_t)y_t - \omega(p_t - y_t) \quad (8.42)$$

$$\implies (G^{-1}(q_t) + 2\omega \operatorname{Id})p_t = (G^{-1}(x_t) + 2\omega \operatorname{Id})y_t \quad (8.43)$$

$$\implies (G^{-1}(q_t) + 2\omega \operatorname{Id})p_t = (G^{-1}(q_t) + 2\omega \operatorname{Id})y_t, \quad (8.44)$$
which, when using the fact that \( G^{-1}(q) + 2\omega \text{ Id} \) defines a linear isomorphism of \( \mathbb{R}^m \), implies that \( p_t = y_t \).

The following theorem shows the relationship between Hamilton’s equations of motion generated by \( H \) in \((q, p)\)-space and the motion generated by the \( \omega \)-expanded Hamiltonian in \((q, x, p, y)\)-space.

**Theorem 8.3.9.** Let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be smooth and let \( H_\omega : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) be the corresponding \( \omega \)-expanded Hamiltonian (definition 8.3.5). Consider the Hamiltonian equations of motion generated by \( H_\omega \) on \( T^*E \) given in lemma 8.3.6 with the initial condition \( q_0 = x_0 \) and \( p_0 = y_0 \). The solution of the system is given by \( q_t = x_t = q_t \) and \( p_t = y_t = p_t \) where \((q_0, p_0) = (q_0, p_0)\) and

\[
\dot{q}_t = \nabla_p H(q_t, p_t) \quad \text{ (8.45)}
\]

\[
\dot{p}_t = -\nabla_q H(q_t, p_t). \quad \text{ (8.46)}
\]

Moreover, \( \lambda = 0 \).

A proof is given in section 8.A.1.12.

**Remark.** The result of theorem 8.3.9 suggests that one could just have well considered eqs. (8.31), (8.33), (8.35) and (8.37) without eq. (8.39) and the Lagrange multipliers fixed at zero. Thus, the usefulness of the formulation involving Lagrange multipliers is not in the identification of the analytical solution in eqs. (8.45) and (8.46), which satisfies the constraint eq. (8.39) by construction, but in the formulation of constraint-preserving numerical integrators. Indeed, Tao [2016] analyzed a numerical integrator for the unconstrained equations of motion in eqs. (8.31), (8.33), (8.35) and (8.37) without Lagrange multipliers to force the solution to lie in \( T^*E \); this integrator is the subject of the subsequent definition.

443
Definition 8.3.10. Let $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ be smooth and let $(q, x, p, y) \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m$. Let $\omega \geq 0$ and $\epsilon \in \mathbb{R}$. Define the following maps

\[
\Phi^1_\epsilon(q, x, p, y) \overset{\text{def.}}{=} \left( q, x + \frac{\epsilon}{2} \nabla_y H(q, y), p - \frac{\epsilon}{2} \nabla_q H(q, y), y \right)
\]

\[
\Phi^2_\epsilon(q, x, p, y) \overset{\text{def.}}{=} \left( q + \frac{\epsilon}{2} \nabla_p H(x, p), x, p, y - \frac{\epsilon}{2} \nabla_p H(x, p) \right)
\]

\[
\Phi^3_\epsilon : \begin{pmatrix} q \\ x \\ p \\ y \end{pmatrix} \overset{\text{def.}}{\mapsto} \frac{1}{2} \begin{pmatrix} q + x \\ p + y \\ q + x \\ p + y \end{pmatrix} + R_\epsilon \begin{pmatrix} q - x \\ p - y \\ q - x \\ p - y \end{pmatrix}
\]

where

\[
R_\epsilon = \begin{pmatrix} \cos(2\omega \epsilon) \text{Id}_m & \sin(2\omega \epsilon) \text{Id}_m \\ -\sin(2\omega \epsilon) \text{Id}_m & \cos(2\omega \epsilon) \text{Id}_m \end{pmatrix}.
\]

The explicit integrator without Lagrange multipliers is defined as the composition

\[
\hat{\Phi}_\epsilon \overset{\text{def.}}{=} \Phi^1_\epsilon \circ \Phi^2_\epsilon \circ \Phi^3_\epsilon \circ \Phi^2_\epsilon \circ \Phi^1_\epsilon.
\]

Thus, $\hat{\Phi}_\epsilon : \mathbb{R}^{4m} \to \mathbb{R}^{4m}$.

Theorem 8.3.11. On the space $\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \cong \mathbb{R}^{4m}$, the explicit integrator without Lagrange multipliers defined in definition 8.3.10 is a symmetric (definition 8.2.24), symplectic (definition 8.2.17), and, given the initial conditions $q_0 = x_0$ and $p_0 = y_0$, is a 2nd-order approximation of the Hamiltonian flow map (definition 8.2.15) $\Phi_t(q_0, q_0, p_0, p_0) \overset{\text{def.}}{=} (q_t, x_t, p_t, y_t)$ where $q_t, x_t, p_t, y_t$ are defined in lemma 8.3.6.

For a proof we refer the interested reader to Tao [2016].

Remark. Let $\hat{\Phi}_\epsilon$ be as in definition 8.3.10. Although the explicit integrator without La-
grange multipliers is a 2\textsuperscript{nd}-order accurate approximation of \( q_t, x_t, p_t, \) and \( y_t \), it differs in the important respect that it does not respect the manifold constraint or the manifold’s hidden constraint (lemma 8.2.19). That is, the map \( \Phi_\epsilon : (q, q, p, p) \mapsto (\tilde{q}, \tilde{x}, \tilde{p}, \tilde{y}) \) may have \( \tilde{q} \neq \tilde{x} \) and \( \tilde{p} \neq \tilde{y} \). This failure to satisfy the manifold constraint will present issues when applying the integrator to Hamiltonian Monte Carlo.

Definition 8.3.12. Let \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) and \( \Phi_\epsilon : \mathbb{R}^{4m} \to \mathbb{R}^{4m} \) be as in definition 8.3.10. Let \( (\alpha, \beta) \in T^*E \). The explicit integrator with Lagrange multipliers \( \mu, \mu' \in \mathbb{R}^m \) is defined as the map \( \hat{\Psi}_\epsilon : T^*E \to T^*E \) with \( \hat{\Psi}_\epsilon(\alpha, \beta) = (\tilde{\alpha}, \tilde{\beta}) \) defined as the solution to the following equations:

\[
\tilde{\beta} = \beta - \begin{pmatrix} \mu \\ -\mu \end{pmatrix} \tag{8.52}
\]

\[
(\tilde{\alpha}, \tilde{\beta}) = \Phi_\epsilon(\alpha, \beta) \tag{8.53}
\]

\[
0 = e(\tilde{\alpha}) \tag{8.54}
\]

\[
\tilde{\beta} = \tilde{\beta} - \begin{pmatrix} \mu' \\ -\mu' \end{pmatrix} \tag{8.55}
\]

\[
0 = \nabla_p H(\tilde{x}, \tilde{p}) - \nabla_y H(\tilde{q}, \tilde{y}), \tag{8.56}
\]

where \( \tilde{\alpha} = (\tilde{q}, \tilde{x}) \in E \) and \( \tilde{\beta} = (\tilde{p}, \tilde{y}) \in T^*_\tilde{\alpha}E \).

Proposition 8.3.13. The map \( \hat{\Psi} \) defined in definition 8.3.12 is a symmetric transformation (definition 8.2.24), manifold-symplectic on \( E \) (definition 8.2.23), and 2\textsuperscript{nd}-order accurate approximation of the Hamiltonian flow map (definition 8.2.15) \( \Phi_t(q_0, q_0, p_0, p_0) \) def. \( = (q_t, x_t, p_t, y_t) \) where \( q_t, x_t, p_t, \) and \( y_t \) are defined in theorem 8.3.9.

A proof is given in section 8.A.1.13.

Remark. Given the restriction that \( q = x \) on \( E \), if we further assume that the hidden constraint implies \( p = y \) then there is a relationship between the Euclidean symplectic
structure $dq \wedge dp$ (definition 8.2.16) on the $(q, p)$-phase space and the manifold symplectic structure $d\alpha \wedge d\beta$ on $E$ (definition 8.2.22). This is the subject of proposition 8.3.14 which shows that the symplectic structures are proportional.

**Proposition 8.3.14.** Let $(\alpha, \beta) \in T^*E$. Identify $\alpha = (q, x)$ and $\beta = (p, y)$. Using the short-hand convention of eq. (8.9), the symplectic structure on $E$ then satisfies $d\alpha \wedge d\beta = 2\ dq \wedge dp$.

A proof is given in section 8.A.1.14.

**Corollary 8.3.15.** Consider $E$ as defined in definition 8.3.1 and suppose $S : \mathbb{R}^{4m} \rightarrow \mathbb{R}$ is a Hamiltonian for which the hidden constraint on $T^*E$ requires $p = y$ (true for the $\omega$-expanded Hamiltonians considered in example 39). Let $\alpha = (q, q) \in E$ and $\beta = (p, p) \in T^*_\alpha E$ and let $(\bar{\alpha}, \bar{\beta})$ be as in definition 8.3.12. Identify $\bar{\alpha} = (\bar{q}, \bar{x})$ and $\bar{\beta} = (\bar{p}, \bar{y})$. Then, if we set $\bar{q} \stackrel{\text{def}}{=} (\bar{q} + \bar{x})/2$ and $\bar{p} \stackrel{\text{def}}{=} (\bar{p} + \bar{y})/2$, we have $d\bar{q} \wedge d\bar{p} = dq \wedge dp$ using the short-hand convention of eq. (8.9).

A proof is given in section 8.A.1.15.

**Lemma 8.3.16.** The map $(q, p) \mapsto (\bar{q}, \bar{p})$ defined in corollary 8.3.15 is invertible.

A proof is given in section 8.A.1.16.

**Remark.** If given $(q, p)$ one computes $(\bar{\alpha}, \bar{\beta})$ defined in definition 8.3.12 with the transformation to $(\bar{q}, \bar{p})$ defined in corollary 8.3.15 then one has constructed an invertible (from lemma 8.3.16) and symplectic (from corollary 8.3.15) map from $\mathbb{R}^m \times \mathbb{R}^m$ to itself. Thus, we may propose this procedure as a solution to the problem statement given at the beginning of section 8.3.1.

### 8.3.2 Methods for identifying the Lagrange Multipliers

If one is to evaluate definition 8.3.12 experimentally, then it becomes necessary to actually compute Lagrange multipliers $\mu, \mu' \in \mathbb{R}^m$. In this section, we discuss how this may be
accomplished. In the case, of $\mu'$, we are able to identify a simple closed-form solution for the Lagrange multipliers as the following result shows.

**Proposition 8.3.17.** Consider computing $\tilde{\beta}$ in eq. (8.55) given $\beta$ in eq. (8.53). For $\epsilon$ sufficiently small, there exist Lagrange multipliers $\mu'$ in eq. (8.55) such that if $\beta = (\tilde{\rho}, \tilde{\gamma})$ then

$$\tilde{\beta} = \left( \frac{\tilde{\rho} + \tilde{\gamma}}{2} \right).$$  \hspace{1cm} (8.57)

A proof is given in section 8.A.1.17.

**Remark.** While for sufficiently small $\epsilon$ there exists a closed-form for the $\tilde{\beta}$ update as discussed in proposition 8.3.17, there is no obvious closed-form for the Lagrange multipliers $\mu$ appearing in eq. (8.55). Therefore, as far as an implementation of the explicit integrator with Lagrange multipliers is concerned (definition 8.3.12), it is necessary to design a numerical sub-routine for finding suitable Lagrange multipliers. Rewriting eqs. (8.52) to (8.54) to make the dependence on the Lagrange multipliers explicit, we have,

$$\bar{\beta}(\mu) \defeq \beta - \left( \begin{array}{c} \mu \\ -\mu \end{array} \right)$$ \hspace{1cm} (8.58)

$$(\tilde{\alpha}(\mu), \tilde{\beta}(\mu)) \defeq \hat{\Phi}_e(\alpha, \bar{\beta}(\mu))$$ \hspace{1cm} (8.59)

$$f(\mu) \defeq e(\tilde{\alpha}(\mu))$$ \hspace{1cm} (8.60)

Thus we may identify suitable Lagrange multipliers by finding the root of the function $f(\mu)$ defined in eq. (8.60). A root finding method such as Newton’s method or a quasi-Newton method (such as Broyden’s method) can therefore be leveraged to identify Lagrange multipliers $\mu$ such that $f(\mu) = e(\alpha'(\mu)) = 0$. A quasi-Newton method may be the more desirable of the two since it will not require higher-order derivative information as
Algorithm 17 A second-order explicit numerical integrator of Hamilton’s equations of motion which is not suitable for Hamiltonian Monte Carlo.

1: **Input**: Initial position variables \((q, x) \in \mathbb{R}^m \times \mathbb{R}^m\) and initial momentum variables \((p, y) \in \mathbb{R}^m \times \mathbb{R}^m\), step-size \(\epsilon \in \mathbb{R}\), and Hamiltonian expansion parameter \(\omega \geq 0\).
2: Compute \((\tilde{q}, \tilde{x}, \tilde{p}, \tilde{y}) = \hat{\Phi}_\epsilon(q, x, p, y)\) where \(\hat{\Phi}_\epsilon\) is defined in eq. (8.51).
3: **Return**: \((\tilde{q}, \tilde{x}, \tilde{p}, \tilde{y})\).

Algorithm 18 Procedure for solving the equation \(f(z) = 0\) via Newton’s method.

1: **Input**: Function \(f : \mathbb{R}^m \to \mathbb{R}^m\), initial guess \(z \in \mathbb{R}^m\), convergence tolerance \(\delta \geq 0\).
2: Set \(\Delta = f(z)\) and \(z' = z\).
3: **While**: \(\Delta > \delta\) compute

\[
\begin{align*}
    z'' &= z' - (\nabla f (z'))^{-1} f(z') \quad (8.61) \\
    \Delta &= \|f(z'')\|_\infty \quad (8.62) \\
    z' &= z'' \quad (8.63)
\end{align*}
\]

4: **Return**: \(z' \in \mathbb{R}^m\).

Algorithm 19 Procedure for solving the equation \(f(z) = 0\) via Broyden’s method.

1: **Input**: Function \(f : \mathbb{R}^m \to \mathbb{R}^m\), initial guess \(z \in \mathbb{R}^m\), convergence tolerance \(\delta \geq 0\).
2: Set \(\Delta = f(z)\) and \(z' = z\).
3: Initialize \(J = \text{Id}\).
4: **While**: \(\Delta > \delta\) compute

\[
\begin{align*}
    z'' &= z' - J^{-1} f(z') \quad (8.64) \\
    \Delta &= \|f(z'')\|_\infty \quad (8.65) \\
    \Delta f &= f(z'') - f(z') \quad (8.66) \\
    \Delta z &= z'' - z' \quad (8.67) \\
    J &\leftarrow J + \frac{\Delta f - J \Delta z \Delta z^\top}{\|\Delta z\|^2} \quad (8.68) \\
    z' &= z'' \quad (8.69)
\end{align*}
\]

5: **Return**: \(z' \in \mathbb{R}^m\).

demanded by Newton’s method.

8.3.3 Algorithms and Approximations

In this section we seek to provide algorithmic descriptions of the mathematical tech-
Algorithm 20 A second-order numerical integrator of Hamilton’s equations of motion constrained to the equality manifold that is suitable for Hamiltonian Monte Carlo.

1: **Input:** Initial position variables \( q \in \mathbb{R}^m \) and initial momentum variables \( p \in \mathbb{R}^m \), step-size \( \epsilon \in \mathbb{R} \), Hamiltonian expansion parameter \( \omega \geq 0 \), Newton convergence tolerance \( \delta \geq 0 \).
2: Set \( x = q \) and \( y = p \).
3: Set \( \alpha = (q, x) \) and \( \beta = (p, y) \) so that \( \alpha \in E \) and \( \beta \in T^*_\alpha E \).
4: Define the function \( f(\mu) \) as in eq. (8.60).
5: Identify the root \( \mu^* \) of \( f \) using algorithm 18 (or algorithm 19) with initial guess \( 0 \in \mathbb{R}^m \) and convergence tolerance \( \delta \).
6: Compute \((\tilde{q}, \tilde{x}, \tilde{p}, \tilde{y}) = \hat{\Phi}^\epsilon(q, x, p + \mu^*, y - \mu^*)\) where \( \hat{\Phi}^\epsilon \) is defined in eq. (8.51).
7: Set \( \tilde{q} = (\tilde{q} + \tilde{x})/2 \). (Note that when \( \delta = 0 \), it follows that \( \tilde{q} = \tilde{x} \) because of the manifold constraint and hence, in this case, \( \tilde{q} = \tilde{x} \).)
8: Set \( \tilde{p} = (\tilde{p} + \tilde{y})/2 \) in accordance with proposition 8.3.17.
9: **Return:** \((\tilde{q}, \tilde{p})\).

Algorithm 21 A biased Markov chain technique for generating samples from a target probability distribution using the explicit integrator in algorithm 17.

1: **Input:** Initial Markov chain position \( q \in \mathbb{R}^m \) and momentum \( p \in \mathbb{R}^m \), a potential energy function \( U : \mathbb{R}^m \to \mathbb{R} \), a Riemannian metric \( G : \mathbb{R}^m \to \text{PD}(m) \), an integration step-size \( \epsilon \in \mathbb{R} \), a number of integration steps \( k \in \mathbb{N} \), a Hamiltonian expansion parameter \( \omega \geq 0 \).
2: Form the Hamiltonian \( H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} \) according to eq. (8.41).
3: With probability \( 1/2 \) set \( \gamma = \epsilon \); otherwise set \( \gamma = -\epsilon \).
4: Initialize the proposal state \((\tilde{q}, \tilde{p}) = (q, p)\).
5: **for** \( i = 1, \ldots, k \) **do**
6: Use algorithm 17 with position variables \((\tilde{q}, \tilde{p})\), momentum variables \((\tilde{p}, \tilde{p})\), step-size \( \gamma \), and Hamiltonian expansion parameter \( \omega \) to compute \((\tilde{q}', \tilde{x}', \tilde{p}', \tilde{y}')\).
7: Update the proposal state \((\tilde{q}, \tilde{p}) = (\tilde{q}', \tilde{p}')\).
8: **end for**
9: Sample \( u \sim \text{Uniform}(0,1) \).
10: Compute the Metropolis-Hastings acceptance rate \( m = \min \{1, \exp (H(q, p) - H(\tilde{q}, \tilde{p}))\} \).
11: **if** \( u < m \) **then**
12: **Return:** The point \((\tilde{q}, \tilde{p})\) as the next Markov chain state.
13: **else**
14: **Return:** The point \((q, p)\) as the next Markov chain state.
15: **end if**
Algorithm 22 A Markov chain technique for generating samples from a target probability distribution using the explicit integrator with Lagrange multipliers in algorithm 20. As $\delta \to 0$, this procedure faithfully satisfies the detailed balance condition.

1: **Input:** Initial Markov chain position $q \in \mathbb{R}^m$ and momentum $p \in \mathbb{R}^m$, a potential energy function $U : \mathbb{R}^m \to \mathbb{R}$, a Riemannian metric $G : \mathbb{R}^m \to \text{PD}(m)$, an integration step-size $\epsilon \in \mathbb{R}$, a number of integration steps $k \in \mathbb{N}$, a Hamiltonian expansion parameter $\omega \geq 0$, a convergence parameter $\delta \geq 0$ for the Lagrange multipliers.

2: Form the Hamiltonian $H : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ according to eq. (8.41).

3: With probability $1/2$ set $\gamma = \epsilon$; otherwise set $\gamma = -\epsilon$.

4: Initialize the proposal state $(\tilde{q}, \tilde{p}) = (q, p)$.

5: **for** $i = 1, \ldots, k$ **do**

6: Use algorithm 20 with position variables $(\tilde{q}, \tilde{q}')$, momentum variables $(\tilde{p}, \tilde{p}')$, step-size $\gamma$, Hamiltonian expansion parameter $\omega$, and convergence tolerance $\delta$ to compute $(\tilde{q}', \tilde{x}', \tilde{p}', \tilde{y}')$.

7: Update the proposal state $(\tilde{q}, \tilde{p}) = (\tilde{q}', \tilde{p}')$.

8: **end for**

9: Sample $u \sim \text{Uniform}(0, 1)$.

10: Compute the Metropolis-Hastings acceptance rate

$$m = \min \{1, \exp (H(q, p) - H(\tilde{q}, \tilde{p}))\}.$$  \hspace{1cm} (8.71)

11: **if** $u < m$ **then**

12: **Return:** The point $(\tilde{q}, \tilde{p})$ as the next Markov chain state.

13: **else**

14: **Return:** The point $(q, p)$ as the next Markov chain state.

15: **end if**

Techniques proposed in the previous sections. Algorithm 17 gives an algorithmic description of the numerical method developed in Tao [2016] which was proposed for use in RMHMC by Cobb et al. [2019]. However, this integrator is neither invertible nor volume-preserving in the sense required for compatibility with the Markov chain transition kernel in definition 8.2.32 except insofar as proposition 8.4.1 is concerned (the subsequent remark explains why this result is not satisfactory).

The technique described mathematically in section 8.3.1 and stated explicitly in definition 8.3.12 provides a means of constructing a reversible (and therefore invertible), volume-preserving, and second-order accurate numerical method for Hamilton’s equations of motion (see proposition 8.3.13). According to corollary 8.2.34, this numerical
method may therefore be used in diffeomorphism Monte Carlo in order to produce a
Markov chain satisfying detailed balance. Algorithm 20 implements the reversible and
volume-preserving numerical integrator; the crucial step of this procedure which distin-
guishes it from algorithm 17 is the requirement to identify constraint-satisfying Lagrange
multipliers. In practice, this is accomplished via Newton’s method, for which pseudo-
code is provided in algorithm 18. The parameter $\delta \geq 0$ in algorithm 20 controls how much
computational effort is expended seeking solutions that exactly satisfy the equality mani-
fold constraint in eq. (8.39): as $\delta \to 0$, the equality manifold constraint must be satisfied
to increasingly fine precision. Alternatively, one may employ Broyden’s method (algo-
rithm 19) to the same end as Newton’s method (algorithm 18), which has an advantage in
that Broyden’s method does not require the exact Jacobian of the function whose root is
sought.

The significance of this framework is that it allows us to place the procedure of Cobb
et al. [2019] on a spectrum controlled by the convergence parameter. When a large
convergence tolerance is employed, virtually no effort is spent seeking an exact root of
eq. (8.60); on the other hand, a minuscule convergence tolerance will ensure that the nu-
merical method faithfully respects the manifold constraints and will be exactly reversible
and volume preserving as required by RMHMC. As noted in the abstract, employing New-
ton’s method to identify suitable Lagrange multipliers may be more computationally bur-
densome than simply using the implicit integrators that the method in Cobb et al. [2019]
sought to replace. Therefore, we do not advocate for Lagrange multipliers as a more effi-
cient alternative to implicitly-defined integrators, but instead to contextualize the explicit
integrator in algorithm 17 as lying at an extremum of a spectrum of numerical methods,
which, as $\delta$ approaches zero, more carefully respect the detailed balance conditions. In
other words, the introduction of Lagrange multipliers only serves to make the algorithm
correct and this is achieved at the cost of making it substantially slower.

With these algorithmic descriptions of the numerical integrators in algorithms 17 and 20
available, we now contextualize them in their capacities as transition mechanisms in Markov chain Monte Carlo. The first of these is algorithm [21] which was proposed by [Cobb et al. [2019]. This algorithm produces biased samples due to violations of detailed balance. Moreover, because algorithm [17] cannot be used to define a diffeomorphism from $\mathbb{R}^m \times \mathbb{R}^m$ to itself, the resulting method is not even compatible with the diffeomorphism Monte Carlo transition kernel given in definition [8.2.32]. On the other hand, algorithm [22] is a correct algorithm which is compatible with diffeomorphism Monte Carlo because algorithm [20] is an exactly reversible, volume-preserving numerical method when $\delta = 0$; when $\delta > 0$, we expect the errors in reversibility and volume preservation to be small whenever $\delta$ is also “small.”

8.4 Related Work

Remark. In [Brubaker et al. [2012], the authors establish ergodicity of the Markov chain in addition to the detailed balance condition. In order to prove ergodicity, it is required that the Markov chain satisfies certain consistency properties in the selection of the Lagrange multipliers. These consistency properties guarantee that reversibility cannot be violated by the selection of distinct, constraint-preserving Lagrange multipliers. In practice, there appears to be an assumption that the Lagrange multipliers are consistently selected, but recently [Lelièvre et al. [2019] gave a rigorous procedure for checking the reversibility of integrators using Lagrange multipliers to satisfy constraints.

[Cobb et al. [2019] proposed to use the explicit numerical integrator defined formally in definition [8.3.10] for Riemannian manifold Hamiltonian Monte Carlo; this integrator cannot be used in definition [8.2.32] because it is not reversible or volume-preserving. However, this integrator can be used to construct a Markov chain satisfying detailed balance in an unsatisfactory sense.
**Proposition 8.4.1** ([Brofos and Lederman](2020b)). Let \( \hat{\Phi}_\epsilon \) be as defined in definition 8.3.10 and let \( \Pi \) be as defined in corollary 8.2.34. Let \((q, p) \sim \Pi\) and \((x, y) \sim \Pi\) be independent. Let \((\tilde{q}, \tilde{x}, \tilde{p}, \tilde{y}) = \hat{\Phi}_\gamma(q, x, p, y)\) be the proposal of a Markov chain where \(\gamma \sim \text{Uniform}(\{-\epsilon, +\epsilon\})\) for some \(\epsilon \in \mathbb{R}\). If the proposal is accepted with probability

\[
\min \left\{ 1, \exp \left( H(q, p) + H(x, y) - H(\tilde{q}, \tilde{p}) - H(\tilde{x}, \tilde{y}) \right) \right\},
\]

then the Markov chain satisfies detailed balance with respect to the product distribution whose density on \(\mathbb{R}^{4m}\) is \((q, x, p, y) \mapsto \pi(q, p)\pi(x, y)\).

**Remark.** The proof of this result follows from the fact that the explicit integrator without Lagrange multipliers (definition 8.3.10) is symmetric and symplectic in the expanded \((q, x, p, y)\)-phase space and hence definition 8.2.32 applies. However, the result is not satisfactory because, with the initial condition that \(q = x\) and \(p = y\) as discussed in theorem 8.3.11, the initial \((q, q, p, p) \in T^*E\) cannot be considered an independent sample from the product distribution whose density is \(\pi(q, p)\pi(x, y)\). Moreover, whenever \(q \neq x\) and \(p \neq y\), there is no reason to expect that the Hamiltonian energies appearing in eq. (8.72) will be conserved by the numerical integrator \(\hat{\Phi}_\epsilon\). Therefore, one expects that, in the context of proposition 8.4.1, the acceptance rate of the Markov chain will be low. Both of these facts limit the usefulness of proposition 8.4.1.

**8.5 Experiments**

We now turn our attention to the experimental evaluation of the modified explicit integrator with Lagrange multipliers in algorithm 20 in which we employ Broyden’s method (algorithm 19) to identify constraint-preserving Lagrange multipliers. This material seeks to verify that the introduction of Lagrange multipliers indeed produces a reversible and volume preserving numerical integrator. In this discussion, we consider variable values of
\( \delta \geq 0 \), the convergence tolerance appearing in algorithm \([20]\). In the case where a “large” threshold (e.g. \( \delta = 1 \times 10^0 \)) is used, the resulting method is essentially algorithm \([17]\) since Newton’s method (algorithm \([18]\)) terminates before a single iteration with the Lagrange multipliers left at zero. On the other hand, if a “small” convergence threshold is used (e.g. \( \delta = 1 \times 10^{-12} \)), then algorithm \([20]\) differs meaningfully from algorithm \([17]\) because of the need to carefully satisfy constraints.

We will evaluate the explicit integrator in the context of Monte Carlo; in these cases we will be interested in the extent to which the integrator is reversible and exhibits volume preservation when the Lagrange multipliers are left at zero. We will assess if this version of the numerical integrator is sufficient for Monte Carlo and present evidence in both directions. As points of comparison, we consider the Euclidean HMC with the standard leapfrog integrator and RMHMC with the generalized leapfrog integrator. We begin, however, without the Monte Carlo context and examine the behavior of the explicit integrator with Lagrange multipliers in its role as a numerical method for approximating the solutions to Hamilton’s equations of motion.

### 8.5.1 Illustration of the Explicit Integrator with Lagrange Multipliers

We discuss in fig. 8.1 an example of a single step of the explicit integrator augmented with Lagrange multipliers. To summarize the procedure and give intuition, the first step is to choose Lagrange multipliers augmenting the momenta such that a single step of the explicit integrator without Lagrange multipliers satisfies the equality constraint. The second step of the integrator is to choose more Lagrange multipliers, again augmenting the momenta, such that the hidden constraint is satisfied; this means that the velocities are equal. As discussed in proposition \([8.3.17]\) for sufficiently small step-sizes, one can choose the Lagrange multipliers so that satisfying the hidden constraint can be realized by averaging.
Figure 8.1: We give an illustration of the explicit integration procedure with Lagrange multipliers. The initial condition of the system is set at $(q, q, p, p) = (3, 3, 0, 0)$. The leftmost plot shows the positional initial condition $(3, 3)$ as a red dot and the constraint $q = x$ as a dashed line. The $(\tilde{q}, \tilde{x})$ produced by $(\tilde{q}(\mu), \tilde{x}(\mu), p'(\mu), y'(\mu)) = \hat{\Phi}(0.3, (3, 3, 0+\mu, 0-\mu))$ is plotted as an orange curve for varying $\mu$ where $\hat{\Phi}$ is defined in definition 8.3.12. One of these $\mu$ satisfies the constraint $\tilde{q} = \tilde{x}$, which we denote by $(\tilde{q}_s, \tilde{x}_s)$. The center plot shows the initial momentum $(0, 0)$ as a red dot and, for varying $\mu$ from the leftmost figure, the resulting $(\tilde{p}(\mu), \tilde{p}(\mu))$ as an orange dotted line. As discussed, one of these Lagrange multipliers $\mu$ satisfies the constraint $\tilde{q} = \tilde{x}$ and the momentum $(\tilde{p}_s, \tilde{y}_s)$ for this Lagrange multiplier is shown as an orange dot. The second step of the integrator is to identify Lagrange multipliers $\mu'$ satisfying the hidden constraint. In the center plot, possible $(\tilde{p} + \mu', \tilde{p} - \mu')$ are shown as a blue line, which intersects the condition $p = y$ at the blue dot. As discussed, for small enough step-size, a suitable Lagrange multiplier $\mu'$ may be found so that the momenta preserving the hidden constraint is the average of $\tilde{p}_s$ and $\tilde{y}_s$. The right-most plot shows the hidden constraint of the system that the velocities $\dot{q}$ and $\dot{x}$ must be equal. The blue line in this plot shows $(\nabla_p H(\tilde{x}_s, \tilde{p}_s + \mu'), \nabla_y H(\tilde{q}_s, \tilde{p}_s - \mu'))$ for varying values of $\mu'$. The initial velocity $(\nabla_p H(q, p), \nabla_y H(q, p))$ is shown as a red dot. The orange dot shows $(\nabla_p H(\tilde{q}_s, \tilde{p}_s), \nabla_y H(\tilde{x}_s, \tilde{p}_s))$, where $\tilde{p}_s = (\tilde{p}_s + \tilde{y}_s)/2$.

the momenta.

8.5.2 Non-Separable Hamiltonians

We consider integrating Hamilton’s equations of motion corresponding to the non-separable Hamiltonian,

$$H(q, p) = \frac{(q^2 + 1)(p^2 + 1)}{2},$$

(8.73)
Figure 8.2: The error in the position variable computed by variations of the explicit and generalized leapfrog integrators in the non-separable Hamiltonian.

Figure 8.3: The deviation in Hamiltonian energy computed by variations of the explicit and generalized leapfrog integrators in the non-separable Hamiltonian.

which was proposed as a test case in [Tao, 2016]. The solution of the position variable in this model has an analytical closed-form,

\[ q_t = q_0 \cn \left( t \sqrt{1 + q_0^2} \left| \frac{q_0^2}{1 + q_0^2} \right| \right), \]  

(8.74)

where \( \cn(\cdot, \cdot) \) is a Jacobian elliptic function. We consider the initial conditions \( q_0 = -3 \) and \( p_0 = 0 \). We use a step-size of 0.01 and 1,000 integration steps. We consider the explicit integrator with convergence thresholds for the Lagrange multipliers of \( 1 \times 10^0 \) or \( 1 \times 10^{-12} \). As a comparison, we also implement the generalized leapfrog integrator with these
same convergence tolerances for the fixed point iterations. We concern ourselves with two aspects of the numerical integration: (i) the accuracy of the solution of the position variable and (ii) the energy conservation properties demonstrated by the numerical method. In fig. 8.2 we show the accumulating error in the position variable; we observe that the two versions of the explicit integrator exhibit nearly indistinguishable errors over the course of the trajectory. The generalized leapfrog integrator with the convergence threshold of $1 \times 10^{-12}$ enjoys a somewhat smaller error in the position variable, whereas the same method with a convergence tolerance of $1 \times 10^{0}$ has much more substantial error over the course of the trajectory. In fig. 8.3 we show the deviation in the Hamiltonian energy for each integrator. We note that the explicit integrator produces the lowest deviations in the Hamiltonian; the degree to which the generalized leapfrog method exhibits energy conservation depends on the convergence threshold used: when using a threshold of $1 \times 10^{-12}$, energy conservation is only slightly worse than for the explicit counterparts; on the other hand, a threshold of $1 \times 10^{0}$ yields an energy drift over the course of the trajectory.

8.5.3 Schwarzschild Geodesic

A more interesting example of a non-separable Hamiltonian derives from Einstein’s theory of general relativity. The Schwarzschild geodesics describe the motion of a particle under the influence of a gravitational field. Within this construction, it is assumed that the mass of the particle is insignificant compared to the mass of the object generating the gravitational field. We consider Schwarzschild geodesics in the plane, which is characterized by three variables: time $\tau$, radius $r$, and angle $\phi$. Each of these variables is equipped with an associated momentum $p_{\tau}, p_{r},$ and $p_{\phi}$, respectively. The Hamiltonian for the Schwarzschild geodesic has the form,

$$H(\tau, r, \phi, p_{\tau}, p_{r}, p_{\phi}) = \frac{1}{2} \left[ \left( 1 - \frac{2}{r} \right)^{-1} p_{\tau}^2 - \left( 1 - \frac{2}{r} \right) p_{r}^2 - \frac{p_{\phi}^2}{r^2} \right]. \quad (8.75)$$
Figure 8.4: We examine the energy conservation properties of the explicit integration methods, the implicit midpoint method, and the generalized leapfrog integrator. We consider the initial position \( \tau, r, \phi = (10, 42, 2) \) with initial momenta \( p_\tau, p_r, p_\phi = (98.2, 0, -45.8) \). We consider integration step-sizes \( t \in \{0.001, 0.01, 0.1\} \) and integrate the dynamics to time ten. In every case, there is a characteristic “hump” in the middle of the trajectory which produces a relatively large deviation from the initial Hamiltonian energy. For \( t = 0.1 \), we observe that explicit integration with \( \mu = 0 \) exhibits an energy differential following the hump.

We evaluate the energy conservation properties of numerical integrators for integrating trajectories corresponding to this Hamiltonian in fig. 8.4. We also examine the total number of gradient evaluations required by each method in fig. 8.5.

### 8.5.4 Banana-Shaped Distribution

The banana-shaped distribution represents a pathology of non-identifiability in Bayesian posteriors. Formally, we consider the following generative model,

\[
\theta_1, \theta_2 \sim \text{i.i.d. Normal}(0, \sigma_\theta^2)
\]

\[
y_1, \ldots, y_n \sim \text{i.i.d. Normal}(\theta_1 + \theta_2^2, \sigma_y^2).
\]

The objective is to sample the posterior distribution \( (\theta_1, \theta_2) | (y_1, \ldots, y_n), \sigma_\theta, \sigma_y \). The likelihood function of \( y_1, \ldots, y_n \) given \( (\theta_1, \theta_2) \) is non-identifiable since the level sets of \( \theta_1 + \theta_2^2 \) are also level sets of the likelihood function. In our experiments we set \( n = 100, \sigma_\theta = 2, \) and \( \sigma_y = 2 \); we use \( \theta_1 = 1/2 \) and \( \theta_2 = \sqrt{1/2} \) as the true parameters when generating the data. This distribution poses unique challenges for the generalized leapfrog integra-
Figure 8.5: We compare the trajectories of the numerical integration strategies using an integration step-size of 0.1 for computing the Schwarzschild geodesics. The generalized leapfrog and explicit integration methods are in close agreement, as shown in the left figure. In contrast, the trajectory computed by the implicit midpoint integrator diverges from these methods. In the right figure, we compare the number of gradient evaluations required in each step of the integrators. From this comparison, we observe that the explicit integrator with Lagrange multipliers requires markedly more gradient evaluations in each step. The implicit midpoint integrator tends to require more gradient evaluations than the generalized leapfrog integrator. However, the explicit integrator with $\mu = 0$ requires only a fixed number (8) of integration steps in each iteration, making it more computationally efficient than the competing methods.

Figure 8.6: *Banana distribution detailed balance errors.* The degree to which volume preservation and reversibility are violated for the explicit integrator with a convergence threshold of $1 \times 10^6$ and $1 \times 10^{-12}$ in the banana-shaped distribution. Decreasing the convergence threshold produces a numerical method that more faithfully reflects the reversibility and volume preservation required for detailed balance.
Figure 8.7: Banana distribution Kolmogorov-Smirnov statistics. Boxplots showing the typical region of Kolmogorov-Smirnov statistics among 10,000 randomly selected one-dimensional subspaces for independent samples and Markov chain samples in the banana-shaped distribution. Even though the explicit integrator with a threshold of $1 \times 10^0$ only enjoys a modest degree of reversibility and volume preservation, it nonetheless exhibits the best ergodicity by this measure.

Figure 8.8: Banana distribution effective sample size. The effective sample size of the RMHMC procedures sampling from the banana-shaped distribution. We generate a distribution of effective sample sizes by splitting a long chain of 1,000,000 samples into twenty contiguous sub-chains of length 50,000 and compute the effective sample size within each sub-chain. We observe that the explicit integrator produces samples that possess larger effective sample sizes than any competing method.

tor; as discussed by Luke Bornn and Julien Cornebise, for large step-sizes, the implicit update to momentum in the generalized leapfrog method may fail to have a solution. This necessitates the use of a small step-size in this integrator.

In this distribution we use the generalized leapfrog integrator with twenty-five inte-
gration steps and an integration step-size of 0.04. This produces an average acceptance rate of eighty-seven percent when using a convergence threshold of $1 \times 10^{-5}$; the particular pathologies of this posterior distribution means that a larger step-size produces a marked decline in acceptance rate, since the implicit update of the momentum in the generalized leapfrog integrator does not converge. We also implement (Euclidean) HMC with a constant mass matrix, which does not involve fixed point iterations, using twenty-five integration steps and a step-size of 0.1; this produces an acceptance rate of eighty-five percent. When implementing the explicit integrator, we consider two convergence tolerances for the Lagrange multipliers of $1 \times 10^0$ and $1 \times 10^{-12}$, and a step-size of 0.1. Either convergence tolerance produces an acceptance rate of ninety-nine percent. We attempt to draw 1,000,000 samples from the target posterior.

We first evaluate the reversibility and volume preservation exhibited by the explicit integrator with the weak and strong thresholds. We expect that a smaller threshold for measuring convergence of the Lagrange multipliers should produce a numerical integrator with superior reversibility and volume preservation. We observe that when using a threshold of $1 \times 10^0$, the median error in reversibility and volume preservation are approximately $1 \times 10^{-3}$. When the threshold is decreased to $1 \times 10^{-12}$, the median error in reversibility is $1 \times 10^{-12}$ while the median error in volume preservation is approximately $1 \times 10^{-8}$. These are visualized in fig. 8.6.

The explicit integrator with Lagrange multipliers is only practical in MCMC if the Lagrange multipliers may be fixed at zero. This approximate method is only justified if the errors in reversibility and volume preservation do not substantively impact the quality of samples produced by the Markov chain. We turn, therefore, to the question, “To what extent do these errors impact the quality of samples generated by the Markov chain?” To assess this, we generate one-million independent samples from the banana-shaped distribution using rejection sampling. We then project these independent samples and samples generated by the Markov chain onto random one-dimensional subspaces and measure

461
the Kolmogorov-Smirnov test statistic of these one-dimensional quantities. When these Kolmogorov-Smirnov statistics are small, it reflects a high degree of fidelity between the independent samples and the samples generated by the chain. We consider 10,000 uniformly random one-dimensional subspaces in our experimentation. In fig. 8.7 we observe that despite the weak convergence threshold used in the explicit integrator, it nonetheless enjoys the best ergodicity when measured according to these Kolmogorov-Smirnov statistics along one-dimensional subspaces. We next consider the separate but related question of the effective sample size of the generated samples; the effective sample size measures the degree to which samples generated by the chain are serially auto-correlated as separate from the closeness to which samples from the chain approximate a target distribution. These results are shown in fig. 8.8. We observe that the combination of adapting proposals to the local geometry, its larger step-size, and its larger acceptance rate produces a significantly larger number of effective samples when using the explicit integrator than either of the leapfrog integrators. This combination of results permits us to conclude that the explicit integrator, even with a very weak threshold, produces samples that are closest to the target density, in addition to yielding samples that exhibit the smallest level of auto-correlation.

8.5.5 Neal’s Funnel Distribution

Neal’s funnel distribution imitates a common attribute of hierarchical Bayesian models. According to [Betancourt and Girolami 2013], in the presence of sparse observations, the posterior distribution tends to resemble a funnel: one region of the posterior being highly concentrated in a small region, the other being diffuse over a large area; both regions have similar probability masses. Neal’s funnel distribution mimics this pattern and is defined
Figure 8.9: *Neal’s funnel distribution detailed balance errors.* The degree to which volume preservation and reversibility are violated for the explicit integrator with a convergence threshold of $1 \times 10^0$ in Neal’s funnel distribution.

Figure 8.10: *Neal’s funnel distribution Kolmogorov-Smirnov statistics.* Boxplots showing the typical region of Kolmogorov-Smirnov statistics among 10,000 randomly selected one-dimensional subspaces for independent samples and Markov chain samples in Neal’s funnel distribution. Even though the explicit integrator with a threshold of $1 \times 10^0$ only enjoys a modest degree of reversibility and volume preservation, it is nonetheless competitive with the generalized leapfrog integrator in this distribution.

According to the following generative model:

\[ v \sim \text{Normal}(0, 3^2) \quad (8.78) \]

\[ x_1, \ldots, x_{10} \sim \text{i.i.d. Normal}(0, \exp(-v)). \quad (8.79) \]
Neal’s funnel distribution is extremely challenging for Euclidean HMC with the leapfrog integrator due to severe multi-scale phenomena when entering or exiting the funnel. Indeed, there is no globally-suitable mass preconditioner that enables Euclidean HMC to sample from this distribution effectively. This encourages one to look toward Riemannian HMC with adaptive preconditioners as an alternative.

In our experimentation, we consider Euclidean HMC with a step-size of 0.1 and eight integration steps. For the Riemannian methods we consider an integration step-size of 0.2 and twenty-five integration steps. When employing the generalized leapfrog method, this produces an acceptance rate of ninety-five percent, while the explicit integrator has an acceptance rate in excess of ninety-nine percent. In the generalized leapfrog integration, we use a threshold of \(1 \times 10^{-5}\), while in the explicit integrator we use a weak convergence threshold of \(1 \times 10^0\). The threshold used in the explicit integrator implies that reversibility and volume preservation hold only approximately; we measure the degree to which these properties are violated in fig. 8.9. We observe that the median error in reversibility and volume preservation is \(1 \times 10^{-3}\) and \(1 \times 10^{-4}\).

As in the case of the banana-shaped distribution, we seek to evaluate the ergodicity
of samples produced by the explicit integrator within the context of methods using the leapfrog and generalized leapfrog methods. Neal’s funnel distribution can be sampled analytically; we therefore consider the projection of the Markov chain and analytical samples onto random one-dimensional subspaces and compute the Kolmogorov-Smirnov statistic of these projections. The product of this analysis is shown in fig. 8.10. We observe that, as anticipated, Euclidean HMC produces samples that exhibit the greatest dissimilarity relative to the analytic samples. The Riemannian methods exhibit much better ergodicity by this metric, with the generalized leapfrog and explicit integrator having qualitatively similar performance. In fig. 8.11 we evaluate the effective sample size of the Markov chains with transitions computed via the leapfrog, generalized leapfrog, and explicit integrators. We observe that the Riemannian methods have the largest effective sample sizes.

8.6 Discussion and Conclusion

This research has examined a modification to the explicit integration strategy used in Hamiltonian Monte Carlo, which was proposed in Cobb et al. [2019]. This modification, based on the theory of Lagrange multipliers, produces a numerical integrator that is symmetric and symplectic in a manner that is compatible with HMC. We found that this integrator has good energy conservation properties, often exceeding the implicit midpoint or generalized leapfrog algorithms on integration tasks involving non-separable Hamiltonians. We subsequently considered applying this modified integrator to Hamiltonian Monte Carlo; in this regime, the symmetry and volume-preservation properties of the numerical integrator are of particular importance. We found that, for large step-sizes, the number of gradient evaluations required by the explicit integrator with Lagrange multipliers is far in excess than that required by the implicit methods. Therefore, using these explicit methods in application to HMC may be inadvisable; or, if they are used, one must be cognizant of the differences in symmetry and volume-preservation between them and competing ge-
ometric integrators. On the other hand, the present work has developed a mathematical theory which provides the explicit integrator (with Lagrange multipliers) with provable symmetry and symplecticness. Therefore, to the extent that the explicit integrator with Lagrange multipliers set to zero complies with the theoretical integrator, one may find the violation of symmetry and symplecticness to be small enough that this variant of the explicit integrator is acceptable for use.
Appendix 8.A

Appendices to Chapter 8

8.A.1 Proofs

8.A.1.1 Proof of Proposition 8.2.3

Proof. We need only verify that $f(v)$ has a Jacobian of full-rank. But this is immediate from $\nabla f(v) = \nabla g(q)$, which has full-rank by definition.

8.A.1.2 Proof of Lemma 8.2.4

Proof. By differentiating the condition $g(q_t) = 0$ we obtain,

$$\frac{d}{dt} g(q_t) = 0 \quad (8.2.1)$$

$$\nabla g(q_t) \frac{d}{dt} q_t = 0. \quad (8.2.2)$$

Hence $\frac{d}{dt} q_t \in T_{q_t} M$. 

□
8.A.1.3 Proof of Proposition 8.2.6

Proof. As in section 8.A.1.2 one need only verify that the Jacobian of $h$ has full-rank. The Jacobian is,

$$ \nabla h(q, v) = \begin{pmatrix} \nabla g(q) & 0 \\ \nabla^2 g(q)[v] & \nabla g(q) \end{pmatrix}. \quad (8.A.3) $$

Using the fact that the rank of a lower-triangular matrix is at least the sum of the ranks of the diagonal blocks, this follows since

$$ \min \{2m, 2k\} \geq \rank(\nabla h(q, v)) \quad (8.A.4) $$

$$ \geq \rank(\nabla g(q)) + \rank(\nabla g(q)) \quad (8.A.5) $$

$$ = 2 \min \{k, m\} \quad (8.A.6) $$

$$ = \min \{2k, 2m\}. \quad (8.A.7) $$

Hence, $\nabla h(q, v)$ has full-rank. \hfill \Box

8.A.1.4 Proof of Lemma 8.2.19

Proof. This follows from differentiating the constraint of the motion:

$$ \frac{d}{dt}g(q_t) = 0 \quad (8.A.8) $$

$$ \implies \nabla g(q_t)\dot{q}_t = 0 \quad (8.A.9) $$

$$ \implies \nabla g(q_t)\nabla_p H(q_t, p_t) = 0. \quad (8.A.10) $$

\hfill \Box
8.A.1.5 Proof of Lemma 8.2.20

Proof. In the proof we will suppress the dependency of \(q_t\) and \(p_t\) on the time variable and simply write \(q\) and \(p\), respectively. By differentiating the constraint of the motion twice with respect to time one obtains

\[
\frac{d^2}{dt^2} g(q) = \frac{d}{dt} [\nabla g(q) \nabla_p H(q, p)] = 0 \tag{8.A.11}
\]

Applying the product rule of differentiation one finds that eq. (8.A.11) is equivalent to,

\[
0 = [\nabla_q \nabla g(q) \cdot \nabla_p H(q, p)] \nabla_p H(q, p) + \nabla g(q) [\nabla_q \nabla_p H(q, p) \nabla_p H(q, p)]
\]

\[
- \nabla^2_p H(q, p) [\nabla_q H(q, p) + \nabla g(q)^\top \lambda(q, p)] \tag{8.A.12}
\]

One can rearrange terms and solve for \(\lambda(q, p)\) provided that \(\nabla g(q) \nabla^2_p H(q, p) \nabla g(q)^\top\) is invertible.

8.A.1.6 Proof of Theorem 8.2.30

Proof. Volume preservation follows from lemma 2.3.52. Symmetry follows from the fact that \(\hat{\Phi}\) is symmetric by assumption and from lemma 2.3.53. The fact that such integrators are first-order accurate is established in lemma 2.3.54. From the symmetry of the integrator, second-order accuracy follows from theorem 2.2.23.

8.A.1.7 Proof of Lemma 8.3.2

Proof. By direct calculation,

\[
\nabla e(\alpha) = \begin{pmatrix} \nabla_q e(q, x) \\ \nabla_x e(q, x) \end{pmatrix} = (\text{Id, } -\text{Id}) \in \mathbb{R}^{m \times 2m}. \tag{8.A.13}
\]
8.A.1.8 Proof of Proposition 8.3.4

Proof. This is a consequence of lemma 8.2.19 applied to $E$. In particular,

\[ \nabla e(\alpha_t)\nabla S(\alpha_t, \beta_t) = 0 \]  \hspace{1cm} (8.A.14)

\[ \Rightarrow \nabla_p S(q_t, x_t, p_t, y_t) = \nabla_y S(q_t, x_t, p_t, y_t). \]  \hspace{1cm} (8.A.15)

However, since $q_t = x_t$ under the equality manifold constraint, eq. (8.A.15) simplifies further to eq. (8.29) as claimed.

8.A.1.9 Proof of Lemma 8.3.6

Proof. This is an immediate application of the definitions from definition 8.2.18 and computing gradients of $\bar{H}_\omega$.

8.A.1.10 Proof of Corollary 8.3.7

Proof. Applying eq. (8.29) to the $\omega$-expanded Hamiltonian and using eq. (8.30) yields,

\[ \nabla_p \bar{H}_\omega(q_t, q_t, p_t, y_t) = \nabla_y H(q_t, q_t, p_t, y_t) \]  \hspace{1cm} (8.A.16)

\[ \Rightarrow \nabla_p H(q_t, p_t) + \omega(p_t - y_t) = \nabla_y H(q_t, y_t) - \omega(p_t - y_t). \]  \hspace{1cm} (8.A.17)
8.A.1.11 Proof of Proposition 8.3.8

Proof. By differentiating condition $q_t = x_t$ from eq. (8.28) with respect to time we obtain $\dot{q}_t = \dot{x}_t$. Recall the relationships deduced in eqs. (8.32) and (8.34):

$$\dot{q} = \nabla_p \tilde{H}_\omega(q, x, p, y) = \nabla_p H(x, p) + \omega(p - y) \quad (8.20)$$

$$\dot{x} = \nabla_y \tilde{H}_\omega(q, x, p, y) = \nabla_y H(q, y) + \omega(y - p). \quad (8.21)$$

We know that $\dot{q} = \dot{x}$ will imply that $p = y$ locally if there is a local bijection between the velocity and the momentum. For fixed $q$ (resp. $x$), by the inverse function theorem, eq. (8.20) (resp. eq. (8.21)) is a local bijection of $\dot{q}$ and $p$ (resp. $\dot{x}$ and $y$) if $\nabla^2 H(x, p) + \omega \text{Id}$ (resp. $\nabla^2 y H(q, y) + \omega \text{Id}$) is invertible at $p$ (resp. $y$). \hfill \Box

8.A.1.12 Proof of Theorem 8.3.9

Proof. Substituting $q_t$ and $p_t$ into eqs. (8.31), (8.33), (8.35), (8.37) and (8.39) yields the following system:

$$\dot{q}_t = \nabla_p H(q_t, p_t) \quad (8.22)$$

$$\dot{q}_t = \nabla_y H(q_t, p_t) \quad (8.23)$$

$$\dot{p}_t = -\nabla q H(q_t, p_t) + \lambda \quad (8.24)$$

$$\dot{p}_t = -\nabla x H(q_t, p_t) - \lambda \quad (8.25)$$

The only feasible choice is $\lambda = 0$ in which case the system reduces to,

$$\dot{q}_t = \nabla_p H(q_t, p_t) \quad (8.26)$$

$$\dot{p}_t = -\nabla q H(q_t, p_t) \quad (8.27)$$

471
This is precisely the system from eqs. (8.45) and (8.46). The constraint $e(q_t, x_t)$ is clearly satisfied. By existence and uniqueness of differential equations, this is the only solution.

8.A.1.13 Proof of Proposition 8.3.13

Proof. This follows as a direct consequence of theorem 8.2.30 using the fact that the Jacobian of the constraint function derived in lemma 8.3.2 has full-rank, the fact that definition 8.3.10 is 2nd-order accurate for the unconstrained system as discussed in theorem 8.3.11, and the construction of $\hat{\Psi}$.

8.A.1.14 Proof of Proposition 8.3.14

Proof. Under the constraints we have $q = x$ and $p = y$ by assumption. Thus,

\[
d\alpha \wedge d\beta = d\begin{pmatrix} q \\ x \end{pmatrix} \wedge d\begin{pmatrix} p \\ y \end{pmatrix} = d\begin{pmatrix} q \\ q \end{pmatrix} \wedge d\begin{pmatrix} p \\ p \end{pmatrix} = \sum_{i=1}^{m} dq_i \wedge dp_i + \sum_{i=1}^{m} dq_i \wedge dp_i = 2 \sum_{i=1}^{m} dq_i \wedge dp_i = 2 dq \wedge dp.
\] (8.A.28)
8.A.1.15 Proof of Corollary 8.3.15

Proof. By direct calculation,

\[
dq \land dp = d \left( \frac{q + \tilde{x}}{2} \right) \land d \left( \frac{p + \tilde{y}}{2} \right) = \frac{1}{4} dq \land (q + x) \land (p + y) = \frac{1}{4} (dq \land dp + dq \land d\tilde{y} + d\tilde{x} \land dp + d\tilde{x} \land d\tilde{y})
\]

(8.A.33)

\[
= \frac{1}{4} d\tilde{q} \land d\tilde{p} + d\tilde{q} \land d\tilde{p} + d\tilde{q} \land d\tilde{p} + d\tilde{q} \land d\tilde{p}
\]

(8.A.34)

\[
= \frac{1}{4} \left( d\alpha \land d\beta + d\alpha \land d\beta \right) = \frac{1}{2} d\alpha \land d\beta
\]

(8.A.35)

\[
= \frac{1}{2} dq \land dp
\]

(8.A.36)

\[
\]

\[
\]

8.A.1.16 Proof of Lemma 8.3.16

Proof. Observe that the map \((q, p) \mapsto (\tilde{q}, \tilde{p})\) can be expressed as the following composition of functions:

\[
(q, p) \mapsto (\alpha, \beta) \mapsto (\tilde{\alpha}, \tilde{\beta}) = \hat{\Psi}_{\epsilon}(\alpha, \beta) \mapsto (\tilde{q}, \tilde{p})
\]

(8.A.39)

Since \((\alpha, \beta), (\tilde{\alpha}, \tilde{\beta}) \in T^*E\) we have \(\alpha = (q, q), \beta = (p, p), \tilde{\alpha} = (\tilde{q}, \tilde{q}),\) and \(\tilde{\beta} = (\tilde{p}, \tilde{p})\).

Hence, in fact, we have \(\tilde{q} = \tilde{q}\) and \(\tilde{p} = \tilde{p}\) when \(T^*E\) is viewed as an embedded submanifold of \(\mathbb{R}^{2m}\). Invertibility then follows as an immediate consequence of the fact that of \(\hat{\Psi}_{-\epsilon} = \hat{\Psi}_{\epsilon}^{-1}\) from proposition 8.3.13 and definition 8.2.24.
8.A.1.17 Proof of Proposition 8.3.17

Proof. When \( \epsilon = 0 \), we know there is a neighborhood of \( 0 \in \mathbb{R}^m \) such that there is a unique Lagrange multiplier satisfying eq. (8.52) from proposition 8.2.31. By inspection, 0 is the unique Lagrange multiplier. Consider the map \( \mu' \equiv \mu'(\hat{\mu}, \hat{y}) = \frac{\hat{y} - \hat{\mu}}{2} \), which equals zero when \( \epsilon = 0 \). By smoothness, we may choose \( \epsilon \) small enough such that \( \mu' \) lies in the neighborhood of zero in which the Lagrange multiplier is uniquely defined. Furthermore, this choice of \( \mu' \) yields

\[
\hat{\beta} = \begin{pmatrix} \hat{p} \\ \hat{y} \end{pmatrix} + \begin{pmatrix} \frac{\hat{y} - \hat{\mu}}{2} \\ -\frac{\hat{y} - \hat{\mu}}{2} \end{pmatrix} = \begin{pmatrix} \frac{\hat{p} + \hat{y}}{2} \\ \frac{\hat{p} + \hat{y}}{2} \end{pmatrix},
\]

which always satisfies the hidden constraint as discussed in corollary 8.3.7. By identification, then, this construction of \( \mu' \) is the unique Lagrange multiplier satisfying eq. (8.23). \( \square \)
Chapter 9

Conclusion

This thesis has carefully examined the central role of the numerical integrator in geometric methods of Markov chain Monte Carlo. In chapter 4, we saw how the choice of numerical integrator can produce a Markov chain with fundamentally distinct properties, such as a higher acceptance rate, superior stability, and computational efficiency. Chapter 5 gave special attention to one aspect of the implementation of numerical integrators that can cause errors in detailed balance above the level of machine precision; we proposed methods by which to make the Markov chain less sensitive to these errors. In chapter 6, we examined the integrator for Lagrangian mechanics, clarified a misconception regarding its error rate, and suggested one mechanism by which its per-step computational burden could be reduced. Chapter 7 examined some conditions under which Riemannian manifold Hamiltonian Monte Carlo or Lagrangian Monte Carlo would be geometrically ergodic, and proposed a modification to these two geometric Markov chain methods that enables them to inherit a geometric ergodicity theory from the manifold Metropolis-adjusted Langevin algorithm. In chapter 8, we saw an example of an integrator that does not satisfy detailed balance with respect to the target density; moreover, we saw how this integrator could be modified using techniques from embedded manifolds and Lagrange multipliers in order to produce a numerical method for which detailed balance does hold.
I wish to conclude this thesis by speculating about topics that could be potentially interesting directions for further research.

**Choice of Metric in RMHMC** Traditionally, RMHMC has been implemented using either the Fisher information metric or the SoftAbs metric. However, more broadly, one has a degree of freedom in choosing the Riemannian metric tensor and these two choices, while common, are not the only ones. I speculate that other choices of metric could have interesting properties, potentially affecting the ergodicity of the resulting Markov chain. *Do other Riemannian metrics, such as the Wasserstein metric, produce geometric Markov chains with any desirable properties relative to the standard choices?*

**Extensions of the SoftAbs Metric** One disadvantage of the integrator employed in Lagrangian Monte Carlo is that it does not appear to be immediately compatible with the SoftAbs metric. This is unfortunate since the SoftAbs metric gives a general purpose method by which to produce a Riemannian metric. Indeed, [Betancourt] [2012] goes to some lengths in order to derive an integration procedure for which employing the SoftAbs metric does not incur greater than cubic computational complexity, noting that a naive implementation would have quartic complexity. *Can one deduce an implementation of the Lagrangian integrator for which Lagrangian Monte Carlo has cubic computational complexity?*

**The Order of Integration in Geometric Monte Carlo** The framework of diffeomorphism Monte Carlo suggests that any smooth function that is invertible can be employed in Markov chain Monte Carlo. In the context of RMHMC and LMC, the integrators employed are second-order. However, one could reduce the order of integration so as to obtain a more computationally expedient method, or increase the order of integration in order to obtain higher acceptance probabilities. *Can one characterize when (or why) second order integration is the best choice, particularly in RMHMC where increasing the order of*
integration also elevates the computational burden?

**Geometric Ergodicity of Geometric Markov Chains**  While it is possible to establish that there exists a geometric rate of convergence for Hamiltonian Monte Carlo under certain conditions, it seems to be more difficult to produce a theory that can accurately compare Euclidean HMC against geometric methods such as RMHMC or LMC. Such analysis may then suggest properties of the Riemannian metric that make it beneficial from an ergodicity perspective, pursuant to the first proposed research direction. One possible mechanism for analysis is to analyze the Gaussian case under the implicit midpoint integrator, which eliminates the accept-reject decision since quadratic Hamiltonians are perfectly conserved. *Can one obtain a theoretical understanding of conditions under which geometric Markov chain methods produce faster convergence toward the target distribution than their non-geometric counterparts?*
Appendix F

The Shadow Hamiltonian of the Implicit Midpoint Integrator

This chapter investigates the shadow Hamiltonian of the implicit midpoint integrator applied to non-canonical Hamiltonian mechanics. As I was working through these calculations, I devised a related quantity, which I call the “holomorphic semi-shadow.” Composition and experimentation are original to me.

Abstract. Shadow Hamiltonians are functions emerging from the backward analysis of numerical integrators that are more precisely conserved along trajectories of the integrator than the original Hamiltonian. This work derives the shadow Hamiltonian of the implicit midpoint integrator in the presence of non-canonical Poisson structures. We then discuss the computation of the shadow Hamiltonian and propose a novel object, the “holomorphic semi-shadow,” which is a fourth-order perturbation of the shadow Hamiltonian (and therefore also closely conserved) with the additional property that it is particularly basic to compute.
F.1 Introduction

Hamiltonian dynamics is a foundational concept in classical mechanics based on the notion of energy [Marsden and Ratiu, 2010]. Hamiltonian dynamics begins by specifying a function \( H : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R} \) called the Hamiltonian which is said to represent the total energy of a system. The two arguments to the Hamiltonian, traditionally denoted \( q \in \mathbb{R}^m \) and \( p \in \mathbb{R}^m \), are called the position and momentum variables, respectively. Taken together, the variable \( z = (q, p) \in \mathbb{R}^{2m} \) is called a position in phase-space. Therefore, the quantity \( H(z) \equiv H(q, p) \) is assumed to represent the total energy in the system when its position is \( q \) and its momentum is \( p \); equivalently, \( H(z) \) is the energy of the system whose phase-space position is \( z \). Hamilton’s equations of motion are then defined by the relationship,

\[
\dot{q} = \nabla_p H(q, p) \quad \text{(F.1)}
\]

\[
\dot{p} = -\nabla_q H(q, p). \quad \text{(F.2)}
\]

Equivalently, in phase-space,

\[
\dot{z} = \begin{pmatrix}
0 & \text{Id} \\
-\text{Id} & 0
\end{pmatrix}
\begin{pmatrix}
\nabla H(z)
\end{pmatrix}
\quad \text{(F.3)}
\]

The matrix \( J_{\text{can}} \in \mathbb{R}^{2m \times 2m} \) appearing in eq. (F.3) is called the canonical symplectic structure. Hamiltonian dynamics have been leveraged to study many natural phenomena such as fluid dynamics [Salmon, 1988], charged particles in a magnetic field [Branson, 2003], quantum mechanics [Binney and Skinner, 2008], the \( n \)-body problem [Meyer and Hall, 2013], the propagation of light [Buchdahl, 1993], rigid body dynamics [Krishnaprasad and Marsden, 1987], and plasma physics [Marsden and Weinstein, 1982], among other important
applications. Within mathematics in particular, Hamiltonian dynamics have important re-
lations with syplectic geometry, Riemannian manifolds and geodesic motion, and Poisson
algebras.

Many interesting problems in Hamiltonian mechanics cannot be solved analytically. This has necessitated the development of numerical integrators that are suitable for approx-
imating Hamilton’s equations of motion; see [Leimkuhler and Reich, 2005], [Hairer et al.,
2006] for detailed discussions. The development of these integrators has focused on the
preservation of important geometrical properties of Hamiltonian mechanics: (i) reversibil-
ity in time (a consequence of the flow property of differential equations; see [Lee, 2003]),
(ii) preservation of the symplectic structure, which implies the conservation of volume
in phase-space (conservation of phase-space volume is also a consequence of Liouville’s
theorem; see [Gibbs, 2010]), and, to the extent possible, (iii) preservation of conserved
quantities of the system (such as the angular momentum in the \( n \)-body problem).

The focus of this work is the implicit midpoint method, a time-reversible and volume-
preserving numerical integrator that is suitable for arbitrary smooth Hamiltonians. The
implicit midpoint integrator also preserves a (relatively) broad class of conserved quan-
tities. However, the most fundamental conserved quantity of Hamiltonian mechanics is
the Hamiltonian energy itself. In all but the most trivial cases, numerical integrators do
not preserve the Hamiltonian energy. Does there exist a function, related to the Hamilto-
nian, which is conserved exactly by the numerical integrator? This question leads to the
concept of the shadow Hamiltonian. A contribution of this work is to identify the shadow
Hamiltonian of the implicit midpoint integrator; our discussion will apply to Hamiltonian
dynamics with the canonical symplectic structure appearing in eq. (F.3) and to a more
general class of dynamics called \textit{non-canonical} Hamiltonian dynamics.

Once the shadow Hamiltonian has been identified, it may be desirable to compute it.
We will discuss a few procedures evaluating the fourth-order truncation of the shadow
Hamiltonian, including methods from automatic differentiation. For certain applications,
however, it may not be the shadow Hamiltonian itself that is of primary interest; rather, it is the property that the shadow Hamiltonian is more precisely conserved by the numerical integrator that is of paramount importance; an example of such an application is shadow Hamiltonian Monte Carlo [Izaguirre and Hampton 2004a]. Based on this observation, another contribution of this work is to discuss the holomorphic extension of the Hamiltonian into the complex domain; this will yield a simple formula to compute a fourth-order perturbation of the shadow Hamiltonian that will also be highly conserved by the integrator. Remarkably, this formula does not require anything more than the capacity to compute the Hamiltonian vector field (necessary for the implicit midpoint integrator) and a holomorphic extension of the Hamiltonian itself.

F.2 Mathematical Background

Let \( q, p \in \mathbb{R}^m \) denote the position and momentum variables of a mechanical system. Recall that \( z = (q, p) \in \mathbb{R}^{2m} \) is called a position in phase-space. We begin by discussing the kinds of Hamiltonian mechanics that we will consider in this work. We first require a suitable definition of a Hamiltonian energy function.

**Definition F.2.1.** Let \( C^2(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}) \) denote the set of twice-differentiable functions from \( \mathbb{R}^m \times \mathbb{R}^m \) to \( \mathbb{R} \). A function \( H \in C^2(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}) \) such that \( H(q, p) \) is the total energy of a mechanical system whose position is \( q \) and momentum is \( p \) is called a Hamiltonian.

For simplicity, we will often identify \( H(z) \) and \( H(q, p) \). Given a Hamiltonian, in order to construct Hamiltonian dynamics we need only specify an additional object called the symplectic structure.

**Definition F.2.2.** Let \( \text{Skew}(n) \) represent the set of skew-symmetric \( n \times n \) matrices with real-valued entries. A matrix \( \mathbb{J} \in \text{Skew}(2m) \) with the additional property that \( \mathbb{J} \) is invertible is called a symplectic structure associated to the phase space \( \mathbb{R}^{2m} \).
Notice that the matrix $J_{\text{can}}$ defined in eq. (F.3) is an example of a symplectic structure on $\mathbb{R}^{2m}$. The symplectic structure and the Hamiltonian energy are combined in order to yield Hamiltonian dynamics according to the following definition.

**Definition F.2.3.** Let $H \in C^2(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R})$ be a Hamiltonian and let $J$ be a symplectic structure associated to the phase space $\mathbb{R}^{2m}$. Then the Hamiltonian dynamics of $H$ and $J$ is the mapping $X_J^H : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ satisfying,

$$
(X_J^H(z))^\top Jv = \nabla H(z)^\top v.
$$

**(F.4)**

for all $v \in \mathbb{R}^{2m}$. Given such a map $X_J^H$, Hamilton’s equations of motion are defined to be

$$
\dot{z} \overset{\text{def}}{=} X_J^H(z) \text{ in which case } X_J^H \text{ is called the Hamiltonian vector field.}
$$

A little bit of linear algebraic manipulations reveals the formula

$$
X_J^H(z) = (-J)^{-1}\nabla H(z).
$$

**(F.5)**

A quick verification that $(-J_{\text{can}})^{-1} = J_{\text{can}}$ shows how the dynamics presented in eq. (F.3) with the canonical symplectic structure follow from the general case in eq. (F.5). Cases where $J \neq J_{\text{can}}$ can correspond to, for instance, motion in a magnetic field or equipping particles with a non-identity mass matrix [Tripuraneni et al. 2017]. In the following we set $B \overset{\text{def}}{=} (-J)^{-1}$ so that $X_J^H(z) = B\nabla H(z)$.

Given a Hamiltonian system with dynamics governed by $\dot{z} = X_J^H(z)$, we may turn our attention to the problem of numerically integrating the dynamics. As described in section F.1, it is desirable for numerical integrators to satisfy three properties reflective of the geometric properties of the differential equation they are integrating. In the following, we formally define these three properties; we begin with the definition of a numerical integrator.

**Definition F.2.4.** Given a Hamiltonian vector field $X_J^H$, a consistent numerical integrator
for $X^j_H$ is a map $\Phi : \mathbb{R} \times \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m}$ (often written as $\Phi(h, z) \equiv \Phi_h(z)$) satisfying

$$\frac{\partial}{\partial h} \Phi_h(z) \lvert_{h=0} = X^j_H(z). \quad (F.6)$$

We now define the important properties of a geometric numerical integrator.

**Definition F.2.5.** A numerical integrator $\Phi_h(z)$ is called time-reversible if $\Phi_h(\Phi_{-h}(z)) = z$.

**Definition F.2.6.** A numerical integrator $\Phi_h(z)$ is called symplectic with respect to the symplectic structure $J$ if

$$\left(\nabla_z \Phi_h(z)\right)^\top J(\nabla_z \Phi_h(z)) = J. \quad (F.7)$$

To see that symplectic numerical integrators preserve volume, compute the determinant of both sides of eq. (F.7) and conclude that $|\det(\nabla_z \Phi_h(z))| = 1$.

To the author’s knowledge, the only numerical integrator satisfying definitions [F.2.5] and [F.2.6] for arbitrary symplectic structures is the implicit midpoint integrator. We now give a formal definition of the implicit midpoint integrator.

**Definition F.2.7.** The implicit midpoint integrator $\hat{z}_h = \Phi_h(z)$ is defined implicitly by the relation

$$\hat{z}_h \overset{\text{def.}}{=} z + hX^j_H \left( \frac{\hat{z}_h + z}{2} \right). \quad (F.8)$$

Because $\hat{z}_h$ appears on both the left- and right-hand sides of eq. (F.8), practical implementations of the implicit midpoint method require fixed point iterations or (quasi-)Newton methods in order to compute the implicit midpoint method’s update.

Another numerical integrator that will appear later in our discussion is the leapfrog integrator. The leapfrog integrator applies when the Hamiltonian is separable. A separable
Hamiltonian is a Hamiltonian that can be written as the sum of a potential energy that depends only on position and a kinetic energy that depends only on momentum; that is, the Hamiltonian can be written as $H(q, p) = U(q) + K(p)$. In this circumstance, the leapfrog integrator with step-size $h$ and initial condition $(q, p)$ is defined by the following series of calculations:

$$\bar{p}_h = p - \frac{h}{2}\nabla U(q)$$  \hspace{1cm} (F.9)

$$\hat{q}_h = q + h\nabla K(\bar{p}_h)$$  \hspace{1cm} (F.10)

$$\hat{p}_h = \bar{p}_h - \frac{h}{2}\nabla U(\hat{q}_h).$$  \hspace{1cm} (F.11)

The position in phase-space $(\hat{q}_h, \hat{p}_h)$ is the output of the leapfrog integrator. The Störmer-Verlet method (or generalized leapfrog integrator) is an implicitly-defined modification to the leapfrog method that allows it to be used with non-separable Hamiltonians [Verlet 1967, Leimkuhler and Reich 2005].

Numerical integrators are consistent with the underlying Hamiltonian vector field in the sense of eq. (F.6). However, numerical integrators are not exact solutions to the underlying differential equation they are attempting to integrate. However, a numerical integrator may be the exact solution to some differential equation. The differential equation that is exactly integrated by numerical integrator is called the shadow (or modified) differential equation for the integrator. Note that the shadow differential equation depends on the integrator and that, in general, different integrators will have distinct shadow differential equations. The concept of a shadow differential equation is formalized in definition F.2.8.

(To simplify the notation, we have suppressed the dependency of the various vector fields on the symplectic structure.)

Definition F.2.8. Let $\Phi_h : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ be a numerical integrator applied to a Hamiltonian vector field $X_H : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$. We want to identify functions $X_H^{[2]}, X_H^{[3]}, \ldots$ with $X_H^{[1]}$: 484
\[ \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m} \] such that the differential equation,

\[
\frac{d}{dt} \tilde{z}(t) = \tilde{X}_H(\tilde{z}(t)) \tag{F.12}
\]

\[
\tilde{X}_H(\tilde{z}) \overset{\text{def.}}{=} X_H(\tilde{z}) + hX^{[2]}_H(\tilde{z}) + h^2X^{[3]}_H(\tilde{z}) + \ldots \tag{F.13}
\]

has solution

\[
\tilde{z}(n h) = \Phi_h \circ \ldots \circ \Phi_h(\tilde{z}_0) \tag{n times} \tag{F.14}
\]

given the initial condition \( \tilde{z}(0) = \tilde{z}_0 \).

Often such an infinite series expansion does not exist but careful study of an appropriate truncation of eq. (F.13) has proved useful to numerical analysts. A thorough introduction to the shadow differential equations can be found in [Hairer et al. 2006]; for now, it will suffice to recall some of the most important results.

**Theorem F.2.9** (Theorem IX.2.2 [Hairer et al. 2006]). Let \( \Phi_h : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m} \) be a symmetric numerical integrator. Then the coefficient functions \( X^{[i]}_H \) of the odd powers of \( h \) (equivalently, \( i \) is even) appearing in eq. (F.13) are all zero.

This greatly simplifies the calculation of the shadow Hamiltonian for symmetric numerical integrators. Indeed, if one has \( X^{[3]}_H \) then the truncation formed by the first two non-zero terms of eq. (F.13) give a formula for the shadow Hamiltonian accurate to fourth-order in \( h \). This will be our approach in section F.3. In the meantime, a general formula exists for the coefficient function \( X^{[3]}_H \) for an arbitrary numerical integrator.

**Proposition F.2.10** (Equation IX.1.4 [Hairer et al. 2006]). Let \( \Phi_h : \mathbb{R}^{2m} \rightarrow \mathbb{R}^{2m} \) be a consistent and symmetric numerical integrator. Assume that a Taylor series expansion of
Φ can be found as

\[ \Phi_h(z) = z + hX_H(z) + \frac{h^2}{2}d_2(z) + \frac{h^3}{6}d_3(z) + \ldots \]  

(F.15)

Then

\[ X^{[3]}_H(z) = \frac{1}{6} \left( d_3(z) - \frac{\partial^3}{\partial h^3} z(h) \right) \bigg|_{h=0}. \]  

(F.16)

where \( z(h) \) is the solution to \( \dot{z} = X_H(z) \) with initial condition \( z(0) = \tilde{z} \).

As we are discussing Hamiltonian dynamical systems, even more can be said for the structure of the shadow differential equation; in particular, it is Hamiltonian itself in the following sense.

**Theorem F.2.11** (Theorem IX.3.5 Hairer et al. [2006]). Let \( \Phi_h : \mathbb{R}^{2m} \to \mathbb{R}^{2m} \) be a symplectic integrator for a Hamiltonian vector field \( X_H(z) = (-J)^{-1}(h) \nabla H(z) \) that is at least first-order accurate. Then the shadow differential equation in eq. (F.13) is of the form,

\[ \tilde{X}_H(z) = \mathbb{B} \nabla \left( H(z) + hH_2(z) + h^2H_3(z) + \ldots \right) \]  

(F.17)

for functions \( H_2, H_3, \ldots \) with \( H_i : \mathbb{R}^{2m} \to \mathbb{R} \).

**Definition F.2.12.** The function \( \tilde{H}(\tilde{z}) \) \( \text{def} \) \( H(\tilde{z}) + hH_2(\tilde{z}) + h^2H_3(\tilde{z}) + \ldots \) is called the shadow Hamiltonian of \( \Phi_h \).

Notice that one could consider truncations of \( \tilde{H}(\tilde{z}) \) to a finite number of terms to obtain the shadow Hamiltonian and shadow differential equation to arbitrary order in \( h \). For our purposes, it will suffice to consider the fourth-order truncation shadow Hamiltonian, which is the subject of section F.3.

---

\(^1\)Theorem IX.3.5 is formulated for objects called “Poisson integrators.” Symplectic integrators with constant symplectic structures (what we consider here) are special kinds of Poisson integrators for “Poisson structures” of the form \( \mathbb{B} = (-J)^{-1} \).
F.3 Identifying the Fourth-Order Shadow Hamiltonian

We are now in a position to identify the shadow differential equation and shadow Hamiltonian to fourth-order in $h$ using the theory presented in section F.2. Lemmas F.3.1 and F.3.2 give the components necessary to find the shadow differential equation using proposition F.2.10.

**Lemma F.3.1.**

$$\left. \frac{\partial^3}{\partial h^3} z(h) \right|_{h=0} = \mathbb{B} \nabla^3 H(\tilde{z}) (\mathbb{B} \nabla H(\tilde{z}), \mathbb{B} \nabla H(\tilde{z})) + \mathbb{B} \nabla^2 H(\tilde{z}) \mathbb{B} \nabla^2 H(\tilde{z}) \mathbb{B} \nabla H(\tilde{z})$$

(A proof is given in section F.7)

**Lemma F.3.2.**

$$\left. \frac{\partial^3}{\partial h^3} \Phi_h(\tilde{z}) \right|_{h=0} = \frac{3}{4} \mathbb{B} \nabla^3 H(\tilde{z}) (\mathbb{B} \nabla H(\tilde{z}), \mathbb{B} \nabla H(\tilde{z})) + \frac{3}{2} \mathbb{B} \nabla^2 H(\tilde{z}) \mathbb{B} \nabla^2 H(\tilde{z}) \mathbb{B} \nabla H(\tilde{z})$$

(A proof is given in section F.8) Combining lemmas F.3.1 and F.3.2 and we have immediately proved, using proposition F.2.10 the following.

**Proposition F.3.3.** The modified differential equation in eq. (F.13) for the implicit midpoint integrator is, to fourth order,

$$\tilde{X}_H(\tilde{z}) = \mathbb{B} \nabla H(\tilde{z}) + \frac{h^2}{6} \left( -\frac{1}{4} \mathbb{B} \nabla^3 H(\tilde{z}) (\mathbb{B} \nabla H(\tilde{z}), \mathbb{B} \nabla H(\tilde{z})) + \frac{1}{2} \mathbb{B} \nabla^2 H(\tilde{z}) \mathbb{B} \nabla^2 H(\tilde{z}) \mathbb{B} \nabla H(\tilde{z}) \right) + O(h^4)$$

Having identified the fourth-order truncation of the implicit midpoint integrator’s shadow
differential equation, we may, by identification, obtain the fourth-order truncation of the corresponding shadow Hamiltonian.

**Theorem F.3.4.**

\[ \tilde{X}_H(\tilde{z}) = B \nabla \left( H(\tilde{z}) - \frac{h^2}{24} X_H(\tilde{z})^\top \nabla^2 H(\tilde{z}) X_H(\tilde{z}) \right) + O(h^4). \] (F.21)

**Corollary F.3.5.** The function,

\[ \tilde{H}^{[4]}(\tilde{z}) = H(\tilde{z}) - \frac{h^2}{24} X_H(\tilde{z})^\top \nabla^2 H(\tilde{z}) X_H(\tilde{z}) \] (F.22)

is the fourth-order truncation of the shadow Hamiltonian for the implicit midpoint integrator.

A proof is given in section F.9. The fourth-order truncation of the shadow Hamiltonian for the implicit midpoint integrator is a second-order (in \( h \)) perturbation of the original Hamiltonian. This result is quite general because it applies for any constant symplectic structure, not just the canonical one. We now give some basic examples and special cases of computing the shadow Hamiltonian of the implicit midpoint integrator.

**Example 40.** Perhaps the best-known integrator of (canonical) Hamiltonian dynamics is the leapfrog integrator, which applies when the Hamiltonian is separable. Consider a separable Hamiltonian of the form,

\[ H(q, p) = U(q) + \frac{1}{2} p^\top M^{-1} p \] (F.23)

The leapfrog integrator has a known fourth-order shadow Hamiltonian of the form,

\[ H(q, p) + \frac{h^2}{12} p^\top M^{-1} \nabla^2 U(q) M^{-1} p - \frac{h^2}{24} \nabla U(q)^\top M^{-1} \nabla U(q). \] (F.24)
If one applied the implicit midpoint integrator to this same separable Hamiltonian, one would compute using theorem [F.3.4] that the corresponding shadow Hamiltonian is,

\[ H(q, p) - \frac{h^2}{24} \left( p^\top M^{-1} \nabla^2 U(q) M^{-1} p + \nabla U(q)^\top M^{-1} \nabla U(q) \right). \] (F.25)

These two shadow Hamiltonians differ in the signs of some terms and in the divisor of \( h^2 \).

**Example 41.** We now consider the application of the implicit midpoint integrator to Hamiltonian dynamics with non-canonical symplectic structure. In section [F.10] we derive the fourth-order truncation of the shadow Hamiltonian for a variant of the leapfrog integrator that can be applied when the motion is under the influence of a magnetic field. In this case, the magnetic symplectic structure is of the form,

\[ J_{\text{mag}} \overset{\text{def}}{=} \begin{pmatrix} G & \text{Id} \\ -\text{Id} & 0 \end{pmatrix} \] (F.26)

where \( G \in \text{Skew}(m) \). We consider a separable Hamiltonian of the form \( H(q, p) = U(q) + \frac{1}{2} p^\top p \). Our main result is the following.

**Theorem F.3.6.** Let \( H(q, p) = U(q) + p^\top p/2 \) and let \( H^{[1]} = U(q) \) and \( H^{[2]} = p^\top p/2 \). The magnetic leapfrog integrator of has a shadow Hamiltonian of the form,

\[ H(q, p) + \frac{h^2}{24} \left( -\frac{1}{24} \nabla U(q)^\top \nabla U(q) + \frac{1}{12} \left( p^\top \nabla^2 U(q)p + \nabla U(q)^\top Gp \right) \right) + \mathcal{O}(h^4). \] (F.27)

If one had instead integrated these Hamiltonian dynamics with the implicit midpoint integrator, these proofs may be independent interest because they leverage completely different theoretical tools in identifying the shadow Hamiltonian than are employed in identifying the shadow Hamiltonian of the implicit midpoint method.
integrator, applying theorem F.3.4 one finds that the fourth-order shadow Hamiltonian is,

\[
H(q, p) = \frac{\hbar^2}{24} \left( p^\top (\nabla^2 U(q) + GG)p + \nabla U(q)^\top \nabla U(q) + \nabla U(q)^\top Gp \right).
\]  

(F.28)

F.4 Computation of the Shadow Hamiltonian

Having now identified the shadow Hamiltonian to fourth-order in \(\hbar\), this section discusses practical methods for computing the truncated shadow Hamiltonian and related quantities. The two “exact” approaches we discuss are direct calculation for an important class of Hamiltonians and calculation via automatic differentiation. A third approach we consider looks at a fourth-order perturbation of the fourth-order truncation of the shadow Hamiltonian, which is obtained by holomorphic extension of the Hamiltonian.

F.4.1 Direct Calculation

An important special variant of the Hamiltonian appearing in, for example, [Girolami and Calderhead 2011], corresponds to the motion of a particle on a Riemannian manifold with a global coordinate system undergoing the influence of a potential. Let \(G : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times m}\) be the Riemannian metric (so that \(G(q)\) is a positive definite matrix for all \(q \in \mathbb{R}^m\)). The Hamiltonian is defined by,

\[
H(q, p) = U(q) + \frac{1}{2} \log \det ((2\pi)^m G(q)) + \frac{1}{2} p^\top G(q)^{-1}p,
\]  

(F.29)
where $U : \mathbb{R}^m \to \mathbb{R}$ is the potential energy function. Besides the Hessian $\nabla^2 U(q)$, computing the Hessian of the Hamiltonian will require the following quantities,

\begin{align}
\nabla^2_M H(q, p) &= G(q)^{-1} \\
\frac{\partial}{\partial q_i} \nabla^2_M H(q, p) &= G(q)^{-1} \frac{\partial}{\partial q_i} G(q) G(q)^{-1} p \\
\frac{\partial^2}{\partial q_i \partial q_j} A(q) &= \text{trace} \left( G^{-1}(q) \frac{\partial}{\partial q_j} G(q) \frac{\partial}{\partial q_i} G(q) + G(q)^{-1} \frac{\partial^2}{\partial q_i \partial q_j} G(q) \right) \\
\frac{\partial^2}{\partial q_i \partial q_j} B(q, p) &= 2 p^\top G(q)^{-1} \frac{\partial}{\partial q_j} G(q) \frac{\partial}{\partial q_i} G(q)^{-1} p + p^\top G(q)^{-1} \frac{\partial^2}{\partial q_i \partial q_j} G(q) G(q)^{-1} p,
\end{align}

where $A(q) = \log \det(G(q))$ and $B(q, p) = p^\top G(q)^{-1} p$. As may be intuitively grasped by these formulas, direct calculation of the shadow Hamiltonian for the implicit midpoint integrator for Hamiltonians of the form in eq. (F.29) can be laborious, involving second-order derivatives of the Riemannian metric. Therefore, we do not recommend the direct computation of the fourth-order shadow Hamiltonian in general.

### F.4.2 Automatic Differentiation

The quantity $\nabla^2 H(z)X_H(z)$ appearing in eq. (F.22) is a Hessian-vector product. A critical observation is that a Hessian-vector product has $2m$-entries whereas directly computing the Hessian requires computing $\binom{2m+1}{2}$ unique quantities. The problem of efficiently computing Hessian-vector products has been of interest to the automatic differentiation community; see [Pearlmutter 1994], [Bishop 2006], [Bradbury et al. 2018].
F.4.3 The Holomorphic Semi-Shadow

Suppose that one is disinclined to compute the higher-order derivatives necessitated in section F.4.1 and that one is loathe to use automatic differentiation as described in section F.4.2. Can one still simply obtain an equation for a quantity that is highly conserved by the implicit midpoint integrator? We can answer in the affirmative if the Hamiltonian is complex analytic. A summary of complex analytic functions may be found in definition 2.1.2 and in the surrounding discussion.

Theorem F.4.1. Let \( H : \mathbb{C}^{2m} \to \mathbb{C} \) be a complex analytic function with the property that \( H : \mathbb{R}^{2m} \to \mathbb{R} \). Let \( z, u \in \mathbb{R}^{2m} \) and \( \epsilon \in \mathbb{R} \). Then

\[
\text{Re} \left( H(z + i\epsilon u) \right) = H(z) - \frac{\epsilon^2}{2} u^\top \nabla^2 H(z) u + \mathcal{O}(\epsilon^4). \tag{F.34}
\]

Corollary F.4.2. The function

\[
\tilde{J}^4_h(z) \overset{\text{def.}}{=} \text{Re} \left( H(z + i\frac{h}{\sqrt{12}} X_H(z)) \right) \tag{F.35}
\]

satisfies \( \tilde{J}^4_h(z) = \tilde{H}^4(z) + \mathcal{O}(h^4) \) where \( \tilde{H}^4 \) is the fourth-order truncation of the shadow Hamiltonian defined in eq. (F.22).

A proof is given in section F.11. Therefore, we have the perhaps surprising result that the shadow Hamiltonian of the implicit midpoint integrator can be computed to fourth order with no more difficulty than computing \( H \) itself provided that the Hamiltonian is complex analytic. Put another way, the complex semi-shadow requires nothing more than quantities already available in implementing the implicit midpoint integrator and the capacity to carry out complex arithmetic. Because of this, we call \( \tilde{J}^4 \) the holomorphic semi-shadow. Both \( \tilde{H}^4 \) and \( \tilde{J}^4 \) are fourth-order (in \( h \)) accurate approximations of the shadow Hamiltonian.
The derivation of the holomorphic semi-shadow is closely related to the complex step approximation from numerical differentiation; for details of the complex step approximation see [Martins et al., 2003]. However, the holomorphic semi-shadow differs from the complex step approximation in at least two ways. First, the complex step method is typically leveraged as a mechanism by which to obtain both an evaluation of a function and its derivative; in the context of the holomorphic semi-shadow, neither of these are of particular interest since the function is just the Hamiltonian and its gradient was presumably already available as a consequence of implementing the implicit midpoint integrator. Rather, the holomorphic semi-shadow requires that we use the real-valued higher order terms in the Taylor expansion beyond the function value and its gradient. The second difference is that the complex step approximation typically utilizes extremely small $\epsilon$ to construct complex-valued perturbations of the input; this is because the complex step approximation does not suffer from the catastrophic cancellation inherent to finite differences and therefore machine-accurate estimates of the gradient can be obtained by taking very small perturbations of the input. For the holomorphic semi-shadow, however, we choose the perturbation size as $h/\sqrt{12}$ so as to precisely match the formula of the fourth-order truncation of the shadow Hamiltonian.

When writing the Hamiltonian so that it is properly extended to the complex domain, some care is required. As an example, many Hamiltonian systems introduce a kinetic energy of the form $K(q, p) = \frac{1}{2}p^\top G(q)^{-1}p$, where $G$ is a positive definite matrix. Considerations from numerical linear algebra would lead us to compute this kinetic energy first by using a Cholesky factorization of $G(q)$ and solving the linear system $G(q)y = p$ for $y$ by forward-backward substitution using the Cholesky factor. However, when perturbed into the complex domain, $G(q + ih\nabla_p H(q, p)/12)$ is unlikely to be conjugate symmetric and therefore the Cholesky factorization does not apply. This issue may be corrected by naively inverting $G(q)$ instead when implementing a function to compute the Hamiltonian energy. Note that this change affects only the computation of the Hamiltonian; computa-
Figure F.1: Results of applying the implicit midpoint integrator to the non-separable Hamiltonian described in section F.5.1. Whether the dynamics are integrated with canonical or non-canonical dynamics, or whether a single-step or unit-time trajectory is computed, we find that the fourth-order truncation of the shadow Hamiltonian is more closely conserved than the Hamiltonian itself. We also observe that the holomorphic semi-shadow is conserved to approximately the same degree as the fourth-order truncation of the shadow Hamiltonian.

One could also use formulas like eq. (F.35) in order to compute quantities appearing in the shadow Hamiltonian of the leapfrog integrator from eq. (F.24). However, the formulas are somewhat more awkward because the signs are not in precise agreement: the perturbative $\epsilon^2$-term appearing in eq. (F.34) is subtracted from the Hamiltonian whereas the perturbative term involving the Hessian of the potential appearing in eq. (F.24) is added to the Hamiltonian.

**F.5 Experiments**

We now consider some numerical illustrations of the shadow Hamiltonian of the implicit midpoint integrator.
Figure F.2: For position and momentum variables in the square $[-3, 3] \times [-3, 3]$, we visualize the Hamiltonian energy, the fourth-order shadow Hamiltonian, and the holomorphic semi-shadow for step-sizes $h = 0.01$ and $h = 0.5$. We show the base-10 logarithm of the absolute difference of the Hamiltonian energy functions. For the small step-size, these energy functions are close. Evident differences between the Hamiltonian and the fourth-order approximations of the shadow Hamiltonian emerge for the larger step-size.

Figure F.3: We show the acceptance probabilities associated to “canonical” HMC and shadow HMC using either the fourth-order truncation of the shadow Hamiltonian or the holomorphic semi-shadow. We find that the two shadow methods give better preservation of the Hamiltonian energy on this logistic regression experiment.
F.5.1 Non-Separable Hamiltonian

Our first example comes from Tao [2016]. Consider the following Hamiltonian,

\[ H(q, p) = \sum_{i=1}^{n} \left( \frac{q_i^2 + 1}{2} \right) \left( \frac{p_i^2 + 1}{2} \right). \]  \hspace{1cm} (F.36)

Such a Hamiltonian is non-separable and therefore necessitates integrators other than the standard leapfrog method; we choose to integrate the Hamiltonian dynamics for the Hamiltonian in eq. (F.36) via the implicit midpoint integrator. We integrate with \( n = 2 \) from the initial position \( q = (-3, 0) \) and \( p = (-2, 1) \). We examine the behavior of the shadow Hamiltonian in four permutations of the following two circumstances: (i) taking a single-step of the integrator versus integrating a trajectory for one unit of time, and (ii) integrating the dynamics with and without the canonical symplectic structure. Our procedure for generating a non-canonical structure matrix \( \mathbb{B} = (-\mathbb{J})^{-1} \) is to skew-symmetrize a matrix of random normal numbers. We consider twenty step-sizes at logarithmically-spaced intervals between \( 1 \times 10^{-5} \) and \( 1 \times 10^0 \). We also compare the fourth order truncation of the shadow Hamiltonian with the holomorphic semi-shadow described in section F.4.3. We report our results in fig. F.1. We find that both the fourth-order truncation of the shadow Hamiltonian and the holomorphic semi-shadow are closely conserved by the implicit midpoint integrator; both functions are more closely conserved than the Hamiltonian energy itself as predicted by the theory.

For the case \( n = 1 \) we can visualize the Hamiltonian, the fourth-order truncation of the shadow Hamiltonian, and the holomorphic semi-shadow. These are shown in fig. F.2 for the step-sizes \( h = 0.01 \) and \( h = 0.5 \); these step-sizes were chosen as representatives of relatively small and large step-sizes for integrating these dynamics. For the small step-size, the Hamiltonian is nearly indistinguishable from the fourth-order truncation of the shadow Hamiltonian and the holomorphic semi-shadow, as predicted by our theory. On the other
hand, for the larger step-size, there are evident differences between the Hamiltonian energy and the two alternative energy functions. However, it is interesting to note that the fourth-order truncation of the shadow Hamiltonian and the holomorphic semi-shadow remain qualitatively similar, which follows from the fact that they are identical to fourth-order in $h$.

F.5.2 Shadow Hamiltonian Monte Carlo

Because the implicit midpoint integrator is symmetric and symplectic, it is suitable to use as a transition operator in Hamiltonian Monte Carlo (HMC). Recently, a variant of HMC has emerged that leverages the shadow Hamiltonian has been developed; see Sweet et al. [2009], Izaguirre and Hampton [2004b], Radivojević and Akhmatskaya [2019]. The role of the shadow Hamiltonian in this method is to serve as the acceptance Hamiltonian as described by Duane et al. [1987] whereas the guidance Hamiltonian gives the dynamics to be integrated by a symmetric and symplectic numerical method. When the guidance and acceptance Hamiltonians are equal, we call the method “canonical” HMC.

The shadow Hamiltonians leveraged in shadow HMC have traditionally corresponded to the leapfrog integrator, which is by far the most common numerical integrator used in HMC. We evaluate how HMC with implicit midpoint integrator compares to its shadow variant. We consider a Bayesian logistic regression problem on a German credit dataset. We formulate a Hamiltonian for this inference task using the approach of Riemannian manifold HMC; see Girolami and Calderhead [2011]. We use ten integration steps and an integration step-size in $\{0.1, 0.5, 0.8, 1.0\}$ and partial momentum refreshment as described in Radivojević and Akhmatskaya [2019]. We measure the energy conservation of the implicit midpoint integrator when the acceptance Hamiltonian is given by the original Hamiltonian (canonical HMC), the fourth-order truncation of the shadow Hamiltonian, and the holomorphic semi-shadow. Using these methods to draw 5,000 samples, we present histograms
of the energy conservation in fig. F.3. While we do observe that the shadow methods yield better energy conservation across each step-size, we note that the implicit midpoint integrator already has very strong energy conservation and stability properties [Pourzanjani and Petzold 2019]. Indeed, using canonical HMC, for the largest step-size, the worst case energy conservation is approximately 0.1; the smallest Metropolis-Hastings acceptance probability that this deviation could produce is $\exp(-0.1) \approx 0.9$. From a practical perspective, therefore, the increase in acceptance probability associated to using shadow sampling with the implicit midpoint integrator appears to be marginal.

One may wonder if it is possible to simply continue to increase the step-size so that the acceptance probability of canonical HMC degrades. There are two issues with this. First, with larger step-sizes, the higher order terms in the shadow Hamiltonian become more relevant so that the fourth-order approximations are not as reliable as they were for smaller step-sizes. The second issue is that the implicit midpoint integrator itself is not impervious to large step-sizes and can begin to diverge in presence of very large steps. Even if larger steps are possible, the additional computational burden of requiring more fixed point iterations to compute the implicit midpoint update will almost surely any benefits of a higher acceptance probability. Therefore, it would seem that the implicit midpoint integrator has defeated the computational advantages of introducing shadow acceptance Hamiltonians into the Markov chain.

## F.6 Conclusion

This work has derived the fourth-order shadow Hamiltonian of the implicit midpoint integrator applied to Hamiltonian dynamics with canonical and non-canonical constant symplectic structures. We discussed means of computing the fourth-order shadow Hamiltonian via either direct gradient evaluations or via automatic differentiation. We also proposed a closely-related quantity called the holomorphic semi-shadow, which is also closely con-

498
served by the implicit midpoint integrator and which is easy to compute provided that the Hamiltonian can be holomorphically extended to complex-valued arguments.

F.7 Proofs Concerning the Analytic Derivatives of the Vector Field

Proof. I will use \( z'(h) \) to denote differentiation of \( z(h) \) with respect to \( h \) with higher derivatives indicated by more primes.

\[
\begin{align*}
  z'(h) &= \mathbb{B} \nabla H(z(h)) \quad \text{(F.37)} \\
  z''(h) &= \mathbb{B} \nabla^2 H(z(h)) z'(h) \quad \text{(F.38)} \\
  z'''(h) &= \mathbb{B} \nabla^3 H(z(h))(z'(h), z'(h)) + \mathbb{B} \nabla^2 H(z(h)) z''(h). \quad \text{(F.39)}
\end{align*}
\]

At \( h = 0 \) we have,

\[
  z'''(0) = \mathbb{B} \nabla^3 H(\tilde{z})(\mathbb{B} \nabla H(\tilde{z}), \mathbb{B} \nabla H(\tilde{z})) + \mathbb{B} \nabla^2 H(\tilde{z}) \mathbb{B} \nabla^2 H(\tilde{z}) \mathbb{B} \nabla H(\tilde{z}). \quad \text{(F.40)}
\]
F.8 Proofs Concerning Derivatives of the Implicit Mid-point Integrator

Proof. Let \( \hat{z}(h) = \Phi_h(\hat{z}) \).

\[
\begin{align*}
\hat{z}'(h) &= \nabla H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) + h \nabla^2 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \hat{z}'(h) \\
\hat{z}''(h) &= \nabla^2 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \hat{z}'(h) + h \left[ \nabla^3 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \left( \frac{\hat{z}'(h)}{2}, \frac{\hat{z}'(h)}{2} \right) + \nabla^2 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \hat{z}''(h) \right] \\
\hat{z}'''(h) &= \nabla^3 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \left( \frac{\hat{z}'(h)}{2}, \frac{\hat{z}'(h)}{2} \right) + \nabla^2 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \hat{z}''(h) \\
&\quad + h \left[ \nabla^4 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \left( \frac{\hat{z}'(h)}{2}, \frac{\hat{z}'(h)}{2}, \frac{\hat{z}'(h)}{2} \right) \right. \\
&\quad \left. + \nabla^3 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \left( \frac{\hat{z}''(h)}{2}, \frac{\hat{z}'(h)}{2} \right) + \nabla^2 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \left( \frac{\hat{z}'(h)}{2}, \frac{\hat{z}'''(h)}{2} \right) \right] \\
&= \frac{3}{4} \nabla^3 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) (\hat{z}'(h), \hat{z}'(h)) + \frac{3}{2} \nabla^2 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \hat{z}''(h) \\
&\quad + h \left[ \nabla^4 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \left( \frac{\hat{z}'(h)}{2}, \frac{\hat{z}'(h)}{2}, \frac{\hat{z}'(h)}{2} \right) + \frac{3}{2} \nabla^3 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \left( \frac{\hat{z}''(h)}{2}, \frac{\hat{z}'''(h)}{2} \right) \right. \\
&\quad \left. + \nabla^2 H \left( \frac{\hat{z}(h) + \hat{z}}{2} \right) \hat{z}'''(h) \right]
\end{align*}
\]
At \( h = 0 \) we have,

\[
\dot{z}'''(0) = \frac{3}{4} \mathbb{B} \nabla^3 H(z_0)(\mathbb{B} \nabla H(z_0), \mathbb{B} \nabla H(z_0)) + \frac{3}{2} \mathbb{B} \nabla^2 H(z_0) \mathbb{B} \nabla^2 H(z_0) \mathbb{B} \nabla H(z_0).
\]

(F.46)

\[\square\]

## F.9 Proofs Concerning the Shadow Hamiltonian of the Implicit Midpoint Integrator

We need only verify

\[
\nabla \left[ X_H(\tilde{z})^\top \nabla^2 H(\tilde{z}) X_H(\tilde{z}) \right] = \nabla \left[ \nabla H(\tilde{z})^\top \mathbb{B}^\top \nabla^2 H(\tilde{z}) \mathbb{B} \nabla H(\tilde{z}) \right]
\]

(F.47)

\[
= \nabla^3 H(\tilde{z}) (\mathbb{B} \nabla H(\tilde{z}), \mathbb{B} \nabla H(\tilde{z}))
\]

(F.48)

\[
+ \nabla^2 H(\tilde{z}) \mathbb{B}^\top \nabla^2 H(\tilde{z}) \mathbb{B} \nabla H(\tilde{z})
\]

(F.49)

\[
= \nabla^3 H(\tilde{z}) (\mathbb{B} \nabla H(\tilde{z}), \mathbb{B} \nabla H(\tilde{z}))
\]

(F.50)

\[
+ 2 \nabla^2 H(\tilde{z}) \mathbb{B}^\top \nabla^2 H(\tilde{z}) \mathbb{B} \nabla H(\tilde{z})
\]

using the skew-symmetry of \( \mathbb{B} \) in the last equality. Hence,

\[
- \frac{h^2}{24} \nabla \left[ X_H(\tilde{z})^\top \nabla^2 H(\tilde{z}) X_H(\tilde{z}) \right] =
\]

\[
= \frac{h^2}{6} \left( -\frac{1}{4} \nabla^3 H(\tilde{z}) (\mathbb{B} \nabla H(\tilde{z}), \mathbb{B} \nabla H(\tilde{z})) + \frac{1}{2} \mathbb{B} \nabla^2 H(\tilde{z}) \mathbb{B} \nabla^2 H(\tilde{z}) \mathbb{B} \nabla H(\tilde{z}) \right)
\]

(F.51)
F.10 The Shadow Hamiltonian of the Magnetic Leapfrog Integrator

Consider a separable Hamiltonian of the form $H(q, p) = U(q) + K(p)$. Recall that the magnetic symplectic structure is a $2m \times 2m$ skew-symmetric matrix of the form,

$$
\Omega_{\text{mag}} = \begin{pmatrix}
G & \text{Id} \\
-\text{Id} & 0
\end{pmatrix}.
$$

where $G$ is a $m \times m$ skew-symmetric matrix. The Hamiltonian equations of motion for the magnetic symplectic structure are defined by

$$
\begin{pmatrix}
\dot{q} \\
\dot{p}
\end{pmatrix} =
\begin{pmatrix}
0 & \text{Id} \\
-\text{Id} & -G
\end{pmatrix}
\begin{pmatrix}
\nabla_q U(q) \\
\nabla_p K(p)
\end{pmatrix}.
$$

When we further assume a quadratic kinetic energy, so that the Hamiltonian becomes $H(q, p) = U(q) + \frac{1}{2} p^\top p$, then there exists a symmetric, symplectic, and explicit integrator for the equations of motion; that integrator was derived in [Tripuraneni et al. 2017]. For my purposes, it isn’t important to know anything about the explicit form of their integrator other than to know that it is motivated as a Strang splitting of the Hamiltonian. Let $H^{[1]}(q, p) = U(q)$ and $H^{[2]}(q, p) = p^\top p/2$. Formally, a single step of their integrator to time $t$ is of the form (see eq. 12 in their paper)

$$
\Phi_t = \Phi_{t/2}^{[1]} \circ \Phi_t^{[2]} \circ \Phi_{t/2}^{[1]}
$$

(F.54)
where $\Phi_t^{[1]}$ is the flow to time $t$ of the differential equation

\[
\begin{pmatrix}
\dot{q} \\
\dot{p}
\end{pmatrix}
= 
\begin{pmatrix}
0 & \text{Id} \\
-\text{Id} & -G
\end{pmatrix}
\begin{pmatrix}
\nabla_q H^{[1]}(q, p) \\
\nabla_p H^{[1]}(q, p)
\end{pmatrix}
\]

(F.55)

and $\Phi_t^{[2]}$ is the flow to time $t$ of the differential equation

\[
\begin{pmatrix}
\dot{q} \\
\dot{p}
\end{pmatrix}
= 
\begin{pmatrix}
0 & \text{Id} \\
-\text{Id} & -G
\end{pmatrix}
\begin{pmatrix}
\nabla_q U(q) \\
0
\end{pmatrix}
\]

(F.56)

This is nothing but the symmetric splitting construction that is a common construction for geometric integrators.

Consider a differential equation of the form,

\[
\dot{z} = B \nabla (H^{[1]}(z) + H^{[2]}(z)).
\]

(F.59)

where $B$ is a skew-symmetric matrix. For HMC, a differential equation of the form of eq. (F.59) captures the common situation of a separable Hamiltonian wherein, if we write $z = (q, p)$, we have

\[
H(q, p) = U(q) + K(p) + H^{[1]}(z) + H^{[2]}(z) .
\]

(F.60)

In such a case, one might write $H(q, p) \equiv H(z) = \frac{1}{2} H^{[1]}(z) + H^{[2]}(z) + \frac{1}{2} H^{[1]}(z)$ in order
to motivate the following integrator.

**Definition F.10.1.** Let $\Phi_h^{[i]} : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ for $i \in \{1, 2\}$. An integrator of the form

$$\Phi_h = \Phi_{h/2}^{[1]} \circ \Phi_h^{[2]} \circ \Phi_{h/2}^{[1]}$$

(F.61)

is called a symmetric Strang splitting.

**Proposition F.10.2.** A symmetric Strang splitting integrator as in definition F.10.1 is symmetric, symplectic, and second-order accurate for the system eq. (F.59) if $\Phi^{[i]}_h(z_0)$ is the exact solution to time $h$ of the system,

$$\dot{z} = \mathbb{B} \nabla H^{[i]}(z)$$

(F.62)

$$z(0) = z_0$$

(F.63)

for $i \in \{1, 2\}$.

**Definition F.10.3.** Let $X : \mathbb{R}^m \to \mathbb{R}^m$. The Lie derivative operator $\mathcal{L}_X$ is defined, for a function $f : \mathbb{R}^m \to \mathbb{R}^n$, by

$$(\mathcal{L}_X f)(z) = (\nabla f \cdot X)(z).$$

(F.64)

**Definition F.10.4.** Let $X : \mathbb{R}^m \to \mathbb{R}^m$ and $Y : \mathbb{R}^m \to \mathbb{R}^m$. The Lie bracket is the commutator of the Lie derivative operators

$$[\mathcal{L}_X, \mathcal{L}_Y] \overset{\text{def.}}{=} \mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X.$$ 

(F.65)

Let $X_{H^{[i]}} \overset{\text{def.}}{=} \mathbb{B} \nabla H^{[1]}$ for $i \in \{1, 2\}$. 

505
**Theorem F.10.5.** Consider a numerical integrator of the form,

\[
\Phi_h = \Phi_{h/2} \circ \Phi_h^{[2]} \circ \Phi_{h/2}.
\] (F.66)

Then the modified differential equation for which \(\Phi_h\) is the exact solution is

\[
\frac{d}{dt} \tilde{z}(t) = \left( \mathcal{L}_{X_{H[1]}} + \mathcal{L}_{X_{H[2]}} + \frac{h^2}{24} [\mathcal{L}_{X_{H[1]}}, [\mathcal{L}_{X_{H[2]}}, \mathcal{L}_{X_{H[1]}}, \mathcal{L}_{X_{H[2]}}]] \
+ \frac{1}{12} [\mathcal{L}_{X_{H[2]}}, [\mathcal{L}_{X_{H[2]}}, \mathcal{L}_{X_{H[1]}}, \mathcal{L}_{X_{H[1]}}]] \right) (\tilde{z}(t)) + \mathcal{O}(h^3)
\] (F.67)

**Proof.** See Chapter 9.4 in *Geometric Numerical Integration* by Hairer, Lubich, and Wanner. \(\square\)

Can we leverage the expression of the modified equation in terms of Lie brackets in order to deduce a corresponding shadow Hamiltonian? Indeed, we can using the correspondence between the Lie bracket and the Poisson bracket.

**Definition F.10.6.** Let \(f\) and \(g\) be real-valued functions on \(\mathbb{R}^{2m}\). The Poisson bracket of \(f\) and \(g\) is defined by \(\{f, g\} (z) = \nabla f(z)^\top \mathbb{B} \nabla g(z)\).

**Proposition F.10.7.** Let \(X_H = \mathbb{B} \nabla H\) and let \(f : \mathbb{R}^{2m} \to \mathbb{R}\). The Poisson bracket and Lie derivative are related by the formula,

\[
\mathcal{L}_{X_H} f = \{f, H\}. \tag{F.68}
\]

Moreover, if \(X_G = \mathbb{B} \nabla G\),

\[
[\mathcal{L}_{X_H}, \mathcal{L}_{X_G}] f = \{f, \{G, H\}\}. \tag{F.69}
\]
Hence the shadow Hamiltonian is obtained by identification as,
\[ \tilde{H} = H^{[1]} + H^{[2]} + h^2 \left( -\frac{1}{24} \left\{ \left\{ H^{[2]}, H^{[1]} \right\}, H^{[1]} \right\} + \frac{1}{12} \left\{ \left\{ H^{[1]}, H^{[2]} \right\}, H^{[2]} \right\} \right) + O(h^4) \]  
(F.70)

This approach has been used in order to derive the shadow Hamiltonian for the leapfrog integrator. We may also leverage this theory in the context of non-canonical HMC in the case where the non-canonical dynamics are integrated using an integrator such as that described in definition [F.10.1] As it turns out, this is the case for magnetic Hamiltonian dynamics wherein the matrix \( \mathcal{B} \) assumes the special form

\[ \mathcal{B} = \begin{pmatrix} 0 & \text{Id} \\ -\text{Id} & \mathbf{G} \end{pmatrix} \]  
(F.71)

where \( \mathbf{G} \) is a skew-symmetric matrix.

**Theorem F.10.8.** Let \( H(q, p) = U(q) + \frac{1}{2} p^\top p \) and let \( H^{[1]} = U(q) \) and \( H^{[2]} = p^\top p/2 \). When \( \mathcal{B} \) has the form of eq. (F.71), then there exist explicit functions \( \Phi^{[1]} \) and \( \Phi^{[2]} \) satisfying the condition of proposition [F.10.2] such that \( \Phi_h = \Phi^{[1]}_{h/2} \circ \Phi^{[2]}_h \circ \Phi^{[1]}_{h/2} \) is a second-order accurate, symmetric and symplectic integrator of the dynamics eq. (F.59).

**Proof.** This is the integrator constructed in [Tripuraneni et al. 2017], which is a generalization of Scovel’s method.

**Theorem F.10.9.** Let \( H(q, p) = U(q) + \frac{1}{2} p^\top p \) and let \( H^{[1]} = U(q) \) and \( H^{[2]} = p^\top p/2 \). The integrator of theorem [F.10.8] has a shadow Hamiltonian of the form,

\[ \tilde{H}(q, p) = H(q, p) + h^2 \left( -\frac{1}{24} \nabla U(q)^\top \nabla U(q) + \frac{1}{12} \left( p^\top \nabla^2 U(q)p + \nabla U(q)^\top \mathbf{G}p \right) \right) + O(h^4). \]  
(F.72)

**Proof.** One just has to compute the Poisson brackets in eq. (F.70).
Proof. Because $H$ is complex analytic it may be expanded in a Taylor series as

$$H(z + i \epsilon u) = H(z) + i \epsilon \nabla H(z)^\top u - \frac{\epsilon^2}{2} u^\top \nabla^2 H(z) u + \mathcal{O}(i \delta^3) + \mathcal{O}(\delta^4).$$  \hfill (F.73)

Extracting the real part gives the result. Recall now that the fourth-order shadow Hamiltonian of the implicit midpoint integrator is the function,

$$\tilde{H}[^4](z) \overset{\text{def.}}{=} H(z) - \frac{\hbar^2}{24} X_H(z)^\top \nabla^2 H(z) X_H(z).$$  \hfill (F.74)

We have

$$\operatorname{Re} \left( H(z + i \frac{\hbar}{\sqrt{12}} X_H(z)) \right) = H(z) - \frac{\hbar^2}{2 \cdot 12} X_H(z)^\top \nabla^2 H(z) X_H(z) + \mathcal{O}(\hbar^4) \hfill (F.75)$$

$$= \tilde{H}[^4](z) + \mathcal{O}(\hbar^4) \hfill (F.76)$$

$\square$
Appendix G

Adaptation of the Independent Metropolis-Hastings Sampler with Normalizing Flow Proposals

This chapter is adapted from Brofos et al. [2021]. This work is a collaboration with Mary-lou Gabrié, Marcus Brubaker, and Roy Lederman. A large proportion of the exposition is due to me, excluding the related work and experimentation on the physical field model, which were contributed by Marlou Gabrié. Research oversight and guidance were provided by Marcus Brubaker and Roy Lederman.

We note that this research complements the preceding research into geometric methods of Markov chain Monte Carlo. Whereas HMC, RMHMC, and LMC are predicated on local information about the posterior distribution, the independent Metropolis-Hastings algorithm discussed in this work facilitates transitions between distant components of the posterior by proposing subsequent states of the Markov chain independently of the current state. In the experiments in the physical field system, this approach is combined with local updates, in the form of Metropolis-adjusted Langevin updates, in order effective explore, and transition between, modes of the posterior distribution.

Abstract. Markov Chain Monte Carlo (MCMC) methods are a powerful tool for computation with complex probability distributions. However the performance of such meth-
ods is critically dependent on properly tuned parameters, most of which are difficult if not impossible to know a priori for a given target distribution. Adaptive MCMC methods aim to address this by allowing the parameters to be updated during sampling based on previous samples from the chain at the expense of requiring a new theoretical analysis to ensure convergence. In this work we extend the convergence theory of adaptive MCMC methods to a new class of methods built on a powerful class of parametric density estimators known as normalizing flows. In particular, we consider an independent Metropolis-Hastings sampler where the proposal distribution is represented by a normalizing flow whose parameters are updated using stochastic gradient descent. We explore the practical performance of this procedure on both synthetic settings and in the analysis of a physical field system and compare it against both adaptive and non-adaptive MCMC methods.

G.1 Introduction

Markov Chain Monte Carlo (MCMC) methods are procedures for generating samples from probability distributions, typically given knowledge of the density of the distribution up to proportionality. These MCMC samplers often depend on parameters; for instance, in the random walk Metropolis procedure on $\mathbb{R}^n$, one may treat the covariance matrix of a normal proposal distribution as a parameter of the method; see, for instance, Haario et al. [2001]. The performance of an MCMC procedure will depend on these parameters. It would be preferable if these parameters could be adapted during sampling, however such adaptations can violate the Markov property of the chain and undermine its convergence to the desired target distribution.

An important variation of MCMC is the independent Metropolis-Hastings sampler. This method samples from a target distribution by first sampling from a auxiliary proposal distribution (independently from the current state of the chain) and accepts or rejects
those proposals according to the Metropolis-Hastings criterion. The effectiveness of this algorithm depends on the ratio of the target density to the ratio of the proposal density [Robert and Casella, 2005]: if the ratio is bounded over the support of the target distribution, the algorithm enjoys a powerful theory of geometric ergodicity. The independent Metropolis-Hastings algorithm is the focus of the present work.

Recently in the machine learning community, normalizing flows have emerged as a powerful mechanism for expressing complex densities, see [Kobyzev et al., 2020, Papamakarios et al., 2021] for a recent reviews. Normalizing flows are defined by a parametric, smooth and invertible function which transforms a simple distribution (e.g., a Gaussian) into a more complex one (e.g., natural images) and uses the change-of-variables formula to exactly determine the resulting probability density function in the complex space. Provided that the family of normalizing flows under consideration is sufficiently expressive, any distribution can be constructed in theory this way. In practice, many normalizing flows exhibit a universal approximation property whereby, given suitable model capacity, they can approximate any distribution arbitrarily well, e.g., [Huang et al., 2018, Jaini et al., 2019]. Indeed, normalizing flows are distinguished among parametric families of distributions by their expressiveness and tractability of sampling and log-density evaluation; the precise attributes that one requires for a proposal distribution in the independent Metropolis-Hastings sampler. By incorporating normalizing flows into the MCMC framework we seek to leverage their expressivity along with the ergodicity of the MCMC procedure in order to produce samples from a target distribution (see fig. G.1). The principle computational challenge associated to normalizing flows is the identification of parameters that produce the best approximation of a target density. Therefore, a question of principle theoretical interest and practical importance is, “During the course of sampling, under what conditions can the parameters of the normalizing flow be continuously adapted?”

The outline of this paper is as follows. In section G.2 we review important concepts from the analysis of Markov chains; we provide the independent Metropolis-Hastings al-
gorithm and state the conditions under which it enjoys geometric ergodicity; we devise a
metric space over transition kernels, which will be important for analyzing notions of con-
tinuity. We review recent experimental works that demonstrated the benefit of normalizing
flow proposals in MCMCs and related theoretical literature in section G.3. In section G.4
we state our theories for the continual adaptation of Markov chains. We begin by con-
sidering deterministic adaptations wherein parameter updates are determined sequentially
and deterministically without regard to the state of the chain; this case can be used to mo-
tivate the adaptation of normalizing flows as a gradient flow. We then proceed to consider
stochastic adaptations wherein the state of the chain and the adaptation of the parameters
of the normalizing flow at the \( n \)th step are not necessarily independent given the history
of the chain up to the \((n - 1)\)th step. This circumstance includes the case wherein the ac-
cepted proposal sampled from the normalizing flow is also used in the computation of the
adaptation, as necessary for the “pseudo-likelihood” algorithm we examine numerically in
section G.5.

G.2 Preliminaries

In giving an overview of Markov chains and their associated theory, we emulate the nota-
tion and presentation of Meyn and Tweedie [1993]. Refer to section F.A.1 for a review of
total variation distances. Throughout, we let \( \mathcal{X} \) denote a set which we equip with its Borel
\( \sigma \)-algebra, denoted \( \mathcal{B}(\mathcal{X}) \). We associate to \( (\mathcal{X}, \mathcal{B}(\mathcal{X})) \) a measure \( \mu : \mathcal{B}(\mathcal{X}) \rightarrow [0, \infty) \)
– satisfying \( \mu(A) \geq 0 \) for all \( A \in \mathcal{B}(\mathcal{X}) \), \( \mu(\emptyset) = 0 \), and the condition of countable add-
tivity – to create the measure space \( (\mathcal{X}, \mathcal{B}(\mathcal{X}), \mu) \). A probability measure is a measure
which satisfies Kolmogorov’s axioms Kolmogorov [1960]. A signed measure relaxes the
condition of non-negativity. If \( X \) is an \( \mathcal{X} \)-valued random variable and \( \Pi \) is a probability
measure on \( (\mathcal{X}, \mathcal{B}(\mathcal{X})) \) we write \( X \sim \Pi(\cdot) \) to mean that for any \( A \in \mathcal{B}(\mathcal{X}) \) we have
\( \Pr [X \in A] = \Pi(A) \). If a probability measure \( \Pi \) has a density with respect to a dominat-
Figure G.1: This work examines the convergence of adaptive Markov chain Monte Carlo algorithms using the independent Metropolis-Hastings algorithm when the proposal distribution is parameterized by a normalizing flow. In this illustration, we seek to draw samples from a target distribution. We begin with an initial parameter $\Theta_0$, which parameterizes a simple proposal distribution, denoted $\Pi_{\Theta_0}$, which is a normalizing flow, and an initial state of the chain $X_0$; a sample from this proposal is accepted or rejected according to the Metropolis-Hastings criterion, yielding a transition to the state $X_1$. The parameters of the normalizing flow are thereafter adapted to produce a new proposal distribution $\Pi_{\Theta_1}$, which we hope is closer to the target distribution. Iterating this procedure we obtain both a sequence of states $(X_n)_{n \in \mathbb{N}}$ and a sequence of normalizing flow parameters $(\Theta_n)_{n \in \mathbb{N}}$. The principle question of this work is to establish when the sequence of states converges to the target density.

In MCMC, we generate a sequence of $\mathcal{X}$-valued random variables, denoted $(X_0, X_1, \ldots)$ that satisfy the Markov property. The transition to state $X_{n+1}$ given $X_n = x_n$ is formally captured by the notion of a transition kernel.

**Definition G.2.1 (Robert and Casella [2005]).** A transition kernel on $\mathcal{X}$ is a function $\mathcal{X} \times \mathcal{B}(\mathcal{X}) \ni (x, A) \mapsto K(x, A)$ that satisfies the following two properties: (i) For all $x \in \mathcal{X}$,
$K(x, \cdot)$ is a probability measure and (ii) For all $A \in \mathcal{B}(\mathcal{X})$, $K(\cdot, A)$ is $\mathcal{B}(\mathcal{X})$-measurable.

Thus, the propagation of the state from step $n$ to step $n+1$ is represented by $X_{n+1} \sim K(x_n, \cdot)$. When considering Markov chains, we will frequently be interested in the $n$-step transition probability measure from some initial state $X_0 = x_0$; we denote this probability measure by $K^n(x_0, \cdot) = \Pr [X_n \in \cdot | X_0 = x_0]$, which has the following expression:

$$K^n(x_0, A) = \int_{\mathcal{X}} \cdots \int_{\mathcal{X}} K(x_0, dx_1)K(x_1, dx_2) \cdots K(x_{n-2}, dx_{n-1})K(x_{n-1}, A).$$

(G.1)

Of principle interest to the theory of Markov chains is the limiting behavior of the $n$-step transition probability measure.

**Definition G.2.2.** The transition kernel $K$ with $n$-step transition law $K^n$ is ergodic for $\Pi$ if, for every $x \in \mathcal{X}$, $\lim_{n \to \infty} \|K^n(x, \cdot) - \Pi(\cdot)\|_{TV} = 0$.

In the sequel, we will require continuity of sequences of transition kernels, which necessitates that we equip the space of transition kernels with a metric. A natural metric considers the worst-case total variation distance between kernels.

**Definition G.2.3.** Two transition kernels $K$ and $K'$ on $\mathcal{X} \times \mathcal{B}(\mathcal{X})$ are equal if $\sup_{x \in \mathcal{X}} \|K(x, \cdot) - K'(x, \cdot)\|_{TV} = 0$.

**Proposition G.2.4.** Let $K$ and $K'$ be transition kernels on $\mathcal{X} \times \mathcal{B}(\mathcal{X})$. Then the function, $d(K, K') = \sup_{x \in \mathcal{X}} \|K(x, \cdot) - K'(x, \cdot)\|_{TV}$ is a distance function on transition kernels.

A proof is given in section F.A.4.

**G.2.2 Independent Metropolis-Hastings**

**Definition G.2.5.** Let $\Pi$ and $\tilde{\Pi}$ be two probability measures on $\mathcal{B}(\mathcal{X})$ with densities with respect to some dominating measure $\mu$ given by $\pi$ and $\tilde{\pi}$, respectively. Consider a Markov
chain \((X_0, X_1, X_2, \ldots)\) constructed via the following procedure given an initial state of the Markov chain \(X_0 = x_0\). First, randomly sample \(\tilde{X} \sim \tilde{\Pi}\). Then set \(X_{n+1} = \tilde{X}\) with probability \(\min\left\{ \frac{\pi(\tilde{X})\tilde{\pi}(X_n)}{\pi(X_n)\tilde{\pi}(X)}, 1 \right\}\) and otherwise set \(X_{n+1} = X_n\). The Markov chain \((X_0, X_1, X_2, \ldots)\) is called the independent Metropolis-Hastings sampler of \(\Pi\) given \(\tilde{\Pi}\).

**Proposition G.2.6.** Let \(K\) denote the transition kernel of the independent Metropolis-Hastings sampler. The stationary distribution of \((X_0, X_1, X_2, \ldots)\) is \(\Pi\) and if there exists a constant \(M \geq 1\) such that \(\frac{\pi(x)}{\tilde{\pi}(x)} \leq M, \ \forall x \in \text{Supp}(\pi),\) then the independent Metropolis-Hastings sampler is uniformly ergodic in the sense that \(\|K^n(x, \cdot) - \Pi(\cdot)\|_{TV} \leq 2 \left(1 - \frac{1}{M}\right)^n\).

For a proof of these results, refer to [Meyn and Tweedie, 1993](#). [Robert and Casella, 2005](#). There is a question of when such a \(M\) as in proposition G.2.6 will exist. Under a compactness condition and assumptions of continuity on both the proposal and target densities, then an affirmative existence result can be given.

**Corollary G.2.7.** If, in addition, \(X\) is a compact set and if \(\pi\) and \(\tilde{\pi}\) are continuous on \(X\), and if \(\text{Supp}(\pi) \subseteq \text{Supp}(\tilde{\pi})\) then there exists such an \(M\) as in proposition G.2.6.

A proof is given in section F.A.5. The transition kernel of the independent Metropolis-Hastings sampler has the form

\[
K(x, dx') = \min \left\{ 1, \frac{\pi(x')\tilde{\pi}(x)}{\pi(x)\tilde{\pi}(x')} \right\} \tilde{\pi}(x') \mu(dx') + \\
\left( 1 - \int_X \min \left\{ 1, \frac{\pi(x)\tilde{\pi}(x)}{\pi(x')\tilde{\pi}(x)} \right\} \tilde{\pi}(x) \mu(dx) \right) \delta_x(dx').
\] (G.2)

The first term in eq. (G.2) is the probability of an accepted transition from \(x\) to the region \(dx'\) whereas the second term is the probability of remaining at \(x\), which only contributes if \(x\) lies in the region \(dx'\).
G.2.3 Normalizing Flows

Normalizing flows are families of parameterizable, smooth bijections from $\mathcal{X}$ into itself for which the bijection, its inverse, and its Jacobian determinant are computationally tractable. Moreover, in order to comprise a practically relevant technique, the family of normalizing flows so described should be able to transform a simple base distribution (such as a standard multivariate Gaussian) into a complex distribution of interest.

Example 42. Let $\mathcal{X} = \mathbb{R}^m$ and consider a simple base probability measure $\Pi_X$ with density $\pi_X$ with respect to Lebesgue measure. Let $X \sim \Pi_X(\cdot)$. Let $\phi_\theta$ be a smooth bijection parameterized by $\theta \in \mathbb{R}^k$. By the change-of-variables formula, the density of $Y = \phi_\theta(X)$ is

$$
\pi_Y(y) = \pi_X(\phi_\theta^{-1}(y)) \cdot \left| \det(\nabla_y \phi_\theta^{-1}(y)) \right|. \quad (G.3)
$$

Moreover, if $\phi_\theta$ meets the computational tractability conditions, one can easily generate samples from $Y$ and compute the probability density $\pi_Y$ at arbitrary locations in $\mathcal{X}$.

A common normalizing flow on Euclidean space is the RealNVP architecture [Dinh et al., 2017], which defines a mapping $\phi(\theta_\mu, \theta_\sigma) : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m}$ as follows. Let $x \in \mathbb{R}^{n+m}$ and let $\mu_\theta : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\sigma_\theta : \mathbb{R}^n \rightarrow \mathbb{R}_+^m$ be parameterized, differentiable functions. Let $y \in \mathbb{R}^{n+m}$ be computed according to:

$$
y = \begin{pmatrix} x_{1:n} \\ x_{(n+1):m} \otimes \sigma_\theta(x_{1:n}) + \mu_\theta(x_{1:n}) \end{pmatrix} \quad (G.4)
$$

where $\otimes$ denotes element-wise multiplication and we have used the notation $x_{a:b}$ to mean $(x_a, x_{a+1}, \ldots, x_{b-1}, x_b)$ for $a$ and $b$ in $\{1, \ldots, n + m\}$. The Jacobian determinant of the
transformation \( x \mapsto y \) is simply \( \prod_{i=1}^{m} \sigma_{\theta_i}^{(i)}(x_{1:n}) \). The inverse transform is

\[
x = \begin{pmatrix}
y_{1:n} \\
(y_{(n+1):m} - \mu_{\theta_n}(y_{1:n})) \odot \sigma_{\theta_n}(y_{1:n})
\end{pmatrix},
\]

where \( \odot \) denoted element-wise division.

### G.2.4 Adaptive Transition Kernels

As alluded to in section G.1 the transition kernel may depend on parameters, denoted by \( \theta \) and taking values in a set \( \mathcal{Y} \). In this case, we express the dependency of the kernel \( K \) on its parameters by writing \( K_{\theta} \). In adaptive MCMC, given a target probability measure \( \Pi \), we seek to strategically construct a sequence of transition kernels \( (K_{\Theta_n})_{n \in \mathbb{N}} \) where \( (\Theta_n)_{n \in \mathbb{N}} \) is a sequence of \( \mathcal{Y} \)-valued random variables. Ideally, the sequence \( (\Theta_n)_{n \in \mathbb{N}} \) will enable sampling from \( \Pi \) that becomes more effective with each step. In the adaptive MCMC framework, the one-step transition laws for \( X_{n+1} \) given \( X_n = x_n \) and \( \Theta_n = \theta_n \) is \( X_{n+1} \sim K_{\theta_n}(x_n, \cdot) \). The \( n \)-step transition law given \( X_0 = x_0 \) and \( (\Theta_0 = \theta_0, \ldots, \Theta_{n-1} = \theta_{n-1}) \) is

\[
K_{(\theta_i)_{i=0}^{n-1}}^n(x_0, A) = \int_{\mathcal{X}} \cdots \int_{\mathcal{X}} K_{\theta_0}(x_0, dx_1) K_{\theta_1}(x_1, dx_2) \\
\quad \cdots K_{\theta_{n-2}}(x_{n-2}, dx_{n-1}) K_{\theta_{n-1}}(x_{n-1}, A),
\]

Therefore, by the law of total expectation, the \( n \)-step transition law given \( X_0 = x_0 \) is

\[
G^n(x_0, A) = \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} K_{(\theta_i)_{i=0}^{n-1}}^n(x_0, A),
\]

where the expectation is computed over the marginal distribution of the parameters. We now give a precise definition for what it means for an adaptive MCMC procedure to be ergodic.

**Definition G.2.8.** The \( n \)-step transition law \( G^n \) is said to be ergodic for the probability measure \( \Pi \) if, for every \( x \in \mathcal{X} \),

\[
\lim_{n \to \infty} \|G^n(x, \cdot) - \Pi(\cdot)\|_{TV} = 0.
\]
The principal theoretical tools of our analysis are the definitions of containment, simultaneous uniform ergodicity, and diminishing adaptation. Diminishing adaptation together with either containment or simultaneous uniform ergodicity implies ergodicity of the adaptive MCMC procedure in the sense of definition G.2.8. The remainder of this section is a review of Roberts and Rosenthal [2007], Bai et al. [2011].

**Definition G.2.9.** The sequence of Markov transition kernels \( \{K_{\Theta_n}\}_{n \in \mathbb{N}} \) is said to exhibit diminishing adaptation if

\[
\lim_{n \to \infty} d(K_{\Theta_{n+1}}, K_{\Theta_n}) = 0
\]

in probability.

**Lemma G.2.10** (Roberts and Rosenthal [2007]). Suppose that \( \Theta_{n+1} = \Theta_n \) w.p. \( 1 - \alpha_n \) and otherwise \( \Theta_{n+1} = \Theta_n' \) where \( \Theta_n' \in \mathcal{Y} \) is any other element of the index set. If \( \lim_{n \to \infty} \alpha_n = 0 \), then \((K_{\Theta_0}, K_{\Theta_1}, \ldots)\) exhibits diminishing adaptation.

**Definition G.2.11.** Define \( W_\epsilon(x, K) = \inf \{n \geq 1 : \|K^n(x, \cdot) - \Pi(\cdot)\|_{TV} < \epsilon\} \). The sequence \((\Theta_n)_{n \in \mathbb{N}}\) is said to exhibit containment if, for every \( \epsilon > 0 \), the sequence

\[
(W_\epsilon(X_0, K_{\Theta_0}), W_\epsilon(X_1, K_{\Theta_1}), \ldots)
\]

is bounded in probability given \( X_0 = x_0 \) and \( \Theta_0 = \theta_0 \), where \( X_{n+1} \sim K_{\Theta_n}(X_n, \cdot) \).

Containment states that for a particular stochastic sequence of adaptations \((\Theta_n)_{n \in \mathbb{N}}\) there is, with arbitrarily high probability, a finite number of steps one may take with any of the parameters in the sequence in order to be arbitrarily close to the target distribution. The following theorems give the relationships between diminishing adaptation, simultaneous uniform ergodicity, containment, and ergodicity of the adaptive MCMC procedure. The proofs of these results may be found in [Roberts and Rosenthal, 2007].

**Theorem G.2.12.** Let \( \{K_\theta\}_{\theta \in \mathcal{Y}} \) be a family of Markov chain transition kernels that are all stationary for the same distribution \( \Pi \). Suppose that the family satisfies definition F.A.15.1 and that the sequence \((\Theta_0, \Theta_1, \ldots)\) satisfies definition G.2.9. Then the chain whose transitions are governed by \( X_{n+1} \sim K_{\Theta_n}(X_n, \cdot) \) is ergodic for the distribution \( \Pi \).
Theorem G.2.13. Let \( \{K_\theta\}_{\theta \in \mathcal{Y}} \) be a family of Markov chain transition kernels that are all stationary for the same distribution \( \Pi \). Suppose that the sequence \((\Theta_0, \Theta_1, \ldots)\) satisfies definitions G.2.9 and G.2.11. Then the chain whose transitions are governed by \( X_{n+1} \sim K_{\Theta_n}(X_n, \cdot) \) is ergodic for the distribution \( \Pi \).

G.3 Related Work

A series of works recently investigated the learning of a proposal distribution for the independent Metropolis-Hastings sampler with normalizing flows, in particular for statistical mechanics field theories. For such models, Albergo et al. [2019] used stochastic independent adaptations models following the optimization of the reverse Kullback-Leibler divergence (KL), as in example 44 of the next section. While this strategy is successful when the target is unimodal, it is known to yield underdispersed approximation of the target distribution and to be prone to mode collapse. Within the framework of variational inference, Naesseth et al. [2020] proposed to address these issues by optimizing instead an approximate forward KL using simple parametric families for the proposal. In this case, adaptations are stochastic and rely on the previous states of the chain to estimate gradients of the approximate forward KL, called “pseudo-likelihood” in example 45 of the present paper. Incorporating normalizing flows, Gabrié et al. [2021] successfully sampled multimodal distributions using an initialization that echoes the containment property. In the context of statistical field theories, Hackett et al. [2021] also demonstrated the need for forward KL training to assist sampling of multimodal distributions while surveying strategies to obtain training samples different from the adaptive MCMC discussed here.

Among the works above, ergodicity was only tested numerically. One exception is Gabrié et al. [2021] where a convergence argument based on a continuous time analysis is developed under the assumption of perfect adaptation. The present paper provides a theoretical framework to analyze for the ergodicity of the methods presented in the body.
of work above. Though our work has focused on establishing ergodicity via the mechanism of Roberts and Rosenthal [2007], we note the work of Andrieu and Éric Moulines [2006], which may be used to establish an ergodicity theory. We concur with the statement in Roberts and Rosenthal [2007] that Andrieu and Éric Moulines [2006] “requir[es] other technical hypotheses which may be difficult to verify in practice” and that diminishing adaptation and containment are “somewhat simpler conditions.” Holden et al. [2009] considered the case of independent adaptations of the independent Metropolis-Hastings algorithm; however, this technique requires that accepted and rejected states be treated identically in the adaptation procedure, so we do not consider it further.

G.4 Analytical Apparatus

We now consider the principle problem of this paper: When can the adaptive independent Metropolis-Hastings sampler with proposal distribution parameterized by a normalizing flow be given an ergodicity theory? We separate our discussion into two components wherein the adaptations are either deterministic or not necessarily independent of the state of the chain.

G.4.1 Deterministic Adaptations

**Theorem G.4.1.** Let $\Pi$ be a probability measure with density $\pi$. Suppose that every $\theta \in \mathcal{Y}$ parameterizes a probability measure $\Pi_\theta$ on $\mathcal{B}(\mathcal{X})$ with density $\tilde{\pi}_\theta$. Suppose that $(\theta_0, \theta_1, \ldots)$ is a deterministic $\mathcal{Y}$-valued sequence. Let $(K_{\theta_n})_{n \in \mathbb{N}}$ be an associated sequence of Markov transition kernels of the independent Metropolis-Hastings sampler of $\Pi$ given $\Pi_{\theta_n}$. Let $K^n(x_0, A)$ denote the $n$-step transition probability from $x_0$ to $A \in \mathcal{B}(\mathcal{X})$ with law eq. (G.6). Then $\Pi$ is the stationary distribution for $K^n$. Suppose further that for each $n \in \mathbb{N}$ there exists $M_n$ satisfying $\pi(x) \leq M_n \tilde{\pi}_{\theta_n}(x)$ for all $x \in \text{Supp}(\pi)$ Then,
\[ \|K^n(x_0, \cdot) - \Pi(\cdot)\|_{TV} \leq 2 \prod_{i=0}^{n-1} \left(1 - \frac{1}{M_i}\right). \]

A proof is given in section F.A.2.

**Example 43.** Let \( \Pi \) be a probability measure with density \( \pi \). Let \( \mathcal{Y} = \mathbb{R}^m \) and suppose that every \( \theta \in \mathcal{Y} \) smoothly parameterizes a probability measure \( \tilde{\Pi}_\theta \) on \( \mathcal{B}(\mathcal{X}) \) with density \( \tilde{\pi}_\theta \) for which \( \text{Supp}(\pi) = \text{Supp}(\tilde{\pi}_\theta) \). Consider the initial value problem

\[ \frac{d}{dt} \theta(t) = -\nabla_{\theta} \mathbb{KL}(\tilde{\pi}_{\theta(t)} \parallel \pi), \quad \theta(0) = \theta_0, \tag{G.8} \]

where \( \theta_0 \in \mathcal{Y} \). Let \((t_0, t_1, \ldots)\) be a deterministic sequence of times and let \( \theta_n = \theta(t_n) \) for \( n \in \mathbb{N} \). Consider the family of Markov chain transition operators of the independent Metropolis-Hastings sampler of \( \Pi \) given \( \tilde{\Pi}_{\theta_n} \) with transition kernels \( K_{\theta_n} \). Then \( \Pi \) is the stationary distribution of the Markov chain whose transitions satisfy \( X_{n+1} \sim K_{\theta_n}(X_n, \cdot) \).

From the condition \( \text{Supp}(\pi) = \text{Supp}(\tilde{\pi}_\theta) \) it follows that \( \Pi \) is the stationary distribution for each \( K_{\theta_n} \). (Note that the condition \( \text{Supp}(\pi) = \text{Supp}(\tilde{\pi}_\theta) \) is required instead of the weaker condition \( \text{Supp}(\pi) \subseteq \text{Supp}(\tilde{\pi}_\theta) \) in order to guarantee that the relative entropy is well-defined.) Since \((\theta_0, \theta_1, \ldots)\) is a deterministic sequence, it follows from theorem G.4.1 that \( \Pi \) is the stationary distribution. The particular mechanism of producing a deterministic sequence was not important; however, the time derivative eq. (G.8) was chosen because it begins to imitate the evolution encountered in normalizing flow loss functions.

**G.4.2 Non-Independent Adaptations**

Notice that the decision to make the adaptation and the subsequent state of the chain dependent is not artificial or contrived; in fact, if such a procedure can be equipped with an ergodicity theory, then the resulting algorithm would have an important computational advantage. Specifically, it would require fewer evaluations of the target density (or the gradient of the target density) than the corresponding procedure with independent adap-
(a) Failure of simultaneous uniform ergodicity
(b) Simultaneous uniform ergodicity via mixing
(c) Containment

Figure G.2: Establishing ergodicity of an adaptive chain sampler with non-independent adaptations requires showing that either the simultaneous uniform ergodicity or containment properties hold. In fig. G.2a, we show how adaptations can produce a catastrophic failure of ergodicity by visualizing the log-densities of the target distribution and the adaptive proposal distributions: the proposal distribution becomes increasingly concentrated, meaning that there is no uniform bound on the $\epsilon$-mixing time. On a compact space, this phenomenon can be defeated by mixing the normalizing flow with a fixed distribution with full support; we visualize this effect in fig. G.2b. An alternative to mixing is to look toward the containment property. When there is a finite upper bound on the difference between the target density and the proposal density, containment is implied for the independent Metropolis-Hastings sampler. In fig. G.2c, containment is visualized as the fact that all of the log-densities of the proposal distributions are uniformly greater than the target log-density translated down by some quantity greater than zero (dotted blue line).

This motivates us to explore this direction. For instance, the following adaptation scheme does not fall into the category of independent adaptations.

**Example 44.** Let $\Pi$ be a probability measure with density $\pi$ on a space $\mathcal{X}$. Let $\mathcal{Y} = \mathbb{R}^m$ and suppose that every $\theta \in \mathcal{Y}$ smoothly parameterizes a probability measure $\tilde{\Pi}_\theta$ on $\mathcal{B}(\mathcal{X})$ with density $\tilde{\pi}_\theta$ for which $\text{Supp}(\pi) = \text{Supp}(\tilde{\pi}_\theta)$. Let $\tilde{X} \sim \tilde{\Pi}_{\theta_{n-1}}$ be the proposal produced by the independent Metropolis-Hastings sampler of $\Pi$ given $\tilde{\Pi}_{\theta_{n-1}}$. Consider the adaptation

$$\theta_n = \theta_{n-1} - \epsilon \nabla_{\theta} \log \frac{\tilde{\pi}_{\theta_{n-1}}(\tilde{X}(\theta_{n-1}))}{\pi(\tilde{X}(\theta_{n-1}))}, \quad (G.9)$$

which can be interpreted as the single-sample approximation of the gradient flow of $\mathbb{KL}(\tilde{\pi}_{\theta_{n-1}} \| \pi)$.

This motivates us to explore this direction. definition G.2.9 and the continuous map-
Lemma G.4.2. Suppose that the map \( \theta \mapsto K_{\theta} \) is continuous and that the sequence \((\Theta_0, \Theta_1, \ldots)\) converges in probability in \( \mathcal{Y} \). Then \((K_{\Theta_0}, K_{\Theta_1}, \ldots)\) exhibits diminishing adaptation.

A proof is given in section F.A.4. We now consider the question of the continuity of the mapping \( \theta \mapsto K_{\theta} \).

Theorem G.4.3. Let \((\theta_1, \theta_2, \ldots)\) be a \( \mathcal{Y} \)-valued sequence converging to \( \theta \). Let \( \pi \) be a probability density function on a space \( \mathcal{X} \) and let \( \tilde{\pi}_\theta \) be a family of density functions on \( \mathcal{X} \) indexed by \( \theta \) such that the map \( \theta \mapsto \tilde{\pi}_\theta \) is continuous. Suppose further that \( \text{Supp}(\tilde{\pi}_\theta) = \mathcal{X} \) for every \( \theta \in \mathcal{Y} \). Let \( \Pi \) be the probability measure on \( \mathcal{B}(\mathcal{X}) \) with density \( \pi \) and let \( \tilde{\Pi}_\theta \) be the probability measure on \( \mathcal{B}(\mathcal{X}) \) with density \( \tilde{\pi}_\theta \). Let \( K_{\theta} \) be the transition kernel of the independent Metropolis-Hastings sampler of \( \Pi \) given \( \tilde{\Pi}_\theta \). Then \( \lim_{n \to \infty} K_{\theta_n} = K_{\theta} \) (i.e. the mapping \( \theta \mapsto K_{\theta} \) is continuous).

A proof is given in section F.A.4. When training normalizing flows, it is typical to apply stochastic gradient descent to the minimization of some loss function. The question of when the iterates of stochastic gradient descent converge is an important question that has been recently treated in the case of non-convex losses. We refer the interested reader to [Bottou, 1999, Mertikopoulos et al., 2020] for conditions and results guaranteeing the convergence of stochastic gradient descent. In practice, the convergence of the sequence of normalizing flow parameters can be further encouraged by a decreasing learning rate schedule. In section F.A.15 we discuss simultaneous uniform ergodicity on compact spaces and give some examples of normalizing flows works in these cases. The condition for geometric ergodicity of the independent Metropolis-Hastings sampler is that there exists \( M \geq 1 \) such that \( \pi(x) \leq M \cdot \tilde{\pi}(x) \) for all \( x \in \text{Supp}(\pi) \) where \( \pi \) is the density of the target distribution and \( \tilde{\pi} \) is the proposal density. By taking the logarithm of both
sides and rearranging we obtain the equivalent inequality, \( \log \pi(x) - \log \tilde{\pi}(x) \leq \log M \) for all \( x \in \text{Supp}(x) \).

**Proposition G.4.4.** Suppose that every \( \theta \in \mathcal{Y} \) parameterizes a probability measure \( \tilde{\Pi}_\theta \) on \( \mathcal{B}(\mathcal{X}) \) with density \( \tilde{\pi}_\theta \). Let \( (\Theta_0, \Theta_1, \ldots) \) be a sequence of \( \mathcal{Y} \)-valued random variables and consider the family of Markov chain transition operators of the independent Metropolis-Hastings sampler of \( \Pi \) given \( \tilde{\Pi}_{\Theta_n} \) with transition kernels \( K_{\Theta_n} \). Suppose that for all \( \delta > 0 \), there exists \( M \equiv M(\delta) \in [1, \infty) \) such that

\[
\Pr \left[ \log \pi(x) - \log \tilde{\pi}_{\Theta_n}(x) < \log M \ \forall \ x \in \mathcal{X} \right] \geq 1 - \delta,
\]

(G.10)

for all \( n \in \mathbb{N} \). Then, \( (\Theta_n)_{n \in \mathbb{N}} \) exhibits containment.

A proof is given in section F.A.6. An even stronger condition than eq. (G.10) is that \( \Pr \left[ |\log \pi(x) - \log \tilde{\pi}_{\Theta_n}(x)| < \log M \ \forall \ x \in \mathcal{X} \right] \geq 1 - \delta \). Thus, we see that containment can be obtained for the transition kernels of the independent Metropolis-Hastings sampler if, for every \( n \), \( \log \tilde{\pi}_{\Theta_n} \) is within \( \log M \) of \( \log \pi \) with probability \( 1 - \delta \). Note that \( M \) does not need to even be close to unity (equivalently, \( \log M \) need not be close to zero) in order for containment to hold; it is sufficient merely that, with high probability, the sequence \( (\Theta_n)_{n \in \mathbb{N}} \) does not produce arbitrarily poor approximations of \( \log \pi \).

The loss functions used in estimating normalizing flows are chosen to encourage closeness of the approximation and the target density. For instance, if one chooses to minimize \( \text{KL}(\tilde{\pi}_\theta || \pi) \) as a function of \( \theta \in \mathcal{Y} \) then \( \text{KL}(\tilde{\pi}_\theta || \pi) = 0 \iff \tilde{\pi}_\theta = \pi \). The minimization of a loss function that encourages the closeness of the approximation and the target density is certainly no guarantee that eq. (G.10) holds; however, it gives an indication that eq. (G.10) might be true. We turn our attention in the next section to the empirical evaluation of adaptive samplers using normalizing flows. Some obstacles that could prevent the conditions of proposition G.4.4 from holding are stated in section F.A.13.
Recently, [Gabrié et al., 2021] proposed to sample from Boltzmann distributions and posteriors over the parameters of physical systems by alternating between an independence Metropolis-Hastings algorithm whose proposal is represented as a RealNVP normalizing flow and local updates computed by the Metropolis-adjusted Langevin algorithm (MALA). In [Gabrié et al., 2021] the authors “demonstrate the importance of initializing the training with some a priori knowledge of the relevant modes.” This incorporation of prior knowledge is done to avert mode-collapse. We can connect knowledge of modes to the property of containment: by ensuring that the proposal density of the independent Metropolis-Hastings sampler places sufficient mass on all modes with high probability, one satisfies containment by proposition G.4.4. The specific training procedure used by these samplers is to adapt parameters of the normalizing flow as

\[ \Theta_{n+1} = \Theta_n + \epsilon \frac{1}{n} \sum_{i=0}^{n} \nabla \log \bar{\pi}_{\Theta_n}(X_i) \]

where \((X_i)_{i=0}^{n}\) are the states of the chain to the \(n\)th step and \((\epsilon_i)_{i=0}^{\infty}\) are a sequence of adaptation step-sizes. Because the states of the chain can only be regarded as approximate samples from the target distribution, we understand this update as seeking to update a “pseudo-likelihood.” Diminishing adaptation of this procedure can be enforced using either lemma G.2.10 or via convergence and continuity using lemma G.4.2. When diminishing adaptation and containment are satisfied, this adaptative algorithm produces an ergodic chain by theorem G.2.13.

### G.5 Experiments

Here we evaluate the adaptive independent Metropolis-Hastings algorithm following the “pseudo-likelihood objective”, with non-independent adaptations, summarized in algorithm 23 in section F.A.14. For the normalizing flow, we use RealNVP coupling layers separated by permutation layers. As a baseline adaptive MCMC technique, we consider the random walk Metropolis method of Haario et al. [2001]; we also compare against Langevin dynamics. To assess the ergodicity of samplers, we compare MCMC samples...
against analytic samples drawn from the target density, except in the case of the physical system wherein we use domain knowledge to compare against Langevin dynamics. Specifically, we choose 10,000 random unit vectors and project the samples of the adaptive chain onto the vector space spanned by the chosen unit vector; we then compare these one-dimensional quantities to the projection of the baseline samples from the target distribution and compute the two-sample Kolmogorov-Smirnov (KS) test statistic \cite{Smirnov1948, CuestaAlbertos2006}. In section \ref{sec:finite-time-degradation}, we show how adaptation can actually degrade sample quality at finite time.

\section*{G.5.1 Affine Flows in a Brownian Bridge}

We consider sampling from a Gaussian process with the following mean and covariance functions: $\mu(t) = \sin(\pi t)$ and $\Sigma(t, s) = \min(t, s) - st$. For $0 < t, s < 1$, covariance function identifies this distribution as a Brownian bridge whose mean is a sinusoid. We seek to sample this Gaussian process at 50 equally spaced times in the unit interval, yield-
ing a fifty-dimensional target distribution. We estimate an affine normalizing flow from a Gaussian base distribution in order to sample from the target. Since the base distribution of the flow is Gaussian, and since affine transformations of Gaussian random variables remain Gaussian, in addition to the pseudo-likelihood training objective, we also consider gradient descent on the exact KL divergence between the target and the current proposal distribution. Minimization of the exact KL divergence is equivalent to maximum likelihood training, and therefore allows us to compare the efficiency lost by training with the pseudo-likelihood objective compared to the true likelihood. To enforce diminishing adaptation, we set a learning rate schedule for the gradient steps on the shift and scale of the affine transformation that converges to zero. In addition to Langevin dynamics, we also consider a preconditioned variant of the Metropolis-adjusted Langevin algorithm that uses the Hessian of the log-density to adapt proposals to the geometry of the target distribution [Girolami and Calderhead, 2011]. Results shown in fig. G.3 demonstrate the advantages of the adaptive independent Metropolis-Hastings samplers.

G.5.2 Two-Dimensional Examples

We use a RealNVP architecture to model a multimodal distribution and Neal’s funnel distribution, both in $\mathbb{R}^2$. The multimodal density is a mixture of two Gaussians with a shared covariance structure given by $\Sigma = \text{diag}(1/100, 1/100)$. The two means of the component Gaussians are $(-2, 2)$ and $(2, -2)$. Neal’s funnel distribution is defined by generating $v \sim \text{Normal}(0, 9)$ and $x \sim \text{Normal}(0, e^{-v})$, which defines a distribution in $\mathbb{R}^2$. To enforce diminishing adaptation, we set a learning rate schedule for the gradient steps on parameters of the RealNVP bijections that converges to zero. Results are shown in fig. G.4. The expressivity of the RealNVP normalizing flow is key to building efficient proposals accommodating for different modes or the challenging structure of the Neal’s funnel.
Figure G.4: Examination of the performance of MCMC methods on sampling from the multimodal and Neal funnel distributions. Both adaptive methods enjoy increasing acceptance rates in the multimodal distribution as a function of sampling iteration, but only the adaptive independent Metropolis-Hastings algorithm exhibits ergodicity for this distribution. Indeed, for the adaptive random walk and Langevin sampling methods, which are based on local updates, the multimodal distribution poses distinct challenges. In fact, both methods get stuck in one of the modes. By contrast, the adaptive independent Metropolis-Hastings samplers exhibit the best ergodicity of all methods considered. In Neal’s funnel distribution, the adaptive independent Metropolis-Hastings algorithm possesses the best ergodicity.

Figure G.5: Results of the $\phi^4$ field experiment. As Langevin dynamics is unable to mix between the two modes, the better ergodicity of the independent Metropolis Hastings algorithm is reflected in Kolmogorov-Smirnov statistics as expected. The single chain Langevin has poorer ergodicity than its parallel chain equivalent, while for the I.M.H. a single chain approaches the ergodicity of the parallel setting. The Effective Sample Sizes are reported for chains of 1000 steps extracted at burn-in, after $4 \times 10^4$ iterations (early) and (late) when the NF proposal acceptance probability has reached 50%. Note that periodic jumps in acceptance correspond to iterations where learning rate was decreased.

G.5.3 Analysis of a Physical Field System

We finally revisit a high-dimensional bi-modal example: the 1d-$\phi^4$ system studied in Gabrié et al. [2021]. The statistics of a field $\phi : [0, 1] \rightarrow \mathbb{R}$ are given by the Boltzmann
weight $e^{-\beta U}$ with the energy functional

$$U(\phi) = \int_0^1 \left[ \frac{a}{2} (\partial_s \phi)^2 + \frac{1}{4a} \left( 1 - \phi^2(s) \right)^2 \right] ds,$$

assuming boundary conditions $\phi(0) = \phi(1) = 0$. We discretize the field at 100 equally spaced locations between 0 and 1, for $a = 0.1$ and $\beta = 20$. States example are plotted in fig. F.A.2 of section F.A.16. The algorithm proposed in Gabrié et al. [2021] is adapted with a learning rate schedule enforcing diminishing adaptations and a mixture transition kernel stochastically choosing from local Langevin updates or proposal sampling from the normalizing flow (section F.A.10 shows that we can expect this mixture kernel to exhibit containment and diminishing adaption). Because the distribution is high-dimensional and multimodal, it is necessary to run multiple parallel walkers initialized around the different modes. In this specific case, the energy and the distributions are even functions of $\phi$. In the experiments, we initialize 100 walkers with uneven proportions in each mode (20-80) and test for the ergodicity of the parallel chains. Results are shown in fig. G.5. Unlike the adaptive independent Metropolis-Hastings samplers, MALA single walkers are stuck in the mode they were initialized in and cannot recover the correct equal weights of the positive and negative mode. Additional details can be found in section F.A.16.

G.6 Conclusion

We have examined the question of when an adaptive independent Metropolis-Hastings sampler can be equipped with an ergodicity theory. We specifically consider the case wherein the proposal distribution is parameterized as a normalizing flow. We have considered the cases of deterministic adaptations, independent adaptations, and non-independent

\footnote{This necessity can be lifted by employing an auxiliary fixed set of “training samples” featuring the two modes, in arbitrary proportions. These samples would drive the learning towards relevant regions, so that a random walker can then inform the adaptation about the relative statistical weights of different modes.}
adaptations. For the non-independent adaptations case, we examine mechanisms by which to enforce the diminishing adaptation and containment conditions that together imply ergodicity. On compact spaces, a stronger condition, simultaneous uniform ergodicity, can be established by mixing the normalizing flow with a fixed distribution whose support matches the support of the target distribution.

**Acknowledgments**

M. G. would like to thank G. Rotskoff and E. Vanden-Eijnden for useful discussions about the physical field system experiments. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1752134. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation. MAB was supported by an NSERC Discovery Grant and as part of the Vision: Science to Applications program, thanks in part to funding from the Canada First Research Excellence Fund.
Appendix F.A

Appendices to Chapter G

F.A.1 Review of Total Variation Distance

Similarity of probability measures can be assessed with respect to several criteria. A ubiquitous notion of distance between probability measures is given by the total variation norm of their difference.

Definition F.A.1.1. Let $\nu_1$ and $\nu_2$ be probability measures on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. Then the total variation distance between $\nu_1$ and $\nu_2$ is defined by,

$$\|\nu_1(\cdot) - \nu_2(\cdot)\|_{TV} = 2 \sup_{A \in \mathcal{B}(\mathcal{X})} |\nu_1(A) - \nu_2(A)|.$$ (F.A.1)

The total variation distance is easily verified to be a proper distance in that it satisfies non-negativity, discernability, symmetry, and the triangle inequality. The total variation distance can therefore be understood as the largest possible disagreement between the probabilities assigned to any measurable set by $\nu_1$ and $\nu_2$. The total variation norm has the following equivalent representations which are occasionally useful.

Proposition F.A.1.2. Within the context of definition [F.A.1.1] the total variation distance
between $\nu_1$ and $\nu_2$ is equivalently expressed as,

$$\|\mu\|_{TV} = \sup_{f \in \mathcal{M}} \left| \int_X f(x) \, \nu_1(dx) - \int_X f(x) \, \nu_2(dx) \right| \quad (F.A.2)$$

where $\mathcal{M} \overset{\text{def}}{=} \{ f : X \to \mathbb{R} \text{ s.t. } |f(x)| \leq 1 \ \forall \ x \in X \}$.

For a proof of this result, and other equivalent characterizations of the total variation distance, we refer the interested reader to Roberts and Rosenthal [2004], Pollard [2001].

**F.A.2 Proofs Concerning Deterministic Adaptations**

**Proposition F.A.2.1** (Roberts and Rosenthal [2007]). Suppose that $(\theta_0, \theta_1, \ldots)$ is a deterministic $\mathcal{Y}$-valued sequence. Let $(K_{\theta_n})_{n \in \mathbb{N}}$ be an associated sequence of Markov transition kernels. If $\Pi$ is stationary for each $K_{\theta_n}$, then $\Pi$ is also the stationary distribution of the Markov chain whose transitions satisfy $X_{n+1} \sim K_{\theta_n}(X_n, \cdot)$.

**Proof.** Let $A \in \mathcal{B}(X)$ and suppose $X_n \sim \Pi$. We will show $X_{n+1} \sim \Pi$.

\[
\Pr \left[ X_{n+1} \in A \right] = \int_X \Pr \left[ X_{n+1} \in A | X_n = x \right] \cdot \Pr \left[ X_n \in dx \right] \quad (F.A.3)
\]

\[
= \int_X \int_Y \Pr \left[ X_{n+1} \in A | X_n = x, \Theta_n = \theta \right] \cdot \delta_{\theta_n}(d\theta) \cdot \Pr \left[ X_n \in dx \right] \quad (F.A.4)
\]

\[
= \int_X K_{\theta_n}(x, A) \cdot \Pr \left[ X_n \in dx \right] \quad (F.A.5)
\]

\[
= \mathbb{E}_{x \sim \Pi} \left[ K_{\theta_n}(x, A) \right] \quad (F.A.6)
\]

\[
= \Pi(A). \quad (F.A.7)
\]

$\square$
Lemma F.A.2.2. For each $n \in \mathbb{N}$,

$$K_{\theta_n}(x, dx') \geq \frac{\pi_{\Pi}(x')}{M_n} \mu(dx'). \quad \text{(F.A.8)}$$

Proof. From eq. (G.2),

$$K_{\theta_n}(x, dx') \geq \min \left\{ 1, \frac{\pi(x')}{\pi(x)} \right\} \tilde{\pi}_{\theta_n}(x') \mu(dx') \quad \text{(F.A.9)}$$

$$= \min \left\{ \tilde{\pi}_{\theta_n}(x'), \frac{\pi(x')}{\pi_n} \tilde{\pi}_{\theta_n}(x) \right\} \mu(dx') \quad \text{(F.A.10)}$$

$$\geq \min \left\{ \tilde{\pi}_{\theta_n}(x'), \frac{\pi(x')}{M_n} \right\} \mu(dx') \quad \text{(F.A.11)}$$

$$= \frac{\pi(x')}{M_n} \mu(dx') \quad \text{(F.A.12)}$$

$$\square$$

Corollary F.A.2.3. For any set $A \in \mathcal{B}(\mathcal{X})$

$$K_{\theta_n}(x, A) \geq \frac{1}{M_n} \Pi(A). \quad \text{(F.A.13)}$$

Proof. Integrate both sides of eq. (F.A.8) over the set $A$. \(\square\)

Proof of theorem G.4.1 From corollary F.A.2.3 it follows that we may express the transition kernel at step $n \in \mathbb{N}$ as

$$K_{\theta_n}(x, A) = \frac{1}{M_n} \Pi(A) + \left( 1 - \frac{1}{M_n} \right) \frac{K_{\theta_n}(x, A) - \frac{1}{M_n} \Pi(A)}{1 - \frac{1}{M_n}} \quad \text{(F.A.14)}$$

$$= \frac{1}{M_n} \Pi(A) + \left( 1 - \frac{1}{M_n} \right) \tilde{K}_{\theta_n}(x, A), \quad \text{(F.A.15)}$$

where $\tilde{K}_{\theta_n}(x, A)$ is another probability measure. With eq. (F.A.15), $K_{\theta_n}$ may be given the following interpretation: With probability $1/M_n$ generate the next state by the dis-
distribution \( \Pi \) and with probability \( 1 - 1/M_n \) generate the next state from the distribution \( \tilde{K}_{\theta_i} \). Given an initial state \( X_0 = x_0 \), consider the Markov chain whose transitions are generated according to \( X_{n+1} \sim K_{\theta_n}(X_n, \cdot) \) with marginal laws \( X_n \sim K^n(x_0, \cdot) \).

From the representation in eq. (F.A.15) and proposition [F.A.2.1], it follows that the total variation distance is zero as soon as we generate the next state from \( \Pi \). Let \( T \) be the random variable representing the first step at which \( X_n \) is generated from \( \Pi \). Then

\[
K^n(x_0, \cdot) = \Pr [T \leq n] \Pi(\cdot) + \Pr [T > n] \tilde{K}^n(x_0, \cdot),
\]

where \( \tilde{K}^n \) is the mixture component of \( K^n \) that is possibly not \( \Pi \). Thus,

\[
\|K^n(x_0, \cdot) - \Pi(\cdot)\|_{TV} = \|\Pr [T \leq n] \Pi(\cdot) + \Pr [T > n] \tilde{K}^n(x_0, \cdot) - \Pi(\cdot)\|_{TV} \quad \text{(F.A.16)}
\]

\[
= \|\Pr [T > n] \tilde{K}^n(x_0, \cdot) - \Pr [T > n] \Pi(\cdot)\|_{TV} \quad \text{(F.A.17)}
\]

\[
= \Pr [T > n] \cdot \|\tilde{K}^n(x_0, \cdot) - \Pi(\cdot)\|_{TV} \quad \text{(F.A.18)}
\]

\[
\leq 2 \prod_{i=0}^{n-1} \left( 1 - \frac{1}{M_i} \right), \quad \text{(F.A.19)}
\]

since \( T > n \) only if we generate the next state from \( \tilde{K}_{\theta_i} \) for \( i = 1, \ldots, n-1 \), each of which occurs with probability \( 1 - 1/M_i \).

**Proposition F.A.2.4.** Suppose that \((\theta_0, \theta_1, \ldots)\) is a deterministic \( \mathcal{Y} \)-valued sequence. Let \((K_{\theta_n})_{n \in \mathbb{N}}\) be an associated sequence of Markov transition kernels. If \( \Pi \) is stationary for each \( K_{\theta_n} \), then

\[
\|K^{n+1}(x_0, \cdot) - \Pi(\cdot)\|_{TV} \leq \|K^n(x_0, \cdot) - \Pi(\cdot)\|_{TV}, \quad \text{(F.A.20)}
\]

where \( K^n \) is defined in eq. (G.6).
Proof.

\[ \| K^{n+1}(x_0, \cdot) - \Pi(\cdot) \|_{TV} = \sup_{f \in M} \left| \int_{\mathcal{X}} f(y) K^{n+1}(x_0, dy) - \int_{\mathcal{X}} f(y) \Pi(dy) \right| \]  
(F.A.21)

\[ = \sup_{f \in M} \left| \int_{\mathcal{X}} f(y) K^{n+1}(x_0, dy) - \int_{\mathcal{X}} f(y) \int_{\mathcal{X}} K_{\theta_n}(w, dy) \Pi(dw) \right| \]  
(F.A.22)

\[ = \sup_{f \in M} \left| \int_{\mathcal{X}} f(y) \int_{\mathcal{X}} K^n(x_0, dw) K_{\theta_n}(w, dy) - \int_{\mathcal{X}} f(y) \int_{\mathcal{X}} K_{\theta_n}(w, dy) \Pi(dw) \right| \]  
(F.A.23)

\[ \leq \sup_{f \in M} \left| \int_{\mathcal{X}} f(w) K^n(x_0, dw) - \int_{\mathcal{X}} f(w) \Pi(dw) \right| \]  
(F.A.24)

\[ = \| K^n(x_0, \cdot) - \Pi(\cdot) \|_{TV}. \]  
(F.A.26)

Definition F.A.2.5. Let \( \Pi \) be a probability measure with density \( \pi \). Suppose that every \( \theta \in \mathcal{Y} \) parameterizes a probability measure \( \tilde{\Pi}_\theta \) on \( \mathcal{B}(\mathcal{X}) \) with density \( \tilde{\pi}_\theta \). Define the ergodic set of \( \Pi \) given \( \mathcal{Y} \) as

\[ Q = \{ \theta \in \mathcal{Y} : \text{there exists } M_\theta < \infty \text{ such that } \pi(x) \leq M_\theta \tilde{\pi}_\theta(x) \ \forall \ x \in \text{Supp}(\pi) \}. \]  
(F.A.27)

The combination of proposition F.A.2.4 and definition F.A.2.5 allows one to give a slight generalization of theorem G.4.1

Corollary F.A.2.6. Let \( \Pi \) be a probability measure with density \( \pi \). Suppose that every \( \theta \in \mathcal{Y} \) parameterizes a probability measure \( \tilde{\Pi}_\theta \) on \( \mathcal{B}(\mathcal{X}) \) with density \( \tilde{\pi}_\theta \). Suppose that \((\theta_0, \theta_1, \ldots)\) is a deterministic \( \mathcal{Y} \)-valued sequence. Let \((K_{\theta_n})_{n \in \mathbb{N}}\) be an associated sequence of Markov transition kernels of the independent Metropolis-Hastings sampler of \( \Pi \) given...
Let $K^n(x_0, A)$ denote the $n$-step transition probability from $x_0$ to $A \in \mathcal{B}(\mathcal{X})$. Then

$$\|K^n(x_0, \cdot) - \Pi(\cdot)\|_{TV} \leq 2 \prod_{i=0}^{n-1} (1 - L_i),$$

where

$$L_i = \begin{cases} \frac{1}{M_i} & \text{if } \theta_i \in Q \\ 0 & \text{otherwise.} \end{cases}$$

**Proof.** The proof proceeds by induction. If $\theta_0 \in Q$ then by argument in the proof of theorem G.4.1 we have

$$\|K^1(x_0, \cdot) - \Pi(\cdot)\|_{TV} = \|K_{\theta_0}(x_0, \cdot) - \Pi(\cdot)\|_{TV} \leq 2 \left(1 - \frac{1}{M_0}\right).$$

If $\theta_0 \notin Q$ then we obtain the vacuously true inequality $\|K^1(x_0, \cdot) - \Pi(\cdot)\|_{TV} \leq 1$. Now assume that $\|K^n(x_0, \cdot) - \Pi(\cdot)\|_{TV} \leq \prod_{i=0}^{n-1} (1 - L_i)$. If $\theta_n \notin Q$ then by proposition F.A.2.4 we have

$$\|K^{n+1}(x_0, \cdot) - \Pi(\cdot)\|_{TV} \leq \|K^n(x_0, \cdot) - \Pi(\cdot)\|_{TV} \leq 2 \prod_{i=0}^{n-1} (1 - L_i) \leq 2 \prod_{i=0}^{n-1} \left(1 - L_i \cdot 1\right),$$

since $L_n = 0$. On the other hand, if $\theta_n \in Q$ then, using the same argument as in the proof...
of theorem G.4.1 the probability that none of \((X_0, \ldots, X_n)\) were drawn from \(\Pi\) is at most
\[
\prod_{i=0}^{n-1} (1 - L_i). \tag{F.A.36}
\]

Correspondingly, the probability that \(X_{n+1}\) is also not drawn from \(\Pi\) is \(1 - 1/M_n\) so that the probability that none of \((X_0, \ldots, X_n, X_{n+1})\) is drawn from \(\Pi\) is at most
\[
\left(1 - \frac{1}{M_n}\right)^{n-1} \prod_{i=0}^{n} (1 - L_i). \tag{F.A.37}
\]

From this the conclusion follows. \(\square\)

## F.A.3 Proofs Concerning Independent Adaptations

**Theorem F.A.3.1.** Let \(\Pi\) be a probability measure with density \(\pi\). Suppose that every \(\theta \in \mathcal{Y}\) parameterizes a probability measure \(\tilde{\Pi}_\theta\) on \(\mathcal{B}(\mathcal{X})\) with density \(\tilde{\pi}_\theta\). Suppose that \((\Theta_0, \Theta_1, \ldots)\) is a stochastic \(\mathcal{Y}\)-valued sequence. Let \((K_{\Theta_n})_{n \in \mathbb{N}}\) be an associated sequence of Markov transition kernels of the independent Metropolis-Hastings sampler of \(\Pi\) given \(\tilde{\Pi}_{\Theta_n}\). Suppose that \(X_n\) and \(\Theta_n\) are independent given the history of the chain to step \(n - 1\).

Let \(G^n(x_0, \cdot)\) be the associated marginal transition law. Then
\[
\|G^n(x_0, \cdot) - \Pi(\cdot)\|_{TV} \leq 2 \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} \left[ \prod_{i=0}^{n-1} (1 - L_i) \right] \tag{F.A.38}
\]
where \(L_i = 1/M_i\) if \(M_i < \infty\) and otherwise \(L_i = 0\).

A proof is given in section F.A.3. This result was previously demonstrated in [Holden et al.][2009], though we have offered a different proof procedure.

**Example 46.** Let \(\Pi\) be a probability measure with density \(\pi\). Let \(\mathcal{Y} = \mathbb{R}^m\) and suppose that every \(\theta \in \mathcal{Y}\) smoothly parameterizes a probability measure \(\tilde{\Pi}\) on \(\mathcal{B}(\mathcal{X})\) with density
\(\pi_\theta\) for which \(\text{Supp}(\pi) = \text{Supp}(\pi_\theta)\). Consider the sequence of updates,

\[
\theta_n = \theta_{n-1} - \epsilon \nabla_\theta \left( \frac{1}{s} \sum_{i=1}^{s} \log \frac{\tilde{\pi}_{\theta_{n-1}}(Y_s(\theta_{n-1}))}{\pi(Y_s(\theta_{n-1}))} \right) \quad \text{(F.A.39)}
\]

where \(Y_1, \ldots, Y_s \overset{\text{i.i.d.}}{\sim} \tilde{\Pi}_{\theta_{n-1}}\). This corresponds to the stochastic gradient approximation of example 43. Consider the family of Markov chain transition operators of the independent Metropolis-Hastings sampler of \(\Pi\) given \(\tilde{\Pi}_{\theta_n}\) with transition kernels \(K_{\theta_n}\). Then \(\Pi\) is the stationary distribution of the Markov chain whose transitions satisfy \(X_{n+1} \sim K_{\theta_n}(X_n, \cdot)\). To see this, let \(\tilde{X}\) be a sample from \(\tilde{\Pi}_{\theta_{n-1}}\) independent of \((Y_1, \ldots, Y_s)\) and let \(U \sim \text{Uniform}(0, 1)\) be independent of both. Then \(X_n = g(x_{n-1}, \theta_{n-1}, \tilde{X}, U)\) where \(g\) is given by,

\[
g(x, \theta, \tilde{x}, u) = \begin{cases} 
\tilde{x} & \text{if } u < \min \left\{ 1, \frac{\pi(\tilde{x})\tilde{\pi}_\theta(x)}{\pi(x)\tilde{\pi}_\theta(\tilde{x})} \right\} \\
x & \text{otherwise}
\end{cases} \quad \text{(F.A.40)}
\]

and \(\Theta_n = f(\theta_{n-1}, Y_1, \ldots, Y_s)\) where \(f\) is given by,

\[
f(\theta, y_1(\theta), \ldots, y_s(\theta)) = \theta - \epsilon \nabla_\theta \left( \frac{1}{s} \sum_{i=1}^{s} \log \frac{\tilde{\pi}_\theta(y_s(\theta))}{\pi(y_s(\theta))} \right). \quad \text{(F.A.41)}
\]

By lemma F.A.3.2, \(\Theta_n\) and \(X_n\) are independent given the history of the chain to step \(n-1\) and therefore, by proposition F.A.11.1, \(\Pi\) is the stationary distribution.

**Lemma F.A.3.2.** Suppose that \((X_1^{(a)}, \ldots, X_r^{(a)})\) and \((X_1^{(b)}, \ldots, X_s^{(b)})\) are two sets of random variables which are independent given the history of the chain to step \(n-1\). Suppose that \(\Theta_n = f(x_{n-1}, \theta_{n-1}, X_1^{(a)}, \ldots, X_r^{(a)})\) and \(X_n = g(x_{n-1}, \theta_{n-1}, X_1^{(b)}, \ldots, X_s^{(b)})\) for two functions \(f\) and \(g\). Then \(X_n\) and \(\Theta_n\) are independent given the history of the chain to step \(n-1\).

**Proof.** The \(\sigma\)-algebra generated by \(\Theta_n\) is a subset of the \(\sigma\)-algebra generated by \((X_1^{(a)}, \ldots, X_r^{(a)})\).
Likewise, the $\sigma$-algebra generated by $X_n$ is a subset of the $\sigma$-algebra generated by $(X_1^{(b)}, \ldots, X_s^{(b)})$. Since $(X_1^{(a)}, \ldots, X_r^{(a)})$ and $(X_1^{(b)}, \ldots, X_s^{(b)})$ are assumed independent given the history of the chain to step $n-1$, the conclusion follows.

**Proof of theorem [F.A.3.1]**

$$
\|G^n(x_0, \cdot) - \Pi(\cdot)\|_{TV} = \| \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, \cdot) - \Pi(\cdot)\|_{TV} \tag{F.A.42}
$$

$$
= 2 \sup_{A \in \mathcal{B}(X)} \left| \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, A) - \Pi(A) \right| \tag{F.A.43}
$$

$$
\leq \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} 2 \sup_{A \in \mathcal{B}(X)} \left| K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, A) - \Pi(A) \right| \tag{F.A.44}
$$

$$
= \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} \| K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, \cdot) - \Pi(\cdot)\|_{TV} \tag{F.A.45}
$$

$$
\leq 2 \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} \prod_{i=0}^{n-1} (1 - L_i) \tag{F.A.46}
$$

where the first inequality can be deduced as follows: By Jensen’s inequality,

$$
\left| \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, A) - \Pi(A) \right| \leq \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} \left| K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, A) - \Pi(A) \right|. \tag{F.A.47}
$$

Moreover,

$$
\left| K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, A) - \Pi(A) \right| \leq \sup_{A \in \mathcal{B}(X)} \left| K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, A) - \Pi(A) \right| \tag{F.A.48}
$$

$$
\implies \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} \left| K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, A) - \Pi(A) \right| \leq \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} \sup_{A \in \mathcal{B}(X)} \left| K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, A) - \Pi(A) \right| \tag{F.A.49}
$$

$$
\implies \sup_{A \in \mathcal{B}(X)} \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} \left| K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, A) - \Pi(A) \right| \tag{F.A.50}
$$

$$
\leq \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} \sup_{A \in \mathcal{B}(X)} \left| K^n_{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, A) - \Pi(A) \right| \tag{F.A.50}
$$
The second inequality follows from corollary [F.A.2.6] as follows:

\[
\| K_n^{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, \cdot) - \Pi(\cdot) \|_{TV} \leq 2 \prod_{i=0}^{n-1} (1 - L_i) \tag{F.A.51}
\]

\[
\implies \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} \| K_n^{(\Theta_0, \ldots, \Theta_{n-1})}(x_0, \cdot) - \Pi(\cdot) \|_{TV} \leq 2 \mathbb{E}_{(\Theta_0, \ldots, \Theta_{n-1})} \left[ n-1 \prod_{i=0}^{n-1} (1 - L_i) \right]. \tag{F.A.52}
\]

F.A.4  Proofs Concerning Continuity of Independent Metropolis-Hastings Transition Kernels

**Theorem F.A.4.1.** Let \( f \) be a continuous function from the metric space \((X, d_X)\) to the metric space \((Y, d_Y)\). If \((X_0, X_1, \ldots)\) is a sequence of \(X\)-valued random variables converging in probability to the random variable \(X\) then \((f(X_0), f(X_1), \ldots)\) is a sequence of \(Y\)-valued converging in probability to \(f(X)\).

**Proof of lemma G.4.2.** Given that \( \theta \mapsto K_\theta \) is continuous, if \((\Theta_0, \Theta_1, \ldots)\) converges in probability to \(\Theta\), we have immediately from theorem [F.A.4.1] that \((K_{\Theta_0}, K_{\Theta_1}, \ldots)\) converges in probability to \(K_{\Theta}\). This means that for all \(\epsilon > 0\) and \(\delta > 0\), there exists \(N(\epsilon, \delta) \in \mathbb{N}\) such that \(\Pr [d(K_{\Theta_n}, K_{\Theta}) < \epsilon] \geq 1 - \delta\) for every \(n \geq N\). For fixed \(\epsilon > 0\) and...
δ > 0, set \( n \geq N(\epsilon/2, \delta) \) so that \( \Pr \left[ d(K_{\Theta_n}, K_{\Theta}) < \epsilon/2 \right] \geq 1 - \delta \). Thus,

\[
\Pr \left[ d(K_{\Theta_n}, K_{\Theta_{n+1}}) < \epsilon \right] \geq \Pr \left[ d(K_{\Theta_n}, K_{\Theta}) + d(K_{\Theta_{n+1}}, K_{\Theta}) < \epsilon \right]
\]

\[
\geq \Pr \left[ d(K_{\Theta_n}, K_{\Theta}) < \epsilon/2 \text{ and } d(K_{\Theta_{n+1}}, K_{\Theta}) < \epsilon/2 \right]
\]

\[
\geq \Pr \left[ d(K_{\Theta_n}, K_{\Theta}) < \epsilon/2 \right] + \Pr \left[ d(K_{\Theta_{n+1}}, K_{\Theta}) < \epsilon/2 \right] - 1
\]

\[
\geq 1 - \delta + 1 - \delta - 1
\]

\[
= 1 - 2\delta.
\]

This establishes diminishing adaptation in the sense of definition G.2.9.

**Proof of proposition G.2.4** To prove symmetry we write,

\[
d(K, K') = \sup_{x \in X} \| K(x, \cdot) - K'(x, \cdot) \|_{TV}
\]

\[
= \sup_{x \in X} \| K'(x, \cdot) - K(x, \cdot) \|_{TV}
\]

\[
= d(K', K).
\]

Identifiability follows from the definition of equality of Markov chain kernels given in definition G.2.3. The triangle inequality is then proven as follows. Let \( K'' \) be another
transition kernel on $\mathcal{X} \times \mathcal{B}(\mathcal{X})$.

\[
d(K, K') = \sup_{x \in \mathcal{X}} \| K(x, \cdot) - K'(x, \cdot) \|_{TV}
\]

\[
\leq \sup_{x \in \mathcal{X}} (\| K(x, \cdot) - K''(x, \cdot) \|_{TV} + \| K''(x, \cdot) - K'(x, \cdot) \|_{TV}) \tag{F.A.62}
\]

\[
\leq \sup_{x \in \mathcal{X}} \| K(x, \cdot) - K''(x, \cdot) \|_{TV} + \sup_{x \in \mathcal{X}} \| K''(x, \cdot) - K'(x, \cdot) \|_{TV}
\]

\[
= d(K, K'') + d(K'', K'). \tag{F.A.64}
\]

In the sequel, we will limit our discussion to the transition kernel of the independent Metropolis-Hastings sampler. Recall that this transition kernel has the following form,

\[
K_{\theta}(x, A) = \int_A \alpha_{\theta}(x, y) \pi_{\theta}(y) \mu(\mathrm{d}y) + \left(1 - \int_{\mathcal{X}} \alpha_{\theta}(x, w) \tilde{\pi}_{\theta}(w) \mu(\mathrm{d}w)\right) 1 \{x \in A\}, \tag{F.A.65}
\]

where

\[
\alpha_{\theta}(x, y) = \min \left\{ 1, \frac{\pi(y) \tilde{\pi}_{\theta}(x)}{\pi(x) \tilde{\pi}_{\theta}(y)} \right\}. \tag{F.A.66}
\]

**Lemma F.A.4.2.** Let $(\theta_1, \theta_2, \ldots)$ be a $\mathcal{Y}$-valued sequence converging to $\theta$. If for all $x \in \mathcal{X}$ and $A \in \mathcal{B}(\mathcal{X})$ we have

\[
\lim_{n \to \infty} \int_A \alpha_{\theta_n}(x, y) \pi_{\theta_n}(y) \mu(\mathrm{d}y) = \int_A \lim_{n \to \infty} [\alpha_{\theta_n}(x, y) \pi_{\theta_n}(y)] \mu(\mathrm{d}y) \tag{F.A.67}
\]

\[
= \int_A \alpha_{\theta}(x, y) \tilde{\pi}_{\theta}(y) \mu(\mathrm{d}y). \tag{F.A.68}
\]

then $\lim_{n \to \infty} K_{\theta_n} = K_{\theta}$. 

542
Proof. By continuity of the distance function,

$$\lim_{n \to \infty} d(K_{\theta_n}, K_\theta) = d( \lim_{n \to \infty} K_{\theta_n}, K_\theta)$$ \hfill (F.A.69)

$$= \sup_{x \in \mathcal{X}} \sup_{A \in \mathcal{B}(\mathcal{X})} \left| \lim_{n \to \infty} K_{\theta_n}(x, A) - K_\theta(x, A) \right|. \hfill (F.A.70)$$

Therefore,

$$\lim_{n \to \infty} K_{\theta_n}(x, A) = \lim_{n \to \infty} \left( \int A \alpha_\theta_n(x, y) \tilde{\pi}_\theta_n(y) \mu(dy) + \left( 1 - \int \mathcal{X} \alpha_\theta_n(x, w) \tilde{\pi}_\theta_n(w) \mu(dw) \right) \mathbf{1}\{x \in A\} \right)$$ \hfill (F.A.71)

$$= \lim_{n \to \infty} \int A \alpha_\theta_n(x, y) \tilde{\pi}_\theta_n(y) \mu(dy) + \lim_{n \to \infty} \left( 1 - \int \mathcal{X} \alpha_\theta_n(x, w) \tilde{\pi}_\theta_n(w) \mu(dw) \right) \mathbf{1}\{x \in A\}$$ \hfill (F.A.72)

$$= \int A \lim_{n \to \infty} \left[ \alpha_\theta_n(x, y) \tilde{\pi}_\theta_n(y) \right] \mu(dy) + \left( 1 - \int \mathcal{X} \lim_{n \to \infty} \left[ \alpha_\theta_n(x, w) \tilde{\pi}_\theta_n(w) \right] \mu(dw) \right) \mathbf{1}\{x \in A\}$$ \hfill (F.A.73)

$$= \int A \alpha_\theta(x, y) \tilde{\pi}_\theta(y) \mu(dy) + \left( 1 - \int \mathcal{X} \alpha_\theta(x, w) \tilde{\pi}_\theta(w) \mu(dw) \right) \mathbf{1}\{x \in A\}$$ \hfill (F.A.74)

$$= K_\theta(x, A). \hfill (F.A.75)$$

Finally,

$$\lim_{n \to \infty} d(K_{\theta_n}, K_\theta) = \sup_{x \in \mathcal{X}} \sup_{A \in \mathcal{B}(\mathcal{X})} \left| \lim_{n \to \infty} K_{\theta_n}(x, A) - K_\theta(x, A) \right|$$ \hfill (F.A.76)

$$= \sup_{x \in \mathcal{X}} \sup_{A \in \mathcal{B}(\mathcal{X})} |K_\theta(x, A) - K_\theta(x, A)| \hfill (F.A.77)$$

$$= 0. \hfill (F.A.78)$$

The following result is called Scheffé’s lemma; see Lebanon [2017], Pollard [2001].
Lemma F.A.4.3. Let \( \pi_n \) be a sequence of probability densities that converge pointwise to another density \( \pi \). Then, let \( \Pi(A) = \int_A \pi(x) \, \mu(dx) \) and \( \Pi_n(A) = \int_A \pi_n(x) \, \mu(dx) \) be the measures whose densities are \( \pi \) and \( \pi_n \) with respect to dominating measure \( \mu \), respectively. Then \( \lim_{n \to \infty} \| \Pi(\cdot) - \Pi_n(\cdot) \|_{TV} = 0 \).

We will also require the following theorem from [Royden 1968, Page 270].

Theorem F.A.4.4. Let \((X, \mathcal{B}(X))\) be a measurable space and let \((\Pi_n)_{n \in \mathbb{N}}\) be a sequence of probability measures converging to the probability measure \( \Pi \). Let \( \alpha_n : X \to \mathbb{R} \) and \( \beta_n : X \to \mathbb{R} \) be two sequences of functions converging pointwise to the functions \( \alpha \) and \( \beta \), respectively. Suppose further that \( |\alpha_n(x)| \leq \beta_n(x) \) for every \( x \in X \) and that,

\[
\lim_{n \to \infty} \int_X \beta_n(x) \, \Pi_n(dx) = \int_X \beta(x) \, \Pi(dx) < \infty. \tag{F.A.79}
\]

Then,

\[
\lim_{n \to \infty} \int_A \alpha_n(x) \, \Pi_n(dx) = \int_A \alpha(x) \, \Pi(dx), \tag{F.A.80}
\]

for \( A \in \mathcal{B}(X) \).

Lemma F.A.4.5. Suppose that for fixed \( x \in X \) the mapping \( \theta \mapsto \tilde{\pi}_\theta(x) \) is continuous, that \( y \in X \), and that \( \text{Supp}(\tilde{\pi}_\theta) = X \) for every \( \theta \in \mathcal{Y} \). Let \((\theta_1, \theta_2, \ldots)\) be a \( \mathcal{Y} \)-valued sequence converging to \( \theta \). Then \( \lim_{n \to \infty} \alpha_{\theta_n}(x, y) \tilde{\pi}_{\theta_n}(y) = \alpha_\theta(x, y) \tilde{\pi}_\theta(y) \) pointwise.
Proof.

\[
\lim_{n \to \infty} \alpha_{\theta_n}(x, y) = \lim_{n \to \infty} \min_{n \to \infty} \left\{ 1, \frac{\pi(y)\tilde{\pi}_{\theta_n}(x)}{\pi(x)\tilde{\pi}_{\theta_n}(y)} \right\} \tag{F.A.81}
\]

\[
= \left( \lim_{n \to \infty} \min_{n \to \infty} \left\{ 1, \frac{\pi(y)\tilde{\pi}_{\theta_n}(x)}{\pi(x)\tilde{\pi}_{\theta_n}(y)} \right\} \right) \tag{F.A.82}
\]

\[
= \left( \lim_{n \to \infty} \min_{n \to \infty} \left\{ 1, \frac{\pi(y)\tilde{\pi}_{\theta_n}(x)}{\pi(x)\tilde{\pi}_{\theta_n}(y)} \right\} \right) \tag{F.A.83}
\]

\[
= \left( \min_{n \to \infty} \left\{ 1, \frac{\pi(y)}{\pi(x)} \lim_{n \to \infty} \tilde{\pi}_{\theta_n}(x) \right\} \right) \tag{F.A.84}
\]

\[
= \left( \min_{n \to \infty} \left\{ 1, \frac{\pi(y)}{\pi(x)} \lim_{n \to \infty} \tilde{\pi}_{\theta_n}(x) \right\} \right) \tag{F.A.85}
\]

\[
= \left( \min_{n \to \infty} \left\{ 1, \frac{\pi(y)}{\pi(x)} \left( \lim_{n \to \infty} \tilde{\pi}_{\theta_n}(x) \right) \frac{1}{\lim_{n \to \infty} \tilde{\pi}_{\theta_n}(y)} \right\} \right) \tag{F.A.86}
\]

\[
= \left( \min_{n \to \infty} \left\{ 1, \frac{\pi(y)}{\pi(x)} \left( \frac{1}{\lim_{n \to \infty} \tilde{\pi}_{\theta_n}(y)} \right) \right\} \right) \tag{F.A.87}
\]

\[
= \left( \min_{n \to \infty} \left\{ 1, \frac{\pi(y)}{\pi(x)} \frac{1}{\tilde{\pi}_{\theta}(y)} \right\} \right) \tag{F.A.88}
\]

\[
= \alpha_{\theta}(x, y). \tag{F.A.89}
\]

The assumption that \( \text{Supp}(\tilde{\pi}_{\theta}) = X \) is used in eq. \([\text{F.A.87}]\). \blacksquare

**Corollary F.A.4.6.** Let \((\theta_1, \theta_2, \ldots)\) be a \(Y\)-valued sequence converging to \(\theta\). Let \(\pi\) be a probability density function on a compact space \(X\) and let \(\tilde{\pi}_{\theta}\) be a family of density functions on \(X\) indexed by \(\theta\) such that the map \(\theta \mapsto \tilde{\pi}_{\theta}\) is continuous (i.e. \(\pi_{\theta_n} \to \pi_{\theta}\)). Assume further that \(\text{Supp}(\tilde{\pi}_{\theta}) = X\) for every \(\theta \in Y\). Let \(x \in X\) be fixed and let \(y \in X\).

Define

\[
\alpha_{\theta}(x, y) = \min \left\{ 1, \frac{\pi(y)\tilde{\pi}_{\theta}(x)}{\pi(x)\tilde{\pi}_{\theta}(y)} \right\}. \tag{F.A.90}
\]

Then,

\[
\lim_{n \to \infty} \int_A \alpha_{\theta_n}(x, y)\tilde{\pi}_{\theta_n}(y) \mu(dy) = \int_A \alpha_{\theta}(x, y)\tilde{\pi}_{\theta}(y) \mu(dy) \tag{F.A.91}
\]
Proof. This follows immediate from theorem F.A.4.4 with \( \beta_n(y) \equiv 1, \alpha_n(y) = \alpha_{\theta_n}(x, y) \)
(which converges pointwise by lemma F.A.4.5) and the measures \( \Pi_n(A) = \int_A \tilde{\alpha}_n(x) \mu(dx) \)
and \( \Pi(A) = \int_A \tilde{\alpha}(x) \mu(dx) \), which converge by lemma F.A.4.3.

Proof of theorem G.4.3. Fix \( x \in \mathcal{X} \) and \( A \in \mathcal{B}(\mathcal{X}) \).
Thus,

\[
\lim_{n \to \infty} K_{\theta_n}(x, A) = \lim_{n \to \infty} \left( \int_A \alpha_{\theta_n}(x, y) \tilde{\pi}_{\theta_n}(y) \mu(dy) + \left( 1 - \int_{\mathcal{X}} \alpha_{\theta_n}(x, w) \tilde{\pi}_{\theta_n}(w) \mu(dw) \right) 1 \{ x \in A \} \right)
\]

\begin{equation}
\tag{F.A.92}
= \lim_{n \to \infty} \int_A \alpha_{\theta_n}(x, y) \tilde{\pi}_{\theta_n}(y) \mu(dy) + \lim_{n \to \infty} \left( 1 - \int_{\mathcal{X}} \alpha_{\theta_n}(x, w) \tilde{\pi}_{\theta_n}(w) \mu(dw) \right) 1 \{ x \in A \}
\end{equation}

\begin{equation}
\tag{F.A.93}
= \int_A \alpha_\theta(x, y) \tilde{\pi}_\theta(y) \mu(dy) + \left( 1 - \int_{\mathcal{X}} \alpha_\theta(x, w) \tilde{\pi}_\theta(w) \mu(dw) \right) 1 \{ x \in A \}
\end{equation}

\begin{equation}
\tag{F.A.94}
= K_\theta(x, A).
\end{equation}

where we have used corollary F.A.4.6 in eq. (F.A.94). The conclusion then follows from lemma F.A.4.2.

F.A.5 Proofs Concerning Simultaneous Uniform Ergodicity on Compact Spaces

Proof of corollary G.2.7. Because \( \mathcal{X} \) is compact and \( \pi \) and \( \tilde{\pi} \) are continuous, we know
that \( \pi \) and \( \tilde{\pi} \) attain maximum and minimum values on \( \mathcal{X} \). Therefore, the ratio \( \pi(x)/\tilde{\pi}(x) \)
(i) does not diverge on \( \text{Supp}(\pi) \) because \( \text{Supp}(\pi) \subseteq \text{Supp}(\tilde{\pi}) \) and (ii) is bounded by

\[
\frac{\max_{x \in \mathcal{X}} \pi(x)}{\min_{x \in \mathcal{X}} \tilde{\pi}(x)},
\]

\begin{equation}
\tag{F.A.96}
\end{equation}

and \( M \) is at most this value, with equality if the maximum of \( \pi \) and the minimum of \( \tilde{\pi} \)
occurs at the same point in \( \mathcal{X} \).
Proof of proposition [F.A.15.2] Define

\[ M_\theta = \max_{x \in \text{Supp}(\pi)} \frac{\pi(x)}{\tilde{\pi}_\theta(x)} \]  \hspace{1cm} (F.A.97)

and recall from corollary [G.2.7] that

\[ \|K_n^\theta(x, \cdot) - \pi\|_{TV} \leq \left(1 - \frac{1}{M_\theta}\right)^n. \]  \hspace{1cm} (F.A.98)

From eq. (F.A.96) and eq. (F.A.208), \( M_\theta \) is bounded as

\[ M_\theta \leq \frac{\max_{x \in \text{Supp}(\pi)} \pi(x)}{\min_{x \in \text{Supp}(\pi)} \tilde{\pi}_\theta(x)} \leq \frac{\max_{x \in \text{Supp}(\pi)} \pi(x)}{\delta} = M_\delta \]  \hspace{1cm} (F.A.99)

The quantity \( M_\delta \) does not depend on \( \theta \in \mathcal{Y} \) and therefore we have, for all \( \theta \in \mathcal{Y} \),

\[ \|K_n^\theta(x, \cdot) - \pi\|_{TV} \leq \left(1 - \frac{1}{M_\delta}\right)^n. \]  \hspace{1cm} (F.A.100)

Using this worst-case bound, we may find an \( n \) satisfying definition [F.A.15.1] for all \( \theta \in \mathcal{Y} \).

\[ \square \]

Proof of lemma [F.A.15.3] Fix \( \theta \in \mathcal{Y} \). Then

\[ \min_{x \in \text{Supp}(\pi)} \tilde{\pi}_\theta^\star(x) = \min_{x \in \text{Supp}(\pi)} \left(\beta \pi_{\Pi^*}^\star(x) + (1 - \alpha)\tilde{\pi}_\theta(x)\right) \]  \hspace{1cm} (F.A.101)

\[ \geq \min_{x \in \text{Supp}(\pi)} \beta \pi_{\Pi^*}^\star(x) \]  \hspace{1cm} (F.A.102)

\[ = \delta. \]  \hspace{1cm} (F.A.103)

The quantity \( \delta \) is greater than zero since \( \beta > 0 \) and \( \text{Supp}(\pi) \subseteq \text{Supp}(\pi_{\Pi^*}^\star) \). Since \( \theta \) was arbitrary, the conclusion follows.

\[ \square \]
F.A.6 Proofs Concerning Containment

Proof of proposition G.4.4

\[
\Pr \left[ \log \pi(x) - \log \bar{\pi}_{\Theta_n}(x) < \log M \ \forall \ x \in \mathcal{X} \right] \geq 1 - \delta \ \forall \ n \in \mathbb{N} \quad \text{(F.A.104)}
\]

\[
\implies \Pr \left[ \frac{\pi(x)}{\bar{\pi}_{\Theta_n}(x)} < M \ \forall \ x \in \mathcal{X} \right] \geq 1 - \delta \ \forall \ n \in \mathbb{N} \quad \text{(F.A.105)}
\]

Then for all \( \epsilon > 0 \) there exists \( N \equiv N(\epsilon, \delta) \in \mathbb{N} \) such that for all \( x \in \mathcal{X} \),

\[
\Pr \left[ \| K_{\Theta_n}^N (x, \cdot) - \Pi(\cdot) \|_{TV} < \epsilon \right] \geq 1 - \delta \ \forall \ n \in \mathbb{N} \quad \text{(F.A.106)}
\]

\[
\implies \Pr [ W_\epsilon(x, K_{\Theta_n}) \leq N ] \geq 1 - \delta \ \forall \ n \in \mathbb{N} \quad \text{(F.A.107)}
\]

\[\square\]

F.A.7 Rao-Blackwellization of the Independent Metropolis-Hastings Algorithm with Deterministic Adaptations

The samples produced by a Markov chain are typically used in the calculation of an empirical expectation. Consider a function \( h : \mathcal{X} \to \mathbb{R} \) and the estimate,

\[
\mathbb{E}_{x \sim \Pi} h(x) \approx \frac{1}{n + 1} \sum_{i=0}^{n} h(X_i) \quad \text{(F.A.108)}
\]

where \( (X_0, X_1, \ldots) \) are the steps produced by a Markov chain which is ergodic for the distribution \( \Pi \). The improvement of such an estimator by the method of Rao-Blackwellization for the non-adaptive independent Metropolis-Hastings method was deduced in [Casella and Robert, 1996]. We next give the generalization of this procedure to the adaptive case. Let \( (\tilde{X}_1(\theta_0), \tilde{X}_2(\theta_1), \ldots, \tilde{X}_n(\theta_n)) \) be the sequence of proposals generated by the adaptive
independent Metropolis-Hastings sampler. Then

\[
\frac{1}{n+1} \sum_{j=0}^{n} h(X_j) = \frac{1}{n+1} \left( h(X_0) + \sum_{j=1}^{n} h(X_j) \cdot 1\{X_j = \bar{X}_j\} + h(X_{j-1}) \cdot 1\{X_j = X_{j-1}\} \right)
\]

\[= \frac{1}{n+1} \sum_{j=0}^{n} h(\bar{X}_j) \cdot \sum_{i=j}^{n} 1\{X_i = \bar{X}_j\}, \tag{F.A.109}\]

where we have set \(\bar{X}_0 = X_0\). The quantity \(\sum_{j=i}^{n} 1\{X_j = \bar{X}_i\}\) is the number of occurrences of the proposal \(\bar{X}_i\) in the chain \((X_i, X_{i+1}, \ldots, X_n)\). The Rao-Blackwell Theorem states that the estimator (which is only a function of \((\bar{X}_0, \bar{X}_1, \ldots, \bar{X}_n)\))

\[
\mathbb{E} \left[ \frac{1}{n+1} \sum_{j=0}^{n} h(\bar{X}_j) \cdot \sum_{i=j}^{n} 1\{X_i = \bar{X}_j\} \mid (\bar{X}_0, \ldots, \bar{X}_n) \right] = \frac{1}{n+1} \sum_{j=0}^{n} h(\bar{X}_j) \cdot \sum_{i=j}^{n} \Pr\left[X_i = \bar{X}_j\right]
\]

\[= \frac{1}{n+1} \sum_{j=0}^{n} h(\bar{X}_j) \cdot \sum_{i=j}^{n} \Pr\left[X_i = \bar{X}_j\right] \tag{F.A.111}\]

has the same expectation as \(\frac{1}{n+1} \sum_{i=0}^{n} h(X_i)\) but lesser variance, making it a preferable estimator, if it can be computed. We therefore turn our attention to the computation of \(\Pr\left[X_i = \bar{X}_j\right]\).

**Theorem F.A.7.1.** For \(i \geq j\)

\[
\Pr\left[X_i = \bar{X}_j\right] = \left( \sum_{l=0}^{j-1} \min \left\{ 1, \frac{\pi(\bar{X}_j)\pi_{\theta_{j-l}}(\bar{X}_l)}{\pi(\bar{X}_l)\pi_{\theta_{j-l}}(X_l)} \right\} \cdot \Pr\left[X_l = \bar{X}_l\right] \right.
\]

\[
\times \prod_{k=l+1}^{j-1} \left( 1 - \min \left\{ 1, \frac{\pi(\bar{X}_k)\pi_{\theta_{k-l-1}}(\bar{X}_l)}{\pi(\bar{X}_l)\pi_{\theta_{k-l-1}}(X_l)} \right\} \right)
\]

\[
\times \prod_{k=j+1}^{i} \left( 1 - \min \left\{ 1, \frac{\pi(\bar{X}_k)\pi_{\theta_{k-l}}(\bar{X}_j)}{\pi(\bar{X}_j)\pi_{\theta_{k-l}}(X_k)} \right\} \right) \tag{F.A.112}\]

549
Proof. For \( i \geq j \),

\[
\Pr[X_i = \tilde{X}_j] = \Pr[X_i = \tilde{X}_j | X_{i-1} = \tilde{X}_j] \cdot \Pr[X_{i-1} = \tilde{X}_j] \quad (\text{F.A.113})
\]

\[
= \Pr[X_j = \tilde{X}_j] \cdot \prod_{k=j+1}^{i} \Pr[X_k = \tilde{X}_j | X_{k-1} = \tilde{X}_j] \quad (\text{F.A.114})
\]

\[
= \Pr[X_j = \tilde{X}_j] \cdot \prod_{k=j+1}^{i} \left( 1 - \min \left\{ 1, \frac{\pi(\tilde{X}_k) \pi_{\theta_{k-1}}(\tilde{X}_j)}{\pi(\tilde{X}_j) \pi_{\theta_{k-1}}(X_k)} \right\} \right) \quad (\text{F.A.115})
\]

Moreover,

\[
\Pr[X_i = \tilde{X}_i] = \sum_{j=0}^{i-1} \Pr[X_i = \tilde{X}_i | X_{i-1} = \tilde{X}_j] \cdot \Pr[X_{i-1} = \tilde{X}_j] \quad (\text{F.A.116})
\]

\[
= \sum_{j=0}^{i-1} \min \left\{ 1, \frac{\pi(\tilde{X}_i) \pi_{\theta_{i-1}}(\tilde{X}_j)}{\pi(\tilde{X}_j) \pi_{\theta_{i-1}}(X_i)} \right\} \cdot \Pr[X_{i-1} = \tilde{X}_j] \quad (\text{F.A.117})
\]

\[
= \sum_{j=0}^{i-1} \min \left\{ 1, \frac{\pi(\tilde{X}_i) \pi_{\theta_{i-1}}(\tilde{X}_j)}{\pi(\tilde{X}_j) \pi_{\theta_{i-1}}(X_i)} \right\} \cdot \Pr[X_j = \tilde{X}_j] \quad (\text{F.A.118})
\]

\[
\cdot \prod_{k=j+1}^{i-1} \left( 1 - \min \left\{ 1, \frac{\pi(\tilde{X}_k) \pi_{\theta_{k-1}}(\tilde{X}_j)}{\pi(\tilde{X}_j) \pi_{\theta_{k-1}}(X_k)} \right\} \right)
\]

Given the initial condition \( \Pr[X_0 = \tilde{X}_0] = 1 \), we obtain a recursive procedure by which to compute \( \Pr[X_i = \tilde{X}_i] \). For instance,

\[
\Pr[X_1 = \tilde{X}_1] = \min \left\{ 1, \frac{\pi(\tilde{X}_1) \pi_{\theta_0}(\tilde{X}_0)}{\pi(X_0) \pi_{\theta_0}(X_1)} \right\} \quad (\text{F.A.119})
\]

\[
\Pr[X_2 = \tilde{X}_2] = \min \left\{ 1, \frac{\pi(\tilde{X}_2) \pi_{\theta_1}(\tilde{X}_0)}{\pi(X_0) \pi_{\theta_1}(\tilde{X}_2)} \right\} \cdot \left( 1 - \min \left\{ 1, \frac{\pi(\tilde{X}_1) \pi_{\theta_0}(\tilde{X}_0)}{\pi(X_0) \pi_{\theta_0}(\tilde{X}_1)} \right\} \right) + \min \left\{ 1, \frac{\pi(\tilde{X}_2) \pi_{\theta_1}(\tilde{X}_1)}{\pi(\tilde{X}_1) \pi_{\theta_1}(\tilde{X}_2)} \right\} \cdot \min \left\{ 1, \frac{\pi(\tilde{X}_1) \pi_{\theta_0}(\tilde{X}_0)}{\pi(X_0) \pi_{\theta_0}(\tilde{X}_1)} \right\} \quad (\text{F.A.120})
\]
and so on. In any case, we obtain the formula, for \( i \geq j \)

\[
\Pr \left[ X_i = \tilde{X}_j \right] = \left( \sum_{l=0}^{j-1} \min \left\{ 1, \frac{\pi(\tilde{X}_j)\tilde{\pi}_{\theta_{j-1}}(\tilde{X}_l)}{\pi(X_l)\tilde{\pi}_{\theta_{j-1}}(X_j)} \right\} \right) \cdot \Pr \left[ X_l = \tilde{X}_l \right] \]

\[
\times \prod_{k=l+1}^{j-1} \left( 1 - \min \left\{ 1, \frac{\pi(\tilde{X}_k)\tilde{\pi}_{\theta_{k-1}}(\tilde{X}_l)}{\pi(X_l)\tilde{\pi}_{\theta_{k-1}}(X_k)} \right\} \right) \quad \text{(F.A.121)}
\]

\[
\times \prod_{k=j+1}^{i} \left( 1 - \min \left\{ 1, \frac{\pi(\tilde{X}_k)\tilde{\pi}_{\theta_{k-1}}(\tilde{X}_j)}{\pi(X_j)\tilde{\pi}_{\theta_{k-1}}(X_k)} \right\} \right)
\]

\[
\square
\]

**F.A.8 Adaptation via the Flow of the KL Divergence**

Consider an adaptive sequence of parameters \( (\theta_0, \theta_1, \ldots) \) which parameterize proposal densities \( (\tilde{\pi}_{\theta_0}, \tilde{\pi}_{\theta_1}, \ldots) \). Given a target distribution \( \Pi \) with density \( \pi \), the efficacy of the adaptive independent Metropolis-Hastings sampler is dictated by the ratio of the target density to the proposal density,

\[
U(\theta) \overset{\text{def.}}{=} \sup_{x \in \text{Supp}(\pi)} \frac{\pi(x)}{\tilde{\pi}_{\theta}(x)}.
\]

The smaller this upper bound, the better the mixing properties of the independent Metropolis-Hastings algorithm with proposal distribution \( \tilde{\Pi}_{\theta_n} \). A central question is whether or not the sequence \( (\theta_0, \theta_1, \ldots) \) actually produces improvements in these upper bounds; i.e. is \( U(\theta_{n+1}) \leq U(\theta_n) \)?

It is important that the bound in eq. (F.A.122) is actually the least upper bound. This is because an arbitrary upper bound may decrease while the least upper bound decreases.

In estimating the parameters of normalizing flows, it is typical that parameters follow, at least approximately, the gradient flow of a prescribed loss function, such as a KL divergence. Since gradient flows are initial value problems with deterministic solutions, by
examining the case wherein adaptations are obtained exactly by gradient flow allows us to bypass the added difficulty of contending with stochastic adaptations.

**Proposition F.A.8.1.** Let $\mathcal{X}$ be a state space and let $\theta \in \mathbb{R}^m$ parameterize a probability measure $\tilde{\Pi}_\theta$ on $\mathcal{B}(\mathcal{X})$ with density $\tilde{\pi}_\theta$. Given a target density $\pi$, consider the function $U : \mathbb{R}^m \to \mathbb{R}$ defined by eq. (F.A.122) and assume further that $U$ is smooth with respect to its argument. Let $L : \mathbb{R}^m \to \mathbb{R}$ be a loss function and consider the gradient flow $\theta_t = -\nabla L(\theta_t)$ given an initial condition $\theta_0$. A sufficient condition that $U(\theta_{t+s}) \leq U(\theta_t)$ is that

$$\nabla U(\theta_{t'}) \cdot \nabla L(\theta_{t'}) \geq 0,$$

where $t' \in (t, t+s)$; i.e., $\nabla L(\theta_{t'})$ is an ascent direction of $U$ at $\theta_{t'}$.

**Proof.** By applying the chain rule,

$$\frac{d}{dt} U(\theta_t) = \nabla U(\theta_t) \cdot \dot{\theta}_t$$

(F.A.124)

$$= \nabla U(\theta_t) \cdot -\nabla L(\theta_t).$$

(F.A.125)

By the fundamental theorem of calculus,

$$U(\theta_{t+s}) - U(\theta_t) = \int_{t}^{t+s} \left( \frac{d}{dt'} U(\theta_{t'}) \right) \left|_{t'=t''} \right. dt''$$

(F.A.126)

$$= -\int_{t}^{t+s} \nabla U(\theta_{t''}) \cdot \nabla L(\theta_{t''}) dt''$$

(F.A.127)

$$\leq 0$$

(F.A.128)

Therefore, $U(\theta_{t+s}) \leq U(\theta_t)$. $\square$

While verifying the conditions of proposition F.A.8.1 in general appears a daunting task, we can do some analysis in simple cases.
Example 47. Consider the problem of sampling \( \text{Normal}(0, 1) \) by adapting a proposal of the form \( \text{Normal}(\mu, \sigma^2) \). Assume further that \( \sigma^2 > 1 \). We can deduce a an upper bound on the ratio of the target density to the proposal density as follows:

\[
\max_{x \in \mathbb{R}} \frac{\exp(-x^2/2)/\sqrt{2\pi}}{\exp(-(x - \mu)^2/2\sigma^2)/\sqrt{2\pi\sigma^2}} = \sigma \max_{x \in \mathbb{R}} \exp\left(-\frac{x^2}{2} + \frac{(x - \mu)^2}{2\sigma^2}\right)
\]

\[
\leq \sigma \exp\left(\frac{\mu^2}{2(\sigma^2 - 1)}\right).
\]

which can be deduced by maximizing \(-\frac{x^2}{2} + \frac{(x - \mu)^2}{2\sigma^2}\) using calculus. The reverse KL divergence between the proposal distribution and the target distribution is seen to be,

\[
\mathbb{K}_L(\text{Normal}(\mu, \sigma^2)\|\text{Normal}(0, 1)) = -\log \sigma + \frac{\sigma^2 + \mu^2}{2} - \frac{1}{2}.
\]

Consider the gradient flow of the reverse KL divergence:

\[
\dot{\mu}_t = -\frac{\partial}{\partial \mu} \mathbb{K}_L(\text{Normal}(\mu, \sigma^2)\|\text{Normal}(0, 1))
\]

\[
= -\mu
\]

\[
\dot{\sigma}_t = -\frac{\partial}{\partial \sigma} \mathbb{K}_L(\text{Normal}(\mu, \sigma^2)\|\text{Normal}(0, 1))
\]

\[
= \frac{1}{\sigma_t} - \sigma_t.
\]

When we specify initial conditions \( \mu_0 \in \mathbb{R} \) and \( \sigma_0 > 1 \), this produces an initial value problem. To verify that adapting by following the gradient flow of KL divergence produces a provable improvement to the upper bound deduced in eq. (F.A.130), it suffices to check that the time derivative of the upper bound in decreasing under the postulated gradient flow.
dynamics. That is,

\[
\frac{d\sigma_t}{dt} \exp\left(\frac{\mu_t^2}{2(\sigma_t^2 - 1)}\right) = \mu_t\sigma_t \exp\left(\frac{\mu_t^2}{2(\sigma_t^2 - 1)}\right) \cdot -\mu_t \\
+ \frac{\exp(\mu_t^2/(2(\sigma_t^2 - 1)))(-\mu_t^2 + 2)\sigma_t^2 + \sigma_t^4 + 1}{(\sigma_t^2 - 1)^2} \cdot \left(\frac{1}{\sigma_t} - \sigma_t\right)
\]

(F.A.136)

\[
= -\frac{(\sigma_t - 1)(\sigma_t + 1)\exp(\mu_t^2/(2(\sigma_t^2 - 1)))}{\sigma_t}
\]

(F.A.137)

It follows that this is a negative quantity if we can establish that \(\sigma_t > 1\). From the initial condition \(\sigma_0 > 1\), it follows that the positive solution of the differential equation \(\dot{\sigma}_t = \frac{1}{\sigma_t} - \sigma_t\) is \(\sigma_t = \sqrt{e^{-2t}(\sigma_0^2 - 1) + 1}\), so we see, indeed, that \(\sigma_t > 1\). Therefore, the upper bound is a decreasing function of \(t\) given the prescribed gradient flow dynamics. The differential equation \(\dot{\mu}_t = -\mu_t\) also has an explicit solution given the initial condition \(\mu_0\), which is \(\mu_t = \mu_0 e^{-t}\). These explicit solutions to the gradient flow of the KL divergence allow us to express the evolution of the upper bound concretely as,

\[
\sqrt{e^{-2t}(\sigma_0^2 - 1) + 1} \exp\left(\frac{\mu_0^2 e^{-2t}}{2e^{-2t}(\sigma_0^2 - 1)}\right) = \sqrt{e^{-2t}(\sigma_0^2 - 1) + 1} \exp\left(\frac{\mu_0^2}{2(\sigma_0^2 - 1)}\right)
\]

(F.A.138)

This is an intriguing formula since it suggests that although the sequence \(M_t\) is decreasing, it decreases only to a non-unit limit \(\exp(\mu_0^2/(2(\sigma_0^2 - 1)))\); indeed, unless \(\mu_0 = 0\), the limit of this upper bound does not approach one. For the purposes of MCMC, this may be acceptable, since uniform ergodicity can be obtained so long as the bound is finite; however, were the upper bound to equal one, this would be optimal. Connecting this back to the question of adaptation, choosing an increasing sequence of times \(t_0 < t_1 < t_2 < \ldots\) and consider using \(\text{Normal}(\mu_{t_n}, \sigma_{t_n}^2)\) as the proposal distribution at step \(n\). The Doeblin
coefficient at step \( n \) is therefore,

\[
L_n = \frac{1}{\sqrt{e^{-2t}(\sigma_0^2 - 1) + 1} \exp\left(\frac{\mu_0^2}{2(\sigma_0^2 - 1)}\right)}.
\]  

(F.A.139)

Finally, let us remark that the undesirable property that the upper bounds do not converge to unity can be easily corrected. The principle issue is that the factors of \( e^{-2t} \) cancel in the exponent. However, consider that instead of the using \((\mu_t, \sigma_t)\) to inform adaptations one instead uses \((\mu_{2t}, \sigma_t)\) so that the mean value is further along in the solution to its initial value problem than the scale. Plugging this into the formula for the upper bound yields,

\[
\sqrt{e^{-2t}(\sigma_0^2 - 1) + 1} \exp\left(\frac{\mu_0^2 e^{-4t}}{2e^{-2t}(\sigma_0^2 - 1)}\right) = \sqrt{e^{-2t}(\sigma_0^2 - 1) + 1} \exp\left(\frac{\mu_0^2 e^{-2t}}{2(\sigma_0^2 - 1)}\right),
\]

(F.A.140)

which converges to unity as \( t \to \infty \) as desired.

Instead of the reverse KL divergence we may consider the forward KL divergence between the target distribution and the proposal.

\[
\mathbb{KL}(\text{Normal}(0, 1) \| \text{Normal}(\mu, \sigma^2)) = \log \sigma + \frac{1 + \mu^2}{2\sigma^2} - \frac{1}{2}.
\]

(F.A.141)

The forward KL divergence produces the following equations of motion.

\[
\dot{\mu}_t = -\frac{\mu_t}{\sigma_t^2}
\]

(F.A.142)

\[
\dot{\sigma}_t = \frac{\mu_t^2 + 1}{\sigma_t^3} - \frac{1}{\sigma_t}
\]

(F.A.143)

Applying the chain rule to eq. (F.A.130) with these equations of motion yields the follow-
ing time derivative of the upper bound,
\[
\frac{d}{dt} \sigma_t \exp \left( \frac{\mu_t^2}{2(\sigma_t^2 - 1)} \right) = \exp \left( \frac{\mu_t^2}{2(\sigma_t^2 - 1)} \right) \frac{-\mu_t^4 \rho^2 + \mu_t^2 \sigma_t^4 - 2\mu_t^2 \rho \sigma_t^6 + \mu_t^2 - \sigma_t^6 + 3\sigma_t^4 - 3\sigma_t^2 + 1}{\sigma_t^4(\sigma_t^2 - 1)^2}.
\]
(F.A.144)

This derivative is less than or equal to zero iff
\[
-\mu_t^4 \rho^2 + \mu_t^2 \rho \sigma_t^6 + \mu_t^2 - \sigma_t^6 + 3\sigma_t^4 - 3\sigma_t^2 + 1 \leq 0 \tag{F.A.145}
\]
\[
\iff \mu_t^2(\sigma_t^2 - 1)^2 \leq \mu_t^4 \sigma_t^2 + (\sigma_t^2 - 1)^3, \tag{F.A.146}
\]
which is true for \(\sigma_t^2 > 1\).

## F.A.9 Adaptation of a Kernel Density Proposal Distribution

A principle difficulty in using normalizing flows as proposal distributions is that it is unclear whether or not a given adaptation of the neural network parameters will provably decrease the upper bound on the ratio of the target density and the normalizing flow density. Analyzing this property theoretically in the case of normalizing flows does not appear to be forthcoming. Nevertheless, we have been able to analyze certain behaviors of proposals based on kernel density estimators. A version of this procedure (to use kernel density estimators as a proposal in independent Metropolis-Hastings) was previously pursued in [Maire et al.] [2019]; however, they do not appear to have looked at the question of when the addition of a new component into the mixture is beneficial, which is the topic under consideration in the sequel.
Let $x_1, \ldots, x_n \in \mathbb{R}^m$. We consider the kernel density estimator computed by,

$$
\tilde{K}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\text{Vol}(\bar{B}_\epsilon(x_i))} 1 \{ x \in \bar{B}_\epsilon(x_i) \}.
$$

(F.A.147)

where $\bar{B}_\epsilon(x)$ is the closed ball of radius $\epsilon$ centered at $x$. For notational convenience, we observe that $V = \text{Vol}(\bar{B}_\epsilon(x)) = \text{Vol}(\bar{B}_\epsilon(x'))$ for any $x, x' \in \mathbb{R}^m$ and we write $m_n(x) = \# \{ i \in (1, \ldots, n) : x \in \bar{B}_\epsilon(x_i) \}$. Therefore, we have the simple expression for the kernel density estimator as $\tilde{K}_n(x) = m_n(x)/nV$.

Suppose that $\Pi$ is a probability measure on $\mathbb{R}^m$ with compactly supported density $\pi : \mathbb{R}^m \to \mathbb{R}^+$. Define,

$$
M_n = nV \sup_{x \in \text{Supp}(\pi)} \frac{\pi(x)}{m_n(x)}.
$$

(F.A.148)

Hence $\pi(x)/\tilde{K}_n(x) \leq M_n$. Given a new observation $x_{n+1} \in \mathbb{R}^m$, we would like to understand conditions under which one can show $M_{n+1} \leq M_n$. This means that the inclusion of a new observation in the kernel density estimate reduces the upper bound on the ratio of the target density and the proposal density. To discuss this, we begin with two definitions.

**Definition F.A.9.1.** The inner bound is defined by,

$$
M'_n = \sup_{x \in \bar{B}_\epsilon(x_{n+1}) \cap \text{Supp}(\pi)} \frac{\pi(x)}{m_n(x)}.
$$

(F.A.149)

**Definition F.A.9.2.** The outer bound is defined by,

$$
M''_n = \sup_{x \in \text{Supp}(\pi) \setminus \bar{B}_\epsilon(x_{n+1})} \frac{\pi(x)}{m_n(x)}.
$$

(F.A.150)

**Lemma F.A.9.3.** Let $A$ and $B$ be sets. Then $\max \{ \sup A, \sup B \} = \sup A \cup B$. 

557
Proof. Since \( A \subset A \cup B \), it is immediate that \( \sup A \leq \sup A \cup B \). Identical reasoning shows that \( \sup B \leq \sup A \cup B \). Therefore, \( \max \{ \sup A, \sup B \} \leq \sup A \cup B \). Now suppose without loss of generality that \( \sup B \geq \sup A \). Then for all \( a \in A \) we have \( a \leq \sup B \); moreover, for all \( b \in B \), \( b \leq \sup B \). Therefore, for all \( x \in A \cup B \), \( x \leq \sup B \).

Thus, \( \sup A \cup B \leq \sup B \), since \( \sup A \cup B \) is by definition the least upper bound. Applying identical reasoning to the case \( \sup A \geq \sup B \) reveals \( \sup A \cup B \leq \max \{ \sup A, \sup B \} \).

Thus, we must have \( \sup A \cup B = \max \{ \sup A, \sup B \} \). \( \square \)

**Corollary F.A.9.4.** Since \( \text{Supp}(\pi) \setminus \bar{B}_\varepsilon(x_{n+1}) = \text{Supp}(\pi) \cap \bar{B}_\varepsilon(x_{n+1})^c \), it follows from the distributive law of set relationships that,

\[
\{ \text{Supp}(\pi) \cap \bar{B}_\varepsilon(x_{n+1}) \} \cup \{ \text{Supp}(\pi) \setminus \bar{B}_\varepsilon(x_{n+1}) \} = \text{Supp}(\pi).
\] (F.A.151)

Applying lemma F.A.9.3 shows that \( nV \max \{ M'_n, M''_n \} = M_n \).

**Lemma F.A.9.5.** The kernel density estimator \( \tilde{K}_{n+1}(x) \) can be written as,

\[
\tilde{K}_{n+1}(x) = \begin{cases} 
\frac{m_n(x)}{(n+1)V} & \text{if } x \notin \bar{B}_\varepsilon(x_{n+1}) \\
\frac{m_n(x) + 1}{(n+1)V} & \text{otherwise}.
\end{cases}
\] (F.A.152)

Proof. This follows immediately from the equation,

\[
\tilde{K}_{n+1}(x) = \frac{1}{(n+1)V} \sum_{i=1}^{n+1} 1 \{ x \in \bar{B}_\varepsilon(x_i) \} \] (F.A.153)

\[
= \frac{1}{(n+1)V} \left( m_n(x) + 1 \{ x \in \bar{B}_\varepsilon(x_{n+1}) \} \right). \] (F.A.154)

\( \square \)
Lemma F.A.9.6. We have,
\[
\sup_{x \in \text{Supp}(\pi) \setminus \bar{B}_t(x_{n+1})} \frac{\pi(x)}{K_{n+1}(x)} = (n + 1)V M''_n. \tag{F.A.155}
\]

Proof. For \(x \in \text{Supp}(\pi) \setminus \bar{B}_t(x_{n+1}),\)
\[
\frac{\pi(x)}{K_{n+1}(x)} = (n + 1)V \frac{\pi(x)}{m_n(x)}. \tag{F.A.156}
\]

Therefore,
\[
\sup_{x \in \text{Supp}(\pi) \setminus \bar{B}_t(x_{n+1})} \frac{\pi(x)}{K_{n+1}(x)} = (n + 1)V \sup_{x \in \text{Supp}(\pi) \setminus \bar{B}_t(x_{n+1})} \frac{\pi(x)}{m_n(x)} \tag{F.A.157}
\]
\[
= (n + 1)V M''_n. \tag{F.A.158}
\]

\qed

Lemma F.A.9.7. We have,
\[
\sup_{x \in \text{Supp}(\pi) \cap \bar{B}_t(x_{n+1})} \frac{\pi(x)}{K_{n+1}(x)} \leq nV Q_n M'_n, \tag{F.A.159}
\]
where
\[
Q_n = \sup_{x \in \text{Supp}(\pi) \cap \bar{B}_t(x_{n+1})} \frac{m_n(x)(n + 1)}{(m_n(x) + 1)n} \leq 1. \tag{F.A.160}
\]

Proof. Within \(\text{Supp}(\pi) \cap \bar{B}_t(x_{n+1})\) we have the bound,
\[
nV \frac{\pi(x)}{m_n(x)} \leq nV M'_n \tag{F.A.161}
\]
\[
\Rightarrow \frac{\pi(x)}{m_n(x)/n} \cdot \frac{m_n(x)/n}{(m_n(x) + 1)/(n + 1)} \leq \frac{m_n(x)/n}{(m_n(x) + 1)/(n + 1)} nV M'_n \tag{F.A.162}
\]
\[
\Rightarrow \frac{\pi(x)}{K_{n+1}(x)} \leq nV \frac{m_n(x)/n}{(m_n(x) + 1)/(n + 1)} M'_n \tag{F.A.163}
\]

559
Taking the supremum on both sides yields,

\[
\sup_{x \in \text{Supp}(\pi) \cap B_r(x_{n+1})} \frac{\pi(x)}{K_{n+1}(x)} \leq nVQ_nM'_n. \tag{F.A.164}
\]

Since \( m_n(x) \leq n \), it follows that \( Q_n \leq 1. \)

**Proposition F.A.9.8.** If \( M''_n > M'_n \), then \( M_{n+1} > M_n \). In this case, the inclusion of the new observation degrades the quality of the proposal distribution.

**Proof.** Since \( M''_n > M'_n \), it follows that

\[
\sup_{x \in \text{Supp}(\pi) \cap B_r(x_{n+1})} \frac{\pi(x)}{K_{n+1}(x)} \leq nVQ_nM'_n \tag{F.A.165}
\]

\[
\leq nVM'_n \tag{F.A.166}
\]

\[
< nVM''_n \tag{F.A.167}
\]

\[
\leq (n + 1)VM''_n \tag{F.A.168}
\]

\[
= \sup_{x \in \text{Supp}(\pi) \setminus B_r(x_{n+1})} \frac{\pi(x)}{K_{n+1}(x)}. \tag{F.A.169}
\]

Hence \( M_{n+1} = (n + 1)VM''_n > nVM''_n = M_n. \)

**Proposition F.A.9.9.** It is necessary and sufficient that \((1 + 1/n)M''_n \leq M'_n\) in order for \( M_{n+1} \leq M_n \).

**Proof.** To establish sufficiency, we have:

\[
M_{n+1} \leq \max \{(n + 1)VM''_n, nVQ_nM'_n\} \tag{F.A.170}
\]

\[
\leq \max \{nVM'_n, nVQ_nM'_n\} \tag{F.A.171}
\]

\[
= nVM'_n \tag{F.A.172}
\]

\[
= M_n \tag{F.A.173}
\]
To show that this is actually necessary, consider $M'_n \leq M'_n < (1 + 1/n)M''_n$. Then, from lemma F.A.9.6 we know,

$$\sup_{x \in \text{Supp}(\pi) \setminus \bar{B}_ε(x_{n+1})} \frac{\pi(x)}{K_{n+1}(x)} = (n + 1)V M''_n.$$  \hfill (F.A.174)

But from lemma F.A.9.7 we have,

$$\sup_{x \in \text{Supp}(\pi) \cap \bar{B}_ε(x_{n+1})} \frac{\pi(x)}{K_{n+1}(x)} \leq nV Q_n M'_n \leq (n + 1)V Q_n M''_n < (n + 1)V M''_n.$$ \hfill (F.A.175, 176, 177)

Now $M_n = nVM'_n < (n + 1)V M''_n = M_{n+1}$. \hfill \square

proposition F.A.9.8 informs us that if the worst case bound is outside of $\bar{B}_ε(x_{n+1})$, then the adaptation has actually made the bound worse. This is because the density outside of $\bar{B}_ε(x_{n+1})$ behaves in a predictable manner: it is decreased by a factor of $n/(n + 1)$. Therefore, if the worst case bound on the ratio occurs outside of $\bar{B}_ε(x_{n+1})$, the ratio can only get worse. At the same time, proposition F.A.9.9 informs us that the worst-case bound on the ratio inside $\bar{B}_ε(x_{n+1})$ must be greater than the worst-case bound outside of $\bar{B}_ε(x_{n+1})$ by at least a factor of $(1 + 1/n)$ in order for the inclusion of the additional observation $x_{n+1}$ to improve the worst-case bound of the ratio of the target density to the kernel density estimator.
F.A.10 Diminishing Adaptation and Containment for Mixture Kernels

Proposition F.A.10.1. Let $\Pi$ be a probability measure with density $\pi$ with respect to measure $\mu$. Let $\theta \in \mathcal{Y}$ and suppose that $\theta$ parameterizes a transition kernel $K_{\theta}$. Let $(\Theta_0, \Theta_1, \ldots)$ be a sequence of $\mathcal{Y}$-valued random variables. Suppose that with probability $1 - \delta$ there exists $M \equiv M(\delta) \in \mathbb{N}$ such that $K_{\Theta_n}(x, dx') \geq \frac{\pi(x')}{M} \mu(dx')$ for every $n = 1, 2, \ldots$. Then $(K_{\Theta_0}, K_{\Theta_1}, \ldots)$ exhibits containment.

Proof. By assumption, with probability $1 - \delta$, there exists $M \equiv M(\delta) \in \mathbb{N}$ such that $K_{\Theta_n}(x, dx') \geq \frac{\pi(x')}{M} \mu(dx')$ for every $n = 1, 2, \ldots$ and any $x \in \mathcal{X}$. Therefore, with probability $1 - \delta$, there exists $M \equiv M(\delta) \in \mathbb{N}$ such that,

$$\|K_{\Theta_n}^m(x, \cdot) - \Pi(\cdot)\|_{TV} \leq \left(1 - \frac{1}{M}\right)^m$$

(F.A.178)

for any $m \in \mathbb{N}$, every $n = 1, 2, \ldots$, and $x \in \mathcal{X}$. Let $\epsilon > 0$ be arbitrary. Then, with probability $1 - \delta$, there exists $N = N(\epsilon, \delta) \in \mathbb{N}$ such that

$$\|K_{\Theta_n}^N(x, \cdot) - \Pi(\cdot)\|_{TV} \leq \epsilon.$$  

(F.A.179)

for every $n = 1, 2, \ldots$ and any $x \in \mathcal{X}$. Namely, the choice

$$N(\epsilon, \delta) = \left\lceil \frac{\log \epsilon}{\log \left(1 - \frac{1}{M(\delta)}\right)} \right\rceil$$

(F.A.180)

suffices. Thus as a special case, with probability $1 - \delta$, there exists $N = N(\epsilon, \delta) \in \mathbb{N}$ such
that

$$\|K_n^\mathcal{N}(X_n, \cdot) - \Pi(\cdot)\|_{TV} \leq \epsilon.$$  \hfill (F.A.181)

for every \(n = 1, 2, \ldots\). Define the function,

$$W_\epsilon(x, \theta) = \inf \{ n \in \mathbb{N} : \|K_n^\mathcal{N}(x, \cdot) - \Pi(\cdot)\|_{TV} \leq \epsilon \}.$$  \hfill (F.A.182)

Hence, with probability \(1 - \delta\), there exists \(N = N(\epsilon, \delta) \in \mathbb{N}\) such that

$$W_\epsilon(X_n, \Theta_n) \leq N$$  \hfill (F.A.183)

for every \(n = 1, 2, \ldots\). This is the containment condition.

**Proposition F.A.10.2.** Let \(\Pi\) be probability measure with density \(\pi\) with respect to measure \(\mu\). Let \(\theta \in \mathcal{Y}\) and suppose that \(\theta\) parameterizes a probability measure \(\tilde{\Pi}_\theta\) with density \(\tilde{\pi}_\theta\). Let \(K_\theta\) be the transition kernel of the independent Metropolis-Hastings sampler of \(\Pi\) given \(\tilde{\Pi}_\theta\):

\[
K_\theta(x, dx') = \min \left\{ 1, \frac{\pi(x')}{\pi(x)} \tilde{\pi}_\theta(x') \right\} \tilde{\pi}_\theta(x') \mu(dx') + \left(1 - \int_{\mathcal{X}} \min \left\{ 1, \frac{\pi(w)}{\pi(x)} \tilde{\pi}_\theta(w) \right\} \tilde{\pi}_\theta(w) \mu(dw) \right) \delta_x(dx').
\]  \hfill (F.A.184)

Let \(K'\) be another transition and consider the transition kernel that is formed by the mixture of \(K_\theta\) and \(K'\): \(\hat{K}_\theta(x, A) = \alpha K_\theta(x, A) + (1 - \alpha) K'(x, A)\) for \(x \in \mathcal{X}\) and \(A \in \mathcal{B}(\mathcal{X})\). Let \((\Theta_0, \Theta_1, \ldots)\) be a sequence of \(\mathcal{Y}\)-valued random variables. If \((K_{\Theta_0}, K_{\Theta_1}, \ldots)\) exhibits diminishing adaptation then so does \((\hat{K}_{\Theta_0}, \hat{K}_{\Theta_1}, \ldots)\). Furthermore, suppose that with probability at least \(1 - \delta\) there exists \(M \equiv M(\delta)\) such that \(K_{\Theta_n}(x, dx') \geq \frac{\pi(x')}{M} \mu(dx')\) for every \(n = 1, 2, \ldots\). Then \((\hat{K}_{\Theta_0}, \hat{K}_{\Theta_1}, \ldots)\) exhibits containment.
Proof.

\[ d(\hat{K}_{\Theta_n}, \hat{K}_{\Theta_{n+1}}) = \sup_{x \in X} \| \hat{K}_{\Theta_n}(x, \cdot) - \hat{K}_{\Theta_{n+1}}(x, \cdot) \|_{TV} \quad \text{(F.A.185)} \]
\[ = \sup_{x \in X} \| \alpha K_{\Theta_n}(x, \cdot) - \alpha K_{\Theta_{n+1}}(x, \cdot) \|_{TV} \quad \text{(F.A.186)} \]
\[ = \alpha \sup_{x \in X} \| K_{\Theta_n}(x, \cdot) - K_{\Theta_{n+1}}(x, \cdot) \|_{TV} \quad \text{(F.A.187)} \]
\[ = \alpha d(K_{\Theta_n}, K_{\Theta_{n+1}}). \quad \text{(F.A.188)} \]

Hence, if \( d(K_{\Theta_n}, K_{\Theta_{n+1}}) \) converges in probability to zero then so does \( d(\hat{K}_{\Theta_n}, \hat{K}_{\Theta_{n+1}}) \).

To show containment, observe that with probability at least \( 1 - \delta \) there exists \( M \equiv M(\delta) \) such that,

\[ \hat{K}_\theta(x, dx') \geq \alpha K_\theta(x, dx') \quad \text{(F.A.189)} \]
\[ \geq \alpha \frac{\pi(x')}{M} \mu(dx'). \quad \text{(F.A.190)} \]

Containment then follows from proposition F.A.10.1. \( \square \)

**F.A.11 Violations of Stationarity**

In general, adaptation of the parameters of the transition kernel will destroy stationarity. However, if the adaptations and the state of the chain enjoy a prescribed independence condition, then stationarity of the target distribution can be conserved.

**Proposition F.A.11.1.** Suppose that \((\Theta_0, \Theta_1, \ldots)\) is a stochastic \(\mathcal{Y}\)-valued sequence. Let \((K_{\Theta_n})_{n \in \mathbb{N}}\) be an associated sequence of Markov transition kernels which produces an \(X\)-valued chain as \(X_{n+1} \sim K_{\Theta_n}(X_n, \cdot)\). Suppose further that \(\Theta_n\) and \(X_n\) are independent given the history of the chain to step \(n - 1\). If \(\Pi\) is stationary for each \(K_{\Theta_n}\), then \(\Pi\) is also the stationary distribution of \((X_n)_{n \in \mathbb{N}}\).
Figure F.A.1: Examinations of the violations of stationarity that result by maximizing the pseudo-likelihood of the accepted samples as an adaptation mechanism. After one-hundred steps of adaptation, one clearly perceives that the distribution of states does \textit{not} follow the target distribution. However, after ten-thousand steps, the distribution of state is closer to the target distribution. We also show the KL divergence between the target distribution at the proposal distribution according to the number of steps of the chain. Results are computed over one-million random simulations.

We first give an illustration of why maximizing the pseudo-likelihood objective may not always be beneficial. In particular, we look for evidence of violations of stationarity; violations of stationarity mean that if one begins with a sample from the target distribution and transforms it according to several steps of the transition kernel with adaptations, then the final state may not be distributed according to the target distribution. This can be interpreted as an undesirable form of sample degradation wherein applications of an adaptive transition kernel move exact samples further from the target distribution.

As a simple example, we consider sampling \textit{Normal}(1, 1/2) using a proposal distribution \textit{Normal}(\mu, \sigma^2); the proposal distribution can be interpreted as a simple normalizing flow consisting of a shift and scale applied to a standard normal base distribution. We consider adapting the parameters of the proposal distribution by computing the maximum likelihood estimates of the mean and standard deviation using the accepted samples. Specifically, let \((X_0, X_1, \ldots, X_n)\) denote the states of the chain to step \(n\); then the parameters of the proposal distribution at step \(n+1\) are \(\mu_{n+1} = (n + 1)^{-1} \sum_{i=0}^{n} X_i\) and \(\sigma_{n+1} = \sqrt{(n + 1)^{-1} \sum_{i=0}^{n} X_i^2 - \mu_{n+1}^2}\). Results for this adaptation mechanism are shown in fig. F.A.1; after one-hundred steps of adaptation, there is a clear violation of stationarity,
but which has nearly vanished by the ten-thousandth step. We observe that adaptations do
tend to reduce the forward KL divergence between the target and the proposal distribution,
with the closeness improving as the number of adaptation steps increases.

F.A.12 Independence and Product Transition Kernels

Let \( \Pi \) be a probability measure of \( \mathbb{R}^m \) with density \( \pi \). Let \( \prod(A_1, \ldots, A_n) = \prod_{i=1}^n \Pi(A_i) \) be the product probability measure on \( (\mathbb{R}^m)^n = \prod_{i=1}^n \mathbb{R}^m \) (the product space of \( n \) copies
of \( \mathbb{R}^m \)). Let \( (x_1, \ldots, x_n) \in (\mathbb{R}^m)^n \) and define the \( i \)th transition kernel by,

\[
K_i((x_1, \ldots, x_n), (dy_1, \ldots, dy_n)) = [\alpha_i(x_i, y_i) dy_i + \beta_i(x_i) \delta_{x_i}(dy_i)] \cdot \prod_{j=1, j \neq i}^n \delta_x(dy_j)
\]

(F.A.191)

\[
\alpha_i(x, y) = \min\left\{ 1, \frac{\pi(y) \tilde{\pi}_i(x|y)}{\pi(x) \tilde{\pi}_i(y|x)} \right\} \tilde{\pi}_i(y|x)
\]

(F.A.192)

\[
\beta_i(x) = \left( 1 - \int_{\mathbb{R}^m} \alpha_i(x, w) \, dw \right)
\]

(F.A.193)

Thus we can understand \( K_i \) as a transition kernel that applies a Metropolis-Hastings
accept-reject decision to the \( i \)th dimension of the posterior using a proposal density \( \tilde{\pi}_i(\cdot|x_i) \) on \( \mathbb{R}^m \) and leaves all other dimensions unchanged. Because of the product structure of the
joint distribution, these Metropolis-Hastings updates preserve the joint distribution even
though the accept-reject decision is computed only on the \( i \)th marginal.

As an example, consider \( K_1 \) and \( K_2 \). What is the composition transition kernel generated
by first applying \( K_1 \) and subsequently applying \( K_2 \)? By the Chapman-Kolmogorov
formula, it is,

\[(K_2 \circ K_1)((x_1, \ldots, x_n), (A_1, \ldots, A_n))\]  \hspace{1cm} (F.A.194)

\[= \int_{\mathbb{R}^m} \cdots \int_{\mathbb{R}^m} K_2((y_1, \ldots, y_n), (A_1, \ldots, A_n))K_1((x_1, \ldots, x_n), (dy_1, \ldots, dy_n))\]  \hspace{1cm} (F.A.195)

\[= \int_{\mathbb{R}^m} \cdots \int_{\mathbb{R}^m} \left( \int_{A_2} \alpha_2(y_2, w) \, dw + \beta_2(y_2) \delta_{y_2}(A_2) \right) \cdot \prod_{j=1,j\neq 2}^{n} \delta_{y_j}(A_j) \]  \hspace{1cm} (F.A.196)

\[\times \left( [\alpha_1(x_1, y_1) \, dy_1 + \beta_1(x_1) \delta_{x_1}(dy_1)] \cdot \prod_{j=1,j\neq 1}^{n} \delta_{x_j}(dy_j) \right)\]

\[= \left[ \int_{A_1} \alpha_1(x_1, w) \, dw + \beta_1(x_1) \delta_{x_1}(A_1) \right] \cdot \left[ \int_{A_2} \alpha_2(x_2, w) \, dw + \beta_2(x_2) \delta_{x_2}(A_2) \right] \cdot \prod_{j=3}^{n} \delta_{x_j}(A_j).\]  \hspace{1cm} (F.A.197)

Notice that the composition kernel assumes a factorized form. Suppose that \((x'_1, \ldots, x'_n) \sim (K_2 \circ K_1)((x_1, \ldots, x_n), \cdot)\). Drawing a sample from this composition kernel can be achieved by setting \(x'_j = x_j\) for \(j = 3, \ldots, n\) and sampling \(x'_1\) and \(x'_2\) independently from the distributions

\[\Pr [x'_i \in A|x_i] = \int_{A} \alpha_i(x_i, w) \, dw + \beta_i(x_i)\delta_{x_i}(A),\]  \hspace{1cm} (F.A.198)

for \(i \in \{1, 2\}\).

The fact that any composition kernel \(K_k \circ \cdots \circ K_1\) has this product distribution form can be established via induction. Assume

\[(K_k \circ \cdots \circ K_1)((x_1, \ldots, x_n), (dy_1, \ldots, dy_n)) = \prod_{i=1}^{k} [\alpha_i(x_i, y_i) \, dy_i + \beta_i(x_i) \delta_{x_i}(dy_i)] \prod_{j=k+1}^{n} \delta_{x_j}(dy_j)\]  \hspace{1cm} (F.A.199)
Then,

\[
(K_{k+1} \circ \cdots \circ K_1)((x_1, \ldots, x_n), (A_1, \ldots, A_n)) = \int_{(\mathbb{R}^m)^n} K_{k+1}((y_1, \ldots, y_n), (A_1, \ldots, A_n))(K_k \circ \cdots \circ K_1)((x_1, \ldots, x_n), (dy_1, \ldots, dy_n))
\]

(F.A.200)

\[
= \int_{(\mathbb{R}^m)^n} \left[ \left( \int_{A_{k+1}} \alpha_{k+1}(y_{k+1}, w) \, dw + \beta_{k+1}(y_{k+1}) \delta_y(A_{k+1}) \right) \prod_{j=1, j \neq k+1}^n \delta_y(A_j) \right] \left[ \prod_{i=1}^k [\alpha_i(x_i, y_i) \, dy_i + \beta_i(x_i) \delta_x(dy_i)] \prod_{j=k+1}^n \delta_x(dy_j) \right]
\]

(F.A.201)

\[
= \left[ \prod_{i=1}^{k+1} \left( \int_{A_i} \alpha_i(y_i, w) \, dw + \beta_i(y_i) \delta_y(A_i) \right) \right] \left[ \prod_{j=k+1}^n \delta_y(A_j) \right].
\]

(F.A.202)

This verifies that the composition kernel has the desired product structure. The special case of \( k = n \) implies,

\[
(K_n \circ \cdots \circ K_1)((x_1, \ldots, x_n), (dy_1, \ldots, dy_n)) = \prod_{i=1}^n [\alpha_i(x_i, y_i) \, dy_i + \beta_i(x_i) \, \delta_x(dy_i)].
\]

(F.A.204)

To sample \((x'_1, \ldots, x'_n) \sim (K_n \circ \cdots \circ K_1)((x_1, \ldots, x_n), \cdot)\), one may simply sample independently from the distributions,

\[
\Pr [x'_i \in A|x_i] = \int_A \alpha_i(x_i, w) \, dw + \beta_i(x_i) \delta_x(A)
\]

(F.A.205)

for \( i = 1, \ldots, n \). Each of these samples may be drawn in parallel because the \( i \)th sample depends only on \( x_i \).
F.A.13 Reasons for Violations of Containment

Inexpressive Family

If the family of proposal densities is not sufficiently expressive, it may be that there does not exist any \( M \geq 1 \) satisfying \( \log \pi(x) - \log \tilde{\pi}_\theta(x) < \log M \) for any \( \theta \). For instance, on \( \mathbb{R}^n \), if the proposal densities have tails that vanish exponentially quickly and the target density’s tails diminish only polynomially, then the family of proposal densities is not sufficiently expressive for proposition G.4.4 to hold; see Jaini et al. [2020] for details on the tail behavior of normalizing flows. On the other hand, if the family of proposal distributions has a universality property (see, inter alia Kobyzev et al. [2020]), then this concern can be alleviated.

Mode Collapse

Proposal densities obtained through the minimization of certain loss functions, including \( \mathbb{KL}(\tilde{\pi}_\theta \| \pi) \), may result in the modes of \( \pi \) not being properly represented in the proposal density \( \tilde{\pi}_\theta \). In the case of \( \mathbb{KL}(\tilde{\pi}_\theta \| \pi) \), the mode-seeking behavior of the loss function can cause the mode-collapse phenomenon, which can invalidate the assumption of proposition G.4.4. One could alleviate this concern by targeting a tempered version of the target density or if one has confidence that mode collapse will not occur (for instance if the target density is unimodal).

Unstable Loss

If the adaptations produced by attempting to minimize the loss function are ill-behaved (for instance if the step-size is too large, leading to divergent adaptations), then the sequence \( (\Theta_n)_{n \in \mathbb{N}} \) may parameterize poor proposal densities which cause the failure of eq. (G.10). If the algorithm relies on the convergence of the sequence \( (\Theta_n)_{n \in \mathbb{N}} \), then ill-behaved adap-
Algorithm 23 Algorithm for sampling from a target distribution by adapting a normalizing flow proposal distribution in the independent Metropolis-Hastings algorithm using the pseudo-likelihood objective.

**Input:** A sequence of step-sizes \((\epsilon_0, \epsilon_1, \ldots)\); a sequence of adaptation probabilities \((\alpha_0, \alpha, \ldots)\) an initial state \(x_0\) and initial parameter \(\theta_0\); a target distribution with density \(\pi : \mathcal{X} \to \mathbb{R}\).

for \(n = 0, 1, 2, \ldots\) do

Sample a proposal state from the current proposal distribution \(\tilde{x}_{n+1} \sim \tilde{\Pi}_{\theta_{n-1}}\).

Generate \(u \sim \text{Uniform}(0, 1)\) and compute the Metropolis-Hastings accept-reject decision.

\[
a \leftarrow u < \min \left\{ 1, \frac{\pi(\tilde{x}_{n+1})\tilde{\pi}_{\theta_n}(x_n)}{\pi(x_n)\tilde{\pi}_{\theta_n}(\tilde{x}_{n+1})} \right\} \quad (\text{F.A.206})
\]

if \(a\) then

Accept the proposal \(x_{n+1} \leftarrow \tilde{x}_{n+1}\).

else

Remain at current state \(x_{n+1} \leftarrow x_n\).

end if

Generate \(u' \sim \text{Uniform}(0, 1)\).

if \(u' < \alpha_n\) then

Update the parameters

\[
\theta_{n+1} \leftarrow \theta_n + \epsilon_n \nabla \log \tilde{\pi}_{\theta_n}(x_k) \quad (\text{F.A.207})
\]

where \(k \sim \text{Uniform}\left(\{0, 1, \ldots, n + 1\}\right)\).

else

Otherwise, keep the current parameters \(\theta_{n+1} \leftarrow \theta_n\).

end if

end for

tations may lead to a violation of ergodicity due to a failure of both containment and diminishing adaptation.

**F.A.14 Pseudo-Likelihood Algorithm**

The pseudo-likelihood training algorithm is presented in algorithm 23.
F.A.15 Simultaneous Uniform Ergodicity on Compact Spaces

Definition F.A.15.1. A family of transition kernels \( \{ K_\theta : \theta \in \mathcal{Y} \} \) is said to exhibit simultaneous uniform ergodicity if, for all \( \epsilon > 0 \), there exists \( n \in \mathbb{N} \) such that \( \| K^n_\theta (x, \cdot) - \Pi(\cdot) \|_{TV} \leq \epsilon \) for all \( \theta \in \mathcal{Y} \) and \( x \in \mathcal{X} \).

Simultaneous uniform ergodicity is a strong condition which states that, no matter which parameter \( \theta \in \mathcal{Y} \) one selects, there is a finite number of steps one can take with that fixed transition kernel in order to become arbitrarily close to the target distribution in total variation. We will see later that this condition can be made to hold for compactly supported target distributions.

We now turn our attention to the question of simultaneous uniform ergodicity.

Proposition F.A.15.2. Let \( \mathcal{X} \) and \( \pi \) satisfy the conditions of corollary G.2.7. Suppose that every \( \theta \in \mathcal{Y} \) parameterizes a probability measure \( \tilde{\Pi}_\theta \) on \( \mathcal{B}(\mathcal{X}) \) whose density \( \tilde{\pi}_\theta \) is continuous and satisfies \( \text{Supp}(\pi) \subseteq \text{Supp}(\tilde{\pi}_\theta) \). Suppose further that for all \( \theta \in \mathcal{Y} \), there exists \( \delta > 0 \) such that

\[
\delta \leq \min_{x \in \text{Supp}(\pi)} \tilde{\pi}_\theta(x) \tag{F.A.208}
\]

Then the family of Markov chain transition operators of the independent Metropolis-Hastings sampler of \( \Pi \) given \( \tilde{\Pi}_\theta \) satisfies the simultaneous uniform ergodicity property.

A proof is given in section F.A.5. Perhaps the most straight-forward mechanism to guarantee eq. (F.A.208) is to consider mixture distributions with a fixed distribution that shares the same support as \( \pi \).

Lemma F.A.15.3. Suppose that \( \Pi^* \) is a distribution on \( \mathcal{X} \) with continuous density \( \pi^*_{\Pi^*} \), such that \( \text{Supp}(\pi) \subseteq \text{Supp}(\pi^*_{\Pi^*}) \). Suppose that every \( \theta \in \mathcal{Y} \) parameterizes a probability measure \( \tilde{\Pi}_\theta \) on \( \mathcal{B}(\mathcal{X}) \) whose density \( \tilde{\pi}_\theta \) is continuous. Consider probability measures \( \tilde{\Pi}_\theta \)
whose densities are constructed from mixtures,

\[ \tilde{\pi}_\theta^*(x) = \beta \pi^*_{\Pi^*}(x) + (1 - \beta) \tilde{\pi}_\theta(x), \tag{F.A.209} \]

where \( \beta \in (0, 1) \). Then \( \tilde{\pi}_\theta^* \) satisfies eq. (F.A.208) with

\[ \delta = \beta \min_{x \in \text{Supp}(\pi)} \pi^*_{\Pi^*}(x). \tag{F.A.210} \]

A proof is given in section F.A.5. A natural choice of \( \Pi^* \) would be the uniform distribution on \( \mathcal{X} \). It is conceivable that one could consider adapting \( \beta \) in the same way that one adapts \( \theta \). However, in order to guarantee that eq. (F.A.210) is greater than zero, one will require the condition that \( \beta \in (\beta_*, 1) \) where \( \beta_* > 0 \).

Example 48. Let \( \Pi \) be a probability measure with density \( \pi \) on a compact space \( \mathcal{X} \). Let \( \mathcal{Y} = \mathbb{R}^m \) and suppose that every \( \theta \in \mathcal{Y} \) smoothly parameterizes a probability measure \( \tilde{\Pi}_\theta \) on \( \mathcal{B}(\mathcal{X}) \) with density \( \tilde{\pi}_\theta \) for which \( \text{Supp}(\pi) = \text{Supp}(\tilde{\pi}_\theta) \). Let \( \tilde{\Pi}^*_\theta \) be as in lemma F.A.15.3. Let \( (\alpha_0, \alpha_1, \ldots) \) be a sequence, bounded between zero and one, converging to zero. Consider the sequence of updates,

\[ \theta_n = \begin{cases} 
\theta_{n-1} - \epsilon \nabla_{\theta} \log \frac{\tilde{\pi}_{\theta_{n-1}}(\tilde{X}(\theta_{n-1}))}{\pi(\tilde{X}(\theta_{n-1}))} & \text{w.p. } 1 - \alpha_{n-1} \\
\theta_{n-1} & \text{otherwise.} 
\end{cases} \tag{F.A.211} \]

where \( \tilde{X} \sim \tilde{\Pi}^*_\theta_{\theta_{n-1}} \). Consider the family of Markov chain transition operators of the independent Metropolis-Hastings sampler of \( \Pi \) given \( \tilde{\Pi}^*_\theta_{\theta_n} \) with transition kernels \( K_{\theta_n} \) where the proposal at step \( n \) is \( \tilde{X} \). Then by theorem G.2.12 the distribution of \( X_{n+1} \sim K_{\theta_n}(X_n, \cdot) \) converges to \( \Pi \).

Examples of compact spaces on which normalizing flows have been applied include the torus, the sphere, the special orthogonal group, and the Stiefel manifold [Rezende].
The reason we were required to invoke a mixture distribution in the adaptation was because it prevented any sequence from becoming arbitrarily ill-suited to sampling the target distribution; the fact that there was a global limit to how bad any proposal distribution could be allowed us to invoke simultaneous uniform ergodicity of the family of distributions.

F.A.16 Experimental Details on Field Experiment

We provide additional details on the experiment presented in section G.5.3.

**Field distribution.** The $\phi^4$ field model is a popular model used to study phase transition in statistical mechanics (see for example [Berglund et al., 2017]). Here we focus on its $1 - d$ version, where the field is defined on the segment $[0, 1]$, and impose Dirichlet boundary conditions $\phi(0) = \phi(1) = 0$. The energy function is an integral over the segment of two terms:

$$U(\phi) = \int_0^1 \left[ \frac{a}{2} (\partial_s \phi)^2 + \frac{1}{4a} (1 - \phi^2(s))^2 \right] ds. \quad (F.A.212)$$

The coupling term $\frac{a}{2} (\partial_s \phi)^2$ encourages the smoothness of the field, while the local potential term $\frac{1}{4a} (1 - \phi^2(s))^2$ favors fields taking values close to 1 or $-1$ over the segment. For large values of the parameter $a$, fields with significant statistical weights will take values close to 0 over the entire segment $[0, 1]$. As $a$ decreases, the system undergoes a phase transition and two distinct modes form concentrating around either $+1$ or $-1$.

Note that here the energy function eq. (F.A.212) is symmetric under the symmetry $\phi \rightarrow -\phi$. We exploit this symmetry to provide high-quality reference samples in the experiments described next. Note however that as a biasing term is added to the energy, the statistical weights of either of the mode becomes unknown.
The numerical experiments described next shows that the adaptive sampler with normalizing flow proposals can recover the relative statistics thanks to efficient mixing, at least at the level of discretization described. Conversely, the energy barrier between the two modes prevents a Langevin sampler from mixing in a reasonable time.

**Numerics.** We sample the field at 100 equally spaced locations between 0 and 1. The RealNVP flow \cite{Dinh2017} we optimize has 5 pairs of affine coupling layers updating each half of the 100 field variables. The scaling and translation transformations of each coupling layer is a 2-hidden-layer perceptron with relu activations and 100 units per layer.

The algorithm minimizing the “pseudo-likelihood” objective, defined in example 45 follows largely the lines of algorithm 23. Being more specific, we collect the states of 100 parallel walkers every 10 sampling iterations and take a gradient step with the corresponding 1000-sample batch. The initial learning rate of $10^{-3}$ is halved every 5000 gradient steps.

We initialize 100 chains: 20 at the uniform value of 1 and 80 at the uniform value of $-1$. Thanks to the adaptation of the normalizing flow, leading to good acceptance as reported in the main text, these chains can easily mix between modes and recover the proper statistical weights of 50/50.
Figure F.A.2: Samples and proposals in the $\phi^4$ field experiment.
Bibliography


Yan Bai, Gareth Roberts, and Jeffrey Rosenthal. On the containment condition for adaptive

Nils Berglund, Giacomo Di Gesù, and Hendrik Weber. An Eyring–Kramers law for the

M. Betancourt. A general metric for riemannian manifold hamiltonian monte carlo, 12
2012.

M. Betancourt, Simon Byrne, Samuel Livingstone, and Mark Girolami. The geometric

M. J. Betancourt and Mark Girolami. Hamiltonian monte carlo for hierarchical models,
2013.

Michael Betancourt. A conceptual introduction to hamiltonian monte carlo, 2017. URL
pages, 42 figures.

Gerard Biau and Laszlo Gyorfi. On a $L_1$-Test Statistic of Homogeneity. *Bulletin of the
doi.org/10.36045/bbms/1170347810.


James Binney and David Skinner. *The Physics of Quantum Mechanics*. Cappella Archive,
2008.

Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science


Nawaf Bou-Rabee and J. M. Sanz-Serna. Geometric integrators and the hamiltonian monte

Nawaf Bou-Rabee and Jesús María Sanz-Serna. Randomized hamiltonian monte carlo.
http://www.jstor.org/stable/26361544

James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dou-
gal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-
Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy pro-


James A. Brofos and Roy R. Lederman. On numerical considerations for riemannian manifold hamiltonian monte carlo, 2021b.


Bin Qian and Liming Wu. $l^1$-uniqueness of the fokker-planck equation on a riemannian manifold, 2013.


